AWARD NUMBER: W81XWH-19-1-0565

TITLE: Defining and Functionally Characterizing the Epigenome in Lethal Prostate Cancer

PRINCIPAL INVESTIGATOR: Mathew Freedman, MD

CONTRACTING ORGANIZATION: Dana-Farber Cancer Institute

REPORT DATE: OCTOBER 2022

TYPE OF REPORT: ANNUAL

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
OCTOBER 2022	ANNUAL	30SEPT2021 - 29SEPT2022

4. TITLE AND SUBTITLE

Defining and Functionally Characterizing the Epigenome in Lethal Prostate Cancer

5a. CONTRACT NUMBER
W81XWH-19-1-0565

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Mathew Freedman, MD

E-Mail: MFREEDMAN@PARTNERS.ORG

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Dana-Farber Cancer Institute
450 Brookline Ave
Boston, MA 02215

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command

Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT: Prostate cancer (PCa) is dependent on the androgen receptor (AR) at all stages of the disease. The centrality of the clinical role of this hormone-driven transcription factor (TF) in PCa renders it an ideal tumor type in which to study epigenetics. Using chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq) in human radical prostatectomy (RP) specimens, we charted the AR cistrome – the universe of all AR binding sites in the genome. We observed that the AR cistrome undergoes significant alterations during the transition from localized to metastatic disease that are strikingly consistent across patients. This finding underlies our hypothesis that aberrant epigenetic signaling helps drive prostate PCa progression and provides the foundation for a deeper interrogation into the PCa epigenome across disease states in vivo. As part of this project, we have begun to define the genome-wide landscape of active enhancers and open across PCa states. The contents of these maps, in turn, guide screens that will identify regulatory elements associated with treatment resistance and key proteins binding to clinically relevant enhancers. Characterizing changes in the epigenome and its associated transcriptional programs will identify new therapeutic targets as well as biomarkers for therapy response and patient prognostication.

15. SUBJECT TERMS

NONE LISTED

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					USAMRDC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU	25	**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

Page

1. INTRODUCTION:

Prostate cancer (PCa) is dependent on the androgen receptor (AR) at all stages of the disease. The centrality of the clinical role of this hormone-driven transcription factor (TF) in PCa renders it an ideal tumor type in which to study epigenetics. Using chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq) in human radical prostatectomy (RP) specimens, we charted the AR cistrome – the universe of all AR binding sites in the genome. We observed that the AR cistrome undergoes significant alterations during the transition from localized to metastatic disease that are strikingly consistent across patients. This finding underlies our hypothesis that aberrant epigenetic signaling helps drive prostate PCa progression and provides the foundation for a deeper interrogation into the PCa epigenome across disease states in vivo. As part of this project, we have begun to define the genome-wide landscape of active enhancers and open across PCa states. The contents of these maps, in turn, guide screens that will identify regulatory elements associated with treatment resistance and key proteins binding to clinically relevant enhancers. Characterizing changes in the epigenome and its associated transcriptional programs will identify new therapeutic targets as well as biomarkers for therapy response and patient prognostication.

2. KEYWORDS:

Prostate cancer; Epigenetics; transcription factor; enhancers; gene regulation; androgen receptor

3. ACCOMPLISHMENTS:

- **What were the major goals of the project?**
  - The ultimate goals of this proposal are to comprehensively characterize the epigenetic landscape in advanced PCa in order to gain insights into key mechanisms driving lethal, treatment-resistant disease. Ultimately, these intergenic sites can be rationally targeted. The anticipated outcomes are that we will identify **areas of vulnerability (Aims 1)** that are **functionally relevant (Aim 2)** and the **transcription factors** that activate them **(Aim 3)**. The aims of the proposal require specific domains of expertise and to address these points we have assembled an outstanding team with the appropriate scientific depth to go from target identification to analysis of function.
  - Aim 1 will generate the most biologically and clinically informative epigenomic ChIP-seq datasets to date in advanced PCa. Aim 1 will also use RNA-seq, combining transcriptional analysis with ChIP to help link the regulatory elements with their target genes. Aim 2 will utilize the powerful tools of genome editing to identify regulatory elements that are functionally relevant in the development of advanced PCa and Aim 3 will identify candidate transcription factors activating disease-specific enhancers. The outcome of this study will be a compendium of candidate regulatory elements that influence PCa progression.
- **What was accomplished under these goals?**

Updates between last year and this year are denoted with lines flanking the text as seen for this sentence.

Aim 1: To characterize the landscape of open chromatin and active enhancers in the progression from hormone-sensitive Prostate Cancer (PCa) to enzalutamide resistant metastatic castration resistant prostate cancer (mCRPC)

Major Task 1: Collect and assemble 150 metastatic tumors

Subtask 1: From the DFCI Gelb Center and the Netherlands Cancer Institute (NKI), identify and collect metastatic prostate cancer biopsies isolated from 50 men with newly diagnosed metastatic prostate cancer (month 1-24)

Subtask 2: From the DFCI Gelb Center and the Netherlands Cancer Institute (NKI), identify and collect metastatic prostate cancer biopsies isolated from 50 men with metastatic castration-resistant prostate cancer prior to initiation of second-line treatment and 50 resistant to enzalutamide. (month 1-24)

- Major subtasks completed. We successfully collected 30 metastatic prostate cancer biopsies from newly diagnosed patients at the DFCI Gelb Center (subtask 1) and 50 biopsy samples from mCRPC patients at the NKI. These include 20 treatment-naive (10 DFCI, 10 NKI), and 20 enzalutamide resistant (10 DFCI, 10 NKI) cases, as well as 40 men with metastatic castration-resistant prostate cancer prior to initiation of second-line treatment (NKI).
- Over the past year, we have continued collecting metastatic prostate cancer samples under DFCI IRB-approved protocol 09-171. We now have a cohort of over 100 samples from DFCI alone to use for analysis in this project.
- In the first year of the reporting period, we focused on our initial cohort of epigenomes, which included samples collected as part of this aim. These initial results were published in *Nature Genetics* in August 2020. **We generated and analyzed 268 epigenomes in specimens derived from human tissue** (Table 1).

Table 1. Specimens in the project to date, specified by tissue and epigenetic mark

	AR	FOXA1	HOXB13	H3K27Ac	H3K4me2	H3K4me3	H3K27me3	ATAC	All marks
Total	151	81	46	150	8	60	110	10	616
Normal prostate epithelium	13*	14	14	37+	4	3	4	4	93
Primary prostate tumor	31*+88(stello, 2018) + 30 (Linder, 2022)	13+35 (Linder 2022)	13	32 + 35 Linder 2022)	4	7+ 50 (Stelloo, 2018)	7+ 95 (Stelloo, 2018	6	416

mCRPC†	15 + 4 (Severson 2021) +2	15 +4	15 +4	17 +29 (Severson, in prep) + 4 (severson, 2021)	0	0	4 (Severson, 2021)	0	65
Median no. ChIP-seq or ATAC-seq peaks (range)	20,619 (1,577–73,723)	37,691 (3,174 – 99,041)	47,338 (1,709 – 90,075)	34,609 (2,337 – 127,042)	69,558 (41,095 – 83,869)	33,215 (28,952 – 38,447)	25,4148 (112,809 – 316,413)	48,139 (25,324 – 60,232)	
* Includes seven normal prostate and 13 primary tumor AR ChIP libraries published previously – Pomerantz et al, *Nat Genet*, 2015.									
+ Includes H3K27Ac ChIP-seq performed in a specimen derived from human fetal urogenital sinus – Guo et al, *PLoS One*, 2012.									
† ChIP-seq experiments performed using PDXs derived from human mCRPC with the exception of two H3K27Ac ChIP-seq specimens derived from patient mCRPC liver biopsies									

Major Task 2: Epigenetic characterization of metastatic samples

Subtask 1: Chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq) for Histone 3 Lysine 27 acetylation (H3K27Ac). (month 3-24)

Subtask completed. From all 60 mCRPC metastatic samples collected from the NKI (Figure 1), H3K27ac ChIP-seq data was generated. After the biopsy was taken, all patients received AR-blocking enzalutamide treatment and were followed over time for response to treatment. After initial filtering on tumor cell percentage and ChIP-seq data passing QC analyses, we have 29 samples for further downstream analyses. H3K27ac ChIP-seq samples from different patients strongly correlated, with equally-sized groups for ‘responders’, ‘non-responders’, and an ‘intermediate’ group.

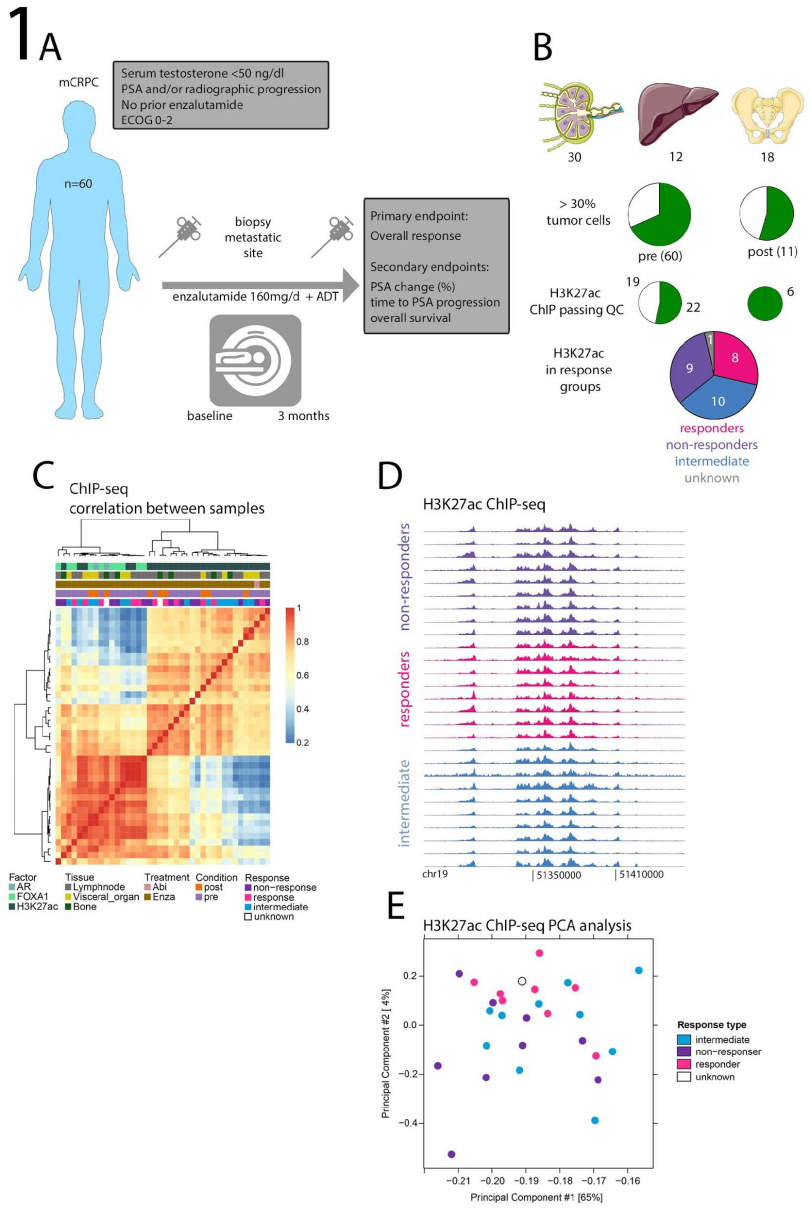


Figure 1: H3K27ac ChIP-seq analyses in mCRPC samples prior and post enzalutamide treatment.
A. Clinical trial design
B. Biopsy sample collection, and filtering steps based on tumor cell percentage and QC.
C. Correlation heatmaps on ChIP-seq data for H3K27ac, AR and FOXA1.
D. Genomebrowser snapshots for H3K27ac ChIP-seq data
E. Principle component analyses for H3K27ac ChIP-seq. Colors indicate response groups.

While overall, H3K27ac profiles were strongly correlating between metastatic samples, a supervised analysis confirmed the presence of 657 H3K27ac sites specifically enriched in the tumors from patients who did not respond to enzalutamide treatment (Figure 2). Based on these data, we hypothesized the 657 active enhancers that demarcate enzalutamide resistance, indicate the acquisition of epigenomic features that render the tumor cell independent on AR action. To test this hypothesis, we analyzed H3K27ac ChIP-seq data from a series of 15 metastatic castration-resistant prostate cancer (mCRPC)-PDX samples, in which the animals were castrated or remained intact. **Also in these PDX models, H3K27ac signal at these 657 enhancers stratified tumors on response to hormonal intervention, in this case castration.**

In our analyses, we also included 4 metastases from one mCRPC patient, collected through rapid autopsy (Severson et al., 2021. https://doi.org/10.1002/1878-0261.12923), where ChIP-seq was performed for H3K27ac, AR, FOXA1 H3K27me3 and CTCF. These data showed a remarkably conserved prostate cancer epigenome between metastases in different organs from within the same patient, and could illustrate little impact of the metastatic site on the epigenome, justifying the combination of different tissues within the same study.

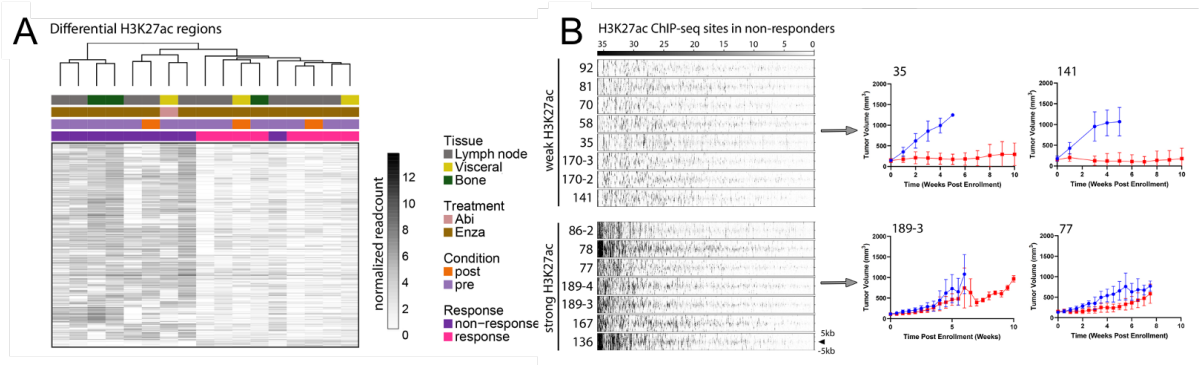


Figure 2: H3K27ac profiling in mCRPC defines enhancers predictive for hormone responsiveness. A. H3K27ac ChIP-seq tracks from mCRPC metastases, from patients before and after Enzalutamide treatment. Distinct H3K27ac profiles were observed that stratified patients on Enzalutamide response.
B. H3K27ac profiles as identified in patients, enable the stratification of mCRPC-patient derived xenografts, on response to castration. Left: grouping of mCRPC PDX, based on H3K27ac. Right: tumor outgrowth, in intact animals (blue) or after castration (red).

Subtask 2: ATAC-seq (month 3-24)

- Due to laboratory shutdowns, we will move forward with ATAC-seq this year.
- In 2020, we published the first datasets generated in this project. We evaluated AR binding in the transition from normal prostate epithelium to localized hormone-sensitive PCa to metastatic castration-resistant disease. **Comparison of the normal prostate, localized hormone-sensitive tumor and metastatic castration resistant PCa (mCRPC) cistromes demonstrated distinct reprogramming of the AR cistrome** (Figure 3). Using a stringent threshold, we identified 17,655 ARBS consistently enriched in the transition from localized PCa to mCRPC (met-ARBS). We similarly performed H3K27Ac ChIP-seq – a mark of active enhancers and promoters – across these clinical states (mCRPC-specific sites are called met-K27ac). Unsupervised principal components analysis of primary tumor versus mCRPC showed clear separation between clinical subtypes (Figure 3). Importantly, genome-wide H3K27Ac in biopsies taken directly from patient mCRPC tumors clustered with the mCRPC PDXs. The majority of met-K27ac peaks overlapped with the met-ARBS peaks (64.9% peak overlap; p-value, <2.2E-16).

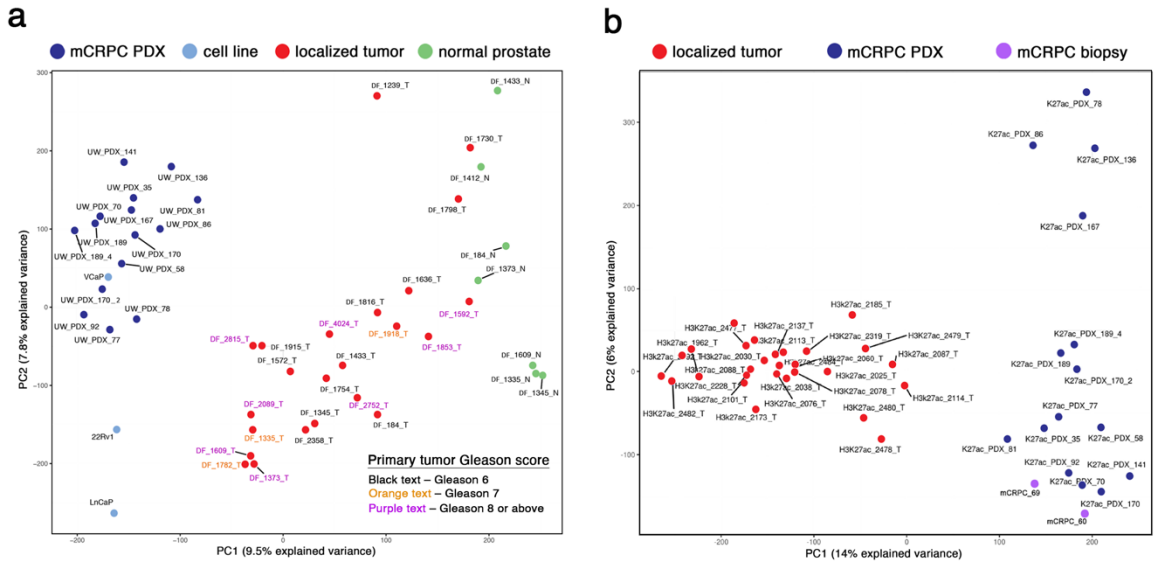


Fig. 3. The AR cistrome and genome-wide H3K27Ac are systematically reprogrammed during prostate cancer progression (a) Principal component analysis (PCA) reveals distinct AR binding patterns across prostate states. Each dot represents the genome-wide AR cistrome in an individual specimen (seven normal prostate epithelium, 23 primary PCa tumors, 15 PDX tumors derived from patient mCRPC, three PCa cell lines derived from metastatic tissue. (b) PCA reveals distinct H3K27Ac binding patterns between primary tumors and mCRPC. Each dot represents genome-wide H3K27Ac signal in an individual subject (24 primary PCa tumors, 15 PDX tumors derived from patient mCRPC, two metastasis specimens biopsied directly from patients with mCRPC).

- To evaluate how well these differential regulatory sites correlate with transcriptional differences, we accessed a publicly available transcriptomics dataset of metastatic prostate versus localized prostate tumor tissue. We rank-ordered differentially expressed genes and then projected onto this distribution the set of transcriptional start sites (TSSs) that contain a met-K27ac site. Transcripts overexpressed in metastases were highly enriched for met-K27ac TSS (p-value, <0.00001; Fig. 4). **The findings demonstrate that newly activated enhancers in mCRPC direct transcript levels of target genes**.

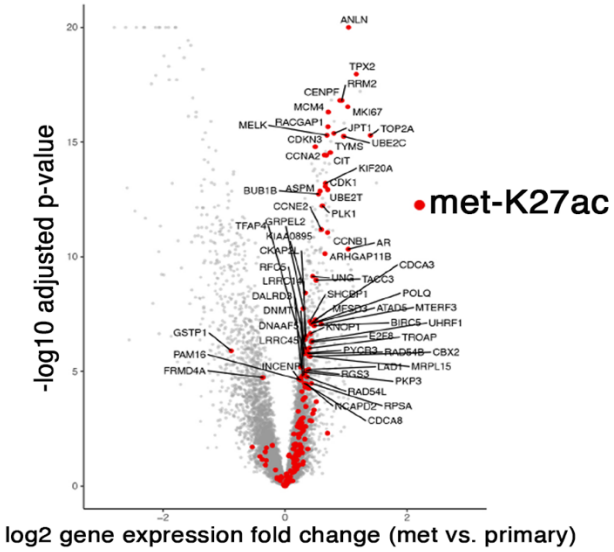


Fig. 4. Newly activated enhancers coincide with genes upregulated in mCRPC. Genes whose expression is upregulated in metastasis compared to primary tumor are enriched for met-K27ac peaks (p-value <0.00001). Each dot represents a gene. Red dots are genes with a met-K27ac in the TSS.

- To test whether other prostate relevant TFs also underwent reprogramming, we performed FOXA1 and HOXB13 ChIP-seq in 14 normal prostate, 13 localized PCa and 15 mCRPC PDX specimens. In stark contrast to AR, the FOXA1 and HOXB13 cistromes demonstrated dramatically less reprogramming during disease progression. Notably, only

306 FOXA1 and 47 HOXB13 peaks were enriched in mCRPC relative to primary disease, compared with 17,655 AR sites. We next focused on the sets of AR sites reprogrammed from normal to primary tumor (n = 9,179, as previously described) and met-ARBS (n = 17,655). Specifically, we evaluated FOXA1 and HOXB13 binding, ATAC-seq, and DNA methylation at these sites. Strikingly, in both normal and primary tumor specimens, FOXA1 and HOXB13 are already present at these 'sentinel' sites where AR is destined to bind (Figure 5). Chromatin was accessible and the DNA was relatively hypomethylated at these loci as well. **The data demonstrate that reprogrammed AR sites during transformation and metastasis are not formed *de novo*, but rather that AR binds to pre-marked, sentinel sites**.

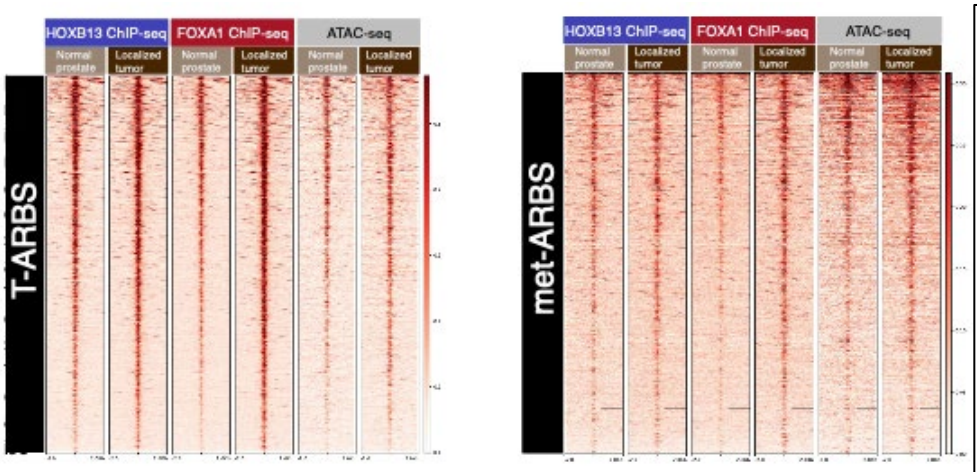


Fig. 5. During PCa progression, AR is re-programmed to pre-marked sites. Heat map indicating transcription factor and ATAC-seq signal intensity in normal prostate epithelium and primary prostate tumor. At left, the 9,179 AR T-ARBS. At right, the 17,655 met-ARBS. Each horizontal line represents a four kilobase (kb) locus. Shade of red reflects average binding intensity at that site across all subjects in the cohort.

- We characterized the TF DNA binding motifs present within met-ARBS, comparing the gained sites to shared AR sites. The most significantly enriched motif associated with met-ARBS was ZEB1 (Zinc Finger E-Box Binding Homeobox 1), a well-described TF involved in mediating epithelial to mesenchymal transition (EMT) in PCa (p=1x10-155)[1,2]. To ascribe putative biological functions to the met-ARBS, the 17,655 met-ARBS were subjected to the Genomic Regions Enrichment of Annotations Tool (GREAT)[3]. Strikingly, the gene ontology (GO) biological processes included "somatic sex determination" (p-value, 1.4 x 10-49), "activation of prostate induction" (p-value, 2.5 x 10-45) and "epithelial cell differentiation involved in prostate gland development" (p-value, 5.0 x 10-20), suggesting that the met-ARBS cistrome is reactivating prostate developmental programs. Similarly, GREAT analysis of met-K27ac revealed multiple GO terms associated with prostate gland organogenesis, such as "epithelial cell maturation involved in prostate gland development" (p-value, 4.9 x 10-38). Next, we investigated similarities between the prostate metastatic epigenome and a large panel of fetal and adult epigenomes. To this end, we assessed the correlation between the set of met-K27ac sites and a series of K27ac epigenomes generated in fetal (N=10 tissue types)[4,5] and adult tissue types (N=27)[4] (Figure 6). The tissues that were most similar to the met-K27ac sites were fetal urogenital sinus (UGS) followed by the fetal tissues most developmentally related to the prostate. sites. **These data indicate that regulatory elements commissioned during prostate cancer progression resurrect prostate-specific fetal tissue developmental programs**. The prostate metastatic epigenomic program is active during development, becomes quiescent in normal prostate and localized prostate tumors, and is reactivated in advanced disease.

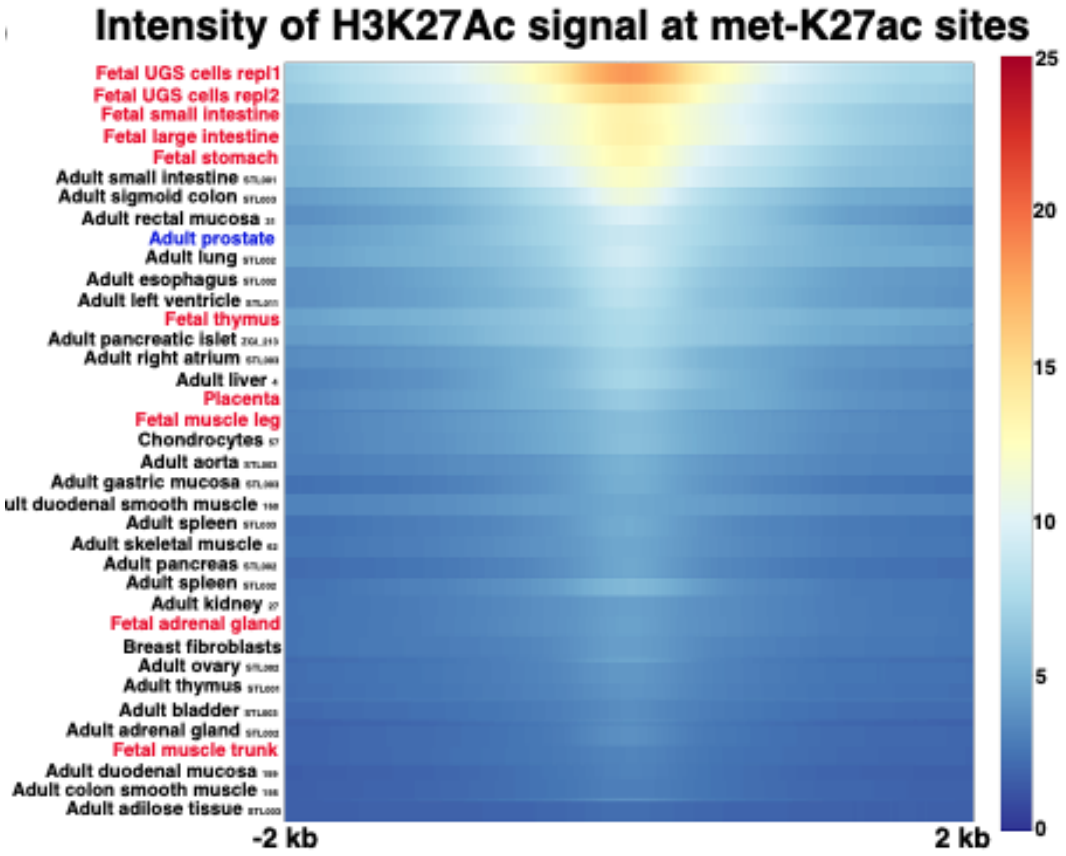


Fig. 6. Regulatory sites activated in mCRPC resurrects prostate developmental programs. Across 37 human adult and fetal cell types, met-K27ac is most strongly associated with fetal urogenital sinus. Cell type listed at left (adult tissues are followed by Roadmap Epigenomics Project identification codes). Urogenital sinus sample was performed in replicate. Heat map indicates H3K27Ac binding intensity met-K27ac sites across a 4 kb interval.

- Previously, we discovered somatic activation of a distal, functionally relevant enhancer that regulates the AR gene[6]. The enhancer region contains recurrent tandem duplications in a whole-genome sequencing (WGS) mCRPC dataset and an H3K27Ac signal that was substantially stronger in mCRPC compared with primary PCa. As part of this project, we ssought to similarly discover other somatically-acquired enhancers in advanced PCa.
- We first intersected the mCRPC-specific H3K27Ac loci with regions containing recurrent structural variants in the WGS dataset from Viswanathan et al, reasoning that recurrent somatic copy number alterations provide a biologically accepted framework for regions under selective pressure[7]. We rank ordered the genomic segments by frequency of overlap between structural variation and met-K27ac sites. Among the top ranked regions were genomic segments containing the genes AR, MYC, FOXA1, HOXB13, and NKX3-1. The genetic regions tended to be large and contained multiple genes. The HOXB13 segment, for example, was 986 kb and contained over 20 genes (Fig. 5).
- To identify enhancer-promoter interactions, we performed H3K27Ac and H3K4me3 HiChIP in LNCaP cells (Fig. 7). Based on looping interaction, co-localization with met-K27ac sites, and recurrence of H3K27Ac signal across a majority of specimens, we prioritized specific candidate enhancers for functional evaluation. Candidate enhancers were functionally evaluated using CRISPR interference (CRISPRi). Site-specific suppression of each putative regulatory element resulted in significantly decreased expression of NKX3-1, HOXB13 and FOXA1. Furthermore, CRISPRi-targeting of each individual enhancer for FOXA1 and HOXB13 decreased LNCaP cell proliferation (Fig. 7).

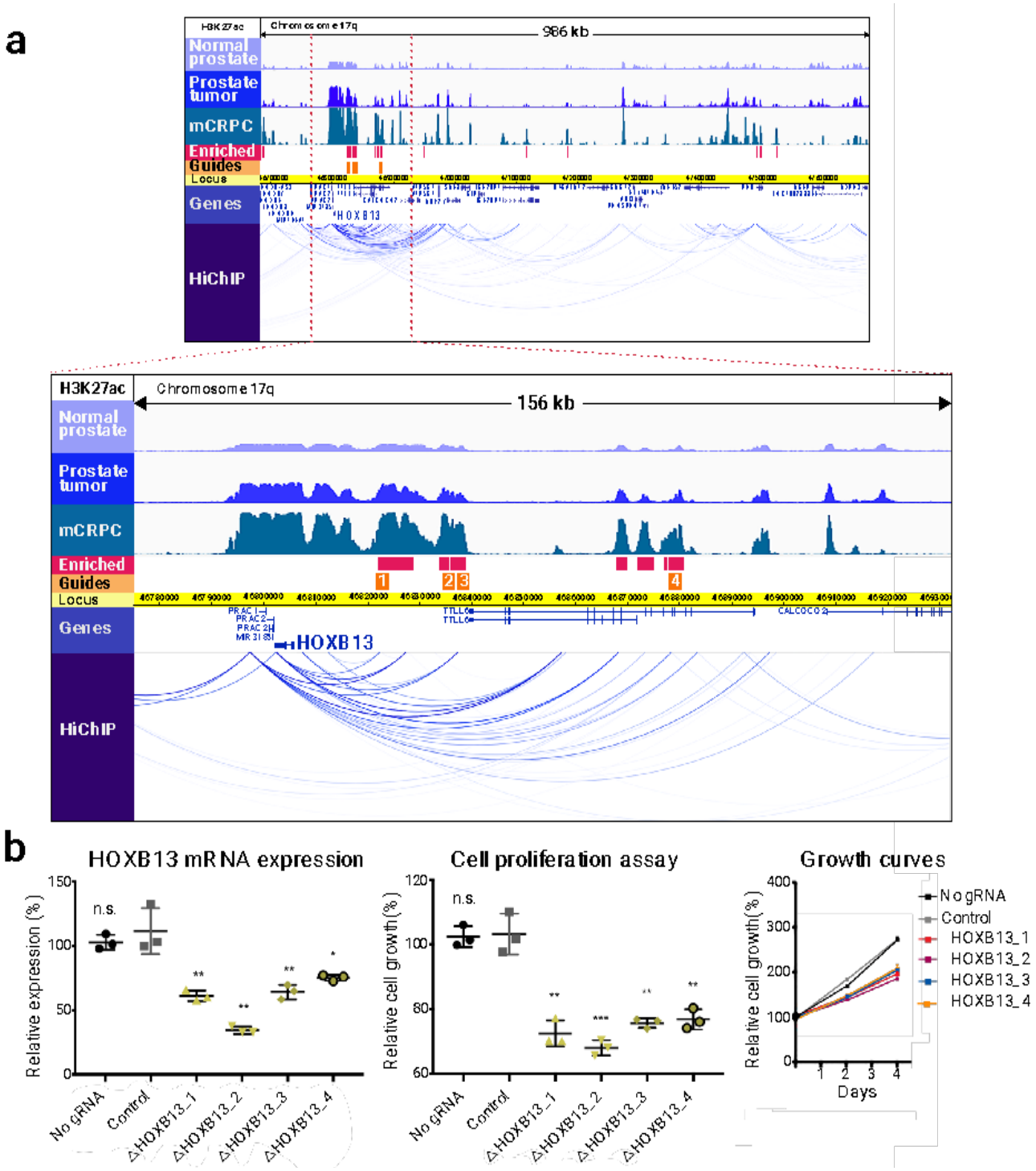


Fig. 7. Functionally relevant mCRPC enhancers are identified by interrogating epigenetic datasets across clinical states. (a) At top, H3K27Ac tracks in a 986 kb region containing HOXB13 identified by integrating ChIP-seq and WGS data, as described. Intensity of ChIP-seq signal was averaged across all DFCI normal prostate, primary prostate tumor and mCRPC specimens, respectively. HiChIP track depicts chromatin looping in the LNCaP cell line. Blue bars show H3K27Ac sites meeting criteria for mCRPC enrichment (met-K27ac). Orange bars depict the locus against which guide RNAs (gRNAs) were designed. Below, magnification of a 156 kb region (bound by red-dotted lines in the upper picture) where met-K27ac and HiChIP signals were strongest. (b) Functional interrogation of candidate metastasis-specific enhancers. Left, LNCaP HOXB13 expression in controls (no gRNA and gRNA targeting unrelated gene HPRT1) and after transduction with each gRNA depicted in (b). Middle and right, LNCaP cell proliferation over the course of four days. Each shape represents an independent experiment, center line indicates mean, error bars indicate ± s.d. Using student's t-test, two-sided – n.s. not significant, *p < 0.05, **p < 0.01, ***p < 0.001.

- **The data indicate that gain of H3K27 acetylation coinciding with somatic DNA amplification identifies *metastasis-specific regulatory elements***. The subsequent Aims

of this project involve deep interrogation of the enhancers discovered using the methodology developed in year 1.

Major Task 3: RNA-seq of metastatic samples. Isolation of RNA, library preparation and Illumina high-throughput sequencing at DFCI (month 3-24)

- For 50 men with metastatic castration-resistant prostate cancer prior to initiation of second-line treatment (NKI), RNA-seq and WGS was performed for 20 samples with sufficient tissue available. For DFCI samples, we will plan to perform RNA-seq in the upcoming year
- We are working with a Dutch consortium (Centre for Personalized Cancer Treatment, sequenced at the Hartwig Medical Foundation) to access whole genome sequencing and RNA-seq on an independent cohort of mCRPC patients.

Specific Aim 2: To perform epigenome-wide Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-based screens across the genome to identify regulatory elements and transcription factors (TFs) associated with enzalutamide resistance in model systems..

Major Task 1: Define functional enhancer landscape for enzalutamide resistance.

Subtask 1: Generate enzalutamide resistant LNCaP clones (month 1-3).

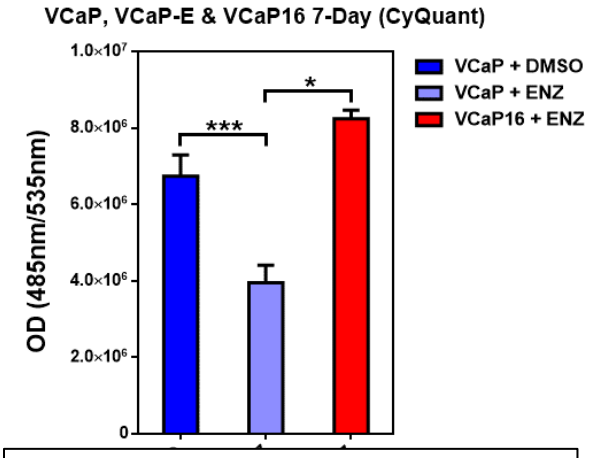


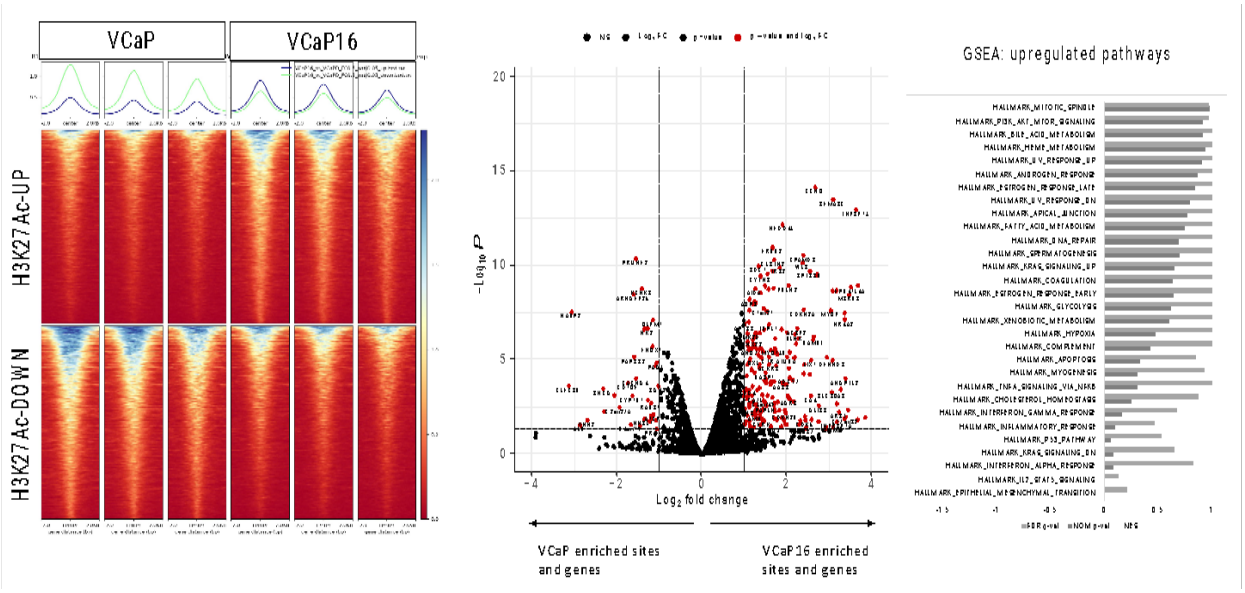
Figure 8: The VCaP-16 enzalutamide resistant cell line proliferates in enzalutamide containing media. Parental VCaP lines are sensitive to enzalutamide (blue bars). VCaP-16 proliferates in the presence of enzautamide (red bar).

Partially completed. Due to COVID with variable access to the laboratory, we were able to generate and comprehensively characterize an enzalutamide resistant clone from the VCaP parental cell line that we term VCaP-16 in collaboration with Dr. Steve Balk's lab. Whereas parental VCaP is sensitive to enzalutamide, VCaP-16 proliferates in enzalutamide containing media similarly to parental VCaP in full media without enzalutamide (Figure 8).

Subtask 2: Perform H3K27 ChIP seq on 5 enzalutamide resistant clones and 5 enzalutamide sensitive clones. (month 3-6).

- H3K27ac ChIP-seq data generated and analyzed in triplicate for VCaP and VCaP-16.
- We performed differential H3K27ac analysis between VCaP and VCaP-16. We identified 6,132 differentially acetylated sites that are more active in VCaP-16 than VCaP and 6,124 sites that are more active in VCaP than VCaP-16 (Figure 9- left panel). These differential epigenetic sites were correlated with RNA-seq data from these two cell lines. The data show a strong correlation between differential epigenetic sites and nearby transcripts (Figure 9 – middle panel). These transcripts are enriched for specific pathways, including epithelial mesenchymal transition (Figure 9 – right panel). These peaks will serve as input for the pooled CRISPR screen.
- Three Enzalutamide resistant prostate cancer cell lines are available and fully annotated in our labs. for the Enzalutamide resistant 42D cell line, 3 replicates of H3K27ac ChIP-

seq has been generated. For the remaining two, e will begin to perform ChIP-seq on them during the next period.



Subtask 3: Conduct pooled CRISPR/Cas9 screen targeting differentially activated enhancers. (month 6-10)

While we were successful in targeting specific AR enhancers (Takeda et al., 2018; Pomerantz et al., 2020), initial CRISPR-Cas9 screening efforts in targeting AR enhancers appeared more challenging. For enhancer screening, we therefore decided to invest in CRISPRi over conventional CRISPR-Cas9, is that the latter requires the presence of an NGG PAM sequence that is found at merely ~25% of AR binding sites. Due to the regional effect of CRISPRi, perturbation of the hormone receptor motif sequence itself is not required to suppress enhancer activity(Pomerantz et al., 2020), and enabling us to cover a larger proportion of all enhancers of interest. Also, as CRISPR technologies are rapidly advancing, new CRISPRi methods are continuously under development with ever-increasing efficiency, which is critical for these high-throughput screens. We are currently re-testing all currently available CRISPRi platforms for enhancer suppression, after which the AR enhancer-suppression screen will be reperformed.

Over the past 1.5 years, we have been working on developing and validating a novel platform to perform high-throughput pooled CRISPRi screening in collaboration with the Nanostring platform. Using Nanostring's single molecule imaging (SMI) platform, we successfully performed a proof-of-concept experiment demonstrating the ability to perform clustered regularly interspaced short palindromic repeats (CRISPR) interference (CRISPRi) using gene expression as a readout. We transduced a pool of guide (gRNAs) against various regulatory elements, including a functionally relevant enhancer of the androgen receptor (AR) that we recently discovered. As a positive control, we designed gRNAs against the HPRT1 gene promoter. We imaged 100,000 cells and measured probes designed against both gRNAs (N=55) and transcripts of interest (N=37) and controls. Thus, 55+37 = 92 objects were measured in every cell.

We evaluated the impact of gRNAs on the HPRT1 gene promoter as well as the AR enhancer. We validated the suppression of HPRT1 transcript levels by qRT-PCR in the bulk population (data not shown). For the SMI platform, we conditioned on cells that carried HPRT1 gRNAs and cells that did NOT carry the HPRT1 gRNAs

and created count distributions for HPRT1 and housekeeping gene expression levels in these cells (Fig. 9). Using SMI, we additionally demonstrated reduction of KLK3, an AR target gene, upon AR enhancer suppression (Fig. 10). This proof-of-concept experiment was performed in prostate cancer cell lines because we had extensive experience with these lines and were using them as positive controls. This novel platform can easily be transitioned to breast cancer cell lines.

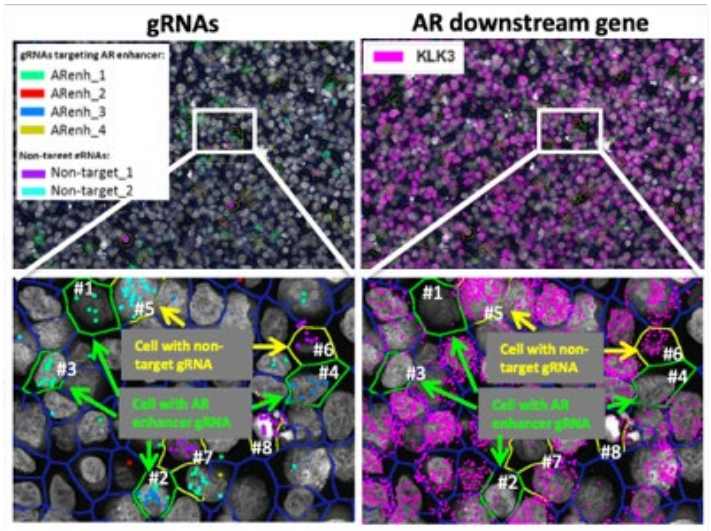


Figure 9: AR enhancer suppression decreases downstream AR target genes. Using Nanostring's single cell SMI platform, we transduced gRNAs against the AR enhancer into the LNCaP PCa cell line. Top row - field of views for gRNA-carrying cells (left), the KLK3 gene (right). It is visually apparent that cells carrying gRNA against the AR enhancer have decreased KLK3 expression. The bar plot quantitates this phenomenon showing the KLK3 expression (normalized to GAPDH) in four cells with gRNA (green) compared to four cells without the gRNA (yellow).

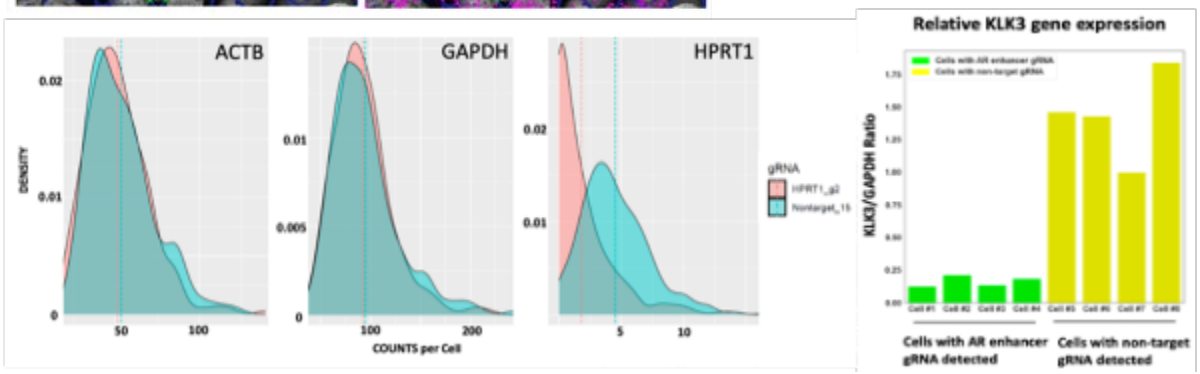


Figure 10: Pooled screening using gene expression as a readout can be performed successfully using the SMI single cell platform. Each plot shows a distribution of HPRT1 gRNA containing cells (orange; N=283) and cells that do not contain HPRT1 gRNAs (green; N=217). In the upper right corner is the transcript that is measured. For the housekeeping genes, ACTB and GAPDH, the distributions are overlapping whereas for HPRT1, the gRNA carrying cells have significantly lower ($P<0.05$) HPRT1 mRNA levels compared to the cells that are not carrying HPRT1 gRNAs.

Subtask 4: Validate gRNAs that score in pooled screen (month 10-12).

- Updates after year 1: postponed due to COVID-19. Subtask will be initiated once the last two enzalutamide resistant clones are available.
- Updates after year 1: postponed due to COVID-19 (see above).

Major Task 2: Identify genes regulated by altered genome (months 12-24)

We identified 17,655 mCRPC-specific AR enhancers in prostate cancer (Pomerantz et al., 2020), a subset of these being not only H3K27ac-positive in this setting, but also amplified in metastatic disease, much analogous to the AR-enhancer we extensively analyzed previously (Takeda et al., 2018). Using H3K27ac HiChIP, we identified the genes that are under direct control of the mCRPC-specific enhancers. Using an ORF-library overexpression screen for induction of castration resistance in vitro (Hwang et al., 2019) individual enhancers that are sufficient to drive resistance in vitro could be identified, and are currently being tested for downstream genetic programs.

Specific Aim 3: To use novel technology (GloPro) to identify the key proteins binding to clinically and functionally relevant enhancers.

Major Task 1: Identify trans-acting factors at the androgen receptor (AR) enhancer

Subtask 1: Optimize guide RNAs (gRNAs) and reaction conditions for localization of dCas9-APEX to AR enhancer in cells

Completed. AR binding sites have been selected based the metastasis-specific enhancers as we identified in our 2020 Nature Genetics paper. For a selection of these, we confirmed direct essentiality of enhancer activity to drive tumor cell proliferation capacity, using CRISPRi.

Subtask 2: Perform streptavidin pulldown and mass spectrometry to identify candidate trans-acting factors

To determine the protein complex composition at the AR enhancer,, we will perform locus-specific proteomics analyses. For this, we will immobilize double-stranded oligonucleotides on beads, and incubate them with nuclear extract from DHT-treated LNCaP prostate cancer cells. After extensive washing to remove unbound proteins, on-bead trypsin digestion is followed by isotope labeling (Nat Protoc. 4, 484-494) and quantitative mass-spectrometry to identify proteins that preferentially bind the regulatory element of interest. As negative control,probes are incubated with lysates from hormone-deprived LNCaPs.These experiments will be performed in close collaboration with mass spectrometry expert Michiel Vermeulen (Radboud University, Nijmegen), who has ample experience in both the required technical and computational expertise to generate and analyse these data.

Subtask 3: Validate binding of individual factors by ChIP-QPCR

Subtask 4: Confirm effect on AR mRNA by suppression of candidate TF by RNAi

- Cell lines with gRNAs and reaction conditions for localization of dCas9-APEX performed in cells
- During the past year, the lead scientist, Dr. Sam Myers, obtained a faculty position at the La Jolla Institute for Immunology. As Dr. Myers started his position in the midst of the pandemic, we are still hopeful that he will remain a collaborator, however we have made contingency plans (see point above with respect to Prof. Dr. Michiel Vermeulen, a recognized expert in this field).

Major Task 2: Identify *trans*-acting factors at enhancer identified by Aims 1 and 2 (months 12-36)

Major Task 3: Corroborate candidate *trans*-acting factors to tumor samples (months 12-36)

- **What opportunities for training and professional development has the project provided?**
  - Over the past year, the project has provided the opportunity to mentor Sylvan Baca, M.D., Ph.D., Talal El Zarif, M.D., two post-docs, and Ji-Heui Seo, Ph.D., a research scientist in the lab of Dr. Freedman. In the group of dr. Zwart at the Netherlands Cancer Institute, three researchers were trained with this project: Yanyun Zhu MSc. , Tesa

Serverson PhD. , and Simon Linder, MSc. Dr. Baca recently secured a highly competitive independent lab position at the Dana-Farber Cancer Institute. He conducted much of the bioinformatics work described above under the mentorship of the PIs on this project. Dr. El Zarif mastered ChIP-seq in human tissue specimens and worked with the bioinformatics team to analyze the data. Dr. Seo was responsible for the development of a novel platform to perform single cell pooled CRISPRi screening. Dr. Severson was promoted to associate staff scientist at the NKI Amsterdam.

- **How were the results disseminated to communities of interest?**
  - The data were presented that the Prostate Cancer Foundation annual Coffey-Holden Conference in 2019. Upon publication of the data, articles describing the findings were published for lay audiences in the *Harvard Gazette* and *Cancer Therapy Advisor.*
  - Upon publication of the data, dr. Zwart was interviewed for the largest Dutch news website (https://www.nu.nl/gezondheid/6066269/dna-handleiding-zorgt-voor-doorbraak-onderzoek-uitzaaiing-prostaatkanker.html) and national radio (https://www.nporadio2.nl/nieuws/29247/onderzoeker-wilbert-zwart-over-nieuwe-ontdekking-op-het-gebied-van-uitgezaaide-kanker).
  - Our 2020 Nature Genetics paper was recognized by the leading journal NEJM as 'clinical implications of basic research' (https://www.nejm.org/doi/full/10.1056/NEJMcibr2030475)
  - Presentation at the Dana-Farber Cancer Institute sponsored Connect:science offering remote seminars available to the entire scientific community during covid.
  - Presentation at the Institute of Oncology Research in Bellinzona, Switzerland

- **What do you plan to do during the next reporting period to accomplish the goals?**
  - Now that we have begun to define the universe of enhancers associated with progression to mCRPC, we will next focus on the epigenetic landscape of enzalutamide-resistant disease. Concurrently, we will functionally characterize the newly discovered non-coding drivers of PCa progression. Our HiChIP-seq data (above) revealed the enhancer landscape in PCa pathogenesis, with both loss and gain of enhancers. We will use these enhancer profiles to perform an in vitro screen in prostate cancer cell lines. To determine the functional significance of these epigenetic alterations, we will perform an unbiased pooled CRISPR/Cas9-based screen to systematically suppress each differentially activated enhancer. Recruitment of a nuclease dead version of Cas9 fused to the Kruppel-associated box repressor domain (dCas9-KRAB) results in loss of H3K27ac and silencing of the enhancer. We will clone a pool of gRNAs targeting each enhancer into a lentiviral expression system. We will use five gRNAs per enhancer targeting the center of each H3K27Ac ChIP-seq peak. We anticipate approximately 4000 enhancers to be differentially activated in tumors versus normal tissue based on our preliminary results, which is well within the range typically done with genome-scale screens involving 20,000 genes and libraries consisting of 100,000 gRNAs. As a result, we will have generated the first comprehensive functional enhancer landscape required for proliferation of primary and metastatic and treatment-resistant PCa.
  - With the identification of the mCRPC-specific cistromes of AR and H3K27ac, computational approaches can be performed to analyze the primary DNA sequence at these specific reprogrammed regulatory regions. Through this approach, candidate transcription factors can be identified that are responsible for driving activity or these mCRPC-specific enhancers. In the next reporting period, we aim to use this approach to identify these potential 'mCRPC-specific acting transcription factors', and perturb their

expression in parental LNCaP, castration resistant derivates and enzalutamide resistant cells, using CRISPR/Cas9 or siRNA approaches.

- Now that we have identified a 'AR-targeted therapy resistance epigenome' based on H3K27ac ChIP-seq analyses, the next step is to validate performance of this signature as response prediction biomarker. Ideally, such a biomarker would be minimally-invasive to limit the impact of samples on the patient, with minimal risk. Therefore, we are analysing MeDIP-seq data from plasma of >100 mCRPC patients, to determine whether our biomarker would be applicable in liquid biopsy samples.
- We will round out our epigenomic dataset by performing ATAC-seq in our set of metastatic samples. These data will provide insights into the accessible chromatin and how or if it changes during important state transitions, such as the development of resistance to anti-androgen therapy.
- We will test regulatory elements that are critical for the acquisition of resistance using a pooled CRISPRi screen. We are excited to move forward with novel technology that we have developed. Using a pooled CRISPRi screen strategy with expression as a readout enables one to interrogate numerous regulatory elements in an cost-efficient manner.

4. **IMPACT:**
   - **What was the impact on the development of the principal discipline(s) of the project?**
     - The epigenome plays foundational roles in prostate cancer development, progression, and treatment. Epigenetic regulatory elements direct the expression or silencing of genes within the vast genome, determining the character and responses of the cell. The self-playing piano is a metaphor for the relationship between the genome and the epigenome. The genome can be considered to be the piano itself, with the potential to play any song. Similarly, every cell in our body contains the same DNA sequence, with the potential to express any of the various sets of genes that characterize each of our highly diverse cell types. The epigenome is represented by the music roll, the spooled sheet whose pattern of holes determines the song to be played. In the cell, epigenomic elements orchestrate a particular gene expression program. Tools to deeply annotate the epigenome in human specimens have only recently matured. Our team has constructed the most extensive library of prostate cancer epigenetic elements to date.
     - DNA sequencing studies across thousands of prostate cancer patients in multiple studies have revealed surprisingly few recurrent genetic mutations. As a result, there are few promising targets in the prostate cancer genome for therapy in a majority of patients. The *epigenome*, on the other hand, is highly dynamic. We have discovered that the prostate epigenome undergoes distinct changes during prostate tumor formation and metastasis. These changes are remarkably consistent across cases. They denote, to our knowledge, the most highly consistent molecular changes in prostate cancer progression. This finding potentially exposes fundamental vulnerabilities in the prostate cell that are universal in the prostate cancer patient population.
     - Our findings also support a long-held hypothesis regarding metastasis formation. We provide striking evidence that prostate cancer cells re-activate the epigenetic paths that the embryonic prostate cell formerly traversed during organ development in the womb. Specifically, the androgen is reprogrammed to sentinel sites in the genome where the mCRPC cell appears to commandeer the regulatory programs of its embryonic ancestors Understanding these epigenomic changes across clinical states presents potential opportunities for clinical translation. For example, the trans-acting factors essential for mCRPC-specific enhancer function may be targeted; or, mCRPC-specific enhancers themselves may be targets for therapy. More fundamentally, as the mechanisms responsible for epigenetic plasticity are better understood, blocking access to latent embryonic programs or "re-reprogramming" the cell to a more differentiated state (e.g., differentiation therapy) may be possible.

- Our findings provide an epigenetics-based biomarker for response to AR-targeted therapy in mCRPC, which would be validated successfully in mCPRC-patient-derived xenografts. The potential implications of these findings are profound, as an 'epigenome-based-biomarker' would allow for tailored treatment selection for mCRPC patients.

- **What was the impact on other disciplines?**
  - Nothing to Report
- **What was the impact on technology transfer?**
  - Nothing to Report.
- **What was the impact on society beyond science and technology?**
  - Through interviews and communication through the lay media, the project had impact on patients diagnosed with metastatic prostate cancer and their families, by communicating the scientific progress has been made for their specific phase of the disease and the international research activities that are being performed to ultimately improve their outcome.

5. **CHANGES/PROBLEMS:**
   - **Changes in approach and reasons for change**
     - The COVID-19 pandemic had a significant impact on our work this year. By the winter/spring 2019-20 we had generated our epigenomics and genomics in human specimens. As we were preparing to perform the functional work outlined in Aim 2 of the project, our laboratories at Dana-Farber Cancer Institute and the National Cancer Institute were closed. Both institutions placed strict restrictions on access to the laboratory for activities unrelated to direct patient care. Our team of investigators gained full access to the laboratory and tools necessary to continue our work by early September. We have resumed the work described above and have begun processing data generated pre-quarantine that have not yet been analyzed.
     - Note that the pandemic affected (and still affects) our Institutional policies including how many people can be at the bench at one time. Therefore, we are operating at ~65% capacity. As can be seen by the data generated, we have done our best to overcome this limitation and will continue to generate and analyze data while adhering to Institute policy.
     - Over the past year, COVID has affected our laboratories, but at a decreased level compared to the prior years. We are operating at ~80% capacity at the moment.
   - **Actual or anticipated problems or delays and actions or plans to resolve them**
     - As discussed in above appropriate sections
   - **Changes that had a significant impact on expenditures**
     - Nothing to Report
   - **Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents**
     - Nothing to Report
   - **Significant changes in use or care of human subjects**
     - Nothing to Report
   - **Significant changes in use or care of vertebrate animals.**
     - Nothing to Report
   - **Significant changes in use of biohazards and/or select agents**
     - Nothing to Report
6. **PRODUCTS:**
   - **Publications, conference papers, and presentations**
     - **Journal publications.**

Nat Genet. 2020 Aug;52(8):790-799. doi: 10.1038/s41588-020-0664-8. Epub 2020 Jul 20.

Prostate cancer reactivates developmental epigenomic programs during metastatic progression

Mark M Pomerantz, Xintao Qiu, Yanyun Zhu, David Y Takeda, Wenting Pan, Sylvan C Baca, Alexander Gusev, Keegan D Korthauer, Tesa M Severson, Gavin Ha, Srinivas R Viswanathan, Ji-Heui Seo, Holly M Nguyen, Baohui Zhang, Bogdan Pasaniuc, Claudia Giambartolomei, Sarah A Alaiwi, Connor A Bell, Edward P O'Connor, Matthew S Chabot, David R Stillman, Rosina Lis, Alba Font-Tello, Lewyn Li, Paloma Cejas, Andries M Bergman, Joyce Sanders, Henk G van der Poel, Simon A Gayther, Kate Lawrenson, Marcos A S Fonseca, Jessica Reddy, Rosario I Corona, Gleb Martovetsky, Brian Egan, Toni Choueiri, Leigh Ellis, Isla P Garraway, Gwo-Shu Mary Lee, Eva Corey, Henry W Long, Wilbert Zwart, Matthew L Freedman

Mol Oncol 2021 Jul;15(7):1942-1955. doi: 10.1002/1878-0261.12923. Epub 2021 Mar 11.

Epigenetic and transcriptional analysis reveals a core transcriptional program conserved in clonal prostate cancer metastases

Tesa M Severson , Yanyun Zhu , Angelo M De Marzo , Tracy Jones , Jonathan W Simons , William G Nelson , Srinivasan Yegnasubramanian , Matthew L Freedman , Lodewyk Wessels , Andries M Bergman , Michael C Haffner , Wilbert Zwart

Nature Communications, in press

Extensive androgen receptor enhancer heterogeneity in primary prostate cancer patients underlies transcriptional diversity and metastatic potential

Jeroen Kneppers , Tesa Severson , Joseph Siefert , Pieter Schol , Stacey Joosten , Ivan Yu , Chia-Chi Huang , Tunc Morova , Umut Altintas , Claudia Giambartolomei , Ji-Heui Seo , Sylvan Baca , Isa Carneiro , Eldon Emberly , Bogdan Pasaniuc , Carmen Jeronimo , Rui Henrique ,Matthew Freedman , Lodewyk Wessels , Nathan Lack , André M. Bergman, Wilbert Zwart

- **Books or other non-periodical, one-time publications.**

  Nothing to Report

- **Other publications, conference papers, and presentations.**

  Nothing to Report

- **Website(s) or other Internet site(s)**

  Nothing to Report

- **Technologies or techniques**

  Nothing to Report

- **Inventions, patent applications, and/or licenses**

  Nothing to Report

- **Other Products**
  The epigenetics data generated by the project thus far has been made publicly available via the NCBI Gene Expression Omnibus upon publication of our manuscript.

7. **PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS**
   - **What individuals have worked on the project?**

8. Name:	Matthew Freedman, M.D.
Project Role:	PI
Researcher Identifier (e.g. ORCID ID):	ORCID ID 0000-0002-0151-1238
Nearest person month worked:	1.2 CM
Contribution to Project:	Dr. Freedman has led overall study design, data generation and data analysis. He directs the database and biobank that provides the materials for the project.
Funding Support:	

9. Name:	Mark Pomerantz, M.D.
Project Role:	PI
Researcher Identifier (e.g. ORCID ID):	ORCID ID 0000-0003-4914-1157
Nearest person month worked:	1.2 CM
Contribution to Project:	Dr. Pomerantz has been involved in overall study design, data generation and data analysis
Funding Support:	

Name:	David Takeda, M.D.
Project Role:	Co-PI
Researcher Identifier (e.g. ORCID ID):	ORCID ID 0000-0002-5986-1169
Nearest person month worked:	0 CM
Contribution to Project:	Dr. Takeda is leading the functional molecular biology experiments
Funding Support:	

Name:	Wilbert Zwart, M.D.
Project Role:	Co-PI
Researcher Identifier (e.g. ORCID ID):	ORCID ID 0000-0002-9823-7289
Nearest person month worked:	1.2 CM
Contribution to Project:	Dr. Zwart has been involved in overall study design, data generation and data analysis
Funding Support:	

- **Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**
- **What other organizations were involved as partners?**
  - The Netherlands Cancer Institute
  - The National Cancer Institute

10. **SPECIAL REPORTING REQUIREMENTS**
    - **COLLABORATIVE AWARDS:** Nothing to Report
    - **QUAD CHARTS:** Nothing to Report
11. **APPENDICES:**

References:

1. Hanrahan K, O'Neill A, Prencipe M, et al. The role of epithelial-mesenchymal transition drivers ZEB1 and ZEB2 in mediating docetaxel-resistant prostate cancer. *Mol Oncol.* 2017;11(3):251-265.
2. Dai Y, Wu Z, Lang C, et al. Copy number gain of ZEB1 mediates a double-negative feedback loop with miR-33a-5p that regulates EMT and bone metastasis of prostate cancer dependent on TGF-beta signaling. *Theranostics.* 2019;9(21):6063-6079.
3. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology.* 2010;28(5):495-501.
4. Roadmap Epigenomics C, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. *Nature.* 2015;518(7539):317-330.
5. Guo C, Liu H, Zhang BH, Cadaneanu RM, Mayle AM, Garraway IP. Epcam, CD44, and CD49f distinguish sphere-forming human prostate basal cells from a subpopulation with predominant tubule initiation capability. *PLoS One.* 2012;7(4):e34219.
6. Takeda DY, Spisak S, Seo JH, et al. A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer. *Cell.* 2018;174(2):422-432 e413.
7. Viswanathan SR, Ha G, Hoff AM, et al. Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. *Cell.* 2018;174(2):433-447 e419.

1. Hanrahan K, O'Neill A, Prencipe M, et al. The role of epithelial-mesenchymal transition drivers ZEB1 and ZEB2 in mediating docetaxel-resistant prostate cancer. *Mol Oncol.* 2017;11(3):251-265.
2. Dai Y, Wu Z, Lang C, et al. Copy number gain of ZEB1 mediates a double-negative feedback loop with miR-33a-5p that regulates EMT and bone metastasis of prostate cancer dependent on TGF-beta signaling. *Theranostics.* 2019;9(21):6063-6079.
3. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology.* 2010;28(5):495-501.
4. Roadmap Epigenomics C, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. *Nature.* 2015;518(7539):317-330.
5. Guo C, Liu H, Zhang BH, Cadaneanu RM, Mayle AM, Garraway IP. Epcam, CD44, and CD49f distinguish sphere-forming human prostate basal cells from a subpopulation with predominant tubule initiation capability. *PLoS One.* 2012;7(4):e34219.
6. Takeda DY, Spisak S, Seo JH, et al. A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer. *Cell.* 2018;174(2):422-432 e413.
7. Viswanathan SR, Ha G, Hoff AM, et al. Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. *Cell.* 2018;174(2):433-447 e419.
1. Hanrahan K, O'Neill A, Prencipe M, et al. The role of epithelial-mesenchymal transition drivers ZEB1 and ZEB2 in mediating docetaxel-resistant prostate cancer. *Mol Oncol.* 2017;11(3):251-265.
2. Dai Y, Wu Z, Lang C, et al. Copy number gain of ZEB1 mediates a double-negative feedback loop with miR-33a-5p that regulates EMT and bone metastasis of prostate cancer dependent on TGF-beta signaling. *Theranostics.* 2019;9(21):6063-6079.
3. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology.* 2010;28(5):495-501.
4. Roadmap Epigenomics C, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. *Nature.* 2015;518(7539):317-330.
5. Guo C, Liu H, Zhang BH, Cadaneanu RM, Mayle AM, Garraway IP. Epcam, CD44, and CD49f distinguish sphere-forming human prostate basal cells from a subpopulation with predominant tubule initiation capability. *PLoS One.* 2012;7(4):e34219.

6. Takeda DY, Spisak S, Seo JH, et al. A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer. *Cell.* 2018;174(2):422-432 e413.
7. Viswanathan SR, Ha G, Hoff AM, et al. Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. *Cell.* 2018;174(2):433-447 e419.