UCSF

UC San Francisco Previously Published Works

Title

An Atlas of Accessible Chromatin in Advanced Prostate Cancer Reveals the Epigenetic
Evolution during Tumor Progression.

Permalink

https://escholarship.org/uc/item/42d651dg

Journal

Cancer Research, 84(18)

ISSN

0008-5472

Authors

Shrestha, Raunak

Chesner, Lisa N

Zhang, Meng

et al.

Publication Date

2024-09-16

DOI

10.1158/0008-5472.can-24-0890

Peer reviewed

An Atlas of Accessible Chromatin in Advanced Prostate Cancer Reveals the Epigenetic Evolution during Tumor Progression

Raunak Shrestha[1,2], Lisa N. Chesner[1,2], Meng Zhang[1,2], Stanley Zhou[3,4], Adam Foye[1,5], Arian Lundberg[1,2,6], Alana S. Weinstein[1,2], Martin Sjöström[1,2], Xiaolin Zhu[1,5], Thaidy Moreno-Rodriguez[1,7], Haolong Li[1,2], SU2C/PCF West Coast Prostate Cancer Dream Team[*], Joshi J. Alumkal[8], Rahul Aggarwal[1,5], Eric J. Small[1,2,5,7], Mathieu Lupien[3,4,9,†], David A. Quigley[1,7,10,†], Felix Y. Feng[1,2,5,7,†]

[1]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA, USA
[2]Department of Radiation Oncology, University of California, San Francisco, San Francisco, CA, USA
[3]Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada
[4]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada
[5]Division of Hematology and Oncology, Department of Medicine, University of California, San Francisco, San Francisco, San Francisco, CA, USA
[6]The Institute of Cancer Research and The Royal Marsden Hospital, London, UK
[7]Department of Urology, University of California, San Francisco, San Francisco, CA, USA
[8]Division of Hematology and Oncology, University of Michigan Rogel Cancer Center, Ann Arbor, MI, USA
[9]Ontario Institute for Cancer Research, Toronto, ON, Canada
[10]Department of Epidemiology & Biostatistics, University of California, San Francisco, San Francisco, CA, USA
[*]A list of members appears in the Supplementary Information
[†]These authors jointly supervised the work

Corresponding Author: Felix Y. Feng, MD, Department of Radiation Oncology, University of California, San Francisco, Box 3110, Room 450, 1450 3rd Street, San Francisco, CA 94158. Phone: 415-502-7222; E-mail: Felix.Feng@ucsf.edu

Running Title: Accessible Chromatin Landscape of Metastatic Prostate Cancer

Conflict of Interest

Abstract

Metastatic castration-resistant prostate cancer (mCRPC) is a lethal disease that resists therapy targeting androgen signaling, the primary driver of prostate cancer. mCRPC resists androgen receptor (AR) inhibitors by amplifying AR signaling or by evolving into therapy-resistant subtypes that do not depend on AR. Elucidation of the epigenetic underpinnings of these subtypes could provide important insights into the drivers of therapy resistance. In this study, we produced chromatin accessibility maps linked to the binding of lineage-specific transcription factors (TF) by performing ATAC sequencing on 70 mCRPC tissue biopsies integrated with transcriptome and whole genome sequencing. mCRPC had a distinct global chromatin accessibility profile linked to AR function. Analysis of TF occupancy across accessible chromatin revealed 203 TFs associated with mCRPC subtypes. Notably, *ZNF263* was identified as a putative prostate cancer TF with a significant impact on gene activity in the double-negative (AR$^-$ neuroendocrine$^-$) subtype, potentially activating *MYC* targets. Overall, this analysis of chromatin accessibility in mCRPC provides valuable insights into epigenetic changes that occur during progression to mCRPC.

Significance

Integration of a large cohort of transcriptome, whole genome, and ATAC-sequencing characterizes the chromatin accessibility changes in advanced prostate cancer and identifies therapy-resistant prostate cancer subtype-specific transcription factors that modulate oncogenic programs.

Introduction

Prostate cancer (PCa) is the second leading cause of cancer-related deaths among men (1). Although PCa is initially responsive to androgen deprivation therapy (ADT), many patients develop resistance and progress to metastatic castrate-resistant prostate cancer (mCRPC). Targeted systemic therapies with second-generation AR-signaling inhibitors (ARSIs), such as abiraterone or enzalutamide, prolong survival and are the standard of care for mCRPC (2–4).

Tumors can develop resistance against ADT and/or ARSI through several distinct mechanisms (5). In most mCRPC, ARSI resistance is achieved through genetic changes that increase AR-signaling (6). Up to 20% of mCRPCs lose complete *AR* dependence and acquire a new cellular phenotype known as treatment-emergent small-cell neuroendocrine (NE) prostate cancer (t-SCNC) or neuroendocrine prostate cancer (NEPC). This AR-NE+ subtype is associated with worse clinical outcomes (7,8). Additional treatment-associated subtypes have been observed, including a double negative subtype (AR-NE-) that bypasses *AR* dependence through FGF/MAPK signaling (9,10) and a double positive subtype (AR+NE+) that gains NE features while maintaining *AR* activity (5). A better understanding of these mCRPC subtypes is foundational for the development of new approaches to overcome resistance.

While *AR* amplification is typical in AR-dependent mCRPCs (11–13) and t-SCNC often harbors *TP53*, *RB1*, and *PTEN* loss (14), other subtypes have no known characteristic genomic alterations. Emerging evidence suggests that epigenetic mechanisms are associated with PCa progression and drug resistance (7,15–17). Specifically, lineage plasticity plays an important role in the development of ARSI resistance (10,18). It has been increasingly recognized that the complex interplay of epigenetic modifications including altered chromatin-binding patterns of transcription factors (TFs), such as *AR* and *FOXA1*, regulate downstream gene activity thereby driving PCa progression (19). Therefore, understanding the chromatin-binding patterns of TFs that are altered in PCa is crucial for the development of effective therapeutic strategies.

The Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) assay has proven to be a very efficient and general epigenetic assay that yields high-quality chromatin signals from small quantities of tissue (20,21). ATAC-seq quantifies chromatin accessibility using transposase enzymes that insert sequencing adapters at accessible chromatin sites. Prior studies (15,21,22) of PCa have mostly used cell line and organoid models, patient-derived xenografts (PDX), or small numbers of tumor tissue biopsies. The Cancer Genome Atlas (TCGA) study performed ATAC-seq on multiple cancer types including 26 localized PCa tumors and revealed cancer-type-specific enrichment of TF binding elements in accessible chromatin regions (21). More recently, using ATAC-seq on CRPC organoids, PDX, and cell lines, Tang et al. (22) identified four mCRPC subtypes and predicted the key TF of each subtype. To the best of our knowledge, the characterization of chromatin accessibility in clinical mCRPC biopsy tissue samples using ATAC-seq has not been conducted to date.

Herein, we describe the first-in-field ATAC-seq study conducted in the largest cohort (n=70) of mCRPC tissue biopsies, to date, from the Stand Up 2 Cancer - Prostate Cancer Foundation West Coast Prostate Cancer Dream Team (WCDT) cohort. Using comprehensive integration of ATAC-seq and RNA-seq from matched tumor samples, we interrogated the changes in chromatin accessibility around regulatory sites to reveal transcriptional regulation associated with mCRPC subtypes. We used computational approaches to produce an exhaustive catalog of TFs that are actively occupied in mCRPC and the transcriptional programs

99   they are predicted to regulate. Finally, we exemplified the use of these new data by characterizing *ZNF263*,
100 a TF previously not associated with PCa biology.

Materials and Methods

Patients and samples
103 Human studies were approved and overseen by the University of California San Francisco Institutional
104 Review Board. All individuals provided written informed consent to obtain fresh tumor biopsies and to
105 perform comprehensive molecular profiling of tumor and germline samples. Fresh-frozen metastatic
106 castration-resistant tissue biopsy samples (n=75) from various anatomic locations representing 69 unique
107 patients (**Supplementary Figure S1** and **Supplementary Table S1**) were collected through a multi-
108 institutional image-guided prospective biopsy trial (NCT02432001) and DNA was extracted as previously
109 described (12,16,17).

ATAC-seq library preparation and high-throughput sequencing
111 The ATAC-seq library preparation was carried out as described in the published method papers by
112 Buenrostro et al. (20) and Corces et al. (23). Briefly, upon thaw, 30 µl of PBS + protease inhibitor was
113 added onto the slide containing the tissue section and subsequently scraped into a 2 ml tube containing 100
114 µl of cold ATAC-Resuspension Buffer (RSB; 0.1% NP40, 0.1% Tween-20, and 0.01% Digitonin) using
115 the tip of a scalpel blade. The sample was incubated on ice for 15 minutes intermittently mixing every 5
116 minutes. After 15 minutes, 1 ml of cold PBS + 0.1% Tween-20 was added into the tube and mixed by
117 inversion, followed by centrifugation at 500 x g for 10 minutes at 4°C. After centrifugation, the supernatant
118 was aspirated, avoiding the pellet containing the cell nuclei in the process. 50 µl of transposition mix [1X:
119 25 µl of 2X TD buffer (20mM Tris-HCl pH 7.6, 10 mM MgCl2, 20% dimethyl formamide), 2.5 µl of
120 transposase, 16.5 µl PBS, 0.5 µl 1% Digitonin, 0.5 µl 10% Tween-20, 5 µl water] was added to the nuclei
121 for resuspension. The reaction was then incubated at 37°C for 30 minutes in a thermomixer with 1000 RPM
122 mixing. After the transposition reaction, the samples were purified using the Qiagen MinElute PCR
123 Purification Kit. Upon elution of the DNA, ATAC-seq libraries were prepared. To minimize PCR biases
124 and duplicates, library preparation was conducted on a real-time qPCR machine, where each sample was
125 pulled off the machine mid-exponential phase. The resulting ATAC-seq libraries were size-selected with
126 Ampure XP beads (Beckman Coulter) for 240-360 bp fragments. Upon successful amplification, an aliquot
127 of the libraries was used for qPCR to calculate the fold enrichment of two accessible chromatin regions
128 over two inaccessible chromatin regions for quality control (i.e. at least 10-fold enrichment). Samples that
129 passed quality control were sequenced on the Illumina NovaSeq 6000 sequencing system.

130 POS (Accessible)
131 *GAPDH* (F) 5'-GCCAATCTCAGTCCCTTCCC-3', (R) 5'-TAGTAGCCGGGCCCTACTTT-3'
132 *KAT6B* (F) 5'-GAAGAGGCGGACCCAGCGGT-3', (R) 5'-TTCCTGCCGGTCATCTCGCTT-3'
133
134 NEG (Closed)
135 *SLC22A3* (F) 5'-GGAGAGGGTGGACAGATTGA-3', (R) 5'-TCAGCCTTGCTGCTACAGTG-3'
136 *QML_93* (F) 5'-CACTGGTTGTCTTTGCAGGA-3', (R) 5'-CCTGGGTCATATTGGGACAC-3'
137

ATAC-seq data processing
139 The ATAC-seq paired-end fastq data was first trimmed to remove the Illumina Nextera adapter sequence
140 using Cutadapt v2.6 (24) with the "-q 10 -m 20" option (**Supplementary Figure S2**). After adapter

141   trimming, FASTQC v0.11.8 (25) was used to evaluate the sequence trimming as well as overall sequence
142   quality. Bowtie2 v2.3.5.1 (26) was then used to align the ATAC-seq reads against the Human reference
143   genome build hg38 using the "--very-sensitive" option. The uniquely mapped reads were obtained in SAM
144   format. Samtools v1.9 (27) was used to convert SAM to BAM file as well as sort the BAM file. Picard tool
145   (https://broadinstitute.github.io/picard) was then used to flag duplicate reads using the MarkDuplicates
146   function with the "REMOVE_DUPLICATES=true" option. The resulting BAM file reads position was
147   then corrected by a constant offset to the read start (positive-stranded +4 bp, negative-stranded -5 bp) using
148   deepTools2 v3.3.2 (28) with the "alignmentSieve –ATACshift" option. This resulted in the final aligned,
149   de-duplicated BAM file that was used in downstream analyses.

150   ATAC-seq peak calling was performed using MACS2 v2.2.5 (29) to obtain narrow peaks with "callpeak -
151   f BAMPE -g hs -qvalue 0.05 --nomodel -B --keep-dup all --call-summits" option. The resulting peaks that
152   mapped to the mitochondrial genome or genomic regions listed in the ENCODE hg38 blacklist
153   (ENCSR636HFF) or peaks that extend beyond the ends of chromosomes were filtered out. The ATAC-seq
154   peaks were annotated with the nearest gene and genomic region where the peak is located using ChIPseeker
155   (30) R-package based on hg38 GENCODE v28 annotations. Possible peak annotations are promoter (±3kb
156   from TSS), exon, 5'UTR, 3'UTR, intron, and distal intergenic.

157   **ATAC-seq quality control**
158   Quality metrics such as Fraction of reads in peak (FRiP) score and fragment length distribution were
159   calculated as described in Corces et al. (21) and Transcription Start Site (TSS) enrichment score was
160   calculated using ATACseqQC version 1.18.1 (31). To ensure the quality of our ATAC-seq dataset, we
161   considered the samples that met the following criteria.

162             $(FRiP\ score > 0.05)\ OR\ (TSS\ Enrich\ score > 8)\ OR\ (ATACseq\ Peak\ counts > 15000)$

163   Five mCRPC samples that failed to satisfy the above criteria were discarded from this study. This resulted
164   in the final set of 70 samples representing 65 unique patients that were used throughout the study.

165   **Consensus ATAC-seq peaks**
166   Non-overlapping unique ATAC-seq narrow peaks regions were obtained from the samples analyzed. Those
167   non-overlapping unique peak regions present in at least two samples were considered consensus peaks.
168   Sequencing reads mapped to the consensus peak regions were counted using the "featurecount" function
169   within Rsubread(32) R-package with the "isPairedEnd=TRUE, countMultiMappingReads=FALSE,
170   maxFragLength=100, autosort=TRUE" option. The read counts of the consensus peaks were normalized
171   with the reciprocal of the size factor and variance-stabilized transform method available in the DESeq2(33)
172   R-package. We note that ATAC-seq peak lengths are highly variable, and so are the lengths of consensus
173   ATAC-seq peaks. Importantly, ATAC-seq read counts tend to be higher for longer peaks which are not
174   corrected using DESeq2. To ensure accurate comparisons of the ATAC-seq peaks, throughout the study
175   comparison is always made between ATAC-seq peaks of the same lengths and never between two peaks of
176   unequal lengths.

177   The ATAC-seq data (read count profiles) of mCRPC samples from our study were combined with those
178   from Tang et al.(22). and adjusted for potential batch effects using the "ComBat" function from the "sva"
179   R-package. Additionally, we attempted to correct potential batch effects in a larger dataset that combined
180   ATAC-seq profiles of benign prostate, localized PCa, mCRPC adenocarcinoma, and t-SCNC/NEPC from
181   different datasets. However, this correction was not possible because some sample phenotypes (covariates)
182   were inseparable from the dataset (batch) they came from.

Differential ATAC-seq and RNA-seq analysis

Differential ATAC-seq and RNA-seq analysis for two groups comparison was conducted using the DESeq2 (33) R-package. The normalized read counts of the consensus peaks were used in the case of the ATAC-seq data. Peaks/genes with Benjamini-Hochberg adjusted $pvalue \leq 0.01$ and $|log_2\ folchange| \geq 1$ were considered statistically significant. For multiple group (3 or more) comparisons of the ATAC-seq dataset, we used the Kruskal-Wallis test and the peaks with $pvalue \leq 0.001$ were considered statistically significant.

Pathway enrichment analysis

To test the association of signaling pathways enriched in the accessible chromatin regions (ATAC-seq peaks), we performed GREAT (34) enrichment analysis using rGREAT R-package. In the case of enrichment analysis of a list of genes, we used a hypergeometric test–based overrepresentation analysis. We used the set of signaling pathways genesets in the Reactome, Hallmark pathway, and GO-Biological Process present in Molecular Signature Database (MSigDB) (35) v7.5.1.

Calculation of AR and NE score

The "singscore" (36) R-package was used to calculate AR and NE scores for each mCRPC sample. The NE score was calculated using the NE genes reported by Beltran et al. (7) and the AR score was calculated using the gene expression profile of the "HALLMARK_ANDROGEN_RESPONSE" geneset from MSigDB (**Supplementary Table S2**).

Transcription factor footprinting

TF footprints were analyzed using Transcription factor Occupancy prediction By Investigation of ATAC-seq Signal (TOBIAS (37)) version 0.12.11. For TF footprinting analysis, we omitted the step of shifting the position of aligned reads in the BAM file in our ATAC-seq data processing pipeline (**Supplementary Figure S2**) as this step was already incorporated within the TOBIAS. We called these BAM files and resulting ATAC-seq peaks, "unshifted". The input data for TOBIAS were prepared as follows. The unshifted ATAC-seq BAM files of all samples within a subtype were merged using the "MergeSamFiles" (Picard) function. A consensus non-overlapping set of unshifted ATAC-seq peaks present in at least two samples in the subtype was generated. A comprehensive list of 541 unique human TFs was compiled by combining TFs from the JASPAR (38) CORE database and refined AR binding motifs (Full Site, Half Site, Lenient Site, and Extended Site) from Wilson et al. (39).

First, the insertion bias of the Tn5 transposase was corrected using the "ATACorrect" function taking the merged unshifted ATAC-seq BAM files and the merged unshifted ATAC-seq peak regions as inputs. The resulting bigWig files were assigned footprinting scores across all accessible chromatin regions using the function "ScoreBigwig". Finally, the scored footprints were matched to the curated list of TF motifs described above, then differential scores for each motif were determined for each subtype comparison using the function "BINDetect" with parameters "--motif-pvalue 0.0001; --bound-pvalue 0.001". TOBIAS categorizes every predicted TF binding site (for each TF motif) into bound and unbound states based on a score threshold per subtype compared. The threshold was set at the level of significance (bound p-value=0.001) of a normal distribution fit to the background distribution of scores.

By utilizing this method, we conducted an analysis of differential TF footprinting. This involved comparing the TF footprints of each mCRPC transcriptional subtype against the others. As a result, we made four comparisons for each subtype, leading to a total of ten comparisons. From each comparison, the subtype-associated TF hits were prioritized based on their differential binding score (s≥|0.1|) and associated q-value

6

225   (top 80% percentile of -log$_{10}$(q-value)). The pairwise differential TF footprint comparison results between
226   mCRPC subtypes were further filtered using their gene expression profiles. TFs not expressed (or with
227   negligible expression level) in relevant subtypes were omitted. These results were visualized in the form of
228   a circularized heatmap. To generate the heatmap, we computed the TF occupancy phenotype score for each
229   mCRPC subtype which is determined by the product of the absolute value of the TF binding score and the
230   absolute value of -log10(qvalue) linked to that specific motif. If a TF is not enriched in a particular subtype,
231   its score is set to 0.

232   For every mCRPC sample, TF footprinting analysis was also conducted in a single sample mode. For this,
233   all TOBIAS TF footprinting functions "ATACorrect", "ScoreBigwig", and "BINDetect" were executed
234   using identical parameters to that of the subtype-level analysis except that only one condition (i.e. respective
235   mCRPC sample) was used in "BINDetect".

Evaluation of the accuracy of TF footprint sites
237   To assess the accuracy of TOBIAS in predicting TF footprint sites, we compared its predictions with the
238   TF-binding sites predicted by ChIP-seq. We utilized publicly available ChIP-seq data for *AR*, *FOXA1*, and
239   *HOXB13* measured in mCRPC from Pomerantz et al. (15). The ChIP-seq TF-binding sites were considered
240   as the reference or the ground truth. The objective was to determine if TF footprints derived from ATAC-
241   seq peak regions could accurately capture the ChIP-seq predicted TF-binding regions.

242   Since, the ChIP-seq experiment on mCRPC was limited to *AR*, *FOXA1*, and *HOXB13*, we restricted our TF
243   footprints evaluation experiment to these three TFs. First, we ran TOBIAS on our ATAC-seq samples using
244   different "--bound-pvalue" (0.5, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001) in the "BINDetect" function.
245   The p-value threshold varies the bound/unbound status assignment for the predicted TF footprint sites. Each
246   run of TOBIAS using different p-value thresholds resulted in a different set of TF footprint sites. These TF
247   footprint sites were compared against those observed in the ChIP-seq TF-binding sites. Overlap of at least
248   1bp between the motif sites from ChIP-seq and ATAC-seq was considered a hit. For each experiment, we
249   computed the predictions' true positive rate (TPR) and false positive rate (FPR). We then generated receiver
250   operating characteristic (ROC) curves and calculated the area under the ROC curve (AUC). We repeated
251   this analysis for every subtype of mCRPC.

Reconstruction of TF-target gene regulatory network
253   The reconstruction of the TF-target gene regulatory network comprises two major steps. (1) Prediction of
254   TF footprints in accessible chromatin using ATAC-seq (linking TF-peaks as described above) and (2)
255   associating ATAC-seq peaks that may potentially regulate expression of individual genes (linking peak-
256   gene). Finally, the TF-peak and peak-gene association results were combined to obtain the TF-gene
257   association.

258   **Peak-to-gene linking predictions**: We used a correlation-based approach to predict potentially causal links
259   between ATAC-seq peaks and gene expression. First, the mCRPC RNA-seq gene expression analysis was
260   restricted to protein-coding genes and filtered out genes not expressed (TPM=0) in more than 25% of all
261   samples. The ATAC-seq peaks were restricted to those with evidence of TF footprints of at least one out of
262   541 TFs measured. Our analysis was further restricted to measure cis-interaction (peak-gene association
263   within the same chromosome). A pairwise Spearman's correlation was measured between every
264   combination of cis-interacting peak-gene pairs across all mCRPC samples. A permutation experiment was
265   performed to evaluate the robustness of our peak-to-gene correlation (see below). The majority of the
266   pairwise peak-gene pairs were random correlations. This warranted an unbiased statistical approach to

267 identify strong correlated or anti-correlated peak-gene pairs. Thus, to identify the most confident set of
268 peak-gene interaction pairs, we interrogated the enrichment of H3K27ac marks on the peak-gene pairs. We
269 used three sets of parameters to scrutinize the peak-gene interaction pairs: Spearman's correlation
270 coefficient (R), the p-value of peak-gene interaction pairs through the permutation experiment, and the
271 distance of the peak to TSS. We used several combinations of values of these three parameters. The peak-
272 gene interaction pairs generated using the parameters, $R \geq |0.4|$, $pvalue \leq 0.05$, and $distance\ to\ TSS \leq$
273 $500kb$ had the highest enrichment of H3K27ac peaks. Thus, we used these parameters to prune off weak
274 peak-gene links.

275 **Permutation experiment**: The gene labels of the RNA-seq gene expression data were randomly permuted,
276 and the pairwise peak-gene Spearman's correlation was measured across all mCRPC samples. This process
277 was repeated 1000 different times. Using the density distribution of all permuted Spearman's correlation
278 coefficients, we calculated p-values for each peak-gene interaction pair.

279 **Reconstruction of the mCRPC subtype-specific regulatory network**: For each mCRPC subtype, we
280 extracted TF associated with the subtype and genes linked to the subtype associated ATAC-seq peaks. This
281 resulted in the TF-target gene links associated with the mCRPC subtype.

282 **Inferring the effects of ZNF263 binding on transcription of downstream target genes**
283 To investigate how *ZNF263* binding influences the expression of the nearest gene in mCRPC, we evaluated
284 the changes in gene expression levels that occurred when *ZNF263* was present or absent in the gene's
285 promoter region. To conduct this analysis, we performed TF footprinting on individual mCRPC samples.
286 For each gene, we grouped the 70 mCRPC samples based on whether *ZNF263* was present or absent in the
287 promoter region of the gene. Genes with at least two samples in each group and at least one sample with
288 expression level log2 TPM > 5 were included for further analysis. Then we performed the Wilcoxon rank
289 sum test and measured the foldchange in the gene expression levels between the two groups. Genes with
290 Benjamini-Hochberg adjusted $pvalue \leq 0.05$ and $|log_2\ folchange| \geq 1$ were considered as statistically
291 significant. The above analysis was also repeated using samples within each subtype (AR+NE-, AR$_{low}$NE-
292 , and AR-NE-). We note that due to limited samples in AR-NE+, AR-NE-, and AR+NE- subtypes, none of
293 the genes were statistically significant.

294 **Tumor Purity Estimation**
295 Tumor purity of the mCRPC samples used in this study was calculated using PURPLE tool
296 (https://github.com/hartwigmedical/hmftools) based on WGS profiles of the corresponding tumor samples.
297 In brief, PURPLE combines B-allele frequency (BAF), read depth ratios, somatic variants, and structural
298 variants to estimate the purity of a tumor sample.

299 **Data Availability**
300 The ATAC sequencing data generated in this study are available from the European Genome-Phenome
301 Archive (EGA) under the accession number EGAS00001006698. The RNA sequencing from the same
302 tumors is available under the accession numbers EGAD00001008487, EGAD00001008991, and
303 EGAD00001009065.

304 The RNA-seq data from the WCDT mCRPC cohort was aligned with STAR and quantified at the gene
305 level for Gencode v28 transcripts as previously described(16). The raw RNA-seq fastq files from PAIR
306 (40) cohort (GSE115414) and CPCG (41) cohort (EGAD00001004424) were processed and analyzed as
307 described above. The mRNA expression data of non-diseased tissues was obtained from the GTEx portal
308 (https://www.gtexportal.org/).

The publicly available ATAC-seq data used in this study were obtained from Pomerantz et al. (15) (Sequence Read Archive (SRA) BioProject accession number PRJNA540151), Cejas et al. (42) (PRJNA691927) and Tang et al. (22) (PRJNA818767). The aligned BAM files from the TCGA localized PCa ATAC-seq data Corces et al. (21) were obtained from the NIH Genomic Data Commons portal (https://portal.gdc.cancer.gov/). These ATAC-seq data were processed and analyzed using the same ATAC-seq data processing pipeline described above.

mCRPC PDX ChIP-seq data for *AR*, *FOXA1*, *HOXB13*, and H3K27ac were obtained from Pomerantz et al. (15) Briefly, Raw ChIP-seq data were downloaded from SRA (PRJNA540151). Reads with base quality scores over 30 across all bases were aligned using bwa-mem v0.7.17 to build hg38. The aligned reads were deduplicated and peaks were called using MACS2 v.2.2.5, with an FDR threshold of 0.01. Peaks in ENCODE hg38 blacklist (ENCSR636HFF) were excluded and only peaks that were enriched at least ten-fold over background were kept for further analysis.

The *ZNF263* ChIP-seq peak genomic regions were obtained from Imbeault et al. (43) (GSE78099), Frietze et al. (44) (GSE19235), and Pope et al. (45) (GSE31477) and *MYC* ChIP-seq peak genomic regions from Barfeld et al. (46) (GSE73994), See et al. (47) (GSE164777), and Guo et al. (48) (GSE157105). The ChIP-seq peak regions were uplifted to hg38 before comparing against the TF footprint regions from mCRPC samples. Further, ChIP-seq profiles of 157 TFs were downloaded from the ChIP-Atlas database (49).

Code Availability
Code used in this manuscript is available at https://github.com/DavidQuigley/WCDT_ATAC_mCRPC

Results

A prospective multi-institution Institutional Review Board-approved study (NCT02432001) obtained fresh-frozen core biopsies of metastases from patients with mCRPC as described previously (12,16,17). To create a rigorous atlas of accessible regulatory DNA elements active in metastatic prostate cancer, we performed ATAC-seq on 70 mCRPC tissue biopsies obtained from various anatomic locations representing 65 unique patients (**Supplementary Figure S1** and **Supplementary Table S1**). The ATAC-seq data was processed using an in-house pipeline (**Methods** and **Supplementary Figure S2**). The sequencing was performed to remarkably high depth, 204-411 million reads (mean 308 million) (**Supplementary Figure S3a**). This resulted in an average of 81,215 (range 22,497-157,071) ATAC-seq peaks per sample marking accessible chromatin regions. We inspected our ATAC-seq dataset using quality control metrics such as the fraction of reads in peak (FRiP) and transcription start site (TSS) enrichment score. Sample sequencing depth was not significantly correlated with the number of peaks detected or any quality metrics. The number of peaks in the sample was significantly correlated with the sample FRiP score (Spearman's correlation coefficient, R=0.79, p-value=5.8×10^{-16}) (**Supplementary Figure S3b-d**). Tumors with higher estimated purity had higher FRiP scores (R=0.4, p-value=4.4×10^{-5}) (**Supplementary Figure S3e**) suggesting that the tumor content in the tissue sample influenced the ATAC-seq sample quality. ATAC-seq peaks have previously been reported to occur most frequently in intronic and distal intergenic regions followed by gene promoter regions, to be enriched at the TSS, and to demonstrate read fragment size periodicity correlated to the integer multiples of the nucleosome (20,21). Our ATAC-seq data were consistent with those reports (**Supplementary Figure S3a, S3f-g**). These findings confirmed the high quality of our ATAC-seq data, consistent with the reports from other groups (20,21).

Chromatin accessibility changes during prostate cancer progression affect stage-specific regulatory elements

To define how chromatin accessibility is altered during prostate cancer progression, we compared the ATAC-seq profile of our mCRPC samples with publicly available ATAC-seq datasets of benign prostate (15) (n=4), localized PCa (21) (n=26), and NE PDX models (42) (n=6 with replicates). The ATAC-seq quality metrics, including the FRiP scores, of our mCRPC samples were comparable to the publicly available datasets (**Supplementary Figure S4**). We first created a consolidated dataset of ATAC-seq peaks by merging the complete dataset. This produced a non-overlapping set of 348,799 consensus genomic regions that were accessible in at least two samples. To test whether this analysis would show systematic differences in ATAC-seq profiles corresponding to tumor stage, we performed principal component analysis (PCA) using the normalized read counts of consensus-accessible regions. Our analysis revealed that mCRPC had distinct chromatin accessibility profiles compared to localized PCa and benign prostate tissue (**Figure 1a** and **Supplementary Figure S5**). Importantly, among the mCRPC cohort, t-SCNC/NEPC samples were found to have distinct chromatin accessibility profiles. Corroborating with the prior observations (15–17), benign prostate and localized PCa had similar chromatin accessibility profiles. A comparison of the accessible chromatin regions identified that mCRPC had the highest number of genomic regions with accessible chromatin conformation (**Figure 1b**). This corresponds to our knowledge of the genome-wide loss of methylation in mCRPC (16,17) and the observation that increased AR expression results in genome-wide chromatin relaxation (50). Accessible ATAC-seq peaks were extensively shared across localized PCa, mCRPC, and t-SCNC/NEPC samples (**Figure 1b**) and were enriched in genes involved in DNA damage repair, apoptosis, and immune system signaling processes (**Supplementary Figure S6**).

Chromatin accessibility is known to significantly affect the transcription of nearby genes (21). We noted there was a robust correlation between the chromatin accessibility of a gene's promoter and the expression of its corresponding gene (mean Spearman's correlation coefficient, R=0.4) (**Supplementary Figure S7a-b**). We found that the differentially expressed genes that were expressed at higher levels were more likely to have accessible chromatin nearby suggesting that changes in chromatin accessibility can contribute to differences in gene expression (**Supplementary Figure S7c**). Motivated by this finding, to identify chromatin variants, the regions of the genome that differ in chromatin accessibility (i.e., differentially accessible ATAC-seq peaks), between various stages of PCa progression, we performed a differential accessibility analysis comparing each stage. We further annotated these peaks with the nearest gene and genomic region where the peak is located (**Methods**). We identified 76,311 (21.9%) peaks that exhibit a significant change in accessibility in at least one stage (**Figure 1c**). The majority of these peaks were exclusively detected in mCRPC including the t-SCNC samples (**Figure 1b-c**). These peaks included promoters of prostate cancer-relevant genes such as *AR*, but most chromatin variants were observed in introns and distal intergenic regions rather than in promoters (**Figure 1d**). ATAC-seq peaks in distal intergenic regions of genes such as *AR*, *CHGA*, *DNMT3A,* and *PIK3R1* were exclusively detected in mCRPC samples. To assess whether chromatin variants were more likely to harbor regulatory DNA, we intersected these regions with measures of H3K27ac, a histone mark associated with active enhancers, in mCRPC (15). We observed increased chromatin accessibility of intergenic regions, introns, and promoters within regions of the H3K27ac signal (**Figure 1e**). We predicted the gene pathways activated by these chromatin variants using the GREAT enrichment tool (34) and found that the chromatin regions exclusively accessible in mCRPC-adeno were enriched in the AR signaling pathway, cell migration, and prostate

10

development processes (**Figure 1f**). Accessible chromatin regions in t-SCNC samples were enriched in neuronal development and differentiation processes. Similarly, chromatin regions accessible in localized PCa were enriched in lipid biosynthetic and metabolic processes and changes in cytoskeleton organization. These results imply that the chromatin variants associated with tumor progression potentially deregulate the oncogenic signaling required for malignant transformation.

Chromatin accessibility in mCRPC is associated with subtypes linked to androgen signaling

To investigate the global patterns of chromatin accessibility in mCRPC, we performed unsupervised hierarchical clustering on ATAC-seq profiles from 70 mCRPC samples. This was achieved by applying pairwise Spearman's correlation to the normalized read counts of consensus accessible peaks. Our analysis revealed three distinct clusters of mCRPC samples (**Figure 2a** and **Supplementary Figure S8**). These clusters were not associated with metastatic tissue sites or alterations of driver genes such as *PTEN*, or *RB1*. Cluster assignments were, however, correlated with tumor AR-pathway and NE scores assessed by gene expression signature analysis (**Methods**, **Figure 2b-c, and Supplementary Table S2**). Tumors in cluster 3 had significantly higher AR scores and lower NE scores (Wilcoxon rank sum test p-value: 3×10^{-8} and 4×10^{-4}), while tumors in cluster 1 had low AR scores but high NE scores (p-value: 6.8×10^{-5} and 5×10^{-4}). Furthermore, leveraging AR and NEPC gene signatures from Labrecque et al. (5), we recently stratified our mCRPC samples into 5 subtypes (AR+NE, AR_{low}NE-, AR+NE+, AR-NE+, and AR-NE-) based on their RNA-seq gene expression profiles (51). Out of 70 mCRPCs analyzed in this study, 26 tumors were classified as AR+NE-, 32 tumors as AR_{low}NE-, 2 tumors as AR+NE+, 4 tumors as AR-NE+, and 6 tumors as AR-NE- (**Supplementary Table S1**). Cluster 1 was associated with the AR-NE+ subtype (Fisher's Exact Test p-value=0.003) whereas cluster 3 was associated with AR+NE- subtype (p-value=0.005). Cluster 2 had mixed sample phenotypes, with almost 58% of the samples in the group associated with AR_{low}NE- subtypes and 25% of the samples associated with AR+NE- subtype. We further observed that clusters 2 and 3 each contained two sub-clusters designated as 2A and 2B, and 3A and 3B respectively. These sub-clusters were primarily distinguished by variations at the tissue level. Notably, both sub-clusters 2B (p-value=0.002) and 3A (p-value=0.002) were enriched with bone metastatic tissue biopsies. A recent study of chromatin availability in cell lines and organoid models by Tang et al. (22) reported the existence of four ATAC-seq subtypes of CRPC: CRPC-AR, CRPC-NE, CRPC-WNT, and stem cell-like (CRPC-SCL) subtypes. To compare these subtypes with our cohort of mCRPC samples, we analyzed ATAC-seq data from both studies together. Unsupervised hierarchical clustering of ATAC-seq profiles showed that the majority of samples in CRPC-AR and AR+NE- subtypes were a part of the same cluster (**Supplementary Figure S9**). Similarly, samples in CRPC-NE and CRPC-WNT subtypes clustered together with the AR-NE+ subtype. Samples in the CRPC-SCL subtype were clustered with AR_{low}NE- and AR-NE- subtypes. These results indicate that the chromatin accessibility in our mCRPC data was linked most strongly with mCRPC transcriptional subtypes.

mCRPC transcriptional subtypes are associated with chromatin variants of prostate cancer signaling pathways

After establishing that androgen signaling was significantly associated with chromatin accessibility status genome-wide, we performed a supervised analysis to identify chromatin loci whose accessibility status was correlated with the five mCRPC transcriptional subtypes from Labrecque et al. (5). This analysis was motivated by our observation that chromatin accessibility proximal to the genes that defined mCRPC transcriptional subtype signatures were correlated with gene expression of the corresponding gene

434 (**Supplementary Figure S10**). Moreover, there was no significant difference between the numbers of peaks
435 and FRiP score among the mCRPC transcriptional subtypes (**Supplementary Figure S11**).

436 We, therefore, tested for differential chromatin accessibility across the five mCRPC subtypes. This analysis
437 identified 6704 ATAC-seq peaks with significant differences in accessibility across all samples in each of
438 the subtypes (**Supplementary Table S3**). AR+NE- tumors most frequently harbored increased
439 accessibility, followed by $AR_{low}NE$- tumors (**Figure 3a**), in PCa-associated genes such as *AR*, *KLK3*,
440 *FOXA1*, *NKX3-1*, *SPOP*, *ZBTB16*, and *NCOA2*. Similarly, regions around several epigenetic drivers of
441 prostate cancer such as *ARID1A*, *SMARCA1*, *KMT2D*, and *KDM6A* were more accessible in AR+NE- and
442 $AR_{low}NE$- tumors compared to the remaining subtypes. Interestingly, most chromatin variants between the
443 mCRPC subtypes were annotated distal to the TSSs (i.e. distal intergenic peaks); about 75% of chromatin
444 variants were located more than 3kb from TSS (**Supplementary Figure S12a**). These regions may
445 represent enhancer regions active in specific subtypes of the disease. One such example was a distal
446 accessible region upstream of the AR that we and others have previously identified as a driver of ADT and
447 ARSI resistance in AR-positive disease (12,13). We additionally identified chromatin variants in distal
448 regions near *NKX3-1* in AR+NE- and $AR_{low}NE$- subtypes, and *GPR37L1* in AR+NE- subtype as compared
449 to the remaining subtypes (**Figure 3b**). *GPR37L1* encodes for a G protein-coupled receptor protein almost
450 exclusively expressed in the nervous system, and studies in murine models have suggested it to have a
451 neuroprotective function (52). *AR* and *NKX3-1* mRNA were expressed in AR+NE- and $AR_{low}NE$- subtypes
452 and *GPR37L1* mRNA was highly expressed in AR-NE+ subtypes (**Supplementary Figure S12b-d**).
453 Moreover, these chromatin variants in *AR*, *NKX3-1*, and *GPR37L1* were correlated with their corresponding
454 gene expression (**Supplementary Figure S12e-g**).

455 To gain insight into the functional role of accessible chromatin, we extracted the chromatin variants across
456 all mCRPC transcriptional subtypes that mapped to the promoter, intron, and distal intergenic regions
457 (**Figure 3c**). Chromatin variants in the promoter, intron, and distal intergenic regions were enriched for the
458 Hallmark Androgen Response pathway (**Figure 3d-f**). Chromatin regions mapped to intron and intergenic
459 were also enriched in oncogenic and proliferative signaling pathways. This underscores the functional
460 importance of intronic and distal intergenic chromatin variants in potentially modulating the epigenetic
461 landscape of mCRPC transcriptional subtypes.

mCRPC transcriptional subtypes are defined by DNA accessibility-guided patterns of transcription factor regulation

464 Following our observation that the regulatory elements are enriched in the accessible chromatin region in
465 mCRPC, we built a comprehensive catalog of transcription factors (TF) occupancy across mCRPC
466 transcriptional subtypes. TF binding to chromatin prevents Tn5 cleavage within the binding site and
467 generates depletion in ATAC-seq coverage known as "TF footprints" (53). We interrogated TF footprint
468 signals in the accessible chromatin of mCRPCs using the TOBIAS (37) software tool. First, to evaluate the
469 reliability of the predicted TF footprints, we compared the *AR*, *FOXA1*, and *HOXB13* footprints predicted
470 in our mCRPC samples against the respective ChIP-seq binding sites observed in mCRPC PDX samples
471 obtained from Pomerantz et al. (15) (**Methods**). Our analysis found that TOBIAS-predicted TF footprint
472 sites were correlated to experimentally observed ChIP-seq binding sites (**Supplementary Figure S13a-f**).
473 For each TF, TOBIAS classifies every predicted TF binding site as a bound or unbound state based on a
474 user-defined footprint score threshold (**Methods**). The bound TF footprint sites were observed to have a
475 depleted ATAC-seq accessibility signal as compared to the unbound sites (**Figure 4a-b**). Moreover, a

476 significant proportion (82-96%) of the predicted TF-bound footprint sites were found to coincide with the
477 regions identified by ChIP-seq peaks as compared to the unbound sites (**Supplementary Figure S13g-i**).
478 This suggests that our predictions of TF footprints are strong and reliable.

479 Next, we performed an unbiased genome-wide active TF occupancy analysis using 541 human TFs
480 (**Supplementary Figure S14**) from the JASPAR (54) database and examined their association with
481 different mCRPC transcriptional subtypes. We performed an unbiased differential TF footprint occupancy
482 analysis comparing each mCRPC transcriptional subtype to the other (**Figure 4c-d** and **Supplementary
483 Figure S15**). Subtype-associated TF hits were prioritized based on their differential binding score and
484 associated probability score (**Methods** and **Supplementary Table S4**). For example, *AR*, *FOXA1*,
485 *HOXB13*, and *NR3C1* were identified to have differential TF footprint occupancy in AR+NE- compared to
486 AR-NE+ or AR-NE- subtypes (**Figure 4c-d** and **Supplementary Table S4**). Similarly, TFs such as *ASCL1*,
487 *NEUROD1*, *SNAI2*, *ID4*, and *NKX3-2* were identified to have differential TF footprint occupancy in AR-
488 NE+ compared to AR+NE- subtype. AR-NE- tumors were enriched for high TF footprint occupancy of
489 *SP1*, *SP2*, *ZNF263*, and *KLF5* (**Figure 4d**).

490 We aggregated the results of significantly differential occupied TFs from all ten pairwise comparisons
491 between the mCRPC subtypes and identified 203 TFs enriched in distinct mCRPC subtypes (**Figure 4e** and
492 **Supplementary Table S5**). To further validate our predictions through computational analysis, we
493 compared our predicted TF footprints to publicly available data on binding sites for 120 TFs obtained using
494 ChIP-seq (**Supplementary Figure S16**). Our analysis revealed that for 94 of these factors, at least 25% of
495 our predicted footprint locations overlapped with the ChIP-seq binding sites. For TFs known to be
496 associated with prostate cancer, such as *FOXA1*, *AR*, *ERG*, *HOXB13*, *NR3C1*, *ASCL1*, and *GATA2*, the
497 overlap between our predicted footprints and the ChIP-seq data was even greater, with over 80%
498 concordance.

499 We observed that TFs such as *AR*, *FOXA1*, *HOXB13*, *GATA2*, *SP1*, *SP2*, and *KLF5* were enriched in
500 multiple subtypes. Eighty-four TFs were associated with both AR+NE- and AR$_{low}$NE- subtypes indicating
501 similarity in gene transcription regulation between these two subtypes. About 80% (20 of 25) of AR-NE-
502 TFs were also associated with the AR$_{low}$NE- subtype corroborating previous (5) observations of
503 transcriptomic similarity between AR$_{low}$NE- and AR-NE-. Interestingly, each subtype of mCRPC shows a
504 distinct affinity towards certain subtype-specific TFs. For example, a large set of 49 TFs, comprising
505 *ASCL1*, *NEUROD1*, *NKX3-2*, and *POU3F2* was found to be exclusively associated with the AR-NE+
506 subtype. *ASCL1* is a pro-neural TF that acts as a driver of the neuronal transcriptional program to support
507 treatment resistance in the AR-NE+ subtype (42,55). Similarly, *POU3F2* is a neural TF that is directly
508 suppressed by *AR* and mediates NE differentiation and treatment resistance in the AR-NE+ subtype (56).
509 *NEUROD1* is a neuronal TF associated with neuronal development in both NEPC and small-cell lung
510 cancer (SCLC) (42). Our results provide a unique opportunity to comprehensively interrogate several TFs
511 associated with mCRPCs.

512 Knowing where TFs bind on the genome is important because it can provide valuable insights into the gene
513 expression regulatory mechanisms. Therefore, we examined the locations of all predicted TF footprints in
514 the genome. We observed that in all TFs, the footprints were identified in both the promoter and distal to
515 the promoter regions of the nearest gene (**Supplementary Figure S17**). In 72% (147 of 203) of the TFs,
516 the majority of the footprints were observed at a distance of more than 3kb from the nearest gene, whereas,
517 in the remaining 28% (56 of 203) of the TFs, the footprints were observed within the promoter region (i.e.,
518 TSS±3kb). Intriguingly, we found that all TFs associated with AR-NE- were preferentially localized in the

519   promoter region. This finding was corroborated by analyzing independent publicly available ChIP-seq
520   profiles of TFs associated with the AR-NE- subtype (**Supplementary Figure S18**). Since the promoters in
521   humans are enriched for the CpG dinucleotide, TFs such as *SP1*, *NRF1*, *ETS*, and many C2H2 zinc finger
522   proteins including KLF-family proteins and *ZNF263* are known to preferentially bind proximal promoter
523   DNA sequences (44,57). Thus, ATAC-seq-based TF footprinting allows us to examine the preferred
524   binding regions of a vast number of TFs.

Identification of the influential transcription factors in mCRPC transcriptional subtypes

526   To identify the functional impact of mCRPC subtype-specific TF binding patterns, we constructed unbiased
527   gene expression networks associated with each TF binding event. We integrated ATAC-seq and matched
528   RNA-seq data from our mCRPC cohort to comprehensively identify correlated or anti-correlated genome-
529   wide cis-interacting peak-gene pairs (**Methods** and **Supplementary Figure S19**). We identified a set of
530   37,865 robust peak-gene pairs (36,616 correlated and 1,249 anti-correlated pairs) consisting of 23,089
531   unique peaks and 7,710 unique genes (**Supplementary Figure S20a-c**). On average, 1.64 (min=1, max=29)
532   genes were associated with a peak (**Supplementary Figure S20d**). Approximately 98% of peaks were
533   correlated with 5 genes or less and 66% of peaks were associated with only 1 gene. Similarly, on average,
534   5 (min=1, max=108) peaks were associated with a gene (**Supplementary Figure S20e**). About 76% of
535   genes were associated with at most 5 peaks and 35% of genes were associated with only 1 peak. Most
536   correlated peaks were proximal to the TSS region of a gene, as compared to its distal region
537   (**Supplementary Figure S20f**). These mCRPC TF network characteristics are consistent with the results
538   obtained from TF networks derived from ATAC-seq in various types of cancer(21,22). We further
539   integrated the TF footprint sites identified in the accessible peak to the genes correlated with the peak to
540   construct a TF-target gene regulatory network and identify the target genes of TFs. Restricting our analysis
541   to the 203 mCRPC-associated TFs, the regulatory network represented 22,608 unique peaks and 7,632
542   unique genes. Based on the ATAC-seq peaks observed in individual mCRPC subtypes and the associated
543   TFs, we derived the mCRPC subtype-specific regulatory network.

544   Taking advantage of these dense regulatory networks, we assessed if we could identify influential TFs
545   potentially regulating the transcriptional programs and driving the mCRPC subtypes. We hypothesized that
546   highly influential TFs regulate the transcriptional activity of a large fraction of the downstream target genes.
547   Thus, we computed the node degree of the TFs in mCRPC subtype-specific regulatory networks (**Figure 5**
548   and **Supplementary Table S6**). We found that several of our top-influential TF hits were well-established
549   drivers of PCa. *FOXA1* and several FOX-family TFs including *FOXC2*, *AR*, *HOXB13*, *GRHL2*, and *SRY*
550   were predicted as the top influential TFs in the AR+NE- subtype. *FOXA1* and *AR* are well-established as
551   drivers of mCRPC (11,12,16,17,58,59). The top influential TFs in the AR-NE+ subtype such as *ASCL1*
552   (42,55), *NEUROD1* (42), *ZEB1* (60), *TCF4* (61), and *SNAI2* (62) are known to promote neuroendocrine
553   differentiation in PCa. The AR+NE+ subtype was enriched with influential TFs found in both AR+ and
554   NE+ subtypes. The Stripe family (63) of TFs including *SP1*, *SP2*, and *KLF5* (64) were predicted to drive
555   both AR$_{low}$NE- and AR-NE- subtypes. This suggests that our analysis nominated consistent TFs hits
556   associated with PCa. Furthermore, this analysis identified several influential TFs that were not previously
557   linked to PCa, including *ZNF263* and *RREB1* in AR-NE- subtype, *ZNF384* and *CDX1* in AR+NE- subtype,
558   and *BACH2* and *ZBTB18* in the AR-NE+ subtype. To explore the regulatory impact of the hits identified
559   through TF footprinting analysis on their target genes, we opted to investigate *ZNF263* since it has not been
560   studied in the context of mCRPC.

ZNF263 regulates the MYC signaling pathway in mCRPC

ZNF263 has been implicated in modulating oncogenic signaling in cancers. For example, *ZNF263* is the most significant TF bound to the endoplasmic reticulum stress-specific super-enhancer and is highly expressed in hepatocellular carcinoma (HCC) (65). *ZNF263* knockdown in HCC cell lines leads to reduced proliferation, apoptosis resistance, and chemoresistance (65). *ZNF263* enhances *EGFR* signaling and the progression of glioblastoma (66). Despite having a KRAB domain that typically facilitates transcriptional repression, *ZNF263* can exert both positive and negative impacts on the transcriptional regulation of the genes it targets (44). However, *ZNF263* has not been studied in the context of prostate cancer.

First, to establish the relevance of *ZNF263* in prostate cancer, we interrogated its mRNA expression in several publicly available datasets. Our analysis of gene expression profiles of non-diseased tissues indicated that *ZNF263* is highly expressed in prostate tissues (**Supplementary Figure 21a**). Moreover, we found elevated expression levels of *ZNF263* in mCRPC as compared to both benign prostate and localized PCa tissue (**Supplementary Figure 21b-c**). Notably, *ZNF263* was expressed in all mCRPC subtypes and there was no difference in expression levels between the subtypes (**Supplementary Figure 21d**). Next, to inspect the reproducibility of our predicted *ZNF263* footprint sites, we compared predicted binding sites in mCRPC to previously published *ZNF263* ChIP-seq profiles measured on human embryonic stem cells and erythroblast cells. We hypothesized that although some *ZNF263* binding sites would be unique to mCRPC, many binding sites would be invariant among cell types, and demonstrating non-random enrichment for experimentally identified binding sites would support the validity of our computational analysis. Indeed, we found that 6-16% of *ZNF263* bound footprint sites were also observed in the ChIP-seq peaks (hypergeometric test p-value $< 2.2 \times 10^{-16}$) (**Supplementary Figure S22a**). Importantly, sites where we predicted TF binding in mCRPC had a greater degree of overlap with the ChIP-seq peaks than sites predicted to be unbound in mCRPC, supporting the robustness of our footprint predictions.

To identify targets of *ZNF263* transcriptional regulation in mCRPC, we assessed variation in gene expression levels in the presence or absence of *ZNF263* binding. As described in the previous section, since *ZNF263* has a strong affinity for binding to the promoter region (**Supplementary Figure S17 and S18**), we focused our analysis on *ZNF263* footprints identified in this region. For each gene, we grouped the mCRPC samples based on whether *ZNF263* was present or absent in the promoter region of the gene. Then we measured the foldchange in the gene expression levels between the two groups. A significant increase in gene expression levels was observed for most genes when *ZNF263* was bound to their promoter region (**Figure 6a**). We conducted an unbiased enrichment analysis to identify common functions of genes whose expression increased when *ZNF263* was bound and identified enrichment in the MYC signaling pathways (**Supplementary Figure S22b**). Conducting the same analysis on individual mCRPC transcriptional subtypes produced comparable findings to the analysis that included all samples (**Supplementary Figure S22c-d**). We observed that *ZNF263* footprints were prevalent in all mCRPC subtypes. However, when compared to other subtypes, the AR-NE- subtype exhibited a higher occupancy of *ZNF263* footprints (**Supplementary Figure S22e**). Interestingly, the target genes of *ZNF263* including *MYC* targets were expressed in AR-NE- subtype as well as many other samples from different other mCRPC subtypes (**Supplementary Figure S23**). These results suggest that *ZNF263* potentially acts as an activator of gene expression in all mCRPC subtypes.

601 Next, to further strengthen our findings, we tested the enrichment of *ZNF263* target genes obtained using
602 TF network analysis. Indeed, MYC signaling was the top-most enriched pathway followed by several other
603 proliferation and oncogenic signaling pathways (**Figure 6b** and **Supplementary Figure S24a**). Conducting
604 the same TF network analysis for *MYC* revealed that the genes predicted to be up-regulated when *MYC* is
605 bound significantly overlap with genes targeted by *ZNF263* (**Figure 6c**). We also investigated genes that
606 are regulated by both *ZNF263* and *MYC*. Interestingly, the top enriched pathways were those involved in
607 MYC and androgen response signaling (**Supplementary Figure S24b**). The evidence presented above
608 suggests that *ZNF263* potentially influences the activity of *MYC* target genes.

609 Further, we inspected the *ZNF263* footprints and their association with that of *MYC*. Despite having a
610 different motif sequence than that of *MYC* (**Supplementary Figure S24c**), approximately 0.35% of all
611 *ZNF263* footprint sites overlap with the footprints of *MYC* and 2.75% of all MYC footprints overlap with
612 that of *ZNF263* (**Figure 6d**). We further excluded these overlapping footprint regions and measured the
613 distance between the remaining *ZNF263* and its nearest *MYC* binding site. We found *MYC*-occupied regions
614 near *ZNF263* binding sites more often in the AR-NE- subtype as compared to the rest of the mCRPC
615 transcriptional subtypes (**Supplementary Figure S24d**). Furthermore, we evaluated if *ZNF263* footprints
616 predicted using ATAC-seq overlap with *MYC* ChIP-seq predicted binding sites. For this, we leveraged
617 publicly available *MYC* ChIP-seq profiles measured on different prostate cancer and osteosarcoma cell
618 lines. We found that about 15-20% of *ZNF263* bound footprint sites overlap with the *MYC* ChIP-seq
619 predicted binding sites (hypergeometric test p-value $< 2.2 \times 10^{-16}$) (**Supplementary Figure S24e**). Thus,
620 implying that *MYC* binds near the binding sites of *ZNF263*.

621 We next tested the hypothesis that *ZNF263* is a co-activator of *MYC* transcriptional targets in mCRPC.
622 *ZNF263* was predicted to bind more gene promoter sites than *MYC* in AR+ and AR- subtypes
623 (**Supplementary Figure S24f**). We tested whether *MYC* targets were differentially expressed when both
624 *ZNF263* and *MYC* were bound in the promoter region, in comparison with *ZNF263* alone or neither protein.
625 Supporting our hypothesis, the presence of *ZNF263* binding increased the expression level of these genes
626 (**Figure 6e**), and the concomitant binding of ZNF263 and MYC further increased the expression level of
627 MYC target genes (**Figure 6f**). These observations are consistent with a model that *ZNF263* collaborates
628 with *MYC* to activate *MYC* targets in mCRPC.

Discussion

630 The prolonged usage of ADT and/or ARSI in advanced prostate cancer leads to the emergence of a diverse
631 spectrum of mCRPC subtypes. While there are emerging genomic and transcriptomic distinctions between
632 the mCRPC subtypes, information regarding the variations in the epigenomic regulatory landscape between
633 the subtypes is scarce. In this study, we present a comprehensive characterization of the chromatin
634 accessibility of mCRPC using integrated analysis of ATAC-seq and RNA-seq from matched samples.
635 Earlier studies (15,21,22) on mCRPC predominantly depended on cell lines, organoids, PDXs, or a limited
636 number of mCRPC tissue biopsies. To date, our study represents the largest group of mCRPC tissue
637 biopsies that have been characterized using ATAC-seq. Here, we show that chromatin accessibility
638 increases during PCa progression to mCRPC. Importantly, we found mCRPC to have unique chromatin
639 accessibility profiles compared to localized PCa and benign prostate. AR signaling is the major driver of
640 mCRPC(11–13,16,17) and our investigation indicated that the functional activity of AR predominantly
641 governs the chromatin accessibility patterns in mCRPC. The ATAC-seq profile of our mCRPC cohort

642 closely recapitulates the chromatin accessibility heterogeneity in the advanced PCa patient population.
643 Furthermore, the subtypes Tang et al. (22) identified through the analysis of ATAC-seq profiles in CRPC
644 cell lines, organoid, and PDX models were similar to the mCRPC subtypes reported in our study.
645 Particularly noteworthy was the finding that the CRPC-WNT and CRPC-SCL subtypes predominantly
646 represented AR- subtypes.

647 This study evaluated differences in chromatin accessibility across 5 mCRPC subtypes. Among these
648 subtypes, AR+NE- and AR-NE+ (t-SCNC/NEPC) subtypes have been extensively studied and well
649 characterized, while AR_{low}NE-, AR+NE+, and AR-NE- subtypes remain relatively less explored. By
650 integrating ATAC-seq and RNA-seq of matched tumors, we established a correlation between the
651 accessibility of a regulatory element to the expression levels of predicted target genes. We then conducted
652 a thorough analysis of TF occupancy signals across the entire accessible genomic regions, leading us to
653 identify 203 TFs that are linked to specific mCRPC subtypes. Some TFs are uniquely enriched in a certain
654 mCRPC subtype whereas others are common to multiple subtypes. We found that a variety of TFs were
655 associated with the t-SCNC phenotype. Furthermore, we found that AR_{low}NE- and AR-NE- have many
656 common TFs, which further supports the previous (5) findings suggesting that these subtypes share common
657 signaling pathways. We observed the presence of numerous TFs associated with AR+NE- in both AR_{low}NE-
658 and AR+NE+ subtypes. Additionally, the AR+NE+ subtype displayed enrichment of several NE-related
659 TFs. Thus, our analysis suggests that AR_{low}NE-, AR+NE+, and AR-NE- subtypes might potentially be an
660 intermediate phenotype between the more extreme AR+NE- and AR-NE+ subtypes. A recent study (5) has
661 reported the existence of multiple mCRPC subtypes within the same metastatic site of a patient supporting
662 mCRPC disease continuum hypothesis. These diverse spectra of mCRPC subtypes are believed to emerge
663 from intertumoral heterogeneity and treatment-induced selective pressures that can change the phenotypic
664 and molecular landscapes of mCRPC (5). This mandates a detailed study interrogating the molecular
665 mechanisms driving these rare intermediate mCRPC subtypes and their clinical outcome for better
666 management of mCRPC.

667 Our analysis identified both established mCRPC-associated TFs as well as TFs that are relatively under-
668 studied in the context of PCa. Our interest in investigating *ZNF263* was driven by several studies linking
669 the altered activity of *ZNF263* to oncogenic processes and chemotherapy resistance in different cancers
670 (65,66). To the best of our knowledge, the role of *ZNF263* in prostate cancer has not been studied to date.
671 Here, we demonstrate that *ZNF263* has a considerable influence on modulating the gene expression in
672 mCRPC and may collaborate with *MYC* in these tumors. Thus, the integration of ATAC-seq and RNA-seq
673 data in our work demonstrates the ability to investigate the effects of TF binding on the activity of their
674 downstream target genes.

675 Our comprehension of gene expression regulation relies heavily on understanding TF binding to regulatory
676 elements (21). Traditionally, ChIP-seq has been the standard method for identifying TF binding sites, but
677 more recently, ATAC-seq has emerged as a promising alternative (20,67). Unlike ChIP-seq, which requires
678 specific antibodies targeting individual TFs, ATAC-seq enables comprehensive genome-wide profiling of
679 footprints of all known TFs within the accessible chromatin regions. Thus, ATAC-seq TF footprint
680 prediction holds promise as a method to screen the genome-wide binding of an extensive range of TFs in a
681 single analysis framework to gain a comprehensive understanding of gene regulation and further reconstruct
682 subtype-specific regulatory networks.

17

Transcription factors are alluring as therapeutic targets because they are master regulators of large gene networks that affect disease outcomes. Although TFs are conventionally considered difficult-to-drug proteins, promising technologies such as PROTAC (68) have enabled the targeted degradation of desired proteins, including TFs. Direct (PROTAC) or indirect inhibition of several TFs is currently being investigated in clinical trials (69). ATAC-seq TF footprinting can aid in the identification of potential therapeutic targets by providing information on the regulatory regions of genes that are accessible to TF binding.

In summary, this study characterizes the changes in chromatin accessibility in advanced PCa. Our results illustrate the importance of studying chromatin shifts at regulatory regions to determine TFs actively occupied in the region, and how they modulate transcriptional programs associated with oncogenic and tumor-suppressive functions. Overall, our findings provide valuable insights into epigenetic changes that occur during mCRPC progression.

Authors' Contributions

R. Shrestha: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **L.N. Chesner**: Data curation, Methodology, Writing – review & editing. **M. Zhang**: Data curation, Methodology, Software, Writing – review & editing. **S. Zhou**: Methodology, Writing – review & editing. **A. Foye**: Data curation, Methodology, Project administration, Resources, Writing – review & editing. **A. Lundberg**: Data curation, Methodology, Software, Writing – review & editing. **A.S. Weinstein**: Data curation, Methodology, Software, Writing – review & editing. **M. Sjöström**: Data curation, Methodology, Software, Writing – review & editing. **X. Zhu**: Data curation, Methodology, Software, Writing – review & editing. **T. Moreno-Rodriguez**: Data curation, Methodology, Software, Writing – review & editing. **H. Li**: Data curation, Methodology, Writing – review & editing. **SU2C/PCF West Coast Prostate Cancer Dream Team**: Resources, Writing – review & editing. **J.J. Alumkal**: Resources, Writing – review & editing. **R. Aggarwal**: Resources, Writing – review & editing. **E.J. Small**: Funding acquisition, Resources, Writing – review & editing. **M. Lupien**: Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing. **D.A. Quigley**: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Writing – review & editing. **F.Y. Feng**: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – review & editing.

Acknowledgments

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin [Internet]. American Cancer Society; 2022 [cited 2022 Dec 23];72:7–33. Available from: https://onlinelibrary.wiley.com/doi/full/10.3322/caac.21708

2. de Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and Increased Survival in Metastatic Prostate Cancer. n engl j med. 2011.

3. Scher HI, Fizazi K, Saad F, Taplin M-E, Sternberg CN, Miller K, et al. Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. New England Journal of Medicine. New England Journal of Medicine (NEJM/MMS); 2012;367:1187–97.

4. Ryan CJ, Smith MR, de Bono JS, Molina A, Logothetis CJ, de Souza P, et al. Abiraterone in Metastatic Prostate Cancer without Previous Chemotherapy. New England Journal of Medicine. Massachusetts Medical Society; 2013;368:138–48.

5. Labrecque MP, Coleman IM, Brown LG, True LD, Kollath L, Lakely B, et al. Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer. Journal of Clinical Investigation. American Society for Clinical Investigation; 2019;129:4492–505.

6. Vlachostergios PJ, Puca L, Beltran H. Emerging Variants of Castration-Resistant Prostate Cancer. Curr Oncol Rep. Current Medicine Group LLC 1; 2017.

7. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med. Nature Publishing Group; 2016;22:298–305.

8. Aggarwal R, Huang J, Alumkal JJ, Zhang L, Feng FY, Thomas G v, et al. Clinical and Genomic Characterization of Treatment-Emergent Small-Cell Neuroendocrine Prostate Cancer: A Multi-institutional Prospective Study. J Clin Oncol [Internet]. 2018;36:2492–503. Available from: https://doi.org/10.1200/JCO.2017.

9. Bluemn EG, Coleman IM, Lucas JM, Coleman RT, Hernandez-Lopez S, Tharakan R, et al. Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling. Cancer Cell. Cell Press; 2017;32:474-489.e6.

10. Westbrook TC, Guan X, Rodansky E, Flores D, Liu CJ, Udager AM, et al. Transcriptional profiling of matched patient biopsies clarifies molecular determinants of enzalutamide-induced lineage plasticity. Nat Commun [Internet]. Nature Research; 2022;13:5345. Available from: http://www.ncbi.nlm.nih.gov/pubmed/36109521

11. Robinson D, Van Allen EM, Wu Y-M, Schultz N, Lonigro RJ, Mosquera J-M, et al. Integrative clinical genomics of advanced prostate cancer. Cell [Internet]. Cell Press; 2015;161:1215–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26000489

12. Quigley DA, Dang HX, Zhao SG, Lloyd P, Aggarwal R, Alumkal JJ, et al. Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. Cell. Cell Press; 2018;174:758-769.e9.

769   13.   Viswanathan SR, Ha G, Hoff AM, Wala JA, Carrot-Zhang J, Whelan CW, et al. Structural
770   Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome
771   Sequencing. Cell. Cell Press; 2018;174:433-447.e19.

772   14.   Park JW, Lee JK, Sheu KM, Wang L, Balanis NG, Nguyen K, et al. Reprogramming normal
773   human epithelial tissues to a common, lethal neuroendocrine cancer lineage [Internet]. Available
774   from: http://science.sciencemag.org/

775   15.   Pomerantz MM, Qiu X, Zhu Y, Takeda DY, Pan W, Baca SC, et al. Prostate cancer reactivates
776   developmental epigenomic programs during metastatic progression. Nat Genet. Nature Research;
777   2020;52:790–9.

778   16.   Zhao SG, Chen WS, Li H, Foye A, Zhang M, Sjöström M, et al. The DNA methylation landscape
779   of advanced prostate cancer. Nat Genet. Nature Research; 2020;52:778–89.

780   17.   Sjöström M, Zhao SG, Levy S, Zhang M, Ning Y, Shrestha R, et al. The 5-
781   Hydroxymethylcytosine Landscape of Prostate Cancer. Cancer Res [Internet]. 2022;82:3888–902.
782   Available from: http://www.ncbi.nlm.nih.gov/pubmed/36251389

783   18.   Rubin MA, Bristow RG, Thienger PD, Dive C, Imielinski M. Impact of Lineage Plasticity to and
784   from a Neuroendocrine Phenotype on Progression and Response in Prostate and Lung Cancers.
785   Mol Cell [Internet]. Cell Press; 2020;80:562–77. Available from:
786   http://www.ncbi.nlm.nih.gov/pubmed/33217316

787   19.   Stelloo S, Nevedomskaya E, Poel HG, Jong J, Leenders GJ, Jenster G, et al. Androgen receptor
788   profiling predicts prostate cancer outcome. EMBO Mol Med. EMBO; 2015;7:1450–64.

789   20.   Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin
790   for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and
791   nucleosome position. Nat Methods. 2013;10:1213–8.

792   21.   Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, et al. The chromatin
793   accessibility landscape of primary human cancers. Science [Internet]. American Association for
794   the Advancement of Science; 2018;362. Available from:
795   http://www.ncbi.nlm.nih.gov/pubmed/30361341

796   22.   Tang F, Xu D, Wang S, Wong CK, Martinez-Fundichely A, Lee CJ, et al. Chromatin profiles
797   classify castration-resistant prostate cancers suggesting therapeutic targets. Science [Internet].
798   American Association for the Advancement of Science; 2022;376:eabe1505. Available from:
799   http://www.ncbi.nlm.nih.gov/pubmed/35617398

800   23.   Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, et al.
801   An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues.
802   Nat Methods. Nature Publishing Group; 2017;14:959–62.

803   24.   Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet
804   J [Internet]. 2011;17:10–2. Available from:
805   https://journal.embnet.org/index.php/embnetjournal/article/view/200

806   25.   Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. 2010.
807   Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

808   26.   Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–
809   9.

810   27.   Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map
811   format and SAMtools. Bioinformatics [Internet]. Oxford Academic; 2009 [cited 2022 Nov
812   9];25:2078–9. Available from:
813   https://academic.oup.com/bioinformatics/article/25/16/2078/204688

814   28.   Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next
815   generation web server for deep-sequencing data analysis. Nucleic Acids Res. Oxford University
816   Press; 2016;44:W160–5.

817   29.   Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis
818   of ChIP-Seq (MACS). Genome Biol [Internet]. 2008;9:R137. Available from:
819   http://www.ncbi.nlm.nih.gov/pubmed/18798982

820   30.   Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation,
821   comparison and visualization. Bioinformatics [Internet]. Oxford Academic; 2015 [cited 2023 Mar
822   18];31:2382–3. Available from:
823   https://academic.oup.com/bioinformatics/article/31/14/2382/255379

824   31.   Ou J, Liu H, Yu J, Kelliher MA, Castilla LH, Lawson ND, et al. ATACseqQC: a Bioconductor
825   package for post-alignment quality assessment of ATAC-seq data. BMC Genomics. 2018;19:169.

826   32.   Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for
827   alignment and quantification of RNA sequencing reads. Nucleic Acids Res [Internet]. Oxford
828   Academic; 2019 [cited 2022 Nov 9];47:e47–e47. Available from:
829   https://academic.oup.com/nar/article/47/8/e47/5345150

830   33.   Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq
831   data with DESeq2. Genome Biol [Internet]. BioMed Central Ltd.; 2014 [cited 2022 Nov 9];15:1–
832   21. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-
833   8

834   34.   McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves
835   functional interpretation of cis-regulatory regions. Nat Biotechnol [Internet]. 2010;28:495–501.
836   Available from: http://www.ncbi.nlm.nih.gov/pubmed/20436461

837   35.   Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular
838   Signatures Database Hallmark Gene Set Collection. Cell Syst. Cell Press; 2015;1:417–25.

839   36.   Bhuva DD, Cursons J, Davis MJ. Stable gene expression for normalisation and single-sample
840   scoring. Nucleic Acids Res. 2020;48:e113.

841   37.   Bentsen M, Goymann P, Schultheis H, Klee K, Petrova A, Wiegandt R, et al. ATAC-seq
842   footprinting unravels kinetics of transcription factor binding during zygotic genome activation.
843   Nat Commun [Internet]. Nature Research; 2020;11:4267. Available from:
844   http://www.ncbi.nlm.nih.gov/pubmed/32848148

845   38.   Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Lemma RB, Turchi L, Blanc-Mathieu
846   R, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding
847   profiles. Nucleic Acids Res. 2022;50:D165–73.

848 39. Wilson S, Qi J, Filipp F V. Refinement of the androgen response element based on ChIP-Seq in
849   androgen-insensitive and androgen-responsive prostate cancer cell lines. Sci Rep [Internet]. Nature
850   Publishing Group; 2016;6:32611. Available from: https://www.nature.com/articles/srep32611

851 40. Pinskaya M, Saci Z, Gallopin M, Gabriel M, Nguyen HT, Firlej V, et al. Reference-free
852   transcriptome exploration reveals novel RNAs for prostate cancer diagnosis. Life Sci Alliance.
853   2019;2.

854 41. Chen S, Huang V, Xu X, Livingstone J, Soares F, Jeon J, et al. Widespread and Functional RNA
855   Circularization in Localized Prostate Cancer. Cell. 2019;176:831-843.e22.

856 42. Cejas P, Xie Y, Font-Tello A, Lim K, Syamala S, Qiu X, et al. Subtype heterogeneity and
857   epigenetic convergence in neuroendocrine prostate cancer. Nat Commun. Nature Research;
858   2021;12.

859 43. Imbeault M, Helleboid PY, Trono D. KRAB zinc-finger proteins contribute to the evolution of
860   gene regulatory networks. Nature. Nature Publishing Group; 2017;543:550–4.

861 44. Frietze S, Lan X, Jin VX, Farnham PJ. Genomic targets of the KRAB and SCAN domain-
862   containing zinc finger protein 263. J Biol Chem [Internet]. 2010;285:1393–403. Available from:
863   http://www.ncbi.nlm.nih.gov/pubmed/19887448

864 45. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, et al. Topologically associating domains are
865   stable units of replication-timing regulation. Nature [Internet]. Nature; 2014 [cited 2023 Mar
866   27];515:402–5. Available from: https://pubmed.ncbi.nlm.nih.gov/25409831/

867 46. Barfeld SJ, Urbanucci A, Itkonen HM, Fazli L, Hicks JL, Thiede B, et al. c-Myc Antagonises the
868   Transcriptional Activity of the Androgen Receptor in Prostate Cancer Affecting Key Gene
869   Networks. EBioMedicine [Internet]. EBioMedicine; 2017 [cited 2023 Mar 27];18:83–93.
870   Available from: https://pubmed.ncbi.nlm.nih.gov/28412251/

871 47. See YX, Chen K, Fullwood MJ. MYC overexpression leads to increased chromatin interactions at
872   super-enhancers and MYC binding sites. Genome Res [Internet]. Genome Res; 2022 [cited 2023
873   Mar 27];32:629–42. Available from: https://pubmed.ncbi.nlm.nih.gov/35115371/

874 48. Guo H, Wu Y, Nouri M, Spisak S, Russo JW, Sowalsky AG, et al. Androgen receptor and MYC
875   equilibration centralizes on developmental super-enhancer. Nat Commun [Internet]. Nat Commun;
876   2021 [cited 2023 Mar 27];12. Available from: https://pubmed.ncbi.nlm.nih.gov/34911936/

877 49. Oki S, Ohta T, Shioi G, Hatanaka H, Ogasawara O, Okuda Y, et al. ChIP-Atlas: a data-mining
878   suite powered by full integration of public ChIP-seq data. EMBO Rep. 2018;19.

879 50. Urbanucci A, Barfeld SJ, Kytölä V, Itkonen HM, Coleman IM, Vodák D, et al. Androgen
880   Receptor Deregulation Drives Bromodomain-Mediated Chromatin Alterations in Prostate Cancer.
881   Cell Rep. Elsevier B.V.; 2017;19:2045–59.

882 51. Lundberg A, Zhang M, Aggarwal R, Li H, Zhang L, Foye A, et al. The genomic and epigenomic
883   landscape of double-negative metastatic prostate cancer. Cancer Res. 2023;

884 52. Meyer RC, Giddens MM, Schaefer SA, Hall RA. GPR37 and GPR37L1 are receptors for the
885   neuroprotective and glioprotective factors prosaptide and prosaposin. Proc Natl Acad Sci U S A

886   [Internet]. Proc Natl Acad Sci U S A; 2013 [cited 2023 Apr 6];110:9529–34. Available from:
887   https://pubmed.ncbi.nlm.nih.gov/23690594/

888   53.   Vierstra J, Stamatoyannopoulos JA. Genomic footprinting. Nat Methods [Internet]. Nature
889   Publishing Group; 2016;13:213–21. Available from:
890   http://www.ncbi.nlm.nih.gov/pubmed/26914205

891   54.   Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Berhanu Lemma R, Turchi L, Blanc-
892   Mathieu R, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor
893   binding profiles. Nucleic Acids Res [Internet]. Nucleic Acids Res; 2022 [cited 2022 Nov
894   29];50:D165–73. Available from: https://pubmed.ncbi.nlm.nih.gov/34850907/

895   55.   Nouruzi S, Ganguli D, Tabrizian N, Kobelev M, Sivak O, Namekawa T, et al. ASCL1 activates
896   neuronal stem cell-like lineage programming through remodeling of the chromatin landscape in
897   prostate cancer. Nat Commun. Nature Research; 2022;13.

898   56.   Bishop JL, Thaper D, Vahid S, Davies A, Ketola K, Kuruma H, et al. The master neural
899   transcription factor BRN2 is an androgen receptor–suppressed driver of neuroendocrine
900   differentiation in prostate cancer. Cancer Discov. American Association for Cancer Research Inc.;
901   2017;7:54–71.

902   57.   FitsGerald PC, Shlyakhtenko A, Mir AA, Vinson C. Clustering of DNA sequences in human
903   promoters. Genome Res [Internet]. Genome Res; 2004 [cited 2023 Mar 21];14:1562–74. Available
904   from: https://pubmed.ncbi.nlm.nih.gov/15256515/

905   58.   Sahu B, Laakso M, Ovaska K, Mirtti T, Lundin J, Rannikko A, et al. Dual role of FoxA1 in
906   androgen receptor binding to chromatin, androgen signalling and prostate cancer. EMBO J.
907   2011;30:3962–76.

908   59.   Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational
909   landscape of lethal castration-resistant prostate cancer. Nature. 2012;487:239–43.

910   60.   Bery F, Cancel M, Guéguinou M, Potier-Cartereau M, Vandier C, Chantôme A, et al. Zeb1 and
911   SK3 Channel Are Up-Regulated in Castration-Resistant Prostate Cancer and Promote
912   Neuroendocrine Differentiation. Cancers (Basel) [Internet]. Cancers (Basel); 2021 [cited 2023 Mar
913   22];13. Available from: https://pubmed.ncbi.nlm.nih.gov/34204608/

914   61.   Lee GT, Rosenfeld JA, Kim WT, Kwon YS, Palapattu G, Mehra R, et al. TCF4 induces
915   enzalutamide resistance via neuroendocrine differentiation in prostate cancer. PLoS One
916   [Internet]. PLoS One; 2019 [cited 2023 Mar 22];14. Available from:
917   https://pubmed.ncbi.nlm.nih.gov/31536510/

918   62.   Cmero M, Kurganovs NJ, Stuchbery R, McCoy P, Grima C, Ngyuen A, et al. Loss of SNAI2 in
919   Prostate Cancer Correlates With Clinical Response to Androgen Deprivation Therapy. JCO Precis
920   Oncol [Internet]. JCO Precis Oncol; 2021 [cited 2023 Mar 22];5:1048–59. Available from:
921   https://pubmed.ncbi.nlm.nih.gov/34322653/

922   63.   Zhao Y, Vartak S v., Conte A, Wang X, Garcia DA, Stevens E, et al. "Stripe" transcription factors
923   provide accessibility to co-binding partners in mammalian genomes. Mol Cell. Cell Press;
924   2022;82:3398-3411.e11.

925   64.   Che M, Chaturvedi A, Munro SA, Pitzen SP, Ling A, Zhang W, et al. Opposing transcriptional
926   programs of KLF5 and AR emerge during therapy for advanced prostate cancer. Nat Commun
927   [Internet]. Nature Research; 2021;12:6377. Available from:
928   http://www.ncbi.nlm.nih.gov/pubmed/34737261

929   65.   Cui J, Liu J, Fan L, Zhu Y, Zhou B, Wang Y, et al. A zinc finger family protein, ZNF263,
930   promotes hepatocellular carcinoma resistance to apoptosis via activation of ER stress-dependent
931   autophagy. Transl Oncol [Internet]. Neoplasia Press, Inc.; 2020;13:100851. Available from:
932   http://www.ncbi.nlm.nih.gov/pubmed/32898766

933   66.   Yu Z, Feng J, Wang W, Deng Z, Zhang Y, Xiao L, et al. The EGFR-ZNF263 signaling axis
934   silences SIX3 in glioblastoma epigenetically. Oncogene [Internet]. Springer Nature;
935   2020;39:3163–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32051553

936   67.   Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible
937   chromatin landscape of the human genome. Nature [Internet]. 2012;489:75–82. Available from:
938   http://www.ncbi.nlm.nih.gov/pubmed/22955617

939   68.   Samarasinghe KTG, Jaime-Figueroa S, Burgess M, Nalawansha DA, Dai K, Hu Z, et al. Targeted
940   degradation of transcription factors by TRAFTACs: TRAnscription Factor TArgeting Chimeras.
941   Cell Chem Biol. Elsevier Ltd; 2021;28:648-661.e5.

942   69.   Henley MJ, Koehler AN. Advances in targeting 'undruggable' transcription factors with small
943   molecules. Nat Rev Drug Discov. Nature Research; 2021. page 669–88.

944

945

Figure Legends

Figure 1. Chromatin accessibility changes during prostate cancer progression affect stage-specific regulatory elements. (a) Principal component analysis (PCA) of the ATAC-seq profiles comparing different stages of prostate cancer including benign prostate, localized prostate cancer (PCa), mCRPC Adeno, and mCRPC t-SCNC. The normalized read counts of these consensus-accessible regions were used for the PCA analysis. Each dot in the plot represents an individual sample. (b) An alluvial plot demonstrating changes in accessible chromatin regions in various stages of prostate cancer. Each bar corresponds to a distinct PCa stage, with the orange and white sections indicating accessible and inaccessible chromatin regions, respectively. The shaded areas connecting the bars represent changes in the accessibility of these chromatin regions. The pink and blue shaded regions respectively represent accessible and inaccessible chromatin regions in mCRPC Adeno. (c) Heatmap representation of the chromatin between different stages. The rows are segregated by the chromatin variants in each stage. (d) The percentage of chromatin variants in mCRPC Adeno. The ATAC-seq peaks are grouped by the genomic regions (promoter, intron, or distal intergenic) to which they are mapped. (e) ATAC-seq profile plot illustrating potential regulatory regions in chromatin variants in mCRPC Adeno. The profile plot represents the overlapping region between the chromatin variants and publicly available H3K27ac ChIP-seq data from mCRPC PDX. (f) Enrichment of chromatin regions exclusively accessible in localized PCa, mCRPC Adeno, or mCRPC t-SCNC against GO Biological Processes.

Figure 2. Chromatin accessibility in mCRPC is associated with subtypes linked to androgen signaling. (a) Unsupervised hierarchical clustering of pairwise sample Spearman's correlation based on the normalized read counts of consensus ATAC-seq peaks of mCRPC. (b-c) Distribution of (b) androgen receptor (AR) pathway score, (c) neuroendocrine (NE) score calculated based on RNA-seq gene expression profiles of mCRPCs classified into individual clusters in Figure 2a. Statistically significant Wilcoxon rank sum test p-values between the clusters are indicated in the plot.

Figure 3. mCRPC transcriptional subtypes are associated with chromatin variants of prostate cancer signaling pathways. (a) Heatmap of chromatin variants between mCRPC transcriptional subtypes. (b) ATAC-seq peaks around *AR*, *NKX3-1*, and *GPR37L1* gene regions. The highlighted vertical strip illustrates the presence of ATAC-seq peaks at the enhancer region. (c) Heatmap representation of the chromatin variants between mCRPC transcriptional subtypes. The rows are segregated by the differential regions mapped to gene promoter, intron, or distal intergenic regions. (d-f) Hallmark pathways enrichment of chromatin variants between the mCRPC transcriptional subtypes mapped to (d) promoter, (e) intron, and (f) distal intergenic regions identified using GREAT enrichment analysis (see Methods section).

Figure 4. mCRPC transcriptional subtypes are defined by DNA accessibility-guided patterns of transcription factor regulation. (a-b) ATAC-seq TF (*AR* and *HOXB13*) footprints signal difference between TF-bound and unbound sites. (c-d) Volcano plot of differential TF footprint occupancy analysis comparing the (c) AR+NE- and AR-NE+ subtypes and (d) AR+NE- and AR-NE- subtypes. Each dot in the plot represents a TF motif. The colored dots indicate a significantly differentially bound TF motif. (Data available as Supplementary Table S4) (e) Heatmap of genome-wide active TF occupancy, determined by TF footprints, associated with different mCRPC transcriptional subtypes. Each rim of the circular heatmap represents an individual mCRPC transcriptional subtype and the sector represents TF. The darker color shade indicates the strong association of the TF with the respective mCRPC subtype. See the Methods

987   section for details on TF occupancy phenotype score calculation.

988   **Figure 5. Identification of the influential transcription factors driving mCRPC transcriptional**
989   **subtypes.** We hypothesized that highly active TF regulate (or influence) gene expression activity of a large
990   fraction of target genes. The plot indicates the top influential mCRPC transcriptional subtype-associated TFs
991   ranked by the number of target genes (based on gene expression) they influence.

992   **Figure 6. *ZNF263* activates *MYC* signaling targets.** (a) The volcano plot depicts the genes that undergo
993   activation or repression upon *ZNF263* binding to their specific promoter region. Each gene is represented
994   by a dot, and the difference in gene expression between samples with and without *ZNF263* in the promoter
995   region was measured as fold change. Additionally, the statistical significance of the difference was
996   evaluated using the Wilcoxon rank sum test to calculate the p-value between the two groups. (b) Over-
997   representation analysis of the predicted *ZNF263* target genes against the Hallmark pathways. (c) Percentage
998   of the predicted target genes of *MYC* that overlap with those of TFs associated with AR-NE- subtype. The
999   overlap of the respective TF target genes with genes in the Hallmark *MYC* targets geneset is illustrated as
1000   the red line. (d) Heatmap of overlapping *ZNF263* and *MYC* footprint sites. The red color highlights the
1001   direct overlap between *ZNF263* and *MYC* footprints. (e) Volcano plot showing the genes that are activated
1002   when *ZNF263* binds to the promoter region compared to genes that are activated with both *ZNF263* and *MYC*
1003   are not bound to the promoter region. (f) Box plot of gene expression foldchange when different
1004   combinations of *ZNF263* and *MYC* bind to the promoter as compared to when both *ZNF263* and *MYC* are
1005   simultaneously absent in the promoter. Each dot represents a gene in Hallmark *MYC* targets geneset.

Figure 1

a

PC3 (6.8%)
PC1 (17%)
PC2 (13.7%)

Benign
localized PCa
mCRPC Adeno
mCRPC t-SCNC
t-SCNC (PDX)

b

Genomic region (in MB)

Benign   localized   mCRPC   mCRPC
         PCa   Adeno   t-SCNC

Closed
Open

c

Chromatin variants

Benign   localized   mCRPC   mCRPC
         PCa   Adeno   t-SCNC

-1 s e 1 -1 s e 1 -1 s e 1 -1 s e 1
ATAC-seq peaks (kbp)

Accessibility
low   high

Benign   localized PCa   s: peak start
mCRPC Adeno   mCRPC t-SCNC   e: peak end

d

mCRPC Adeno

% of chromatin variants

promoter   intron   distal intergenic

e

ATAC-seq profile (BPM)

-3 -2 -1 center 1 2 3
Distance from peak center (kbp)

promoter   intron   distal intergenic

f

localized   mCRPC   mCRPC
PCa   Adeno   t-SCNC

AR signaling pathway
Cellular response to testosterone stimulus
Transcriptional regulation
Epithelial cell apical/basal polarity
Apical/basal cell polarity
ER signaling pathway
Cell migration
Integrin biosynthetic process
Epithelial cell differentiation
Prostate gland development
Cytoskeleton organization
Neuron projection guidance
Steroid metabolic process
Cell morphogenesis involved in differentiation
Cell morphogenesis involved in neuron differentiation
Chondroitin sulfate proteoglycan biosynthetic process
Lipid biosynthetic process
UDP-glucuronate metabolic process
Neuron differentiation
Neuron development
Cell morphogenesis
Neuron projection morphogenesis
Axonogenesis

Enrichment
Score
$[-\log_{10}(FDR)]$
100
80
20
10

Figure 2

Figure 3

a

RIF1
GPR37L1
CCND1
BRAF
NKX3-1
SPOP
NCOA2
KLK3
ETV1
ARID1A
SMARCA1
FOXA1
AR
KDM6A
KMT2D

Chromatin variants

samples

Normalized
ATAC-seq counts

-2   0   2

AR+NE-
AR$_{low}$NE-
AR+NE+
AR-NE+
AR-NE-

b

AR   NKX3-1   GPR37L1

chromatin variants

AR+NE-
AR$_{low}$NE-
AR+NE+
AR-NE+
AR-NE-
H3K27ac

ATAC-seq signal
(Bins Per Million mapped reads)

ChrX (Mb)   Chr8 (Mb)   Chr1 (Mb)

67.0   67.5   23.67   23.68   202.08   202.13

c

AR+NE-   AR$_{low}$NE-   AR+NE+   AR-NE+   AR-NE-

promoter

intron

distal intergenic

Chromatin variants

-1 s e 1 -1 s e 1 -1 s e 1 -1 s e 1 -1 s e 1

ATAC-seq peaks (kbp)

s: peak start
e: peak end

Accessibility

low   high

d

promoter

Androgen Response
Coagulation
Fatty Acid Metabolism
Estrogen Response Late
Peroxisome
Spermatogenesis

0   4   8   12

-log$_{10}$(Bonferroni pvalue)

e

intron

Androgen Response
Estrogen Response Early
IL2-STAT5 Signaling
TGF-Beta Signaling
Fatty Acid Metabolism
Mitotic Spindle
Coagulation
Bile Acid Metabolism
Apoptosis
Notch Signaling
Peroxisome

0   4   8   12

-log$_{10}$(Bonferroni pvalue)

f

distal intergenic

Androgen Response
Coagulation
Unfolded Protein Response
UV Response Up
Adipogenesis

0   4   8   12
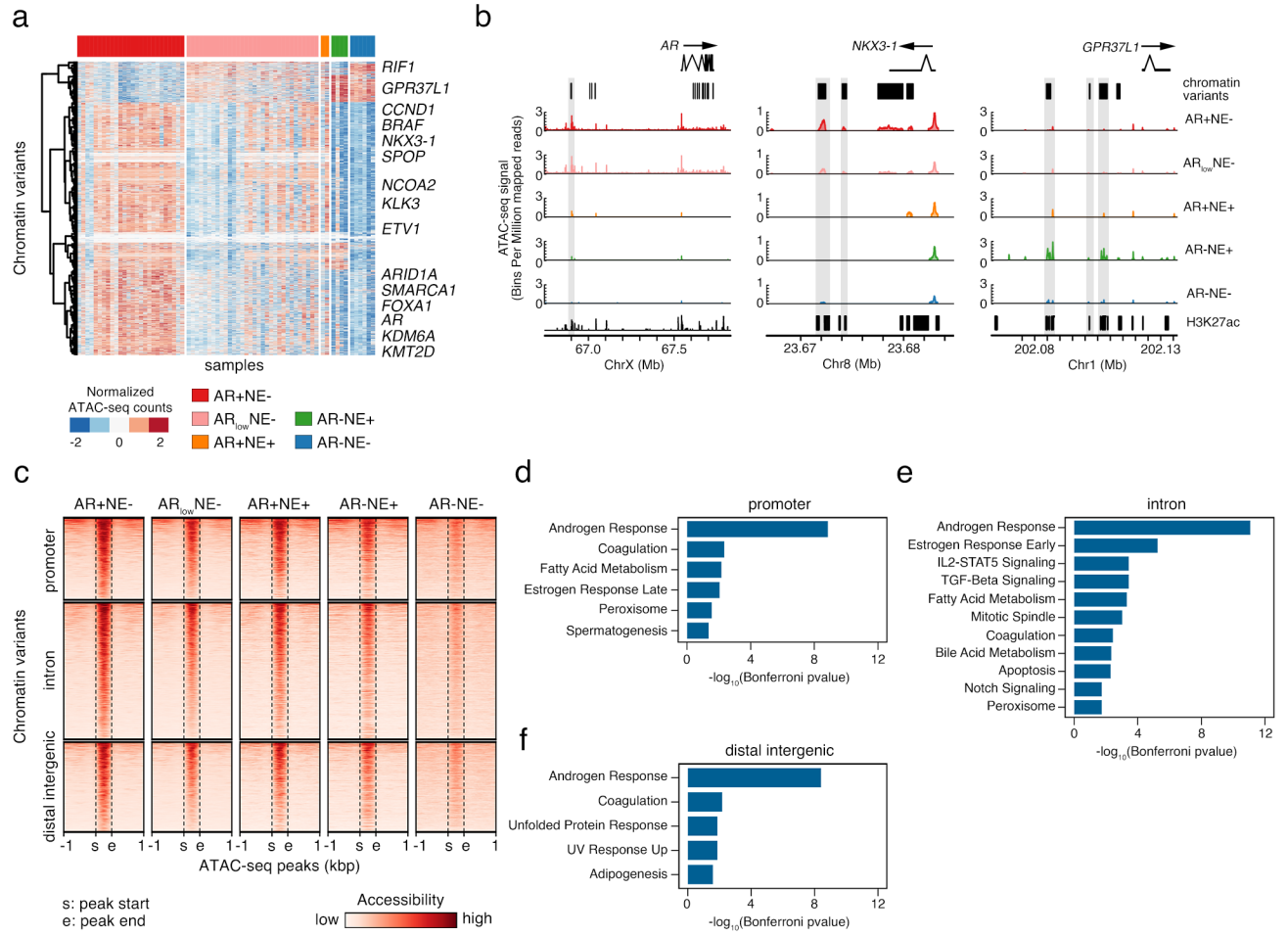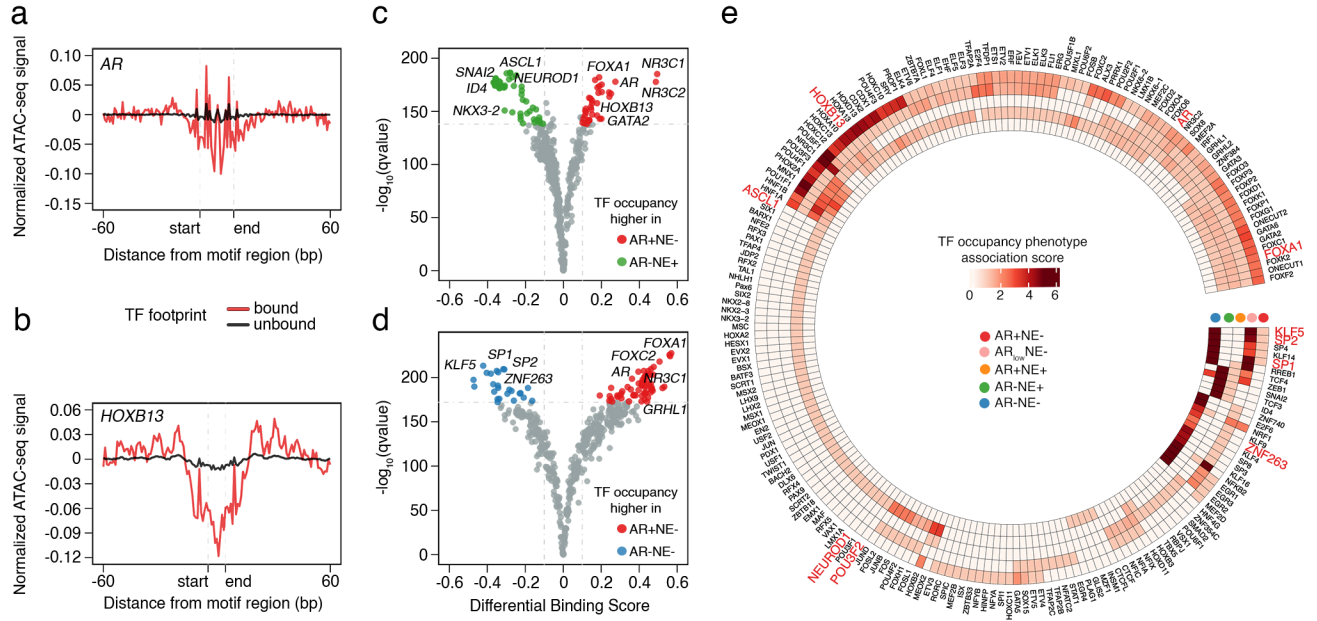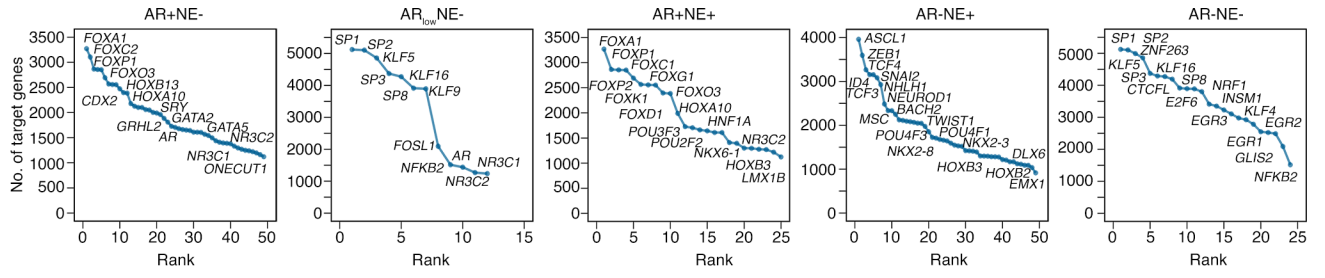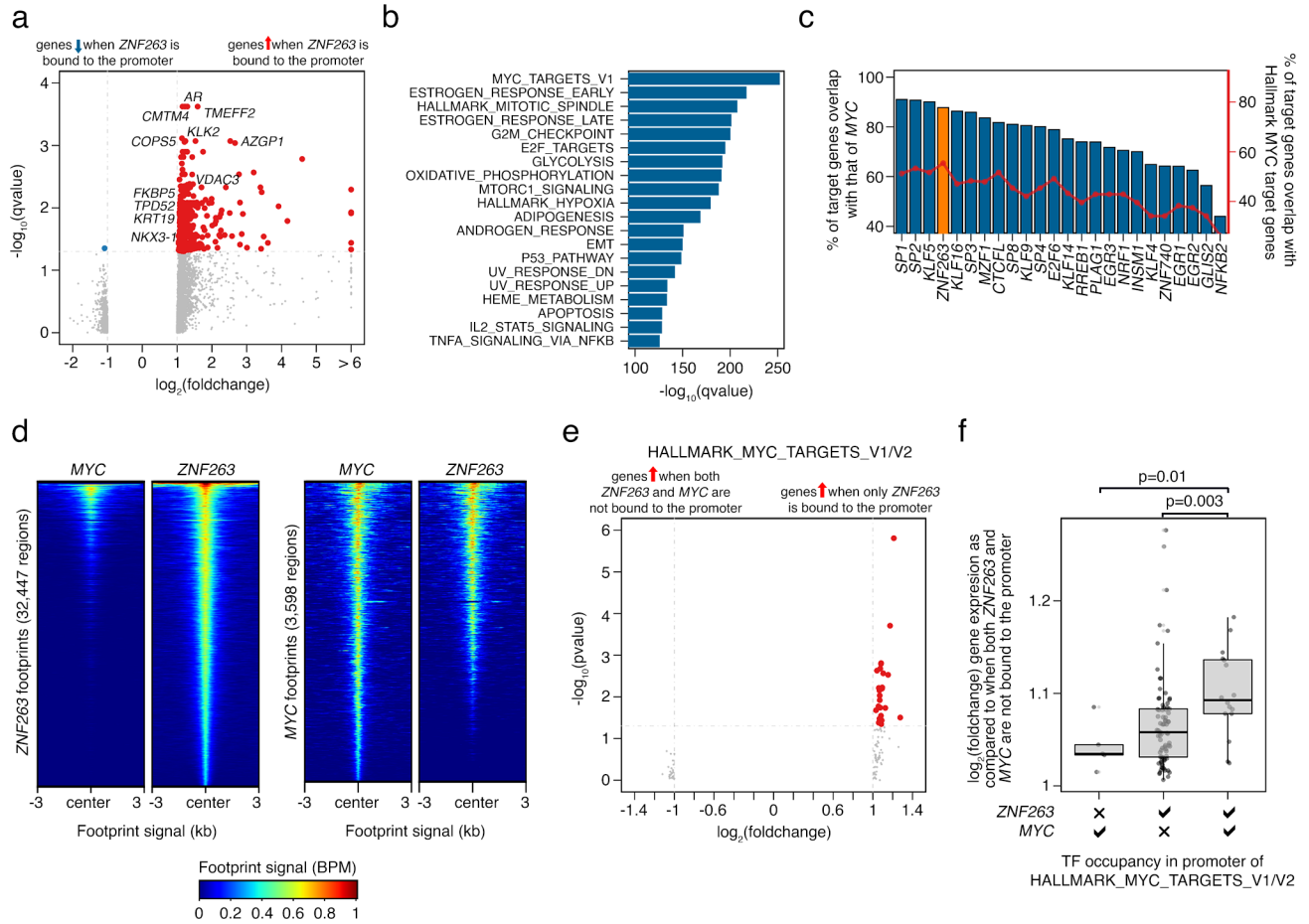
-log$_{10}$(Bonferroni pvalue)

Figure 4

Figure 5

Figure 6

Parsed Citations

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin [Internet]. American Cancer Society; 2022 [cited 2022 Dec 23];72:7-33. Available from: https://onlinelibrary.wiley.com/doi/full/10.3322/caac.21708
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

2. de Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and Increased Survival in Metastatic Prostate Cancer. n engl j med. 2011.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

3. Scher HI, Fizazi K, Saad F, Taplin M-E, Sternberg CN, Miller K, et al. Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. New England Journal of Medicine. New England Journal of Medicine (NEJM/MMS); 2012;367:1187-97.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

4. Ryan CJ, Smith MR, de Bono JS, Molina A, Logothetis CJ, de Souza P, et al. Abiraterone in Metastatic Prostate Cancer without Previous Chemotherapy. New England Journal of Medicine. Massachusetts Medical Society; 2013;368:138-48.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

5. Labrecque MP, Coleman IM, Brown LG, True LD, Kollath L, Lakely B, et al. Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer. Journal of Clinical Investigation. American Society for Clinical Investigation; 2019;129:4492- 505.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

6. Vlachostergios PJ, Puca L, Beltran H. Emerging Variants of Castration-Resistant Prostate Cancer. Curr Oncol Rep. Current Medicine Group LLC 1; 2017.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

7. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med. Nature Publishing Group; 2016;22:298-305.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

8. Aggarwal R, Huang J, Alumkal JJ, Zhang L, Feng FY, Thomas G v, et al. Clinical and Genomic Characterization of Treatment-Emergent Small-Cell Neuroendocrine Prostate Cancer: A Multi-institutional Prospective Study. J Clin Oncol [Internet]. 2018;36:2492-503. Available from: https://doi.org/10.1200/JCO.2017.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

9. Bluemn EG, Coleman IM, Lucas JM, Coleman RT, Hernandez-Lopez S, Tharakan R, et al. Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling. Cancer Cell. Cell Press; 2017;32:474-489.e6.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

10. Westbrook TC, Guan X, Rodansky E, Flores D, Liu CJ, Udager AM, et al. Transcriptional profiling of matched patient biopsies clarifies molecular determinants of enzalutamide-induced lineage plasticity. Nat Commun [Internet]. Nature Research; 2022;13:5345. Available from: http://www.ncbi.nlm.nih.gov/pubmed/36109521

Pubmed: Author and Title
Google Scholar: Google Scholar Search

11. Robinson D, Van Allen EM, Wu Y-M, Schultz N, Lonigro RJ, Mosquera J-M, et al. Integrative clinical genomics of advanced prostate cancer. Cell [Internet]. Cell Press; 2015;161:1215-28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26000489
Pubmed: Author and Title
Google Scholar: Google Scholar Search

12. Quigley DA, Dang HX, Zhao SG, Lloyd P, Aggarwal R, Alumkal JJ, et al. Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. Cell. Cell Press; 2018;174:758-769.e9.
Pubmed: Author and Title
Google Scholar: Google Scholar Search

13. Viswanathan SR, Ha G, Hoff AM, Wala JA, Carrot-Zhang J, Whelan CW, et al. Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. Cell. Cell Press; 2018;174:433-447.e19.
Pubmed: Author and Title
Google Scholar: Google Scholar Search

14. Park JW, Lee JK, Sheu KM, Wang L, Balanis NG, Nguyen K, et al. Reprogramming normal human epithelial tissues to a common, lethal neuroendocrine cancer lineage [Internet]. Available from: http://science.sciencemag.org/

15. Pomerantz MM, Qiu X, Zhu Y, Takeda DY, Pan W, Baca SC, et al. Prostate cancer reactivates developmental epigenomic programs during metastatic progression. Nat Genet. Nature Research; 2020;52:790-9.
Pubmed: Author and Title
Google Scholar: Google Scholar Search

16. Zhao SG, Chen WS, Li H, Foye A, Zhang M, Sjostrom M, et al. The DNA methylation landscape of advanced prostate cancer. Nat Genet. Nature Research; 2020;52:778-89.
Pubmed: Author and Title
Google Scholar: Google Scholar Search

17. Sjostrom M, Zhao SG, Levy S, Zhang M, Ning Y, Shrestha R, et al. The 5- Hydroxymethylcytosine Landscape of Prostate Cancer. Cancer Res [Internet]. 2022;82:3888-902. Available from: http://www.ncbi.nlm.nih.gov/pubmed/36251389
Pubmed: Author and Title
Google Scholar: Google Scholar Search

18. Rubin MA, Bristow RG, Thienger PD, Dive C, Imielinski M. Impact of Lineage Plasticity to and from a Neuroendocrine Phenotype on Progression and Response in Prostate and Lung Cancers. Mol Cell [Internet]. Cell Press; 2020;80:562-77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33217316
Pubmed: Author and Title
Google Scholar: Google Scholar Search

19. Stelloo S, Nevedomskaya E, Poel HG, Jong J, Leenders GJ, Jenster G, et al. Androgen receptor profiling predicts prostate cancer outcome. EMBO Mol Med. EMBO; 2015;7:1450-64.
Pubmed: Author and Title
Google Scholar: Google Scholar Search

20. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10:1213-8.
Pubmed: Author and Title
Google Scholar: Google Scholar Search

21. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, et al. The chromatin accessibility landscape of primary human cancers. Science [Internet]. American Association for the Advancement of Science; 2018;362. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30361341

Pubmed: Author and Title
Google Scholar: Google Scholar Search

22. Tang F, Xu D, Wang S, Wong CK, Martinez-Fundichely A, Lee CJ, et al. Chromatin profiles classify castration-resistant prostate cancers suggesting therapeutic targets. Science [Internet]. American Association for the Advancement of Science; 2022;376:eabe1505. Available from: http://www.ncbi.nlm.nih.gov/pubmed/35617398

Pubmed: Author and Title
Google Scholar: Google Scholar Search

23. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods. Nature Publishing Group; 2017;14:959-62.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

24. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J [Internet]. 2011;17:10-2. Available from: https://journal.embnet.org/index.php/embnetjournal/article/view/200

Pubmed: Author and Title
Google Scholar: Google Scholar Search

25. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

26. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357-

Pubmed: Author and Title
Google Scholar: Google Scholar Search

9.

27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics [Internet]. Oxford Academic; 2009 [cited 2022 Nov

9. ;25:2078-9. Available from: https://academic.oup.com/bioinformatics/article/25/16/2078/204688

Pubmed: Author and Title
Google Scholar: Google Scholar Search

28. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. Oxford University Press; 2016;44:W160-5.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

29. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol [Internet]. 2008;9:R137. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18798982

Pubmed: Author and Title
Google Scholar: Google Scholar Search

30. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics [Internet]. Oxford Academic; 2015 [cited 2023 Mar

18. ;31:2382-3. Available from: https://academic.oup.com/bioinformatics/article/31/14/2382/255379

Pubmed: Author and Title
Google Scholar: Google Scholar Search

31. Ou J, Liu H, Yu J, Kelliher MA, Castilla LH, Lawson ND, et al. ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. BMC Genomics. 2018;19:169.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

32. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res [Internet]. Oxford Academic; 2019 [cited 2022 Nov 9];47:e47-e47. Available from: https://academic.oup.com/nar/article/47/8/e47/5345150
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

33. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol [Internet]. BioMed Central Ltd.; 2014 [cited 2022 Nov 9];15:1-

21. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-

34. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol [Internet]. 2010;28:495-501. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20436461
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

35. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. Cell Press; 2015;1:417-25.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

36. Bhuva DD, Cursons J, Davis MJ. Stable gene expression for normalisation and single-sample scoring. Nucleic Acids Res. 2020;48:e113.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

37. Bentsen M, Goymann P, Schultheis H, Klee K, Petrova A, Wiegandt R, et al. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Nat Commun [Internet]. Nature Research; 2020;11:4267. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32848148
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

38. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Lemma RB, Turchi L, Blanc-Mathieu R, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2022;50:D165-73.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

39. Wilson S, Qi J, Filipp F V. Refinement of the androgen response element based on ChIP-Seq in androgen-insensitive and androgen-responsive prostate cancer cell lines. Sci Rep [Internet]. Nature Publishing Group; 2016;6:32611. Available from: https://www.nature.com/articles/srep32611
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

40. Pinskaya M, Saci Z, Gallopin M, Gabriel M, Nguyen HT, Firlej V, et al. Reference-free transcriptome exploration reveals novel RNAs for prostate cancer diagnosis. Life Sci Alliance. 2019;2.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

41. Chen S, Huang V, Xu X, Livingstone J, Soares F, Jeon J, et al. Widespread and Functional RNA Circularization in Localized Prostate Cancer. Cell. 2019;176:831-843.e22.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

42. Cejas P, Xie Y, Font-Tello A, Lim K, Syamala S, Qiu X, et al. Subtype heterogeneity and epigenetic convergence in neuroendocrine prostate cancer. Nat Commun. Nature Research; 2021;12.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

43. Imbeault M, Helleboid PY, Trono D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature. Nature Publishing Group; 2017;543:550-4.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

44. Frietze S, Lan X, Jin VX, Farnham PJ. Genomic targets of the KRAB and SCAN domain- containing zinc finger protein 263. J Biol Chem [Internet]. 2010;285:1393-403. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19887448
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

45. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, et al. Topologically associating domains are stable units of replication-timing regulation. Nature [Internet]. Nature; 2014 [cited 2023 Mar 27];515:402-5. Available from: https://pubmed.ncbi.nlm.nih.gov/25409831/

46. Barfeld SJ, Urbanucci A, Itkonen HM, Fazli L, Hicks JL, Thiede B, et al. c-Myc Antagonises the Transcriptional Activity of the Androgen Receptor in Prostate Cancer Affecting Key Gene Networks. EBioMedicine [Internet]. EBioMedicine; 2017 [cited 2023 Mar 27];18:83-93. Available from: https://pubmed.ncbi.nlm.nih.gov/28412251/

47. See YX, Chen K, Fullwood MJ. MYC overexpression leads to increased chromatin interactions at super-enhancers and MYC binding sites. Genome Res [Internet]. Genome Res; 2022 [cited 2023 Mar 27];32:629-42. Available from: https://pubmed.ncbi.nlm.nih.gov/35115371/

48. Guo H, Wu Y, Nouri M, Spisak S, Russo JW, Sowalsky AG, et al. Androgen receptor and MYC equilibration centralizes on developmental super-enhancer. Nat Commun [Internet]. Nat Commun; 2021 [cited 2023 Mar 27];12. Available from: https://pubmed.ncbi.nlm.nih.gov/34911936/

49. Oki S, Ohta T, Shioi G, Hatanaka H, Ogasawara O, Okuda Y, et al. ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. 2018;19.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

50. Urbanucci A, Barfeld SJ, Kytola V, Itkonen HM, Coleman IM, Vodak D, et al. Androgen Receptor Deregulation Drives Bromodomain-Mediated Chromatin Alterations in Prostate Cancer. Cell Rep. Elsevier B.V.; 2017;19:2045-59.
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

51. Lundberg A, Zhang M, Aggarwal R, Li H, Zhang L, Foye A, et al. The genomic and epigenomic landscape of double-negative metastatic prostate cancer. Cancer Res. 2023;
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

52. Meyer RC, Giddens MM, Schaefer SA, Hall RA. GPR37 and GPR37L1 are receptors for the neuroprotective and glioprotective factors prosaptide and prosaposin. Proc Natl Acad Sci U S A [Internet]. Proc Natl Acad Sci U S A; 2013 [cited 2023 Apr 6];110:9529-34. Available from: https://pubmed.ncbi.nlm.nih.gov/23690594/

53. Vierstra J, Stamatoyannopoulos JA. Genomic footprinting. Nat Methods [Internet]. Nature Publishing Group; 2016;13:213-21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26914205
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

54. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Berhanu Lemma R, Turchi L, Blanc-Mathieu R, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res [Internet]. Nucleic Acids Res; 2022 [cited 2022 Nov
  Pubmed: Author and Title
  Google Scholar: Google Scholar Search

29];50:D165-73. Available from: https://pubmed.ncbi.nlm.nih.gov/34850907/

55. Nouruzi S, Ganguli D, Tabrizian N, Kobelev M, Sivak O, Namekawa T, et al. ASCL1 activates neuronal stem cell-like lineage programming through remodeling of the chromatin landscape in prostate cancer. Nat Commun. Nature Research; 2022;13.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

56. Bishop JL, Thaper D, Vahid S, Davies A, Ketola K, Kuruma H, et al. The master neural transcription factor BRN2 is an androgen receptor-suppressed driver of neuroendocrine differentiation in prostate cancer. Cancer Discov. American Association for Cancer Research Inc.; 2017;7:54-71.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

57. FitsGerald PC, Shlyakhtenko A, Mir AA, Vinson C. Clustering of DNA sequences in human promoters. Genome Res [Internet]. Genome Res; 2004 [cited 2023 Mar 21];14:1562-74. Available from: https://pubmed.ncbi.nlm.nih.gov/15256515/

58. Sahu B, Laakso M, Ovaska K, Mirtti T, Lundin J, Rannikko A, et al. Dual role of FoxA1 in androgen receptor binding to chromatin, androgen signalling and prostate cancer. EMBO J. 2011;30:3962-76.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

59. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. Nature. 2012;487:239-43.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

60. Bery F, Cancel M, Gueguinou M, Potier-Cartereau M, Vandier C, Chantome A, et al. Zeb1 and SK3 Channel Are Up-Regulated in Castration-Resistant Prostate Cancer and Promote Neuroendocrine Differentiation. Cancers (Basel) [Internet]. Cancers (Basel); 2021 [cited 2023 Mar 22];13. Available from: https://pubmed.ncbi.nlm.nih.gov/34204608/

61. Lee GT, Rosenfeld JA, Kim WT, Kwon YS, Palapattu G, Mehra R, et al. TCF4 induces enzalutamide resistance via neuroendocrine differentiation in prostate cancer. PLoS One [Internet]. PLoS One; 2019 [cited 2023 Mar 22];14. Available from: https://pubmed.ncbi.nlm.nih.gov/31536510/

62. Cmero M, Kurganovs NJ, Stuchbery R, McCoy P, Grima C, Ngyuen A, et al. Loss of SNAI2 in Prostate Cancer Correlates With Clinical Response to Androgen Deprivation Therapy. JCO Precis Oncol [Internet]. JCO Precis Oncol; 2021 [cited 2023 Mar 22];5:1048-59. Available from: https://pubmed.ncbi.nlm.nih.gov/34322653/

63. Zhao Y, Vartak S v., Conte A, Wang X, Garcia DA, Stevens E, et al. "Stripe" transcription factors provide accessibility to co-binding partners in mammalian genomes. Mol Cell. Cell Press; 2022;82:3398-3411.e11.

Pubmed: Author and Title
Google Scholar: Google Scholar Search

64. Che M, Chaturvedi A, Munro SA, Pitzen SP, Ling A, Zhang W, et al. Opposing transcriptional programs of KLF5 and AR emerge during therapy for advanced prostate cancer. Nat Commun [Internet]. Nature Research; 2021;12:6377. Available from: http://www.ncbi.nlm.nih.gov/pubmed/34737261

Pubmed: Author and Title
Google Scholar: Google Scholar Search

65. Cui J, Liu J, Fan L, Zhu Y, Zhou B, Wang Y, et al. A zinc finger family protein, ZNF263, promotes hepatocellular carcinoma resistance to apoptosis via activation of ER stress-dependent autophagy. Transl Oncol [Internet]. Neoplasia Press, Inc.; 2020;13:100851. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32898766

Pubmed: Author and Title
Google Scholar: Google Scholar Search

66. Yu Z, Feng J, Wang W, Deng Z, Zhang Y, Xiao L, et al. The EGFR-ZNF263 signaling axis silences SIX3 in glioblastoma epigenetically. Oncogene [Internet]. Springer Nature; 2020;39:3163-78. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/32051553
   Pubmed: <u>Author and Title</u>
   Google Scholar: <u>Google Scholar Search</u>

67. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature [Internet]. 2012;489:75-82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22955617
   Pubmed: <u>Author and Title</u>
   Google Scholar: <u>Google Scholar Search</u>

68. Samarasinghe KTG, Jaime-Figueroa S, Burgess M, Nalawansha DA, Dai K, Hu Z, et al. Targeted degradation of transcription factors by TRAFTACs: TRAnscription Factor TArgeting Chimeras. Cell Chem Biol. Elsevier Ltd; 2021;28:648-661.e5.
   Pubmed: <u>Author and Title</u>
   Google Scholar: <u>Google Scholar Search</u>

69. Henley MJ, Koehler AN. Advances in targeting 'undruggable' transcription factors with small molecules. Nat Rev Drug Discov. Nature Research; 2021. page 669-88.