

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
<input type="checkbox"/>	<input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input type="checkbox"/>	<input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> A description of all covariates tested
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input type="checkbox"/>	<input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
<input type="checkbox"/>	<input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Fluorescence images were captured using a Leica TCS SP5 Confocal Laser Scanning Microscope (Leica Microsystems) and image acquisition was performed with Leica Application Suite Advanced Fluorescence (LAS AF v2.6.0). H&E images were captured using Olympus BX 61 VS Slide Scanner (Olympus) and image acquisition was performed with Olympus VS-ASW (v2.5) software. Tumor images were captured through a stereo microscope (Olympus, SZX16) with digital camera (Olympus, DP71) and image acquisition was performed with Olympus DP Controller (v3.3.1.292). Western blots were imaged using a ChemiDoc MP Imaging System (Biorad, 17001402) or exposed to X-Ray films (Research Products International, 248300). Luminescence was measured using GloMax® Explorer multimode plate reader (Promega, v3.1.0). Flow sorting was performed on an BD Influx™ cell sorter (BD Biosciences, X64650000124). Flow Cytometer data was collected with FACSCanto II Flow Cytometer (BD Bioscience). Single cells or single nuclei were loaded onto Chromium X Controller (10X Genomics) for library preparation. Tissue microarray slides were imaged in the Vectra Polaris Automated Quantitative Pathology Multispectral Imaging System (Akoya Biosciences) and post-acquisition were processed by InForm software (Akoya Biosciences, v3.1), and the tiles were stitched using Halo AI (Indica Labs, v3.6). Tissue segmentation of the images was performed using a deep learning classifier by training the algorithm “DenseNet V2” from the Halo AI plug-in (Indica Labs, v3.6). To detect proliferation and apoptosis in tumor sections, tiled scans of whole tumor tissue sections in high-resolution were acquired using a Leica STELLARIS 5 Confocal Microscope (Leica Microsystems) and image acquisition was performed with Leica Application Suite X (LAS X v4.5.0).
Data analysis	All software and software versions used to analyze data are described in the Methods in the relevant sections. The following is a list of software used: ImageJ (NIH, v1.52K), Prism (GraphPad Software, v10.5.0), BD FACS™ Software (BD Biosciences, v1.2.0.142), FlowJo (BD, v10.8.2). The initial processing of raw single cell sequencing data was processed using Cell Ranger (10x Genomics) with different versions: single-cell RNA sequencing (scRNA-seq) using Cell Ranger (v2.1.1: MJ002 and MJ004; v3.0.2: MJ005, MJ007, MJ008 and MJ012; v5.0.1: MJ014 and MJ015), single-cell ATAC sequencing (scATAC-seq) using Cell Ranger ATAC (v1.0.1) and single-nuclei multiome ATAC and RNA sequencing using Cell Ranger ARC (v1.0.0: MJ018, MJ019, MJ020, MJ021, MJ022 and MJ023; v2.0.2: MJ024 and MJ025). scRNA-seq and snRNA-seq were

analyzed independently using scanpy (v1.9.1) for Python (v3.9). Doublet detection and removal was performed by Scrublet (v0.2.2) algorithm and cells that have passed quality control was subsampled using the subsample function implemented in scanpy (v1.9.1). A sample-specific regulatory network (interactome) was reverse engineered using ARACNe algorithm (version ARACNe-AP). Mouse prostate cancer scRNA-seq and snRNA-seq gene expression profiles were scaled independently and transformed to protein activity profiles using VIPER algorithm implemented in pyVIPER (vipier-in-python v1.0.9). All samples were merged to be projected on a 2-dimensional plane using Diffusion Maps implemented in scanpy (v1.9.1). To recover cell identities, clusters of cells that share the same regulatory programs were identified using acdc_py (v1.1.0). Cut&Tag data analysis was performed using the following packages: cutadapt (v3.6), BOWTIE2 (v2.4.2), sambamba (v1.0.1), SEACR (v1.3), MACS2 (v2.2.8), bedtools (v2.27.1), MEME suit (v5.5.7), deepTools (v3.5.5) and IGV (v2.13.0). Coverage tracks were generated by the bamCoverage function of deepTools (v3.5.5). Heatmap and enrichment plots of histone mark peaks were generated using the computeMatrix and plotHeatmap function implemented in deepTools (v3.5.5). To integrate CUT&Tag and VIPER analysis of snRNA-seq data, histone mark count matrices were processed using limma voom (v3.54.2). Linear models were fitted for each gene based on the voom transformed data with moderated t-statistics computed using the eBayes method from limma. Top-ranked genes were extracted using the topTable function. Differential gene expression was analyzed in Seurat (v4.1.3), with parameter test.use set to DESeq2 (v1.28.0) in the FindMarkers function. To integrate CUT&Tag and bulk RNA-seq data, RNA-seq reads were mapped to the genome using HISAT2 (v2.1.0). The mapped reads count of each gene was measured by featureCounts (v1.6.1). The RNA-seq reads count matrix was combined with the CUT&Tag signal reads count matrix for all gene loci in R (v4.1.2). Multi-drug synergy was calculated based on Bliss reference model using SynergyFinder (v3.0). For estimation of outlier measurements, the cNMF algorithm implemented in SynergyFinder (v3.0) was utilized. To evaluate the association of NSD2 with overall survival in mCRPC, transcriptomes were aligned to the human reference genome (GRCh37/hg19) using TopHat2 (v2.0.7). Gene expression was calculated using Cufflinks (v2.2.1). The computational analysis of multiplex images was performed by machine learning using Halo AI (Indica Labs, v3.6). To identify and count proliferating or apoptotic cells on tumor sections, a machine learning classifier was trained using QuPath (v0.5.1). Computer code to reproduce some of the panels in the manuscript is available at <https://github.com/VasciaveoLab/nsd2-paper-at-nature>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

- To determine the identity of cell clusters defined by the VIPER algorithm based on differential activity of regulatory proteins, we inferred VIPER activity on clusters using a published NEPC gene signature. A 29 gene NE signature was used to examine the correlation of NSD2 expression with NE gene expression in the Royal Marsden Hospital (RMH) cohort and as well as the Prostate Cancer Foundation-Stand Up to Cancer (PCF-SU2C) cohort. To evaluate differences in gene expression of NSD1, NSD2, and NSD3 in human prostate tumor samples, we analyzed bulk RNA-seq from 49 CRPC patients (15 CRPC-NE and 34 CRPC-Adeno) in a published dataset. The NEPC gene signature and related dataset is deposited in dbGap phs000909.v.p1 and is accessible through the cBioPortal for Cancer Genomics (www.cbioportal.org).
- To evaluate the association of NSD2 with overall survival in mCRPC, two independent mCRPC biopsy RNA-seq cohorts were used: 1) A cohort of 141/159 mCRPC transcriptomes generated by the SU2C/PCF Prostate Cancer Dream Team. This dataset is available at www.cbioportal.org and in GitHub https://github.com/cBioPortal/datahub/tree/master/public/prad_su2c_2019. 2) A cohort of 94/95 mCRPC transcriptomes from patients treated at the Royal Marsden Hospital. This dataset has been deposited in the European Genome-phenome Archive (EGA) with accession number EGAS50000001269.
- We also analyzed a published single-cell RNA-seq dataset. Expression data from tumor cells were downloaded from GEO with accession number GSE264573 and supplementary file msk.integrated.remove.cellcycle.tumor.cells.
- CUT&Tag motif analysis was performed using the Simple Enrichment Analysis (SEA) function from MEME suit (v5.5.7) with the JASPAR 2022 Core motif database. The dataset was downloaded from JASPAR: <https://jaspar2022.genereg.net/downloads/>. Coverage tracks were generated by the bamCoverage function of deepTools with bin size as 50 bp, with problematic ENCODE regions (the ENCODE blacklist for mm10). The mm10 blacklist was downloaded from: <https://www.encodeproject.org/files/ENCFF547MET/>.
- A sample-specific regulatory network (interactome) was reverse engineered using ARACNe-AP and Regulatory proteins (RP) were selected into manually curated protein sets, including Transcription Factors (TF), co-Transcription Factors or chromatin remodeling enzymes, based on the Gene Ontology (GO) identifiers GO:0003700 and GO:0003712.
- Gene Set Enrichment Analysis (GSEA) on pseudo-bulk from snRNA-seq data of various treatment conditions was performed using a canonical AR target signature, which can be found in the online Methods Summary from: <https://www.nature.com/articles/nature13229#Sec2>.
- Raw sequencing data and count matrixes from the current work have been deposited to the Gene Expression Omnibus (GEO) under accession number GSE237197.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

Paraffin blocks of tumor samples from 64 male patients treated for localized or metastasis prostate cancer at Weill Cornell Medical Center between 1997-2019 were made into tissue microarrays (TMAs).

Population characteristics

Prostate cancer patients either underwent androgen deprivation therapy (ADT) or without ADT. Information for prostate cancer samples is shown in Supplementary Table 4.

Recruitment

All prostate cancer patients were seen at Weill Cornell Medical Center during their clinical care. There is unlikely to be any selection bias.

Ethics oversight

All studies were conducted under protocols approved by Weill Cornell Medical Center. All prostate cancer patients or families provided informed consent for research use of biospecimens and clinical data under an institutional approved protocol (IRB #1008011210).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences

☐ Behavioural & social sciences

☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

1. The number of NPP53 mice used for organoid establishment was based on the objective of capturing the broad spectrum of phenotypic heterogeneity observed in human Castration-Resistant Prostate Cancer (CRPC).
2. To compare mean fluorescence intensity of histone marks between neuroendocrine (NE) and non-neuroendocrine (nonNE) cells, we found that a sample size of 20-80 observations per group were able to yield consistent results in t-tests.
3. To compare the expression of NSD1, NSD2 and NSD3 across different cell subtypes in human treatment-resistant prostate cancer, we analyzed 20.1K CRPC, 11.6K CRPC-NE, and 4.0K treatment naïve/CSPC single cells across 21 patients and tissue samples, resulting in strong statistical power in comparing gene expression across subtypes.
4. Western blots comparing the levels of histone mark and histone methyltransferase expression were performed on 4 NE and 4 nonNE organoids. The four independent biological replicates provide a robust assessment of the true biological variation. These experiments were repeated three times on three different batches of samples and consistent results were obtained.
5. Analysis of NSD2 and H3K36me2 levels in a prostate cancer tissue microarray (TMA) were performed using multiplex imaging combined with unbiased machine learning at the single-cell level. The collection of samples contained 33 primary PCa, 6 de novo NEPC, 18 mCRPC, and 6 CRPC-NE patients, which yields statistical significance for group comparisons. Although the statistical power is reduced for de novo NEPC, which has a relatively smaller sample size, it does not affect the major conclusions of the current manuscript.
6. We performed Kaplan-Meier plots of overall survival in two independent human mCRPC patient cohorts, RMH and PCF-SU2C. The sample sizes in the RMH cohort (n=28) and PCF-SU2C (n=27) provides sufficient statistical power for comparisons.
7. In our Gene Set Enrichment Analyses (GSEA) of sgControl and sgNSD2 treated organoids or empty vector (EV) and H3.3K36M transfected organoids, the nominal P-value was determined by comparing the observed enrichment score to the null distribution generated through 1,000 random permutations of gene labels. A sample size of 1,000 permutations is generally considered to have a significance level of $\alpha = 0.05$.
8. For comparisons of organoid growth curves and dose-response curves, sample sizes were chosen to generate at least 3 biological replicates for each condition, which yielded sufficient data for statistical analysis.
9. To evaluate tumor response to AR inhibition in NSD2 knock-out (sgNSD2) and control (sgControl) or H3.3K36M and empty vector (EV) transfected xenografts, we compared tumor growth curves in cohorts of xenografted mice treated with either enzalutamide or vehicle control, with each cohort having 5 mice, which was sufficient to achieve statistical significance for large treatment effects.
10. To test whether NSD2i effectively targets NSD2, NSD2i and DMSO control treatments were tested on 5 different human CRPC organoid lines of different subtypes. The five independent biological replicates provide a robust assessment of the true biological variation.
11. For a pilot study of NSD2i dosing, three doses of NSD2i and vehicle control were tested on xenografted mice using two mice for each treatment condition, which provided an approximate, non-statistical impression of potential dose-response trends.
12. Bliss synergy analysis was performed using bootstrapping, in which replicates were drawn from the original dataset with each simulation having its own Bliss synergy score, which creates a distribution of scores instead of a single value. This resampling technique generates a robust statistical assessment of the synergy score.
13. To test the efficacy of NSD2i and enzalutamide co-treatment in inhibiting tumor growth in vivo, we compared tumor growth curves between cohorts of xenografted mice under different treatment conditions, with each cohort containing 6 mice, which was sufficient to achieve statistical significance and to reduce the number of mice needed for experimentation.
14. To compare the expression of NSD1, NSD2, and NSD3 in a human clinical dataset, sample sizes of 15 CRPC-NE and 34 CRPC-Adeno samples were sufficient to create statistically informative box plots.
15. Quantitative comparisons of H3K36me2, H3K27me3, and H3K27ac CUT&Tag signals at genomic domains marked by H3K36me2 were performed in four NE and four nonNE organoid lines. The analysis of 4 biological replicates exceeds the general recommendation for three biological replicates for these types of analyses.
16. To detect correlation between NSD2 expression and a neuroendocrine signature, we performed a Spearman correlation test using two independent CRPC cohorts, with the sample size of 159 in the PCF-SU2C cohort exceeding the recommended 149 samples for a moderate correlation ($r = 0.3$) [1] and reaching statistical significance ($p = 3 \times 10^{-5}$). The lack of statistical significance ($p = 0.1$) in the RMH dataset is likely due to the smaller number of samples ($n = 95$) in this cohort [1].
17. Western blot analyses of histone mark and histone methyltransferase levels after NSD2 knock-out were performed on 4 NPPO organoid lines (source data shown in Supplementary Figure 1), exceeding the recommended minimum of three for reliable and statistically valid comparisons. The effect of H3.3K36M transfection were examined in 4 NPPO organoid lines, but only 2 NPPO organoid lines could be analyzed due to the death of other 2 organoid lines after H3.3K36M transfection. The sample size of two in this study represents the minimum of two biological replicates for a reliable comparison.
18. CUT&Tag comparison of histone marks in NSD2 knock-out or H3.3K36M transfected organoids were performed on two biological replicates. Statistical analysis was performed on 20,818 tested genomic regions between sgCtrl and sgNsd2 or between empty vector (EV) and

H3.3K36M-transfected NPPO organoids. The experiment was repeated 3 times with reproducible results.

19. Comparison of the percentages of proliferating and apoptotic cells under four treatment conditions in MSKPCa10, MSKPCa14, and WCM1262 grafts was performed using 6 mice/group, exceeding the recommended minimum of three for reliable and statistically valid comparisons.

Reference

1. Bujang, M.A. (2024). An elaboration on sample size determination for correlations based on effect sizes and confidence interval width: a guide for researchers. *Restor Dent Endod*, 49(2), e21. doi: 10.5395/rde.2024.e21.

Data exclusions

The exclusion criteria for data was established prior to the experiment. In mouse tumor xenografts, the criterion for tumor volume was ~250 mm³ at week two of grafting. If the tumor volume was not in the range (240-260 mm³), the animal was excluded from drug treatment experiments. In human tumor xenografts, the criterion for tumor volume was ~80 mm³ at week two of grafting. If the tumor volume was not in the range (60-100 mm³), the animal was excluded from drug treatment experiments. In the human tissue microarray analysis, a single CRPC-NE sample that stained negative for Chromogranin A and Synaptophysin was excluded for further analysis.

Replication

All experiments described in this manuscript include sufficient biological replicates to draw statistically meaningful conclusions. All experiments have been repeated at least three times, and the results are reproducible. All western blots, H&E, and immunofluorescence staining have been performed at least three times on different batches of samples to ensure that the results are reproducible. The representative images shown in the manuscript correspond to reproducible general conclusions.

Randomization

Animals were randomly assigned to groups for treatment. Randomization was applied in designing the grid layout for human tissue microarray. Cores were dispersed randomly in the grid and, if taking multiple cores per case, each core from a case was allocated to a different recipient block.

Blinding

The investigators were blinded to data collection and data analysis. All histological analyses were blinded.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | | |
|-------------------------------------|-----------------------------------------------------------------|
| n/a | Involved in the study |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | | |
|-------------------------------------|----------------------------------------------------|
| n/a | Involved in the study |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> ChIP-seq |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

Antibodies used

All antibodies used are described in the Methods in the relevant sections and are summarized in Supplementary Table 6. The following is the list of antibodies used: rabbit anti-H2AK119ub (Cell Signaling Technology; cat# 8240; clone D27C4; lot# 8; 1:400); rabbit anti-H2BK120ub (Cell Signaling Technology; cat# 5546; 1:400); rabbit anti-H3K4me1 (Cell Signaling Technology; cat# 5326; clone D1A9; lot# 5; 1:400); rabbit anti-H3K4me3 (Active motif; cat# 39159; 1:400); rabbit anti-H3K9me2 (Cell Signaling Technology; cat# 4658; clone D85B4; lot# 10; 1:400); rabbit anti-H3K9me3 (Active motif; cat# 39765; lot# 8210001; 1:400); rabbit anti-H3K18ac (Abcam; cat# ab1191; lot# GR186537-1; 1:400); rabbit anti-H3K27ac (Active motif; cat# 39133; lot# 6921014; 1:400); rabbit anti-H3K27ac (Cell Signaling Technology; cat# 8173; clone D5E4; lot# 6; 1:100); rabbit anti-H3K27me3 (Cell Signaling Technology; cat# 9733; clone C36B11; lot# 16&19; 1:400); rabbit anti-H3K36me2 (Cell Signaling Technology; cat# 2901; clone C75H12; lot# 5; 1:400); rabbit anti-H3K36me2 (Cell Signaling Technology; ab176921; clone EPR16994(2); lot# GR252916-7; 1:100); rabbit anti-H3K36me3 (Active Motif; cat# 61101; lot# 28818005; 1:100); rabbit anti-H3K79me2 (Abcam; cat# ab3594; lot# GR3231418-1; 1:400); rabbit anti-H4K16ac (Cell Signaling Technology; cat# 13534; clone E2B8W; lot# 3; 1:400); mouse anti-5-methylcytosine (Active Motif; cat# 39649; clone 33D3; lot# 339118025; 1:400); rabbit anti-5-hydroxymethylcytosine (Active Motif; cat# 39769; lot# 21518003; 1:400); rabbit anti-Chromogranin A (Abcam; cat# ab15160; lot# GR3205971-2; 1:400); mouse anti-HNF-3 beta/FoxA2 (Novus Biologicals; cat# H00003170-M12; clone 6C12; lot# HB231-6C12; 1:400); mouse anti-FOXA2 (Abcam; cat# ab60721; clone 7E6; lot# GR3357851-1; 1:400); rabbit anti-FOXA2 (Abcam; cat# ab108396; clone EPR4465; lot# GR211960-5; 1:400); mouse anti-Synaptophysin (BD Biosciences; cat# 611880; clone 2/Synaptophysin; lot# 2045364; 1:100); rabbit anti-Androgen Receptor (Abcam; cat# ab133273; clone EPR1535(2); lot# GR3271456-1; 1:100); chicken anti-Vimentin (Abcam; cat# ab24525; lot# GR3305913-2; 1:400); rat anti-Ki67 (Thermo Fisher Scientific; cat# 14-5698-82; clone SolA15; lot# 4328926; 1:400); rat anti-Cytokeratin 8 (Developmental Studies Hybridoma Bank; cat# TROMA-I; clone TROMA-I; 1:100); mouse anti-WHSC1/NSD2 (Abcam; cat# ab75359; clone 29D1; lot# GR3393997-4; 1:200); rabbit anti-BSD antibody (Abcam; cat# ab38307; lot# GR23921-47; 1:400); mouse anti-HA-Tag (Cell Signaling Technology; cat# 2367; clone 6E2; lot# 5; 1:400); Mouse anti-CD56 (Cell Signaling Technology; cat# 3576; Clone 123C3; lot# 9; 1:200); Rabbit anti-CD56 (Cell Signaling Technology; cat# 99746; Clone E7X9M; lot# 3; 1:400); Rabbit anti-Cleaved Caspase-3 (Asp175) (Cell Signaling Technology; cat# 9579; Clone D3E9; lot# 1; 1:400); Mouse anti-Cytokeratin 8+18 antibody (abcam; cat# ab17139; Clone

5D3, lot# 1055063-6; 1:400); Rabbit anti-Cytokeratin 8+18 antibody (abcam; cat# ab53280, Clone EP1628Y, lot# CR3176229-1; 1:400); Rabbit anti-WHSC1L1/NSD3 antibody (Cell Signaling Technology; cat# 92056, Clone D4N9N, lot# 1; 1:400); Rabbit anti-Rb antibody (Cell Signaling Technology; cat# 9313, Clone D20, lot# 7; 1:400); Rabbit anti-Phospho-Rb (Ser807/811) antibody (Cell Signaling Technology; cat# 8516, Clone D20B12, lot# 11; 1:400); Rabbit anti-Histone H3 (abcam; cat# ab1791; 1:1000); Mouse anti- β -Actin (ACTB) Antibody (Millipore Sigma; cat# A2228, Clone AC-74; 1:400); Normal Rabbit IgG (Cell Signaling Technology; cat# 2729; lot# 9; 1:100); Guinea Pig anti-Rabbit IgG (Heavy & Light Chain) (Antibodies-online; cat# ABIN101961; lot# 43586; 1:100); Cy5.5[®] Conjugated mouse anti-HA-Tag (Cell Signaling Technology; Cat. #62145; Clone 6E2; 1:400); Alexa Fluor[™] 488 Goat anti-Rabbit IgG (H +L) (Thermo Fisher Scientific; cat# A-11008; 1:400); Alexa Fluor[™] 488 Goat anti-Mouse IgG (H+L)(Thermo Fisher Scientific; cat# A28175; 1:400); Alexa Fluor[™] Plus 555 Goat anti-Mouse IgG (H+L) (Thermo Fisher Scientific; cat# A32727; 1:400); Alexa Fluor[™] 555 Goat anti-Rabbit IgG (H+L) (Thermo Fisher Scientific; cat# A-21428; 1:400); Alexa Fluor[™] 555 Goat anti-Rat IgG (H+L) (Thermo Fisher Scientific; cat# A-21434; 1:400); Alexa Fluor[™] 555 Goat anti-Chicken IgY (H+L) (Thermo Fisher Scientific; cat# A-21437; 1:400); Alexa Fluor[™] 647 Goat anti-Chicken IgY (H+L) (Thermo Fisher Scientific; cat# A-21449; 1:400).

Validation

We used commercial antibodies. All of these were quality-controlled and validated by the manufacturer. Validation statements can be found on the manufacturer's website: cellsignal.com, activemotif.com, abcam.com, novusbio.com, bdbiosciences.com, thermofisher.com, dshb.biology.uiowa.edu, sigmaaldrich.com. No further validation was performed.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

Mouse tumor organoid lines were established from male Nkx3.1CreERT2/+; Ptenflox/flox; TrpP53flox/flox; Rosa26-EYFP (NPP53) mice or Trp53flox/flox; Rb1flox/flox; Ptenflox/flox (TKO) mice. The human MSKPCa10, MSKPCa14 and MSKPCa2 organoid lines were provided by Dr. Yu Chen (Memorial Sloan Kettering Cancer Center). The human WCM1262 and WCM154 organoid lines were provided by Dr. Himisha Beltran (Dana-Farber Cancer Institute, Harvard Medical School). Drosophila Schneider 2 (S2) cells were purchased from Thermo Fisher Scientific (cat# R69007). The Gibco[™] Drosophila S2 cells was derived from a primary culture of late stage (20-24 hours old) Drosophila melanogaster embryos. Many features of the S2 cell line suggest that it is derived from a macrophage-like lineage.

Authentication

The genetic background of mouse organoid lines has been verified by genotyping. The genetic background of human organoid lines was determined by MSK-IMPACT sequencing and confirmed to match parental tumor features. The histological features of mouse and human organoid lines have been evaluated by a pathologist and characterized using specific markers (Supplementary Table 2). Each lot of Gibco[™] Drosophila S2 cells is tested for cell growth and viability post-recovery from cryopreservation. Additionally, the Master Seed Bank has been tested for contamination of bacteria, yeast, mycoplasma, and virus, and has been characterized by isozyme and karyotype analysis.

Mycoplasma contamination

All cell line used were tested negative for mycoplasma contamination.

Commonly misidentified lines (See [ICLAC](#) register)

None.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

The Nkx3.1CreERT2/+; Ptenflox/flox; TrpP53flox/flox; Rosa26-EYFP (NPP53) mice maintained on a mixed C57BL/6-129Sv background (3-5 months of age) were tamoxifen induced for 4 consecutive days. The survival time of tumor-bearing NPP53 mice in this study ranged from 228 to 435 days after tamoxifen induction (tumor bearing duration: 228-435 days). In addition, adult male NOD/SCID (NOD.CB17-Prkdc Scid/J) mice (18-23 body weight; 6-8 weeks of age) were used in this study.

Wild animals

No wild animals were used in this study.

Reporting on sex

Although the prostate is biologically a male organ, the findings in this study have broad implications that are not limited to tumors of the male genitourinary system. All experimental animals used to generate tumors were male in order to provide a natural androgen-regulated microenvironment. No sex-disaggregated data were collected and analyzed in this study.

Field-collected samples

No field-collected samples were used in this study.

Ethics oversight

All experiments using animals were performed according to protocols approved by the Institutional Animal Care and Use Committee (IACUC) at Columbia University Irving Medical Center (USA).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

- ☒ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

The raw and processed Cut&Tag data were deposited to GEO under accession number GSE308575.

Files in database submission

GSM9248517 NPPO-1NE, H3K36me2, Cut&Tag
GSM9248518 NPPO-2, H3K36me2, Cut&Tag
GSM9248519 NPPO-4, H3K36me2, Cut&Tag
GSM9248520 NPPO-6, H3K36me2, Cut&Tag
GSM9248521 NPPO-1nonNE, H3K36me2, Cut&Tag
GSM9248522 NPPO-7, H3K36me2, Cut&Tag
GSM9248523 NPPO-8, H3K36me2, Cut&Tag
GSM9248524 NPPO-9, H3K36me2, Cut&Tag
GSM9248525 NPPO-1NE, H3K27me3, Cut&Tag
GSM9248526 NPPO-2, H3K27me3, Cut&Tag
GSM9248527 NPPO-4, H3K27me3, Cut&Tag
GSM9248528 NPPO-6, H3K27me3, Cut&Tag
GSM9248529 NPPO-1nonNE, H3K27me3, Cut&Tag
GSM9248530 NPPO-7, H3K27me3, Cut&Tag
GSM9248531 NPPO-8, H3K27me3, Cut&Tag
GSM9248532 NPPO-9, H3K27me3, Cut&Tag
GSM9248533 NPPO-1NE, H3K27ac, Cut&Tag
GSM9248534 NPPO-2, H3K27ac, Cut&Tag
GSM9248535 NPPO-4, H3K27ac, Cut&Tag
GSM9248536 NPPO-6, H3K27ac, Cut&Tag
GSM9248537 NPPO-1nonNE, H3K27ac, Cut&Tag
GSM9248538 NPPO-7, H3K27ac, Cut&Tag
GSM9248539 NPPO-8, H3K27ac, Cut&Tag
GSM9248540 NPPO-9, H3K27ac, Cut&Tag
GSM9248541 NPPO-7, IgG1, Cut&Tag
GSM9248542 NPPO-7, IgG2, Cut&Tag
GSM9248543 NPPO-7, IgG3, Cut&Tag
GSM9248544 NPPO-1NE-sgControl, IgG, Cut&Tag
GSM9248545 NPPO-1NE-sgControl, H3K36me2, Cut&Tag
GSM9248546 NPPO-1NE-sgControl, H3K36me3, Cut&Tag
GSM9248547 NPPO-1NE-sgControl, H3K27me3, Cut&Tag
GSM9248548 NPPO-1NE-sgNsd2, H3K36me2, Cut&Tag
GSM9248549 NPPO-1NE-sgNsd2, H3K36me3, Cut&Tag
GSM9248550 NPPO-1NE-sgNsd2, H3K27me3, Cut&Tag
GSM9248551 NPPO-2-sgControl, IgG, Cut&Tag
GSM9248552 NPPO-2-sgControl, H3K36me2, Cut&Tag
GSM9248553 NPPO-2-sgControl, H3K36me3, Cut&Tag
GSM9248554 NPPO-2-sgControl, H3K27me3, Cut&Tag
GSM9248555 NPPO-2-sgNsd2, H3K36me2, Cut&Tag
GSM9248556 NPPO-2-sgNsd2, H3K36me3, Cut&Tag
GSM9248557 NPPO-2-sgNsd2, H3K27me3, Cut&Tag
GSM9248558 NPPO-4-EV, IgG, Cut&Tag
GSM9248559 NPPO-4-EV, H3K36me2, Cut&Tag
GSM9248560 NPPO-4-EV, H3K36me3, Cut&Tag
GSM9248561 NPPO-4-EV, H3K27me3, Cut&Tag
GSM9248562 NPPO-4-H3.3K36M, H3K36me2, Cut&Tag
GSM9248563 NPPO-4-H3.3K36M, H3K36me3, Cut&Tag
GSM9248564 NPPO-4-H3.3K36M, H3K27me3, Cut&Tag
GSM9248565 NPPO-6-EV, IgG, Cut&Tag
GSM9248566 NPPO-6-EV, H3K36me2, Cut&Tag
GSM9248567 NPPO-6-EV, H3K36me3, Cut&Tag
GSM9248568 NPPO-6-EV, H3K27me3, Cut&Tag
GSM9248569 NPPO-6-H3.3K36M, H3K36me2, Cut&Tag
GSM9248570 NPPO-6-H3.3K36M, H3K36me3, Cut&Tag
GSM9248571 NPPO-6-H3.3K36M, H3K27me3, Cut&Tag
GSM9248572 NPPO_1NE, DMSO, Cut&Tag
GSM9248573 NPPO_1NE, KTX, Cut&Tag
GSM9248574 NPPO_1NE, IgG, Cut&Tag

Genome browser session (e.g. [UCSC](http://genome.ucsc.edu))

http://genome.ucsc.edu/s/Jia%20Jessie%20Li/NSD2_CRPC%2DNE

Methodology

Replicates

Experiments were performed three times and consistent results were obtained.

Sequencing depth

Sequencing was performed on pooled libraries. Cut&Tag detecting various histone marks in mouse NE and non-NE Npp53 organoids: 10M reads per individual library; 35 libraries; 350M reads in total; 2x150bp paired-end sequencing. Cut&Tag comparing the level of histone marks in sgNsd2 vs. sgControl or H3.3K36M vs. Control organoids: 12M reads per individual library; 28 libraries; ~350M reads in total; 2x150bp paired-end sequencing. Cut&Tag comparing the level of H3K36me2 mark in DMSO and NSD2i treated NPPO-1NE organoids: 12M reads per individual library; 3 libraries; 36M reads in total; 2x150bp paired-end sequencing.

Antibodies	Anti-rabbit H3K36me2 (Abcam; cat# ab176921; clone EPR16994(2); lot# GR252916-7); Anti-rabbit H3K27me3 (Cell Signaling; cat# 9733; clone C36B11; lot# 16); Anti-rabbit H3K36me3 (Active Motif; cat# 61101; lot# 28818005); Anti-rabbit H3K27ac (Cell Signaling; cat# 8173; clone D5E4; lot# 6); Normal Rabbit IgG (Cell Signaling; cat# 2729; lot# 9); Guinea Pig anti-Rabbit IgG (Heavy & Light Chain) (Antibodies-online; cat# ABIN101961; lot# 43586).
Peak calling parameters	Peak calling was performed using SEACR (H3K36me2) or MACS2 (H3K27ac) with IgG input as control.
Data quality	CUT&Tag reads were mapped to the mouse genome assembly mm10 using Bowtie2 (v2.4.2). Potential PCR duplicates were removed by the function markup of sambamba (v1.0.1).
Software	cutadapt (v3.6) Bowtie2 (v2.4.2) sambamba (v1.0.1) SEACR (v1.3) MACS2 (v2.2.8) bedtools (v2.27.1) MEME suit (v5.5.7) deepTools (v3.5.5) IGV (v2.13.0)

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For flow sorting experiments, organoids were incubated with prewarmed TrypLE at 37 °C for 10 minutes. It was then 1:10 neutralized with PBS and 5% CS-FBS. Spin down at 1000 rpm for 1 minute, organoids were resuspended with PBS and dissociated into single cells by gentle pipetting. Filter through 40 µm cell strainer (Corning, 431750) three times. Spin down at 1000 rpm for 5 minutes. Cell pellets were resuspended with PBS and 2% CS-FBS. After filtering through a Falcon™ Tube with 35 µm cell strainer cap (Corning, 352235), cell number was counted in a TC20 automated cell counter. Adjust the volume to make the final cell concentration of 5000 cells/µl. For flow cytometry analysis, organoids were dissociated into single cells using protocols mentioned above. Resuspend cells in 100 µl 4% paraformaldehyde per 1 million cells. Fix for 15 minutes at room temperature (RT). Neutralize with 1ml PBS and centrifuge to collect cells. Wash with PBS once. Resuspend cells in 0.5 ml PBS. Permeabilize cells by adding 0.5 ml 1% Triton-X 100 slowly to the cells, while gentle vortexing, to a final concentration of 0.5% Triton-X 100. Permeabilize for 10 minutes at RT. Wash cells by centrifugation in 10ml PBS to remove Triton-X 100. Antibody was diluted in 0.5% BSA PBS buffer. Resuspend cells in 100 µl 1:50 diluted Cy5.5® Conjugated mouse anti-HA-Tag antibody (Cell Signaling Technology; Cat. #62145; Clone 6E2). Incubate for 1 hour in the dark at RT. Wash by centrifugation in 0.5% BSA PBS buffer twice. Resuspend cells in 300 µl 0.5% BSA PBS buffer and filter through a Falcon™ Tube with 35 µm cell strainer cap.
Instrument	Flow sorting was performed on an BD Influx™ cell sorter (BD Biosciences, X64650000124). Flow Cytometer data was collected with FACSCanto II Flow Cytometer (BD Bioscience).
Software	Flow sorting data was collected and analyzed using BD FACS™ Software (BD Biosciences, version 1.2.0.142). Flow Cytometer data was analyzed using FlowJo (BD, version 10.8.2).
Cell population abundance	Cell purity was assessed following flow sorting by single-cell RNA sequencing.
Gating strategy	<p>To sort the NE and non-NE population from NPPO-1 organoids, gating was first done on forward scatter (FSC) and side scatter (SSC) to exclude debris. Doublets were excluded by gating on trigger pulse width against FSC height. Individual NE and non-NE tumor cells were sorted based on scatter parameters. In general, NE tumor cells have less internal complexity (granularity) than non-NE tumor cells and exhibit SSC of a lower intensity. This allows NE and non-NE tumor cells to be distinguished.</p> <p>To isolate RFP-positive (RFP+) single cells from CRISPR/Cas9-mediated gene knockout organoids, fluorescence-activated cell sorting (FACS) was conducted to sort RFP+ cells. Gating was first done on FSC and SSC to exclude debris. Doublets were excluded by gating on trigger pulse width against FSC height. The sorting was performed with the laser set at PE channel. Flow cytometry was performed to analyze the percentage of HA-Tag-positive cells from organoids transfected with HA-tagged H3.3K36M transgene. Gating was first done on FSC and SSC to exclude debris. Doublets were excluded by gating on trigger pulse width against FSC height. The percentage of HA-Tag+ cells were identified by setting the laser at APC channel and collecting Cy5.5 positive events.</p>
<input checked="" type="checkbox"/> Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.	