Supplemental Information

N-Myc induces an EZH2-mediated transcriptional program driving Neuroendocrine Prostate Cancer

Etienne Dardenne[1,11], Himisha Beltran[2,3,4,11], Matteo Benelli[5], Kaitlyn Gayvert[6,7], Adeline Berger[1], Loredana Puca[4], Joanna Cyrta[1,4], Andrea Sboner[1,4,6,7], Zohal Noorzad[1], Theresa MacDonald[1], Cynthia Cheung[1], Ka Shing Yuen[1], Dong Gao[8], Yu Chen[3,8,9], Martin Eilers[10], Juan-Miguel Mosquera[1,4], Brian D. Robinson[1,4], Olivier Elemento[2,4,6], Mark A. Rubin[1,2,4,6], Francesca Demichelis[4,5] and David S. Rickman[1,2,4,#]

[1]*Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, 10065.*
[2]*Meyer Cancer Center, Weill Cornell Medicine, New York, New York, 10065.*
[3]*Department of Medicine, Weill Cornell Medicine, New York, New York, 10065.*
[4]*Englander Institute for Precision Medicine, Weill Cornell Medicine and New York-Presbyterian Hospital, New York, NY, 10065.*
[5]*Centre for Integrative Biology, University of Trento, Trento, 38123, Italy.*
[6]*Institute for Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, 10065.*
[7]*Tri-Institutional Training Program in Computational Biology and Medicine, New York, NY 10065.*
[8]*Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065.*
[9]*Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, 10065.*
[10]*Theodor Boveri Institute and Comprehensive Cancer Center Mainfranken, Biocenter, University of Würzburg, Am Hubland, 97074 Würzburg, Germany.*
[11] These authors contributed equally to this work.

Correspondance: David S. Rickman, dsr2005@med.cornell.edu, *Weill Cornell Medicine, New York, New York, 10065, USA*

Running Title: N-Myc drives Neuroendocrine Prostate

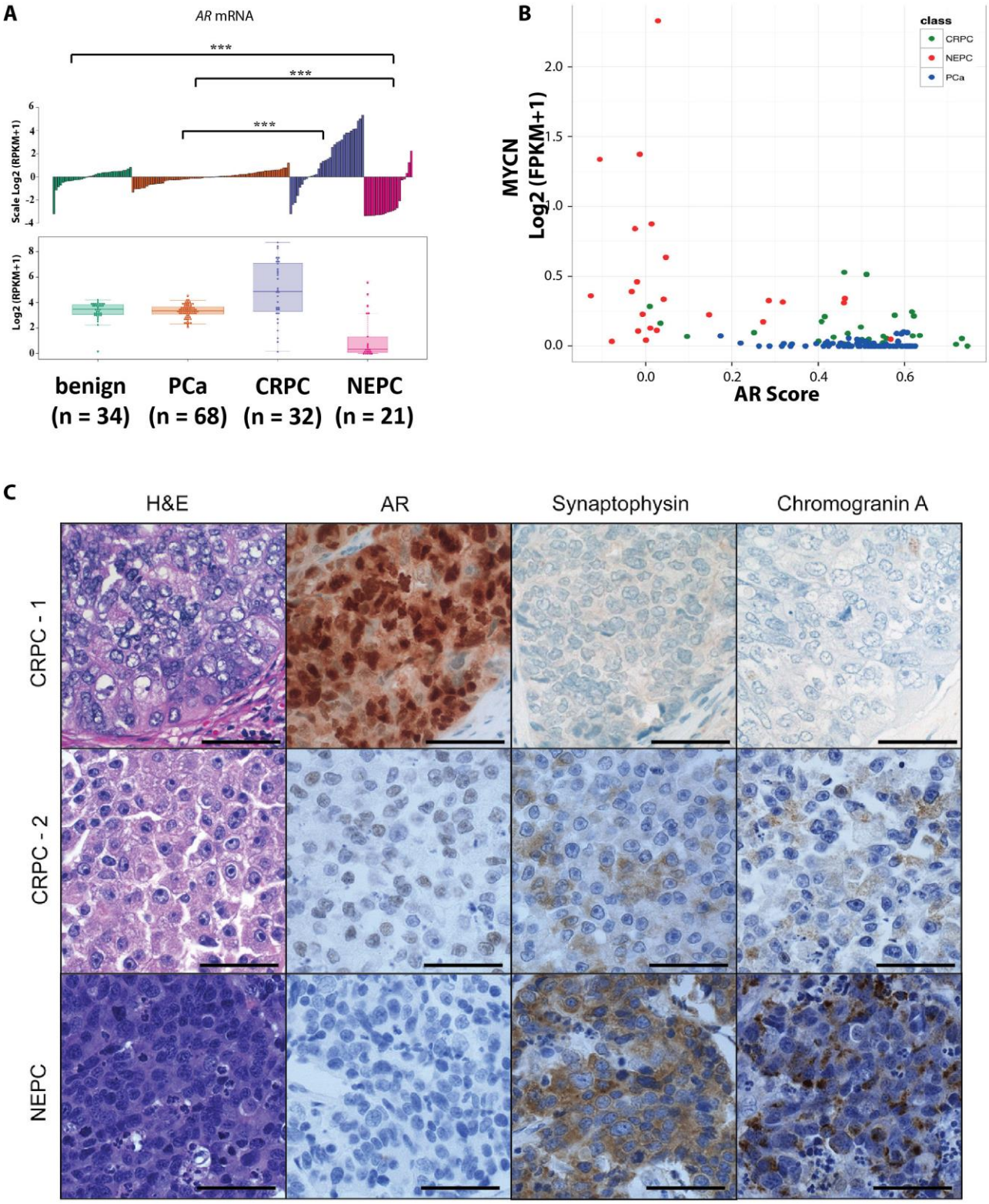

Figure S1. Related to Figure 1. Characterization of N-Myc expressing patients samples. A. *AR* mRNA level in 34 benign prostate, 68 prostate adenocarcinoma (PCa), 32 castrate resistant prostate adenocarcinoma (CRPC) and 21 neuroendocrine prostate

cancer (NEPC) clinical samples. **B.** Scatter plot of samples in the same patient cohort in **A** based on *MYCN* mRNA and AR score. **C.** CRPC-1, CRPC-2 and NEPC correspond to cases previously illustrated in Figure 1, with low, intermediate and high *MYCN* expression by RNA ISH. In this panel, CRPC-1 demonstrates strong AR expression by IHC, while the neuroendocrine markers synaptophysin and chromogranin A are negative. CRPC-2, a case of aggressive castrate-resistant prostate cancer that progressed despite multiple therapies, has weak AR expression and focal neuroendocrine positivity. NEPC is AR-negative, and strongly positive for synaptophysin and chromogranin A. (original magnification, 40x for both H&E and immunohistochemistry, scale bar = 50 um).

Table S1, related to Figure 2. Genotype and phenotype assessment of different GEM mice. Provided as an Excel file.

Table S2, related to Figures 2, 3, 5 and Figure S2,3,4. N-Myc gene signatures of mouse RNAseq data. Provided as an Excel file.

Table S3, related to Figures 2, 3, 4, 5 and Figure S2, 3, 4. Summary of GSEA analyses of mouse and cell line N-Myc gene signatures. Provided as an Excel file.

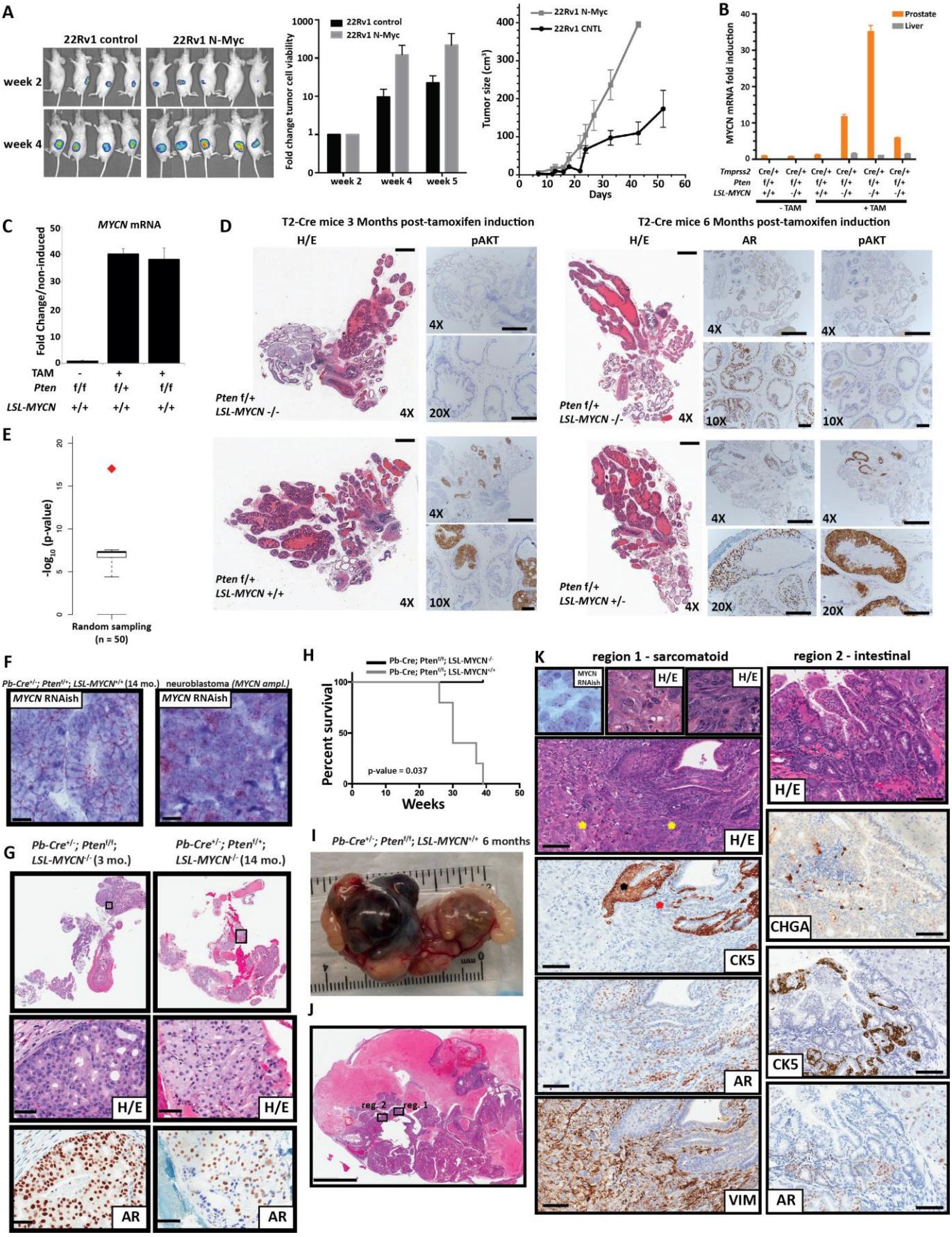
A
22Rv1 control
22Rv1 N-Myc
week 2
week 4
Fold change tumor cell viability
22Rv1 control
22Rv1 N-Myc
week 2
week 4
week 5
22Rv1 N-Myc
22Rv1 CNTL
Tumor size (cm³)
Days
B
MYCN mRNA fold induction
Prostate
Liver
Tmprss2
Pten
LSL-MYCN
- TAM
+ TAM
C
MYCN mRNA
Fold Change/non-induced
TAM
Pten
LSL-MYCN
D
T2-Cre mice 3 Months post-tamoxifen induction
H/E
pAKT
Pten f/+
LSL-MYCN -/-
Pten f/+
LSL-MYCN +/+
T2-Cre mice 6 Months post-tamoxifen induction
H/E
AR
pAKT
Pten f/+
LSL-MYCN -/-
Pten f/+
LSL-MYCN +/-
E
-log10 (p-value)
Random sampling
(n = 50)
F
Pb-Cre+/-; Ptenf/+; LSL-MYCN+/+ (14 mo.)
neuroblastoma (MYCN ampl.)
MYCN RNAish
MYCN RNAish
G
Pb-Cre+/-; Ptenf/f; LSL-MYCN-/- (3 mo.)
Pb-Cre+/-; Ptenf/+; LSL-MYCN-/- (14 mo.)
H/E
H/E
AR
AR
H
Pb-Cre; Ptenf/f; LSL-MYCN-/-
Pb-Cre; Ptenf/f; LSL-MYCN+/+
Percent survival
p-value = 0.037
Weeks
I
Pb-Cre+/-; Ptenf/f; LSL-MYCN+/+ 6 months
J
reg. 2
reg. 1
K
region 1 - sarcomatoid
region 2 - intestinal
MYCN RNAish
H/E
H/E
H/E
H/E
CK5
AR
VIM
CHGA
CK5
AR

Figure S2. Related to Figures 2 and 3. N-Myc over-expression is associated with aggressive, NEPC-like prostate cancer. A. Left: Representative bioluminescent images of mice with 22Rv1 xenografts with and without N-Myc at the indicated time point Right: Average and standard deviation of quantification of luminescence measured at the indicated time point for each group of mice. **B.** RT-QPCR of *MYCN* in 5 weeks post-tamoxifen induction in prostate and liver at the indicated genotype. **C.** Top: RT-qPCR of *MYCN* mRNA in T2-Cre$^{+/+}$ mouse prostates at 3 months post tamoxifen-induction. **D.** Shown are photomicrographs of a 4 month old, 3 months post-induction (left) *Tmprss2$^{Cre/Cre}$* (T2); *Pten$^{f/+}$; R26$^{+/+}$* or *T2; Pten$^{f/+}$; R26$^{MYCN/MYCN}$* or 6 month (right) *T2; Pten$^{f/+}$; R26$^{+/+}$ or T2; Pten$^{f/+}$; R26$^{MYCN/+}$* mouse prostates. Left: H&E staining (4X) and p-AKT IHC staining (Low magnification 4X (scale bar = 600um), high magnification 10X (scale bar = 200um) and 20X (scale bar = 200um)) of mouse prostate at 3 months post-tamoxifen induction at the indicated genotypes. Right: H&E staining (left) and AR (middle) or p-AKT (right) IHC staining at 6 months post-tamoxifen induction. **E.** Significance of classification of Ward's hierarchical clustering of the normalized FPKM values of the 779 prioritized genes across 203 CRPCs Box plot shows the distribution of p-values obtained by the random sampling strategy used to test the class segregation power (see **Supplemental Experimental Procedures**). Red diamond refers to the p-value obtained by the prioritized gene list. **F.** RNA *in situ* hybridization (RNAish, red chromogen) of *MYCN* RNA in the NEPC foci from the 14-month old *Pb-Cre$^{+/-}$; Pten$^{f/-}$; LSL-MYCN$^{-/-}$*mice and from an *MYCN*-amplified Neuroblastoma (Scale bar = 50 uM). **G.** Low (7x) and high (40X) magnification H&E staining or AR IHC staining prostates from 3-month old *Pb-Cre$^{+/-}$; Pten$^{f/f}$; LSL-MYCN$^{-/-}$* or 14-month old *Pb-Cre$^{+/-}$; Pten$^{f/-}$; LSL-MYCN$^{-/-}$* mice (Scale bar = 50 uM). **H.** Survival curve of mice with the indicated genotype. **I.** photomicrograph of the tumor taken just after tumor resection. **J.** Low magnification photomicrograph H&E image of 2 representative divergent differentiated divergent regions (4x original magnification, scale bar = 6000uM). **K.** Photomicrograph images of the 2 representative divergent differentiated divergent regions shown in **(J)** observed in the, highly proliferative (Ki67 IHC) invasive prostate carcinoma that developed in a 6-month-old *Pb-Cre$^{+/-}$; Pten$^{f/f}$; LSL-MYCN$^{+/+}$* mouse. Regions 1: sarcomatoid differentiation. Shown are *MYCN* RNAish in invasive carcinoma (left) and spindled (center) and pleomorphic epithelioid (right) cytomorphology (40X (region 1, scale bar = 100uM) and

20x original magnification (region 2, scale bar = 200uM) that correspond to left and bottom areas of the H&E image below (yellow asterisks). Below: IHC of sarcomatoid areas are cytokeratin 5 (CK5) negative (red asterisk), vimentin (Vim) positive that are adjacent to basal cells and squamous differentiation (black asterisk) mPIN glands; Region 2: intestinal differentiation with scattered chromogranin A positive cells (original magnification 20x, scale bar = 200uM).

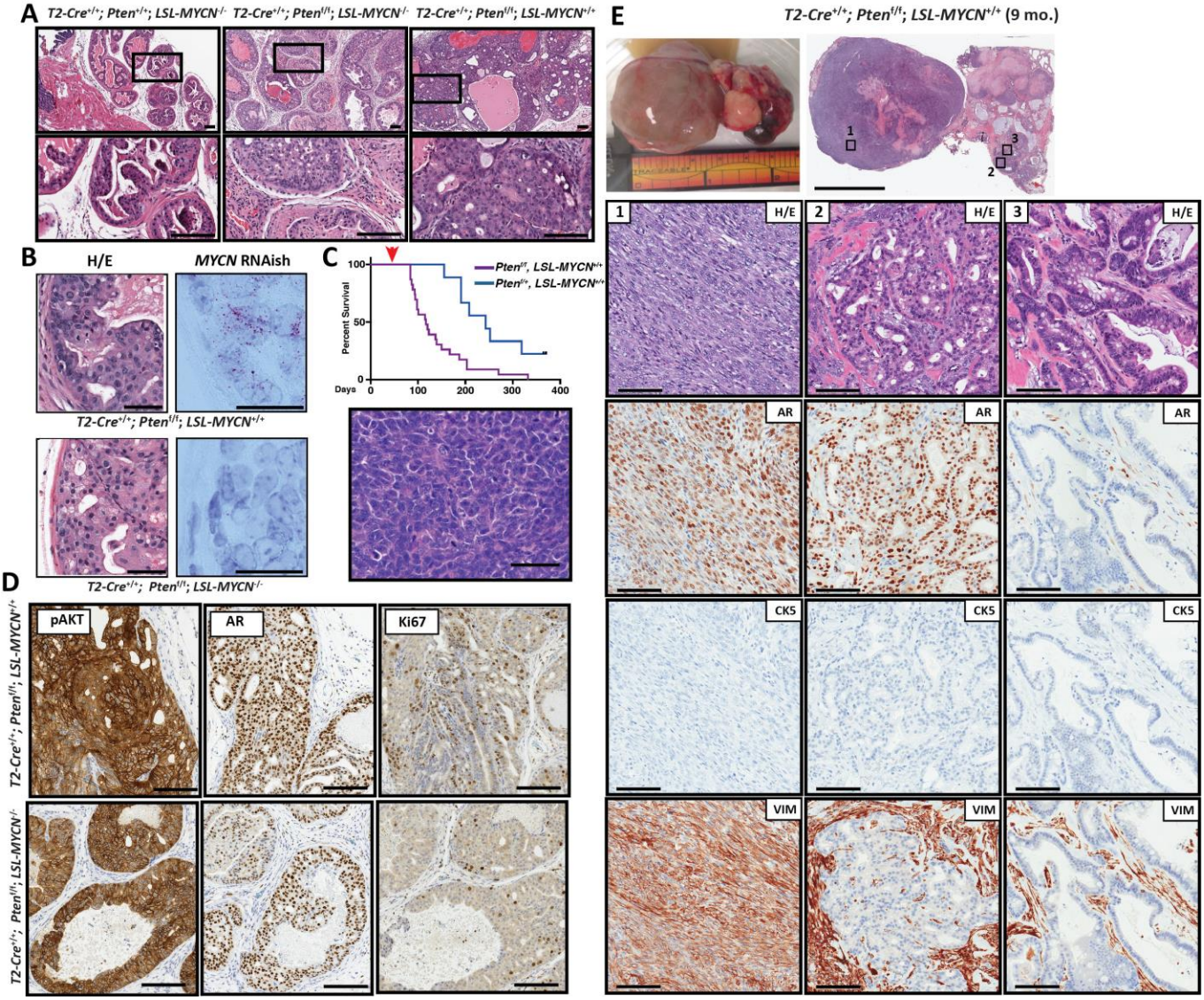

Figure S3. Related to Figure 3. Mouse N-MYC signature is clinically relevant. A. Shown are photomicrographs of H&E staining from 4 month old normal (T2-Cre$^{+/+}$; Pten$^{+/+}$; *LSL-MYCN*$^{-/-}$ or 3 months post-induction T2-Cre$^{+/+}$; Pten$^{f/f}$; *LSL-MYCN*$^{-/-}$ or T2-Cre$^{+/+}$; Pten$^{f/f}$; *LSL-MYCN*$^{+/+}$ mouse prostates. H&E staining (Low magnification 10X, high magnification 40X (scale bar = 100um)) of mouse ventral prostates for the indicated genotype. **B.** mHGPIN lesions are shown in the H&E image and matching MYCN RNA ISH images at lower (40X, scale bar = 50 um) and higher (100X, scale bar = 25 um) magnification. **C.** Kaplan Meyer survival curve for Tmprss2-Cre mice at the indicated genotype post-tamoxifen induction (red arrow). The median survival is 117 days for Pten$^{f/f}$; *LSL-MYCN*$^{+/+}$ (n = 23) and 243 days for the Pten$^{f/+}$; *LSL-MYCN*$^{+/+}$ mice (n=9). bottom: H&E staining of a liver cancer foci from a 6 months post-tamoxifen induction

Pten$^{f/f}$; *LSL-MYCN*$^{+/+}$ mice (Scale bar = 50 uM). **D.** Shown are photomicrographs of a 4 month old, 3 months post-induction, T2-Cre$^{+/+}$; Pten$^{f/f}$; *LSL-MYCN*$^{-/-}$ or T2-Cre$^{+/+}$; Pten$^{f/f}$; *LSL-MYCN*$^{+/+}$ mouse prostates. From left to right: p-AKT, androgen receptor (AR) or Ki-67 IHC staining (magnification 40X). **E.** Low (scale bar = 2mm) H&E staining and high magnification (scale bar = 100um) H&E staining or AR, cytokeratin 5 (CK5) and vimentin IHC staining of a 9 months post-induction T2-Cre$^{+/+}$; Pten$^{f/f}$; *LSL-MYCN*$^{+/+}$ mouse prostate. The indicated regions are representative areas that show AR-positive prostatic adenocarcinoma (**region 2**) and foci of poorly differentiated carcinoma with divergent differentiation (e.g. sarcomatoid (**region 1**) or intestinal (**region 3**)).

Table S4, related to Figure 4 and Figure S4. N-Myc gene signatures cell line RNAseq data. Provided as an Excel file.

Table S5, related to Figure 4 and Figure S4. nMotif analysis at these 175 base pair AR enhancer regions. Provided as an Excel file.

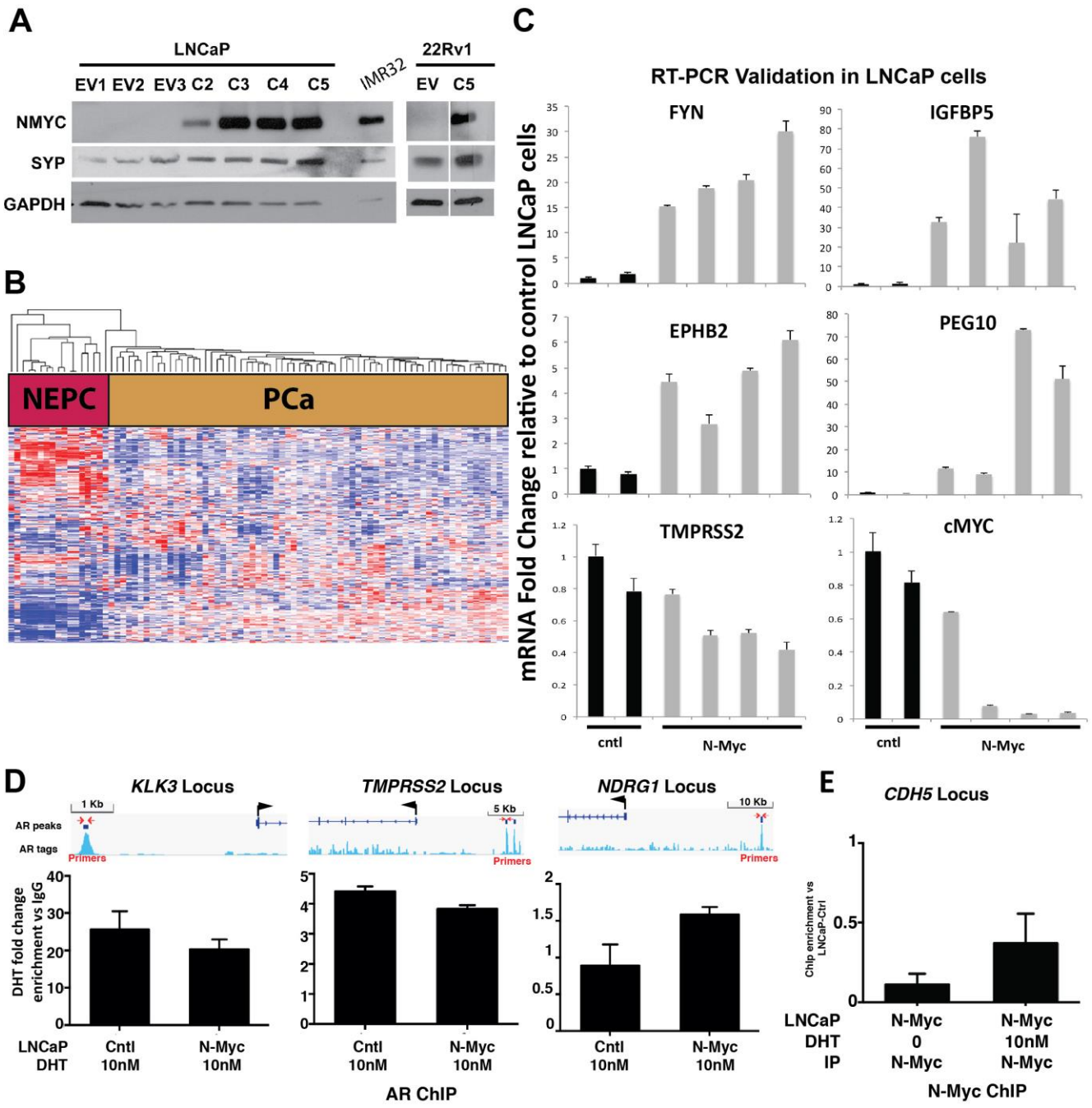

Figure S4. Related to Figure 4. N-Myc signature in LNCaP cells is clinically relevant and exhibits a dramatic down-regulation of AR signaling. A. Western blot analysis of N-Myc in LNCaP-Cnt (Ca and Cb) and 4 independent LNCaP-N-Myc (C2-C5), 22Rv1 control (EV) or 22Rv1-N-Myc (C5) cell populations (IMR32 is used as a positive control). **B.** Dendrogram of NEPC and PCa clinical samples shown in **Figure 1A** clustered using on LNCaP-N-Myc signature of 277 genes (Benjamini-Hochberg adjusted

p-value < 0.01). The majority (70%, 195 genes) of the differential genes significantly distinguished NEPC and PCa patient tumor samples using available RNA-seq data from the WCM patient cohort (adjusted p-value < 0.01). **B.** RT-PCR validation of the indicated genes in the different biological replicate LNCaP cell lines. Values are normalized to a control gene and to the control LNCaP cells. **D.** ChIP-PCR for N-Myc and AR at known AR enhancers for indicated genes following 48 hours growth in charcoal stripped serum with or w 24 hours of 10 nM DHT in LNCaP-Cntl or LNCaP-N-MYC cells. **E.** ChIP-PCR for N-Myc and control region at the *CDH5* promoter locus. IgG ChIP-PCR values were bellow the detection range and therefore not shown.

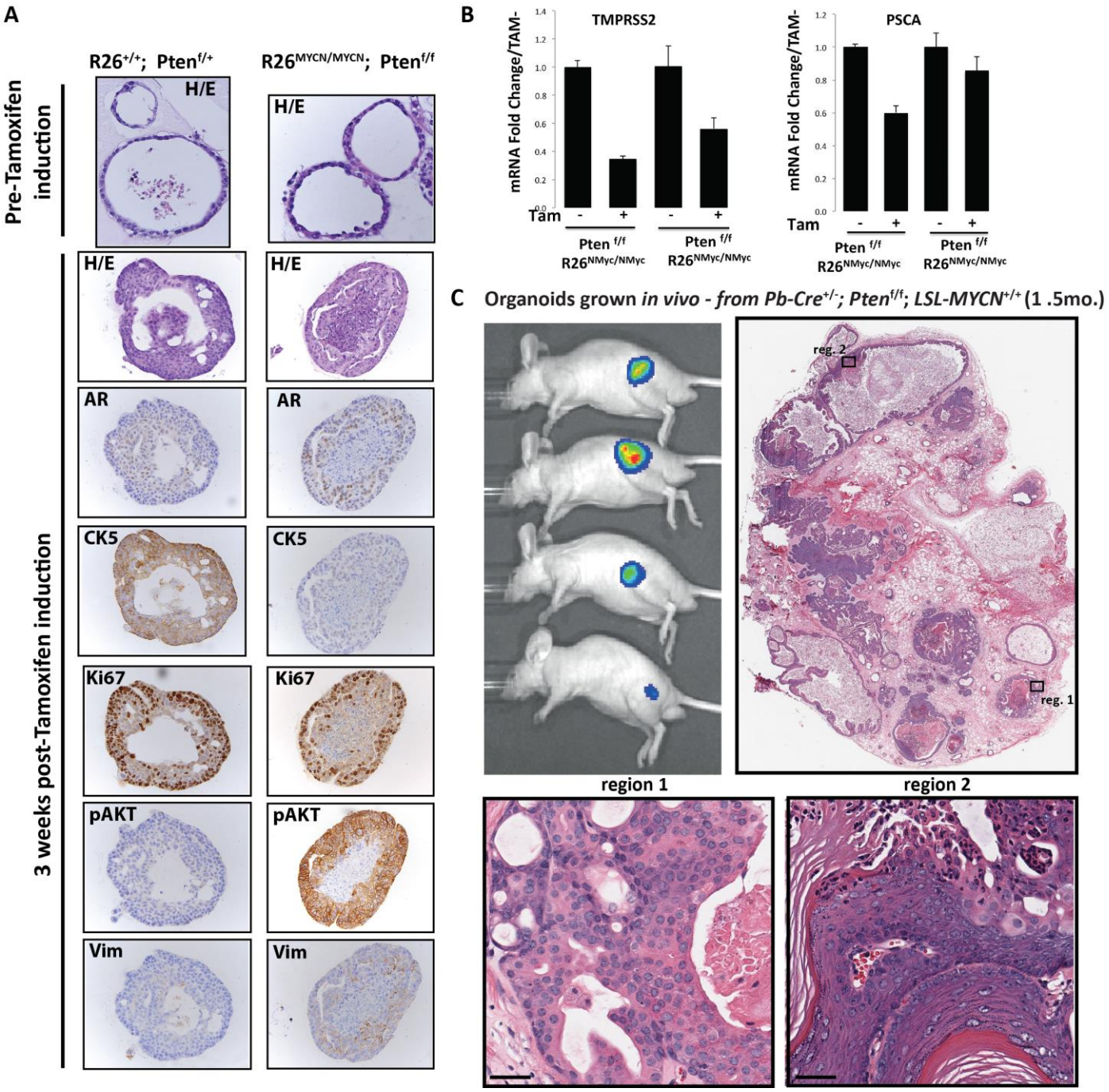

Figure S5. Related to Figure 5. Characterization of prostatic organoids culture. A. From top to bottom: Microphotograph of H&E staining pre-tamoxifen and 3 weeks post-tamoxifen induction. Androgen receptor (AR), cytokeratin 5 (CK5), Ki-67, p-AKT and Vimentin (Vim) IHC staining for the indicated genotypes (Magnification = 20X). **B**. RT-QPCR of Tmprss2 and PSCA at 3 weeks post-tamoxifen induction at the indicated genotype. Values are normalized to a control gene and to non-induced organoid for the given genotype. **C**. Top left: Representative images of Pb-Cre$^{+/-}$; Pten$^{f/f}$; LSL-MYCN$^{+/+}$

organoids generated from 1.5-month old mouse and grown *in vivo* as allografts 3 months post subcutaneous injection showing bioluminescent-positive foci. Top right: H&E staining of a low magnification microphotograph (2x) of the organoid allograft tumor has been harvested the same day of the bioluminescent imaging. Below: High magnification of different regions highlighting the different pathologies found in the organoids allograft (region 1, adenocarcinoma; region 2, squamous differentiation). Scale bar = 100 uM.

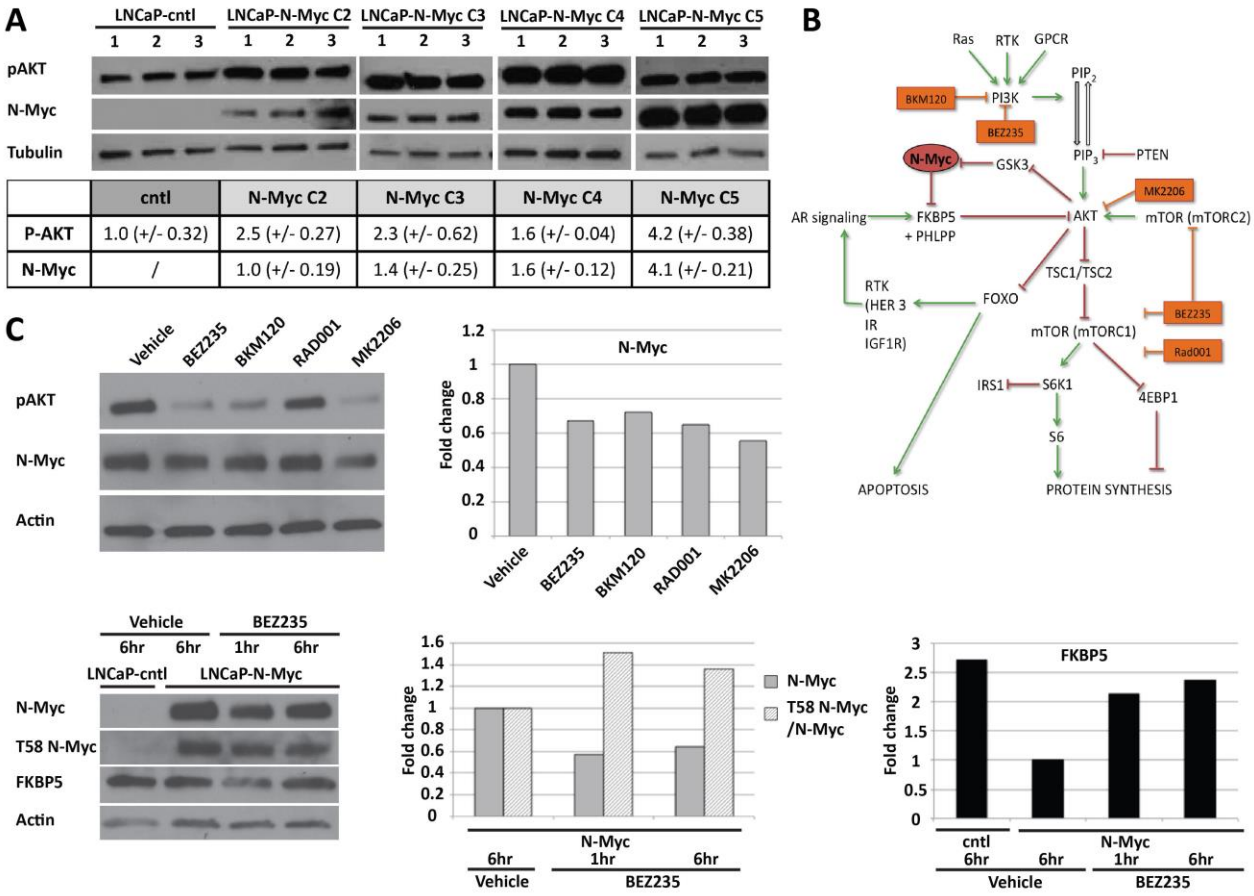

Figure S6. Related to Figure 6. N-Myc over-expression is associated with increased AKT activity. A. Western blot analysis and quantification of p-AKT and N-Myc in 4 independent LNCaP-N-MYC cell populations compared to LNCaP-cntl (n = 3 replicate dishes per cell population). **B.** Schematic representation of N-Myc and PI3K/AKT pathway interaction, highlighting the targets of PI3K/AKT pathway inhibitors used in this study. **C. Upper panel:** Western blot analysis of p-AKT and total N-Myc in LNCaP-N-MYC cells treated 6hr with BEZ235 (500 nM), BKM120 (1 μM), RAD001 (100 nM) or MK2206 (1 μM), compared to cells treated with vehicle (left). Quantification of fold-change in N-Myc expression in LNCaP-N-MYC cells treated with PI3K/AKT pathway inhibitors compare to cells treated with vehicle. The normalization has been done on Actin expression (right). **Lower panel:** Western blot analysis of total N-Myc, T58-phosphorylated N-Myc and FKBP5 in LNCaP-N-MYC cells compare to LNCaP-cntl, after 1hr or 6hr of BEZ235 (500 nM) treatment (left). Quantification of fold-change in protein expression of N-Myc-T58N-Myc/N-Myc ratio and FKBP5 in LNCaP-cntl and LNCaP-

NMYC treated cells compare to LNCaP-N-MYC untreated cells. The normalization has been done on Actin expression (right).

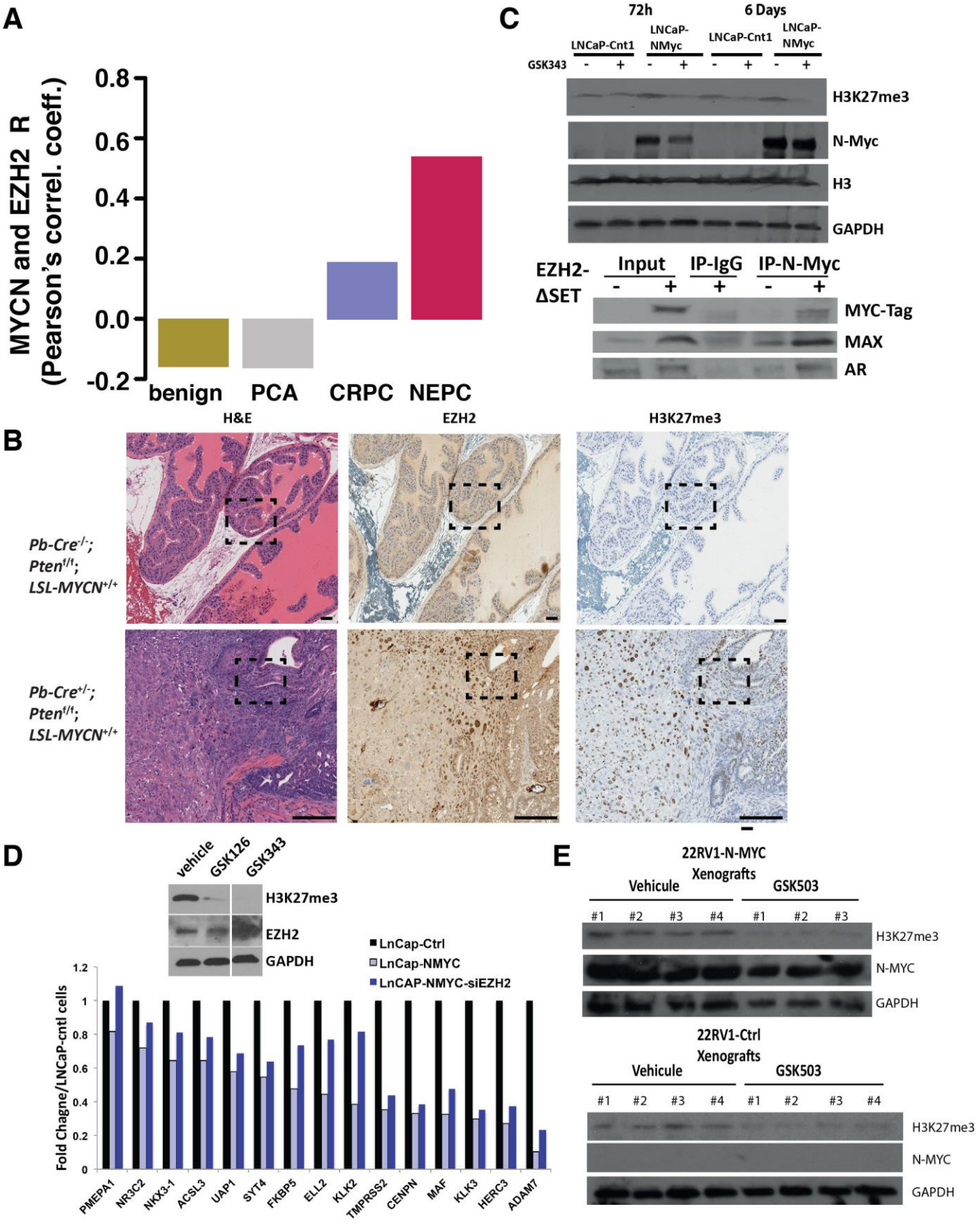
A
MYCN and EZH2 R
(Pearson's correl. coeff.)
0.8
0.6
0.4
0.2
0.0
-0.2
benign
PCA
CRPC
NEPC
C
72h
6 Days
LNCaP-Cnt1
LNCaP-NMyc
LNCaP-Cnt1
LNCaP-NMyc
GSK343
- + - + - + - +
H3K27me3
N-Myc
H3
GAPDH
EZH2-ΔSET
Input
IP-IgG
IP-N-Myc
- + + - +
MYC-Tag
MAX
AR
B
H&E
EZH2
H3K27me3
Pb-Cre-/-; Ptenf/f; LSL-MYCN+/+
Pb-Cre+/-; Ptenf/f; LSL-MYCN+/+
D
vehicle
GSK126
GSK343
H3K27me3
EZH2
GAPDH
LnCap-Ctrl
LnCap-NMYC
LnCAP-NMYC-siEZH2
Fold Chagne/LNCaP-cntl cells
1.2
1
0.8
0.6
0.4
0.2
0
PMEPA1
NR3C2
NKX3-1
ACSL3
UAP1
SYT4
FKBP5
ELL2
KLK2
TMPRSS2
CENPN
MAF
KLK3
HERC3
ADAM7
E
22RV1-N-MYC Xenografts
Vehicule
GSK503
#1 #2 #3 #4 #1 #2 #3
H3K27me3
N-MYC
GAPDH
22RV1-Ctrl Xenografts
Vehicule
GSK503
#1 #2 #3 #4 #1 #2 #3 #4
H3K27me3
N-MYC
GAPDH

Figure S7. Related to Figure 7. N-Myc interacts with EZH2 to drive transcriptional program. A. Pearson's correlation coefficients between the gene expression level of *MYCN* and EZH2 marker genes in the indicated clinical samples. **B.** From top to bottom: H&E staining, EZH2 (BD 612666) and H3K27me3 (Cell signaling 9733) IHC at the indicated genotypes taken from 6 month old mice (Magnification = 10X and 40X (scale bar = 100um)). **C.** Western Blot analysis of H3K27me3, N-MYC and total H3 in LNCaP-Cntl and LNCaP-N-Myc cells after 72hours or 6 days of EZH2 inhibitor GSK343 (5μM). Below: Co-immunoprecipitation of myc-tagged SET deleted mutant EZH2, Max upon N-Myc pull down in LNCaP-N-Myc cells and in LNCaP-N-Myc cells following transfection of the Myc-tagged SET domain-deletion EZH2 mutant (top right) or 6-day treatment of either the EZH2 inhibitors GSK126 or GSK343 (bottom). **D.** Nanostring mRNA fold change of examples of N-Myc downregulated genes and that depend on EZH2. Inset: Western Blot analysis of H3K27me3, EZH2 and N-Myc in LNCaP-N-Myc cells after 6 days of EZH2 inhibitor GSK126 or GSK343 (5μM). These cells were used for the RNAseq data shown in **Figure 7**. **E.** Western Blot analysis of H3K27me3, EZH2 and GAPDH from protein lysates extracted from the indicated xenografts following the treatment of vehicle or GSK503 as described in **Figure 7**.

Supplemental Experimental Procedures

Pathologic Evaluation

Hematoxylin and eosin (H&E)–stained slides were reviewed by study pathologists with expertise in uropathology and evaluation of prostate cancer mouse models (J.M.M., B.D.R., and M.A.R.). Pathologic evaluation of human samples included the Proposed 2015 WHO Morphologic Classification of Prostate Cancer with Neuroendocrine Differentiation (Epstein et al., 2016).

MYCN RNA in situ hybridization (ISH) and signal quantification

The single-color chromogenic QuantiGene® ViewRNA ISH Tissue Assay (Affymetrix, Santa Clara, CA), uses pairs of specially designed oligonucleotide probes that, through sequence-specific hybridization, recognize both the specific target MYCN RNA sequence and the signal amplification system. Briefly, the signal amplification system consists of the pre-amplifier, amplifier and enzyme-conjugated label probe, which assemble into a tree-like complex through sequential hybridization. Signal amplification occurs at target sites bound by probe pairs only. Nonspecific off-target binding by single probes does not result in signal amplification. The MYCN probe used for this assay was MYCN, HUMAN .440ml - ViewRNA TYPE 1, Product code VA1-18174-01 (Affymetrix, Santa Clara, CA). All steps of MYCN RNA ISH staining of the slides were performed manually, optimized in both human and mouse tissues. Briefly, 5µm formalin-fixed, paraffin-embedded unstained tissue sections were mounted on positively charged glass slides, deparaffinized in xylene and rehydrated through a series of alcohols. Conditions included 18 minutes of pretreatment and 18 minutes of Protease. The rehydrated sections were treated with 3% hydrogen peroxide at room temperature (RT) for 10 minutes to block endogenous peroxidase. Sections were then boiled in 1x citric buffer (10nmol/L Na-citrate, pH 6.0) for 15 minutes and incubated with protease (2.5mg/ml; Sigma Aldrich) at 40^{O}C for 30 minutes. The slides were hybridized sequentially with target probes (20nmol/L) in hybridization buffer A (6x saline sodium citrate [SSC] buffer [1x SSC is 0.15mol/L NaCl and 0.015 mol/L Na-citrate], 25% formamide, 0.2% lithium dodecyl sulfate (LDS) and blocking reagents) at 40^{O}C for 2 hr., signal pre-amplifier in hybridization buffer B (20% formamide, 5xSSC, 0.3% LDS, 10% dextran sulfate and blocking reagents) at 40^{O}C for 30 min., amplifier in hybridization buffer B at 40^{O}C for 30

minutes and horseradish peroxidase or alkaline phosphatase-labeled probes in hybridization buffer C (5x SSC, 0.3% LDS and blocking reagents) at 40^{O}C for 15 minutes. Hybridization signals were detected under bright-field microscope as red colorimetric staining (using Fast Red chromogen, BioCare Biomedical, Concord, CA) followed by counterstaining with hematoxylin. Red signals were granular and discrete, corresponding to individual RNA targets. Automated signal quantification was performed using the RNAscope® SpotStudio v1.0 Software (Definiens AG, München, Germany). The controls used for each RNA ISH experiment were: 1) MYCN overexpressing tissue (human neuroblastoma, amplification demonstrated by exome sequencing and validated by fluorescence *in situ* hybridization [FISH], and high MYCN expression by RNA-seq); 2) Quantigene ViewRNA ISH Tissue Control Kit (1-Plex), which includes positive controls GAPD, ACTB, UBC, and PPIB (Affymetrix, Santa Clara, CA); 3) and negative controls (E.coli, K12 dapB) (Affymetrix, Santa Clara, CA).

Principal Components Analysis of CPRC, NEPC and NB RNA-seq data

In total 701 RNA-seq samples were considered, including 178 CRPC-Adeno, 25 CRPC-NE and 498 Neuroblastoma (NB) from Zhang et al. (Zhang et al., 2015) including 92 and 401 NB samples annotated as *MYCN* amplified and *MYCN* wt, respectively (*MYCN* annotation is not available for 5 samples). NB processed RNA-seq data by TopHat2/Cufflinks pipeline were downloaded from the GEO database (www.ncbi.nlm.nih.gov/geo/) with accession number GSE49711. Principal Component Analysis (PCA) of the mouse model signature (n = 779) was performed using the princomp function of R “stats” package (https://cran.r-project.org/). The cov.rob function of the “MASS” R package (Venables and Ripley, 2002) was used to compute the covariance matrix.

Clinical sample RNA-seq data processing

In total, 203 castration resistant tumors were considered. Hematoxylin and eosin stained slides were reviewed for all cases by study pathologists (MAR, BR, JMM) for histologic evaluation and classified based on morphology as adenocarcinoma or NEPC based on a published pathologic classification system (Epstein, et al, AJSP 2014). Reads were mapped to the human genome reference sequence (hg19/GRC37) using STAR v2.3.0e

(Dobin et al., 2013), and the resulting alignment files were converted into Mapped Read Format (MRF) using RSEQtools (Habegger et al., 2011). Quantification of gene expression was performed by RSEQtools using GENCODE v19 (http://www.gencodegenes.org/releases/19.html) as reference gene annotation set. A composite model of genes based on the union of all exonic regions from all gene transcripts was used, resulting in a set of 20,345 protein-coding genes. Expression levels were estimated as FPKM.

AR Signaling and Integrated NEPC score.

For each sample, AR signaling was assessed based on the expression levels of 30 genes previously reported (Hieronymus et al., 2006) by means of correlation against LNCaP cells signal (Beltran et al., 2016) and referred to as "AR signaling". The Integrated Neuroendocrine Prostate Cancer (NEPC) score estimates the likelihood of a test sample to be NEPC and it's computed based on a set of 70 genes (Beltran et al., 2016). The gene set stems from the integration of differentially deleted/amplified and/or expressed and/or methylated genes in NEPC versus CRPC.

NMYC correlation analysis

R MYCN was computed as the Pearson's correlation coefficient between the expression level of *MYCN* and a set of 9 NE genes, including 5 upregulated (*CHGA*, *SYP*, *ENO2*, *NCAM*, *AURKA*) and 4 downregulated (*KLK3*, *NKX3-1*, *AR*, *RB1*) genes. Fisher r-to-z transformation by "paired.r" function of the "psych" R package (https://cran.r-project.org/web/packages/psych/) was used to assess the significance of the difference between R MYCN of CRPC and NEPC

Statistics of class segregation on expression data.

To test for NEPC cluster enrichment, we applied hyper-geometric test by considering either the pathology classification or the Integrated NEPC score values. To estimate the power of the mouse model signature in segregating CRPC samples, we compared the signature statistics against the statistics of 50 sets of randomly sampled genes of the same size (n=779 genes). For each set, we built the 2x2 contingency tables obtained by counting the number of CRPC and NEPC samples falling in the two main clusters

identified by the Ward's hierarchical clustering of the normalized FPKM values (Pearson's correlation was used as distance measure). P-values by Fisher Exact Test were used as class segregation statistics estimator. **Figure S7** shows the distribution of the p-values obtained by the random sampling strategy together with the p-value estimated by using the signature.

Genetically engineered mouse lines

We crossed transgenic mouse lines that carry an integrated CAG-LSL-*MYCN* gene at the *ROSA26* (*R26*) locus (kind gift from Johannes Schulte (Althoff et al., 2015)) with mice that have an integrated Cre recombinase either following the mouse *Tmprss2* promoter (Gao et. al., manuscript in preparation) or with a probasin promoter that have been previously described (Chen et al., 2013). Both lines of mice have crossed with mice containing a floxed *Pten* locus that have previously described (Chen et al., 2013). Following removal of the LSL cassette by Cre, a chicken actin promoter drives N-Myc expression in these models. All lines of mice are bred on the same mixed genetic background (C57/Bl6/129x1/SvJ).

Mouse prostate organoids cultures

Mouse prostate organoids cultures were generated and maintained from an adapted protocol that was recently described (Gao et al., 2014; Karthaus et al., 2014). Briefly, murine prostates from 3 months old mice were dissected and enzymatically digested with collagenase type II (GIBCO) and TrypLE (GIBCO). Cells were embedded in growth factor reduced Matrigel (Corning) to a 1:5 ratio and overlayed with ADMEM/F12 medium supplemented with B27 (Life technologies), 10 mM HEPES (Life technologies), Glutamax (Life technologies), Penicillin/Streptomycin (Life technologies, Primocin (Invivogen) and the following growth factors: EGF 50 ng/ml (Peprotech), R-spondin1 and Noggin conditioned medium and the TGF-β/Alk inhibitor A83-01 500 nM (Santa-Cruz), Rock Inhibitor (Selleck Chemical) 10μM, n-acetylcysteine (N-AC) (Sigma-Aldrich) 1.25 mM and Dihydrotestosterone (DHT) (Sigma-Aldrich) was added at 1 nM. Organoids were treated with 1 μM of enzalutamide for 7 days. For the AKT inhibitor, organoids are dissociated and seeded as single cells. At day 4, organoids are formed and treated with the drug during 72 hours.

Bioinformatic Analysis of mouse, cell line, and EZH2 targeting RNAseq data

Sequencing reads were aligned using STAR (Dobin et al., 2013) to hg19. For each sample, HTSeq (Anders et al., 2015) and Cufflinks (Trapnell et al., 2012) were then used to generate read counts and FPKM respectively. For the LNCaP and 22Rv1 samples, the gene counts and DESeq2 (Love et al., 2014) were used to identify differentially expressed genes, defined to be those with adjusted $p < 0.10$ and absolute $\log_2$ fold change > 1. Similarly, differentially expressed exons were found using DEXSeq (Anders et al., 2012) and identifying exons with $p < 0.10$ and absolute $\log_2$ fold change > 1. For the mouse samples, Cuffdiff (Trapnell et al., 2012) was then used to identify differentially expressed genes, defined to be those with $p < 0.05$. For the siEZH2 and GSK343 RNA-seq, fold change was calculated relative to siRNA Ctrl and vehicle respectively and differentially expressed genes were defined to be those with absolute $\log_2$ fold change > 0.5. The hypergeometric test and Gene Set Enrichment Analysis (GSEA)(Subramanian et al., 2005) was used to identify enriched signatures using hallmark and c2 canonical pathways collection in the MSigDB database (Liberzon et al., 2011).

Bioinformatic Analysis of H3K27me3 ChIP-seq

Sequencing reads were aligned to hg19 using the Burrows-Wheeler Alignment (bwa) Tool (Li and Durbin, 2009). Genes were considered direct targets if there was more read coverage by H3K27me3 than input at promoter regions, which were defined to be 2kb upstream of the transcription start site of RefSeq transcripts. Coverage plots were generated using the Integrative Genomics Viewer (IGV) (Robinson et al., 2011). H3K27me3 levels were normalized by input and differences in log2 fold change at N-Myc signature genes were statistically tested using the Mann–Whitney U test. Promoters with more than 2-fold increase relative levels of H3K27me3 were considered to be differentially expressed and were tested for enriched signatures using the hypergeometric test and the MSigDB database (Liberzon et al., 2011).

Cell lines, transfection and drug treatment

All cell lines were purchased from ATCC and maintained according to the manufacturers' protocols. LNCaP and 22RV1 were culture in RPMI supplemented with 10% of FBS (Life

technology) and a cocktail of penicillin/streptomycin (Life technology). All the siRNA using in this studies were purchase from Dharmacon as On-target plus siRNA and transfected as a final concentration of 60nM with lipofectamine RNAimax (Life technology). Cells were treated with GSK126 (BioVision 2282-5), GSK343 (Sigma Aldrich SML0766) and GSK503 at the indicated concentration for 6 days. BKM120 and RAD001 were kindly provided by Dr. Lewis Cantley, WCM. Cells were treated with 1 μM of BKM120 and 100 nM of RAD001 for 24h for figure 6A (Western Blot). Cells were treated with 100nM of MLN8237 and the indicated concentration of BEZ235 (Selleck Chemicals) for 72h for the drug response.

Co-immunopreciptiation

Cell lysates were prepared by lysing cells in buffer containing 50 mM Tris (pH 7.5), 120 mM NaCl, 1% NP-40, 5 mM EDTA and protease and phosphatase inhibitors (Thermo Scientific). Protein G magnetic beads (Life technology) were cross-linked to 5 ug of anti-N-MYC antibody (Santa-Cruz sc53993) with Bis(sulfosuccinimidyl)suberate (BS^3) following manufacturer instruction. 500 ug of total protein extract was incubated the cross-linked beads overnight and washed 5 times using lysis buffer. Input, IgG and IP were next resolved to SDS–PAGE, transferred onto a PVDF membrane (Millipore) and incubated overnight at 4°C with primary antibodies. The antibodies used were: -N-Myc, Santa Cruz sc53993 (B8.4.B, N-Myc), Androgen Receptor (Millipore 06-680), Aurora A/AIK (Cell Signaling 3092S), Following four washes with TBS-T, the blot was incubated with horseradish peroxidase-conjugated secondary antibody and immune complexes were visualized by enhanced chemiluminescence detection (ECL plus kit, GE Healthcare, UK). GSK126 (BioVision 2282-5) and GSK343 (Sigma Aldrich SML0766) were added at 3 μM for 72h prior the co-immunoprecipitation.

Immunoblot Analysis

Protein lysates were prepared in the RIPA buffer supplemented with protease inhibitor cocktail and phosphatase inhibitors (Thermo Scientific). The total protein concentration of the soluble extract was determined using the BCA protein assay Kit (Thermo Scientific). Each protein sample (50ug) was resolved to SDS–PAGE, transferred onto a PVDF membrane (Millipore) and incubated overnight at 4°C with primary antibodies. The

antibodies used were: -N-Myc, Santa Cruz sc53993 (B8.4.B,N-Myc), Androgen Receptor (Millipore 06-680), Aurora A/AIK (Cell Signaling 3092S), pTEN ((D4.3) XP Cell Signaling 9188S), P-Akt (S473) (Cell Signaling 9271S), Phospho-S6 Ribosomal Protein (Ser235/236) (Cell Signaling 2211), Synaptophysin (Thermo fisher PA1-1043), Tubilin (Epitomics 1878-1), H3K27me3 (Abcam 6002), EZH2/KMT6 (Abcam 3748), GAPDH (Millipore AB2302). Following four washes with TBS-T, the blot was incubated with horseradish peroxidase-conjugated secondary antibody and immune complexes were visualized by enhanced chemiluminescence detection (ECL plus kit, GE Healthcare, UK).

Chromatin Immunoprecipitation

Fifty million LNCaP-N-Myc or LNCaP-ctrl cells were washed in PBS twice and then fixed using 1% formaldehyde for 8 minutes at room temperature and quenched 5 min using 125 mM glycine. The cells were centrifuged and the cell pellet was resuspended in 2 milliliters of dilution buffer (10mM Tris HCL pH8.0, 0.1% SDS, 1.0mM EDTA with protease and phosphatase inhibitors). Protein-bound chromatin was fragmented by sonication for 15 minutes (Covaris series E220). Equal volumes of chromatin were immunoprecipitated with either mouse anti-N-Myc (Santa Cruz sc53993 (B8.4.B,N-Myc) or Androgen Receptor (Millipore 06-680) and mouse or rabbit IgG as a negative control. Following extensive washing the DNA was eluted using 100 mM NaHCO3 and 1% SDS and the crosslinks were reversed using 300 mM NaCl at 65°C for 16 hours. The eluted DNA was purified using Qiagen PCR Qiaquick kit following manufacturer's protocol. For qPCR amplification we used the ABI 7500fast system and the relative standard curve method in a 96-well format. We designed primer sets that target each of the AR enhancer of KLK3 (PSA), TMPRSS2 and NDRG1. Two microliters of either eluted DNA or a 1:10 dilution of the input chromatin preparation from each cell line was assayed in order to calculate the percentage of enrichment. Input DNA was also analyzed at 5 concentrations (0.004 ng – 40ng) to generate the standard curve per primer pair and per 96-well plate. All reactions were run in triplicates. The following primers were used in this study: *KLK3* enhancer F-GCC TGG ATC TGA GAG AGA TAT CAT C, R-ACA CCT TTT TTT TTC TGG ATT GTT G, (chr19:50,850,923-50,851,007; amplification 85pb) ; *NDRG1* enhancer F-GAG GAA GGA TGT GTG TGC AA, R-AAT GGA GAC TTG GCT GGA GA,

(chr8:133,322,405-133,322,539 amplification 135 pb); *TMPRSS2* enhancer F-TGG AGC TAG TGC TGC ATG TC, R-CTG CCT TGC TGT GTG AAA AA, (chr21:41,521,697-41,521,870, amplification 174 pb). *CDH5* promoter F-AGC CAG CCC AGC CCT CAC R-CCT GTC AGC CGA CCG TCT TTG, (chr16:66,366,528-66,366,676, 149 pb) AR peaks and total AR tags were generated using GEO dataset GSE27824 (Wang et al., 2011) align to human genome hg18. The N-Myc ChIP-PCR values from the LNCaP-N-Myc cells were normalized to the values obtained from the input sample and further normalized by dividing the input-normalized N-Myc ChIP-PCR values obtained from the control cells. The same normalization were performed for the IgG ChIP-PCR values from both cell lines.

Bioinformatic Analysis of N-Myc motifs

We defined N-Myc motifs to be the canonical E-Box CACGTG and the non-canonical E-Boxes CANNTG motifs, within a 175 pb window, along with the two reported position count matrix motifs in the aggregate motif database *MotifDB (MotifDb: An Annotated Collection of Protein-DNA Binding Sequence Motifs. In, (R package version 1.12.0)*: M3642_1.02 and M6352_1.02. Both of these N-Myc motifs are originally reported and can be found in the Catalog of Inferred Sequence Binding Preferences Database (http://cisbp.ccbr.utoronto.ca/). We queried for the N-Myc motifs in EZH2 (GSM969570) (Xu et al., 2012) and AR (GSM686917) (Wang et al., 2011) ChIP-seq peaks. Significance was assessed using a permutation test, in which the motif enrichment was compared it to that in 500 randomly permutated peak sets. We created the randomly permutated peaks by sampling from 145 previously annotated ENCODE ChIP-seq narrow peaks (http://physiology.med.cornell.edu/faculty/elemento/lab/CS_files/Encode_hg18.tar.gz), in which the peak number and width of the true ChIP-seq dataset were preserved.

Quantitative RT-PCR (qRT-PCR) and multiplex qRT-PCR Nanostring assay

For cell line, RNA extraction was performed using Trizol (Life technology) following manufacturer recommendation. For mouse tissues frozen embedded prostate H&E were assessed for pathology and normal, HGPIN or prostate cancer lesions were core from the frozen blocks and the RNA extracted using RNA simply Kit (Promega). Quantitative

PCR was performed using SYBR Green dye on Applied Biosystems 7500 Real Time PCR system (Applied Biosystems, Foster City, CA) following manufacturer instruction. Primers used in this study are available upon request. For the Nanostring assay we used 200ng input RNA was used for each sample. Samples were run on the NanoString nCounter® Analysis System according to the manufacturers directions. Briefly, total RNA was hybridized with capture codeset overnight at 65°C, then placed on the Prep Station at max sensitivity (3 hours) to let samples being affixed to cartridges. Cartridges were then removed from the Prep station and scanned on the Digital Analyzer at 555 fields of view. Raw Nanostring gene expression data were normalized using negative controls, positive controls, and 6 housekeeping genes via the nSolver™ analysis software version 2. The 180-gene Nanostring assay includes 30 genes that are implicated in AR signaling, 61 genes associated with NEPC clinical samples, 12 positive or negative control genes and 6 housekeeping genes.

Immunohistochemistry (IHC)

Formalin fixed paraffin embedded (FFPE) tissue sections were de-paraffinized and endogenous peroxidase was inactivated. Antigen retrieval was accomplished by the Bond Epitope Retrieval Solution 1 (ER1) at 99-100°C for 30 minutes (Leica Microsystems). Following retrieval, the sections were incubated sequentially with the primary antibody for 25 minutes, post-primary for 15 minutes and polymer for 25 minutes ending with colorimetric development with diaminobenzidine (DAB) for 10 minutes (Bond Polymer Refine Detection; Leica Microsystems). Antibodies used were: -Keratin 5 (AF 138) - for IHC Biolegend PRB-1609, Ki67 polyclonal VWR (Vector) VP-K451, AR (Androgen receptor) Abcam: ab108341, Vimentin Santa Cruz: sc-7557, p-AKT Cell Signaling: 4060

Xenografts

1 million 22RV1-Ctrl or 22RV1-N-Myc cells were injected into NU/J mice (Jackson Laboratories, Bar Harbor, Maine). These cells were engineered to express luciferase using a transposable element vector (kind gift from John Ohlfest, University of Minnesota Medical School, Minneapolis, MN). Tumor volume was measured every week. Mice were imaged every weeks: on the day of imaging, we injected (intraperitoneal) 50-microliters

of D- Luciferin (75mg/kg) into anesthetized mice. Ten minutes later the mice animals were placed on their ventral side and bioluminescence images were acquired with the IVIS Imaging System (Xenogen). Analysis was performed using LivingImage software (Xenogen) by measurement of the average photon flux (measured in photons/s/cm2/steradian) within a region of interest. For all *in vivo* studies, GSK503 or vehicle was administered intraperitoneally at a dose volume of 0.2 ml per 20 g body weight in 20% captisol adjusted to pH 4–4.5 with 1 N acetic acid. 5 nude mice were randomized in each treatment group before the initiation of dosing and GSK503 treatment was initiated once the tumour volumes were approximately 150 mm^3. For castration study, the scrotum approach was chosen. Mice were castrated once the tumour volumes were approximately 150 mm^3. Mice were anesthetized with 4% isoflurane; the skin over the scrotum was disinfected by 70% ethanol and Betadine solution; a 0.5-cm incision was made over the scrotum; the testes were exposed by pulling the adipose tissue; the blood vessels supplying the testis are cauterized; the incision edges are closed with sterilized wound clips; suture removal is performed between post operative days 10 -14. For tail vein injection study, nude mice were warned for 10-15 min using a heat lamp. Mice were placed into a rodent restrainer; the tail is wiped with alcohol and 1 million of 22RV1-Ctrl or 22RV1-N-Myc cells were injected. Mice were imaged every week.

In situ proximity ligation assay

LNCaP-N-MYC cells were treated with for 72h with 100nM of MLN8237 or Vehicule. The in situ proximity ligation assay (PLA) was conducted using the Duolink II Kit (Olink Bioscience) as recommended by the manufacturer. Fixed and saturated cells were incubated with antibodies against N-MYC (sc53993 (B8.4.B, N-Myc)) and Aurora-A (Cell signaling 4718S). Interactions were revealed using secondary antibodies coupled to specific PLA DNA probes that hybridized and were enzymatically joined when located in close proximity. After rolling circle amplification, each interaction generated a fluorescent spot that was analyzed by confocal microscopy.

Supplemental References

Althoff, K., Beckers, A., Bell, E., Nortmeyer, M., Thor, T., Sprussel, A., Lindner, S., De Preter, K., Florin, A., Heukamp, L. C.*, et al.* (2015). A Cre-conditional MYCN-driven neuroblastoma mouse model as an improved tool for preclinical studies. Oncogene *34*, 3357-3368.

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. Genome Res *22*, 2008-2017.

Beltran, H., Prandi, D., Mosquera, J. M., Benelli, M., Puca, L., Cyrta, J., Marotz, C., Giannopoulou, E., Chakravarthi, B. V., Varambally, S.*, et al.* (2016). Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med *22*, 298-305.

Chen, Y., Chi, P., Rockowitz, S., Iaquinta, P. J., Shamu, T., Shukla, S., Gao, D., Sirota, I., Carver, B. S., Wongvipat, J.*, et al.* (2013). ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. Nat Med *19*, 1023-1029.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Epstein, J. I., Egevad, L., Amin, M. B., Delahunt, B., Srigley, J. R., Humphrey, P. A., and Grading, C. (2016). The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. Am J Surg Pathol *40*, 244-252.

Gao, D., Vela, I., Sboner, A., Iaquinta, P. J., Karthaus, W. R., Gopalan, A., Dowling, C., Wanjala, J. N., Undvall, E. A., Arora, V. K.*, et al.* (2014). Organoid cultures derived from patients with advanced prostate cancer. Cell *159*, 176-187.

Habegger, L., Sboner, A., Gianoulis, T. A., Rozowsky, J., Agarwal, A., Snyder, M., and Gerstein, M. (2011). RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries. Bioinformatics *27*, 281-283.

Hieronymus, H., Lamb, J., Ross, K. N., Peng, X. P., Clement, C., Rodina, A., Nieto, M., Du, J., Stegmaier, K., Raj, S. M.*, et al.* (2006). Gene expression signature-based chemical genomic prediction identifies a novel class of HSP90 pathway modulators. Cancer Cell *10*, 321-330.

Karthaus, W. R., Iaquinta, P. J., Drost, J., Gracanin, A., van Boxtel, R., Wongvipat, J., Dowling, C. M., Gao, D., Begthel, H., Sachs, N.*, et al.* (2014). Identification of multipotent luminal progenitor cells in human prostate organoid cultures. Cell *159*, 163-175.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739-1740.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Venables, W. N., and Ripley, B. D. (2002). Modern Applied Statistics with S. , (New York: Springer).

Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M. U., Ohgi, K. A.*, et al.* (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature *474*, 390-394.

Xu, K., Wu, Z. J., Groner, A. C., He, H. H., Cai, C., Lis, R. T., Wu, X., Stack, E. C., Loda, M., Liu, T.*, et al.* (2012). EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science *338*, 1465-1469.