Supplemental information

Mesenchymal and stem-like prostate cancer linked to therapy-induced lineage plasticity and metastasis

Hyunho Han, Yan Wang, Josue Curto, Sreeharsha Gurrapu, Sara Laudato, Alekya Rumandla, Goutam Chakraborty, Xiaobo Wang, Hong Chen, Yan Jiang, Dhiraj Kumar, Emily G. Caggiano, Monica Capogiri, Boyu Zhang, Yan Ji, Sankar N. Maity, Min Hu, Shanshan Bai, Ana M. Aparicio, Eleni Efstatiou, Christopher J. Logothetis, Nicholas Navin, Nora M. Navone, Yu Chen, and Filippo G. Giancotti
Supplemental information

Mesenchymal and stem-like prostate cancer linked to therapy-induced lineage plasticity and metastasis

Hyunho Han, Yan Wang, Josue Curto, Sreeharsha Gurrapu, Sara Laudato, Alekya Rumandla, Goutam Chakraborty, Xiaobo Wang, Hong Chen, Yan Jiang, Dhiraj Kumar, Emily G. Caggiano, Monica Capogiri, Boyu Zhang, Yan Ji, Sankar N. Maity, Min Hu, Shanshan Bai, Ana M. Aparicio, Eleni Efstatiou, Christopher J. Logothetis, Nicholas Navin, Nora M. Navone, Yu Chen, and Filippo G. Giancotti
**A**

Gene Atlas Analysis

- **C1. Fetal Liver** ($p < 0.0001$)
- **C2. Bone Marrow** ($p < 0.0001$)
- **C3. Prostate** ($p < 0.0001$)
- **C4. Lymph Node** ($p < 0.0001$)

**B**

**PCO-94 compiled dataset (PDX/Cell/Organoid, n=94)**

- MDA-Pca PDX, n=20
- MSKCC Pca Organoid/ODX, n=10 (Cell 2014)
- Broad Institute Pca Cell Line, n=8 (CCLE)
- LuCaP PDX, n=48 (GSE93809)
- LuCaP Organoid, n=8 (GSE113741)

**C**

**PCO-94 PAM clustering**

Consensus Plots

- **C1**
- **C2**
- **C3**
- **C4**

**D**

Consensus CDF

- **K = 2**
- **K = 3**
- **K = 4**
- **K = 5**
- **K = 6**
- **K = 7**
- **K = 8**
- **K = 9**
- **K = 10**

**E**

**Top 20 Genesets**

- Hallmark: Androgen Response
- WANG: Response to Androgen up
- PID AR TF Pathway
- DQANE: Response to Forskolin up
- NELSON: Response to Androgen up

**F**

**Cluster**

- **C1**
- **C2**
- **C3**

**G**

**Cluster**

- **C1**
- **C2**
- **C3**

**Figure S1. Han et al. 2021**
Supplementary Figure 1.

Related to Figure 1. Intrinsic transcriptomic subtyping and heterogeneity of PCOs.

A. Unsupervised clustering of the SU2C/PCF M-CRPC transcriptome dataset identifies four clusters corresponding to the sites of origin of the samples. The whole SU2C/PCF dataset (samples n=118, transcripts n=50,331, (Robinson et al., 2015)) was downloaded from cBioPortal.org. After batch correction (batch 1: BROAD INSTITUTE; batch 2: U. MICHIGAN), the dataset was subjected to unsupervised clustering using Partition Around Medoids (PAM) method (left dot plot). The top genes upregulated in the four clusters were: *DEFB1*, *C2orf70*, *AKR1B10*, *METTL7B* and *SERPINA1* in Cluster 1; *AC020663.1, AL590710.1, RP11-83M16.5, RP11-136F16.1* and *AC004815.1* in Cluster 2; *C9orf91, ERGIC1, STK39, HOMER2, ACACA in Cluster 3*; and *CXCR3, CD79A, IGHV4-61* and *IGLV3-25* in Cluster 4. Gene Atlas Analysis was performed by “Gene Enrichment” command of ASAP (Automated Single-cell Analysis Pipeline, (Gardeux et al., 2017)), computing overlap of reference datasets from BioGPS Human Cell Type and Tissue Gene Expression Profiles (Su et al., 2004). Cluster 1: Fetal liver (Odd Ratio (OR): 30.0, adj. p = 0). Cluster 2: Bone marrow (OR: 68.3, adj. p = 7.5e-5). Cluster 3: Prostate (OR: 24.7, adj. p = 1.2e-9). Cluster 4: Lymph node (OR: 545.1 adj. p = 9.7e-10). Sites of origin of the samples are shown below (right dot plot). C# = Cluster #.

B. Compilation of the tumor cell intrinsic PCO-94 dataset. The names, no. of samples and sources of the five datasets that were merged are indicated. MDA PCa PDX: microarray data of MDA PCa PDX, including tumors of same origin but grown in castrated and uncastrated hosts (133-4_cas1,2; 180-30_cas1,2) (Tzelepi et al., 2012); MSKCC PCa organoid/ODX (organoid-derived xenograft): mRNA expression (RNA Seq FPKM) data available for 10 of 12 PCa organoids (Gao et al., 2014). CCLE PCa Cell lines gene expression (RNA Seq RPKM) data (Release date: 14-Feb-2018. Broad Institute). LuCaP M-CRPC PDXs custom Agilent 44k whole genome expression microarray (includes early/late and castration-resistant passages. GSE93809. (Nguyen et al., 2017)). LuCaP PDX-derived organoids (RNA Seq TPM) data (includes two repeats. GSE113741(Beshiri et al., 2018)).

C, D. PCO-94 dataset consensus plots. Comparison of PAM clustering (K=3) and consensusclustering (K=2-5) results. Algorithm: SOM (Self-Oganizing Map); distance measure: Pearson (C). Consensus Cumulative Distribution Function (CDF) from K=2 to K=10. The proportion ofambiguous clustering (PAC) score was measured by the ΔCDF from u1(0.1) to u2(0.9) (left). PAC score bar graph sorted by the lowest PAC, indicating K=3 as optimal n of clusters (right)(D).

E. Volcano plots of top 20 positively (right, red) and negatively (left, blue) correlated gene sets along the principal coordinates 1 and 2 shown in Figure 1A. Pearson correlation coefficients of ssGSEA scores and coordinate values are shown. The C2
“curated”, the C5 “Gene Ontology” and the H “hallmark” gene sets (from the mSigDB collections) were used in the analysis.

F, G. Heatmap of normalized Enrichment Scores (NES) of the three clusters from Figure 1C. Hallmarks (+) enrichment: NES > +1.0 in one cluster only. Hallmarks (-) enrichment: NES < -1.0 in one cluster only (F). Prostate carcinogenesis gene sets: genes up/downregulated in PC vs benign prostate tissue (Liu et al., 2008; Tomlins et al., 2007; Wallace et al., 2008) or in PC cell lines but not in normal prostate epithelial or stromal cells (Yegnasubramanian et al., 2008). (bottom) Genes upregulated in oncogene activated PC tumor models (Acevedo et al., 2007; Azare et al., 2007; Liu et al., 2008) (G).
**A**

PCO samples, stratified by cluster

- **ARPC**
- **MSPC**
- **NEPC**

*Relative fraction*

- **Cluster 1**
- **Cluster 2**
- **Cluster 3**

Pure samples (largest fraction > 0.75)

Mixed samples (largest fraction < 0.75)

---

**B**

CCLE Cell Lines

- **ARPC**
- **MSPC**
- **NEPC**

---

**C**

LuCaP PDXs and Organoids

---

**D**

Organoids vs ODXs

- **ARPC**
- **MSPC**
- **NEPC**

---

**E**

Microarray vs RNA-seq

- **ARPC**
- **MSPC**
- **NEPC**

---

Figure S2, Han et al. 2021
Supplementary Figure 2.

Related to Figure 1. Intrinsic transcriptomic subtyping and heterogeneity of PCOs.

A. Relative fraction of ARPC, MSPC and NEPC in PCO samples (n=94). Samples were stratified by PAM cluster and sorted by largest fraction size. Samples were considered pure if the largest fraction was > 0.75 (white horizontal line) and mixed otherwise.

B. Relative fraction of ARPC, MSPC and NEPC in CCLE prostate cancer cell lines. Note that all samples are assigned as pure.

C. Hierarchical clustering and dendrogram of LuCaP PDXs and Organoids (n=56). Samples were aligned with the bar graph illustrating the relative fraction prediction from CIBERSORT analysis. Note that all repeat samples, early and late passage pairs and castration-resistant derivatives cluster together in each category. It was previously suggested that LuCaP PDX 86.2, classified as mixed, may be a mosaic sample because the ERG expression appeared after the first four passages (Nguyen et al., 2017).

D, E. Comparison of relative fraction prediction from pairs of tumor organoids and ODXs (organoid-derived xenografts) (D) and microarray data versus RNA-seq data of same PDXs (E). No statistically significant difference was noted (chi-square test). ns = not significant.

MSK-PCa#.o indicates organoids and MSK-PCa#.m organoid-derived mouse xenografts.
**Figure S3. Han et al. 2021**

### A. CIBERSORT Immune Cell Deconvolution (FHCRC + SU2C-PCF + UCSF)

- **Cluster 1**
  - SU2C-PCF (n=118)
  - FHCRC (n=171)
  - UCSF (n=99)

- **Cluster 2**
  - Cluster 3

- **Cluster 4**
  - Cluster 5

### B. Canonical EMT Markers

- **SU2C-PCF (Robinson et al., 2015)**
  - mRNA expression (log2)
  - Relative fraction

### C. Dysfunction Score

- **SU2C-PCF**
  - ARPC
  - MSPC
  - NEPC

### D. TIDE Score

- **HUGO_ANTI-PD1_NONRESPONDER_UP**
  - mRNA expression (log2)
  - Relative fraction

### E. RNAi dependency score

- **SU2C-PCF**
  - ARPC
  - MSPC
  - NEPC

### F. Relative fraction

- **22RV1**
- **VCAP**
- **MDAPCA2B**
- **LNCAP**
- **PC3**
- **DU145**
- **NCIH660**
- **MED12**
- **TEAD1**
- **SOX2**
- **PAX7**
- **KLF4**

### G. Chromosome 3q26

- **PRKCI**
- **SOX2-OT**

### H. Figure S3. Han et al. 2021
Supplementary Figure 3.

Related to Figure 2. Biological and clinical characteristics of M-CRPC subtypes.

**A.** Clustering of the M-CRPC patient datasets according to the PCO subtypes and calculation of the proportion of each subtype in each clinical sample. Three-dimensional plots of the PCoA analysis of the M-CRPC datasets. Clustering by Partition Around Medoids (PAM) (K=3) (left). Relative fractions of ARPC, MSPC and NEPC were calculated for each cluster by using CIBERSORT. n% indicates the proportions of each cluster in the three datasets. Clusters were numbered by their size. Note that cluster 1 of each dataset has the largest ARPC, cluster 2 MSPC and cluster 3 NEPC (right). Consistent with the results of CIBERSORT analysis, the separation between cluster 1 and 2 was less clear than that of cluster 3 versus the remainder in the PCoA plot.

**B.** Expression of canonical EMT genes in ARPC, MSPC and NEPC cases. P values were calculated by one-way ANOVA. Multiple comparison by Dunnett’s test. Original gene expression data from the SU2C-PCF dataset (Robinson et al., 2015).

**C.** Comparison of the classification of M-CRPC based on spontaneous intrinsic clustering of PCO datasets (Figure S2A) to the previously classification proposed by Dr. Nelson and colleagues (Bluemn et al., 2017; Su et al., 2019).

**D.** CIBERSORT deconvolution analysis of immune cell subsets in each subtype of M-CRPC. Immune cell population deconvolution analysis (LM22) was performed as previously described (Su et al., 2019). Pure ARPC, MSPC and NEPC samples from panel A were merged. P value calculated by one-way ANOVA. Immune cell subsets, which did not differ significantly between subtypes were not included in the graph. Box and whiskers plot (min to max). *p < 0.05, **p < 0.01, ***p < 0.001.

**E.** Tumor Immune Dysfunction and Exclusion (TIDE) score (Jiang et al., 2018) of the three subtypes in each M-CRPC dataset. The graphs represent the CD8+ T-Cell Dysfunction scores of each sample according to subtype and dataset. P values were calculated by one-way ANOVA. Multiple comparison by Dunnett’s test. *p < 0.05, **p < 0.01, ***p < 0.001.

**F.** Enrichment plots of the HUGO_ANTI-PD1_NONRESPONDER_UP gene set in MSPC as compared to the other two subtypes. GSEA was performed in the three M-CRPC datasets and the PCO-94 dataset. Gene set was generated by selecting genes upregulated in metastatic melanoma anti-PD-1 non-responders as compared to responders (Mann-Whitney test p < 0.01) (Hugo et al., 2016). NES = normalized enrichment score.

**G.** Copy number alterations at chromosome 3q26 in MSK-PCa Organoids and CCLE PC cell lines. Segment view by IGV (hg19). Relative fraction of ARPC, MSPC and
NEPC is shown on the right.

**H.** The Cancer Dependency Map RNAi dependency score of ARPC cell lines (LNCaP, MDAPCA2B, 22RV1 and VCAP), MSPC cell lines (DU145 and PC3) and the NEPC cell line (NCI-H660). The top 2 genes scoring in each subtype are shown.
**A**

Beltran et al (2016)

![Heatmap showing RB1 proportion across histologies](image)

**B**

![Heatmap showing staining for various markers](image)

**C**

- **FENG_CCLE_RB1_LOST_UP**
  - NES: 2.36
  - \( p: 0.00 \)

- **FENG_CCLE_RB1_LOST_DN**
  - NES: -1.99
  - \( p: 0.00 \)

Figure S4. Han et al. 2021
Supplementary Figure 4.

Related to Figure 2. Biological and clinical characteristics of M-CRPC subtypes

A. Deconvolution analysis of the CRPC-adeno + CRPC-NE dataset (Beltran et al., 2016). CRPC-adeno = prostate adenocarcinoma and CRPC-NE = neuroendocrine carcinoma. Genetic alteration of \( RB1 \) and histological subtype are indicated on the top.

B. Immunohistochemistry heatmap of PTEN, p53, RB1, AR, PSMA, ERG, vimentin and chromogranin A (CgA) in the PDXs grouped by transcriptomic subtype. Original data from the Movember GAP1 PDX project: An international collection of serially transplantable prostate cancer PDX models (Navone et al., 2018). Histology of the parental tumors are shown in the top row.

C. GSEA of the FENG_CCLE_RB1_LOST_UP and DOWN signatures in NEPC vs ARPC and MSPC in the PCO-94 dataset. A pan-cancer transcriptomic signature predicting biallelic loss of \( RB1 \) and comprising genes upregulated (left) or downregulated in \( RB1 \) mutant tumors (right) was used (Cheng et al., 2019). NES = normalized enrichment score.
Figure S5. Han et al. 2021
Supplementary Figure 5.

Related to Figure 2. Biological and clinical characteristics of M-CRPC subtypes

A, B. Proportion of ARPC, MSPC, and NEPC in the metastases of patients stratified by status of exposure to abiraterone and enzalutamide. Deconvolution analysis (A) and AR score and NEPC score (B) in the newer SU2C-PCF M-CRPC dataset (Abida et al., 2019). Samples were stratified by their NHAs exposure status – exposed, naïve, on treatment or unknown. Samples with neuroendocrine histology are excluded from the dataset. P values were calculated by one-way ANOVA. Multiple comparison by Dunnett’s test (mean ± S.D, ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001).

C. Survival analysis of on-treatment patients. Samples were classified by deconvolution and purity analysis (cut-off value: 0.6). MSPC cases were compared to non MSPC cases. Log-rank test.

D. Percentage of mutations and copy number alterations (CNAs) of TP53 and PTEN in primary adenocarcinoma subtypes. Chi-square test. ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001.

E, F. Percentage of Tumor stage (E) and Nodal stage (F) in primary adenocarcinoma subtypes. If radical prostatectomy data were available, pathologic T/N stage was used. Chi-square test. ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001.

G-I. Kaplan-Meier plots of progression-free survival in the CPC-GENE (G), TGCA (H) and DKFZ early onset (I) data sets. Log-rank (Mantel-Cox) test for comparison of survival curves. Progression was defined as biochemical recurrence after primary therapy in each study. In DKFZ, PCA034 was excluded from analysis (T01,2,4,6 classified as mixed, T03,5 as ARPC). All other sample assignments were consistent in each patient. Survival data were acquired from cBioPortal.org.

J. Proportion of intrinsic subtypes in primary prostate adenocarcinoma samples from the indicated four datasets. Cut-off value (0.6) for deconvolution and purity analysis was the same as in M-CRPC datasets. Note that NEPC is zero, likely due to the histologic criteria used for acquisition of the samples and the rarity of primary de novo neuroendocrine carcinoma in the prostate. n = number of samples with RNA expression data. ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001.
**Figure S6. Han et al. 2021**

- **A** Cell Growth and Apoptosis.
- **B** Chromatin Accessibility.
- **C** G1/S transition and mTORC1 signaling.
- **D** VCaP + enzalutamide.
- **E** VCaP + enzalutamide.
- **F** Chromatin Accessibility.
- **G** Major peak sites.
- **H** Chromatin Accessibility.
- **I** Chromatin Accessibility.
- **J** Chromatin Accessibility.
- **K** ssGSEA score.
Supplementary Figure 6.

Related to Figure 3. Exposure to enzalutamide induces E/M and stemness traits and confers tumor initiation capacity.

A. LNCaP cells were treated with enzalutamide for 0 to 14 days. Cell growth was determined by counting viable cells at the indicated time points of exposure to the drug (upper). Early apoptosis (annexin v+, PI+ cells) and late apoptosis (annexin v+ PI- cells) were determined by flow cytometry (lower). Mean ± S.D, n = 3 technical replicates.

B. BrdU incorporation in LNCaP cells treated with enzalutamide for 0, 3 and 7 days. Paraformaldehyde fixed cells were used for immunostaining of BrdU (green) and DAPI (blue)(scale bar = 200 um). Numbers indicate the percentage of BrdU positive cells at each time point (mean ± S.D, n = 3 technical replicates).

C. PCoA of the transcriptomes of LNCaP cells treated with enzalutamide for 0 to 14 days and from the PCO-94 dataset. X axis: drug incubation times. Y axis: coordinate 2 values. Three samples per time point. ssGSEA scores of the top two gene sets (G1/S transition and mTORC1 signaling) correlating with coordinate 2.

D, E. Immunofluorescent staining and immunoblotting of VCaP cells treated with enzalutamide for the indicated times. Cells were stained with antibodies to AR, ITGB4 and EpCAM and counterstained with DAPI (scale bar = 50 um) (D) or subjected to immunoblotting with antibodies to AR, AR-v7, TMPRSS2, and EMT and stemness markers (E). n = 3 technical replicates.

F-H. ATAC-seq analysis of LNCaP cells treated with enzalutamide for 0 to 14 days. Principal Component Analysis (PCA). Duplicates per time point. X axis component 1 is indicated as “Major” peaks, and Y axis component 2 as “Minor” peaks in subsequent analyses. Inner donutplot: total variability explained (%). Major peaks were primarily defined by the variability of PCA component 1 (F). Pie chart of the assignment of major peaks to the indicated genomic regions (G). Enhancer sites and promoter regions in “major” peak summit. Average of duplicates (H).

I. Shift of promoter histone marks during enzalutamide treatment of LNCaP cells. H3K4me3 (upper) and H3K27me3 (lower) peaks of genes upregulated in ARPC, MSPC and NEPC (defined in Figure 1A) were determined by ChIP-seq. Distance = distance from transcription start site. Mean read density, n = 3 technical replicates.

J. H3K4me3 and H3K27me3 signals over the promoters of EMT and stemness-related genes in LNCaP cells treated with enzalutamide (0 to 14 days). Mean read density, n = 3 technical replicates.

K. ssGSEA score heatmap of LNCaP cells treated with enzalutamide. Day 0 samples and day 14 samples. The hallmark gene sets with signal/noise ratio > 2.0 were used.
Figure S7. Han et al. 2021
Supplementary Figure 7.

Related to Figure 4. NRG1 rescues reprogrammed LNCaP cells from quiescence and enables metastasis.

A. Single cell RNA-seq analysis of LNCaP cells treated with enzalutamide for the indicated times. UMAP plot of merged data.

B. Principal gene expression features of LNCaP clusters ordered by deduced evolutionary trajectory.

C. Heatmap of ssGSEA scores in each cluster. Columns aligned by pseudotime. Prostate luminal progenitor gene sets from Henry et al., 2018. Others: Hallmark gene sets.

D. Violin plots of ssGSEA scores. Enrichment of hallmark EMT, Hedgehog and Ras, embryonic diapause, prostate luminal progenitor, hallmark E2F target, DNA-repair and Myc target gene sets in each cluster. P values were calculated by one-way ANOVA. Multiple comparison by Dunnett’s test. *p < 0.05, **p < 0.01, ***p < 0.001.

E. Violin plots of EPCAM and ITGB4 gene expression in cluster 1 to 6. P values were calculated by one-way ANOVA. Multiple comparison by Dunnett’s test. *p < 0.05, **p < 0.01, ***p < 0.001.

F. FACS analysis of LNCaP cells treated with enzalutamide for the indicated times. Two populations defined by EpCAM and CD104 expression were identified: EpCAM^{HIGH}CD104^{LOW} and EpCAM^{LOW}CD104^{HIGH}. The images are representative of three independent replicates.

G. Reprogrammed LNCaP cells are not rescued by EGF, FGF, or the combination. LNCaP^{GFP} cells were treated with enzalutamide for 7 days and, when indicated, for 5 additional days with enzalutamide and 20 ng/mL recombinant human EGF, 40 ng/mL of FGF, or the combination. Cell numbers were estimated by the intensity of GFP fluorescence. Welch’s t-test. **p < 0.01, ns = not significant.

H. Gene expression heatmap of androgen response ssGSEA score, AR targets and prototypical hybrid E/M markers of LNCaP cells treated or not with enzalutamide or enzalutamide plus rhNRG1.

I. Immunoblotting of AR, AR targets and prototypical E/M markers in LNCaP cells treated or not with enzalutamide or enzalutamide plus rhNRG1. β-Actin was used as loading control. Representative WB of 3 independent experiments.

J. Volcano plot of the correlation between the ssGSEA score of the cluster_3_up gene set and the Z score of RPPA protein expression in the TCGA prostate
adenocarcinoma dataset.

**K.** Double immunofluorescent staining of AR and ITGB4 in LNCaP and re-LNCaP-NRG1 cells. The images are representative of 3 independent replicates (scale bar = 100 um).

**L.** The graph indicates the percentage of AR+, AR low and AR+ tumor cells in the adrenal gland metastases generated by reprogrammed LNCaP cells (left). mean ± S.D. Representative areas from AR- (mouse #181) and AR+ metastases (#182) are shown (right) (scale bar = 100 um).
Figure S8. Han et al. 2021
Supplementary Figure 8.

Related to Figure 5. Enzalutamide-induced transcriptional inactivation of TP53 is mediated by BRCA1 and E2F1 and confers hybrid E/M and stemness traits.

A. Top 10 gene sets enriched in MSPC as compared to ARPC in the PCO-94 dataset. Genesets are from the C6 Oncogenic Signatures (n=189) of the MSigDB Collection.

B. Enrichment plots of P53_DN.V1_UP and P53_DN.V2_UP gene sets in MSPC vs ARPC in the PCO-94 dataset.

C. Scatter plot of TP53 mRNA expression z-score and AR activity z-score of samples with wild type TP53 (no mutation, diploid) in the SU2C-PCF dataset (n=58) and FHCRC dataset (n=72). AR activity was measured by ssGSEA of the HALLMARK_ANDROGEN_RESPONSE gene set.

D. Heatmap of TP53 target genes by category. FPKM values from RNA-seq of LNCaP cells treated with 10 µM enzalutamide for 0, 1, 3, 5, 7 or 14 days. Target gene list acquired from Fischer M. Census and evaluation of p53 target genes (Fischer, 2017).

E. Profile of ChIPseq peaks for BRCA1 and E2F1 on the promoter of TP53. ChIPseq data were acquired from Remap 2020 of the ENCODE project and aligned around the TP53 transcriptional start site (TSS).
Figure S9. Han et al. 2021
Supplementary Figure 9.

Related to Figure 5. Enzalutamide-induced transcriptional inactivation of TP53 is mediated by BRCA1 and E2F1 and confers hybrid E/M and stemness traits.

A. ChIP-qPCR analysis of BRCA1 and control immunoglobulin G (IgG) binding to the TP53 promoter in LNCaP cells. Schematics of the oligonucleotide primers at the promoter and exon7 (top). Occupancy of the promoter by BRCA1 (left). Occupancy of the promoter by BRCA1 in sh-control and sh-BRCA1 LNCaP cells (right). Error bars indicate mean ± SD. **p<0.01, ***p<0.005, ****p<0.001, n = 3 technical replicates.

B. ChIP-qPCR analysis of E2F1 and control immunoglobulin G (IgG) binding to the TP53 promoter in LNCaP cells. Schematics of the oligonucleotide primers at the promoter and exon7 (top). Occupancy of the promoter by E2F1 (left). Occupancy of the promoter by E2F1 in si-control and si-E2F1 LNCaP cells (right). Error bars indicate mean ± SD. **p<0.01, ***p<0.005, ****p<0.001, n = 3 technical replicates.

C, D. ChIP-qPCR analysis of the enrichment of H3K4me3, H3K27me3 and Ser2-P RNA Pol II at the TP53 promoter in sh-BRCA1 (C) or si-E2F1 LNCaP cells (D) as compared to controls. Error bars denote mean ± SD. **p<0.01, ***p<0.005, ****p<0.001, n = 3 technical replicates.

E. Relative levels of Brca1 (left) and Trp53 (right) mRNAs in Pten-P8−/− cells stably transduced with the indicated Brca1 short hairpin RNAs. GAPDH served as internal control. n = 3 technical replicates.

F, G. ChIP-qPCR analysis of the enrichment of H3K4me3, H3K27me3, Ser2-P RNA Pol II and Brca1 at the TP53 promoter in shBRCA1 (F) or enzalutamide treated Pten-P8−/− cells (G). Error bars denote mean ± SD. **p<0.01, ***p<0.005, ****p<0.001, n = 3 technical replicates.

H. Immunoblotting with antibodies to p53, AR, and selected EMT markers in Pten-P8−/− cells stably transduced with the indicated Trp53 hairpins. Rho-GDI served as loading control. Representative WB of 3 independent experiments.

I. Quantification (left) and representative images (right) of the tumor spheres formed by control and Trp53-silenced Pten-P8−/− cells at day 10 after seeding. The indicated cells were plated in triplicate in 24 well ultra-low attachment plates at a seeding density of 3,000 cells/well. Error bars, mean ± SD of triplicate experiments, **p<0.01, ***p<0.005, ****p<0.001 two-tailed Student t test, scale bar = 500 um.

J. Quantification of Matrigel invasion assay. Trp53-silenced Pten-P8−/− cells were subjected to Matrigel invasion assay in response to NRG1. Indicated cells were
plated in triplicates in 24 well Matrigel coated chambers at a seeding density of 10,000 cells/well and counted 96 hrs after seeding with and without NRG as attractant. The error bars represent the SD of triplicate experiments. **p<0.01, ***p<0.005, ****p<0.001 two-tailed Student t test.
Supplementary Figure 10.

Related to Figure 6. Inactivation of BMP-SMAD signaling promotes anti-androgen resistance.

A. Scatter plots of the correlation between ssGSEA scores of the hallmark androgen response and LDN193189 response in the SU2C-PCF dataset (left) and UCSF dataset (right).

B. Expression levels of the indicated BMP-SMAD signaling components measured by cDNA microarray in late passage LNCaP cells transduced with sh-Ctrl and sh-AR RNAs. Data from GSE22483 (Gonit et al., 2011). mean ± S.D.

C. Cell growth assay of LNCaP cells treated with or without LDN193189 for the indicated times. mean ± S.D, ns: not significant.

D. Immunoblotting of P-SMAD1/5/8 and total SMAD6 overexpressing LNCaP cells treated or not with BMP4 (100 ng/mL, 30 mins) or pretreated with anti-NEO1 blocking antibody (5 µg/mL), control IgG or LDN193189 at 25 and 50 nM for 2 hours and then treated with BMP4. Representative WB of n = 3 technical replicates.

E. Immunoblotting of AR and NEO1 in total lysates of LNCaP cells transfected with control or human Neo1 specific siRNA (SMART pool) at the indicated time points. Representative WB of n = 3 technical replicates.

F. Immunoblotting of AR, P-SMAD1/5/8, and total SMAD1 in total lysates of LNCaP cells pretreated with anti-NEO1 blocking antibody (5 µg/mL) or control IgG and treated with BMP4 (100 ng/mL, 30 mins). Representative WB of n = 3 technical replicates.

G, H. ClusterProfiler dot plot shows the GSEA of EMT (G) and stemness and lineage signatures (H) enriched in control as compared to LDN193189 treated LNCaP cells.

I, J. Quantification (I) and representative images (J) of tumor organoid formation by control and TP53-silenced LNCaP cells. The indicated cells were pretreated with or without 50 nM LDN193189 and/or rh-NRG1 for 8 days. 100 –10,000 dissociated cells were plated in ultra-low attachment 96 well plates. Number of organoids per field (left) and diameter in micrometers (right) were measured at day 10. Scale bar=100 µm. Error bars, mean ± SD of triplicate experiments, **** p<0.0001 two-tailed Student t test.

K. Relative cell growth of sh-Co., sh-TP53#1 and sh-TP53#2 LNCaP cells pretreated or not with 50 nM LDN-193189 for 8 days and then cultured in 10% FBS (left), 10% CSS and 10 µM enzalutamide (middle) or 10% CSS and 25 µM enzalutamide for additional 9 days (right). Mean ± S.D of triplicate experiments, ****p < 0.0001 two-
tailed Student t test.

L. Clusterprofiler dot plot shows the GSEA of the top 10 signatures from the Hallmark gene sets in control as compared to LDN193189 treated LNCaP cells (left) or in control as compared to enzalutamide treated LNCaP cells (7 days) (right). The X axis indicates the ES"represents the GSEA enrichment score (ES). The Y axis shows the names of the signatures. Dot size represents the -log10 (FDR_q_value + 0.001). Dot color represents the significance.
Supplementary Figure 11.

Related to Figure 7. Preclinical efficacy of combination therapies for mixed ARPC/MSPC and MSPC.

A. Predicted drug sensitivity heatmaps of the MSPC samples from the FHCRC and the UCSF datasets. Heatmap shows the scores of sensitivities to the pan-HER inhibitor neratinib or the HER1/2 inhibitor lapatinib and the FGF inhibitor AZD4547 or nintedanib (top) and levels of expression of the HER1-4 mRNAs (bottom). Spearman correlation coefficients between the indicates rows(right).

B. Cell viability dose-response curves of control LNCaP and re-LNCaP-NRG1 cells to lapatinib, MLN0128, AZD4547 or Erlotinib. Error bars, mean ±SD of 8 technical replicates, ****p<0.0001 Welch's unequal variances t-test of AUC.

C. Numbers, sizes and total burden (total sum of sizes) of metastases generated by re- LNCaP-NRG1 injected i.c. in castrated male NSG mice, which were then treated with enzalutamide at 10 mg/kg and/or neratinib at 40 mg/kg daily for 4 weeks. v = vehicle; e = enzalutamide; n = neratinib. *p < 0.05, ***p < 0.001, ****p < 0.0001.

D. Immunohistochemical staining of AR and P-HER2 in the metastases generated by re-LNCaP-NRG1 cells in mice treated with enzalutamide, neratinib, or the combination. Percentages of nuclear AR positivity are indicated (scale bar = 100 um).

E. Immunoblotting with antibodies to P-AKT (S473), total AKT, P-ERK1/2 and total ERK in re-LNCaP-NRG1 cells treated or not with 2 ng/ml NRG1 in the presence of the indicated concentrations of neratinib or MKK2206 for 4 hours. Representative WB of n = 3 technical replicates.

F. Immunoblotting with antibodies to P-AKT (S473), total AKT, P-S6 and total S6 protein in re-LNCaP-NRG1 cells treated or not with 2 ng/ml NRG1 in the presence of the indicated concentrations of the mTOR kinase inhibitor MLN0128 or AZD8055 for 4 hours. Representative WB of n = 3 technical replicates.

G. Immunoblotting with antibodies to total and phosphorylated HER2/3 and components of the ERK1/2, AKT and mTOR kinase pathway in re-LNCaP-NRG1 cells stimulated with NRG1 and treated for 1 or 2 days with 50 nM neratinib (nera), 0.5 μM MK2206 (MK) or 100 nM MLN0128(MLN). Representative WB of n = 3 technical replicates.

H, I. Cell viability dose response curve of LNCaP cells treated with 50 nM MLN1028 in combination with the indicated doses of neratinib (H) and DU145 cells treated with 100 nMneratinib in combination with the indicated doses of MLN128 (I). Error bars, mean ±SD of 8 replicates, **** p<0.0001 Welch's unequal variances t-test of AUC.

J. Bioluminescent imaging of the metastases generated by DU145 cells injected i.c. in castrated male NSG mice treated with 40 mg/kg neratinib, 0.3 mg/kg MLN0128, or
the combination daily for 4 weeks. The graph shows the total normalized photon flux at the indicated times. n = 5 mice per group, mean ± S.D, *p < 0.05, ****p < 0.0001, two-tailed Student t test.

K. Total normalized photon flux of the DU145 metastases at 4 weeks in castrated male NSG mice treated with daily neratinib at 40 mg/kg, MLN0128 at 0.3mg/kg, or the combination for 4 weeks. n = 5 mice per group, mean ± IQR, *p < 0.05, ***p < 0.001, ****p < 0.0001.
| Antibodies | REAGENT | SOURCE | IDENTIFIER |
|------------|---------|--------|------------|
| ITGB4 (Western blot, IHC) | Cell Signaling | Cat# 14803; RRID: AB_2798620 |
| AR (Western blot) | Cell Signaling | Cat# 5153; RRID: AB_10691711 |
| AR (Western blot) | Santa Cruz | Cat# Sc-816; RRID: AB_630864 |
| AR (IHC) | Abcam | Cat# ab133273; RRID: AB_11156085 |
| Synaptophysin | Abcam | Cat# ab32127; RRID: AB_2286949 |
| ZO-1 (D7D12) | Cell Signaling | Cat# 8193; RRID: AB_10898025 |
| p-AKT S473 | Cell Signaling | Cat# 4060; RRID: AB_2315049 |
| AKT-pan | Cell Signaling | Cat# 4691; RRID: AB_915783 |
| p-AKT T308 | Cell Signaling | Cat# 13038; RRID: AB_2629447 |
| p-PRAS40 S183 | Cell Signaling | Cat# 5936; RRID: AB_10838139 |
| PRAS40 | Cell Signaling | Cat# 2691; RRID: AB_2225033 |
| p-4EBP1 T37/46 | Cell Signaling | Cat# 2855; RRID: AB_560835 |
| 4EBP1 | Cell Signaling | Cat# 9452; RRID: AB_331692 |
| p-ERK1/2 T202/Y204 | Cell Signaling | Cat# 4370; RRID: AB_2315112 |
| ERK1/2 | Cell Signaling | Cat# 4695; RRID: AB_390779 |
| p-S6 S235/236 | Cell Signaling | Cat# 4858; RRID: AB_916156 |
| S6 | Cell Signaling | Cat# 2217; RRID: AB_331355 |
| Beta-Actin | Santa Cruz | Cat# sc-47778; RRID: AB_2714189 |
| KLK3 | Cell Signaling | Cat# 5365; RRID: AB_2797609 |
| E-cadherin | Cell Signaling | Cat# 3195; RRID: AB_2291471 |
| Vimentin (Western blot) | Cell Signaling | Cat# 5741; RRID: AB_10695459 |
| Vimentin (IHC) | Agilent (DAKO) | Cat# M7020; RRID: AB_2304493 |
| Vimentin (IHC) | Abcam | Cat# GR12L-100UG; RRID: AB_212942 |
| Fibronectin (3F12) | Life Technologies | Cat# MA514737; RRID: AB_10985252 |
| NSE | Abcam | Cat# ab53025; RRID: AB_881756 |
| Chromogranin A | Abcam | Cat# ab15160; RRID: AB_301704 |
| p53 | Cell Signaling | Cat# 9282; RRID: AB_331476 |
| p53 (7F5) (Western blot) | Cell Signaling | Cat#2527; RRID: AB_10695803 |
| BRCA1 | Santa Cruz Biotechnology | Cat# sc-6954; RRID: AB_626761 |
| E2F1 | Cell Signaling | Cat# 3742; RRID: AB_2096936 |
| Neogenin | Sigma-Aldrich | Cat# HPA027806; RRID: AB_10601006 |
| HER3 | Cell Signaling | Cat# 12708; RRID: AB_2721919 |
| Phosho-HER3 (Tyr1289) | Cell Signaling | Cat# 4791; RRID: AB_2099709 |
| Phosho-HER2 (Tyr1248) | Cell Signaling | Cat# 2247; RRID: AB_331725 |
| Antibody Name                  | Source  | Catalog Number | RRID          |
|-------------------------------|---------|----------------|---------------|
| HER2 (44E7)                   | Cell Signaling | Cat# 2248, RRID: AB_2099242 |
| Phospho-53BP1 (Ser1778)       | Cell Signaling | Cat# 2675, RRID: AB_490917 |
| Phospho-Smad1 (Ser463/465)/ Smad5 (Ser463/465)/ Smad8 (Ser462/468) | Cell Signaling | Cat# 13820, RRID: AB_2493181 |
| NKX3-1                        | Cell Signaling | Cat# 83700, RRID: AB_2800027 |
| BMPR1B                        | Abcam    | Cat# ab175385-100ul ; RRID: AB_2827412 |
| BMPR2                         | Abcam    | Cat# 130206; RRID: AB_11155873 |
| SMAD1                         | Abcam    | Cat# ab33902, RRID: AB_777975 |
| Rho GDlalpha (A-20)           | Santa Cruz | Cat# sc-360, RRID: AB_2227516 |
| BMPR2                         | Abcam    | Cat# ab92323, RRID: AB_10585592 |
| CD326 (EpCAM) Mouse anti-Human, PE (1B7) | eBioscience™ | Cat# 12-9326-42; RRID: AB_11044497 |
| CD104 (Integrin beta 4) Rat anti-Human, eFluor® 660 (439-9B) | eBioscience™ | Cat# 50-1049-82; RRID: AB_11219277 |
| BV421™ anti-human CD44        | BioLegend | Cat# 338809; RRID: AB_2562405 |
| PerCP/Cy5.5 anti-human PSMA/FOLH1 | BioLegend | Cat# 342512; RRID: AB_2563245 |
| H3K27me3 (ChIP)               | Millipore | Cat# 07-449; RRID: AB_310624 |
| H3K4me3 (ChIP)                | Cell Signaling | Cat# 9751S; RRID: AB_2616028 |
| SMAD1/5/8 (ChIP)              | Santa Cruz | Cat# sc-6031x; RRID: AB_785721 |
| pSMAD1/5/9 (ChIP)             | Cell Signaling | Cat# 13820S; RRID: AB_2493181 |
| AR (ChIP)                     | Millipore | Cat# 06-680; RRID: AB_310214 |
| Pol II (ChIP)                 | BioLegend | Cat# 664912; RRID: AB_2650945 |
| Pol II pSer2 (ChIP)           | Abcam    | Cat# ab5095; RRID: AB_304749 |
| BRCA1 (ChIP)                  | Bethyl   | Cat# A300-000A; RRID: AB_67367 |
| E2F1 (ChIP)                   | Millipore | Cat# 05-379; RRID: AB_2096772 |

**Table S8** List of the antibodies, Related to Key resources table.
| Oligonucleotides | Source | Identifier |
|------------------|--------|------------|
| **REAGENT**      | **SOURCE** | **IDENTIFIER** |
| Human TP53 short hairpin #1 (TGCTGTGACATGACGCAGCTTATTTTAC AATAAATAGTGAACAGAGTAGTTTATTGTAA AATAAGAGACGCTTACTGCTCCTCGGA) | Pelossof, R., Fairchild, L. et al. (2017) | N/A |
| Human TP53 short hairpin #2 (TGCTGTGACATGACGCAGCTTACTACTA AATAAATAGTGAACAGAGTAGTTTATTGTAA AATAAGAGACGCTTACTGCTCCTCGGA) | Pelossof, R., Fairchild, L. et al. (2017) | N/A |
| Human BRCA1 short hairpin #1 (TGCTGTGACATGACGCAGCTTACTACTA AATAAATAGTGAACAGAGTAGTTTATTGTAA AATAAGAGACGCTTACTGCTCCTCGGA) | Pelossof, R., Fairchild, L. et al. (2017) | N/A |
| Human BRCA1 short hairpin #2 (TGCTGTGACATGACGCAGCTTACTACTA AATAAATAGTGAACAGAGTAGTTTATTGTAA AATAAGAGACGCTTACTGCTCCTCGGA) | Pelossof, R., Fairchild, L. et al. (2017) | N/A |
| Human non-targeting control short hairpin (TGCTGTGACATGACGCAGCTTACTACTA AATAAATAGTGAACAGAGTAGTTTATTGTAA AATAAGAGACGCTTACTGCTCCTCGGA) | Pelossof, R., Fairchild, L. et al. (2017) | N/A |
| Silencer® Select Validated Human BRCA1 siRNA [s457] | ThermoFisher Scientific | 4390824 |
| Silencer® Select Validated Human E2F1 siRNA [s4405] | ThermoFisher Scientific | 4390824 |
| Silencer® Select Validated Human MYBL2 siRNA [s9118] | ThermoFisher Scientific | 4392420 |
| Silencer® Select Validated Human SNAI2 siRNA [s13127] | ThermoFisher Scientific | 4390824 |
| Silencer® Select Validated Human RUNX2 siRNA [s2457] | ThermoFisher Scientific | 4392420 |
| Silencer® Select Validated Human HOXC9 siRNA [s6827] | ThermoFisher Scientific | 4392420 |
| Silencer® Select Negative Control No. 1 siRNA | ThermoFisher Scientific | 4390843 |
| Silencer® Select Validated Human NEO1 siRNA [s9452] | ThermoFisher Scientific | 4392420 |
| Silencer® Select Validated Human BMPR1B siRNA [s2042] | ThermoFisher Scientific | 4392420 |
| Silencer® Select Validated Human BMPR2 siRNA [s2044] | ThermoFisher Scientific | 4427038 |
| Silencer® Select Validated Human BMPR2 siRNA [s2045] | ThermoFisher Scientific | 4427038 |
| RGMA Taqman probe | ThermoFisher Scientific | Hs00297192_mq |
| RGMb Taqman probe | ThermoFisher Scientific | Hs00543559_m1 |
| BMPR2 Taqman probe | ThermoFisher Scientific | Hs00176148_m1 |
| BMPR1B Taqman probe | ThermoFisher Scientific | Hs01010965_m1 |
| BMPR1A Taqman probe | ThermoFisher Scientific | Hs01034913_g1 |
| NEO1 Taqman probe | ThermoFisher Scientific | Hs00933950_m1 |
| SMAD1 Taqman probe | ThermoFisher Scientific | Hs00195432_m1 |
| Oligonucleotide | Supplier          | Product Code         |
|----------------|-------------------|----------------------|
| SMAD5 Taqman probe | ThermoFisher Scientific | HS00195437_m1 |
| SMAD9 Taqman probe | ThermoFisher Scientific | HS00931723_m1 |
| PUMA/BBC3 Taqman probe | ThermoFisher Scientific | HS00248075_m1 |
| CDKN1A Taqman probe | ThermoFisher Scientific | HS00355782_m1 |
| P53 Taqman probe | ThermoFisher Scientific | HS01034249_m1 |
| E2F1 Taqman probe | ThermoFisher Scientific | HS00153451_m1 |
| BRCA1 Taqman probe | ThermoFisher Scientific | HS01556193_m1 |
| MYBL2 Taqman probe | ThermoFisher Scientific | HS00942540_m1 |
| GAPDH Taqman probe | ThermoFisher Scientific | HS99999905_m1, HS02786624_g1 |
| 18S Taqman probe | ThermoFisher Scientific | HS99999901_s1 |
| ChIP qPCR: Human AR promoter 1F | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human AR promoter 1R | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human AR promoter 2F | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human AR promoter 2R | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human AR negative control-F | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human AR negative control-R | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human TP53 promoter-F | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human TP53 promoter-R | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human TP53 exon-F | SYBR green (Sigma) | N/A |
| ChIP qPCR: Human TP53 exon-R | SYBR green (Sigma) | N/A |
| ChIP qPCR: Mouse Trp53 promoter-F | SYBR green (Sigma) | N/A |
| ChIP qPCR: Mouse Trp53 promoter-R | SYBR green (Sigma) | N/A |

**Table S9** List of the oligonucleotides, Related to Key resources table.