SUPPLEMENTARY MATERIALS

Identifying Phased Mutations and Complex Rearrangements in Human Prostate Cancer Cell Lines through Linked-Read Whole Genome Sequencing

Minh-Tam Pham[1,2,4,5], Anuj Gupta[4], Harshath Gupta[1,4], Ajay Vaghasia[1,4,5], Alyza Skaist[4], McKinzie A. Garrison[1,3,4,6], Jonathan B. Coulter[2,4], Michael C. Haffner [4,7,8], S. Lilly Zheng[9], Jianfeng Xu[9], Christina DeStefano Shields[1,4], William B. Isaacs[1,2,5], Sarah J. Wheelan[1,3,4], William G. Nelson[1,4,5], Srinivasan Yegnasubramanian[1,4,5]

Affiliations:

[1]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[2]Department of Urology, James Buchanan Brady Urological Institute, Johns Hopkins University School of Medicine, Baltimore, MD, USA.

[3]Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[4]Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, MD, Baltimore, USA

[5]Cellular and Molecular Medicine Graduate Program, Johns Hopkins University School of Medicine, MD, Baltimore, USA

[6]Biochemistry, Cellular and Molecular Biology Graduate Program, Johns Hopkins University School of Medicine, MD, Baltimore, USA

[7]Division of Human Biology and Clinical Research, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[8]Department of Pathology, University of Washington, Seattle, WA, USA

[9]Program for Personalized Cancer Care, Northshore University Health System

Supplementary Methods

Pubmed Query of Prostate Cancer Cell lines**:**

To assess the number of manuscripts citing the use of all of the 20 distinct prostate cancer cell lines reported in references (1, 2, 3), we carried out the following query in Pubmed on July 2, 2021:

"DU-145"[All Fields] OR "DU145"[All Fields] OR "PC3"[All Fields] OR ("pc 3 cells"[MeSH Terms] OR ("pc 3"[All Fields] AND "cells"[All Fields]) OR "pc 3 cells"[All Fields] OR "pc 3"[All Fields]) OR ("lncap"[All Fields] OR "lncaps"[All Fields]) OR "ARCAP"[All Fields] OR "DUCAP"[All Fields] OR "LAPC-4"[All Fields] OR "LAPC4"[All Fields] OR "MDA-PCA-2A"[All Fields] OR "MDA-PCA-2B"[All Fields] OR "NCI-H660"[All Fields] OR "22RV1"[All Fields] OR "1013L"[All Fields] OR "PC-346C"[All Fields] OR "PC346C"[All Fields] OR "PSK-1"[All Fields] OR "UM-SCP-1"[All Fields] OR "VCAP"[All Fields]

This resulted in identification of 23,029 citations. We also queried each cell line individually, and noted that the top 7 cited cell lines were PC3, LNCaP, DU145, 22RV1, VCaP, LAPC-4, MDA-PCA-2B, which collectively accounted for 22,920 citations which was determined by the following Pubmed query:

"DU-145"[All Fields] OR "DU145"[All Fields] OR "PC3"[All Fields] OR ("pc 3 cells"[MeSH Terms] OR ("pc 3"[All Fields] AND "cells"[All Fields]) OR "pc 3 cells"[All Fields] OR "pc 3"[All Fields]) OR ("lncap"[All Fields] OR "lncaps"[All Fields]) OR "LAPC-4"[All Fields] OR "LAPC4"[All Fields] OR "MDA-PCA-2B"[All Fields] OR "22RV1"[All Fields] OR "VCAP"[All Fields]

Cell line culture methods

The cancer cell lines included three widely-used androgen deprivation sensitive cell lines (LAPC4, VCaP, and LNCaP), and their respective androgen deprivation insensitive subclones LAPC4-CR (4), VCaP-CR , LNCaP_Abl (5) and LNCaP_C42b (6). Additional androgen receptor (AR) positive cell lines included CWR22Rv1, an androgen deprivation sensitive line derived from a xenograft model established from a primary adenocarcinoma, and MDA-PCa-2b derived from a patient of African American ancestry (1). Remaining cell lines included two commonly used AR-negative prostatic carcinoma lines, PC3 and DU145 (1). Frozen pellets were used for mycoplasma testing and STR genotyping (GenePrint 10 System) performed by the Johns Hopkins Genetic Resource Core Facility. STR percent match was calculated using the second Master vs query algorithm (7) (based on ATCC query STR data. With cell lines that do not have ATCC query STR data, the STR profile of the parental cell line was used for reference. All cell lines were mycoplasma free and passed the a priori defined STR genotype threshold of 75%.See Table S1

Analysis of Allele Fraction and Copy number

Cell lines in the WI dataset were cultured per manufacturer's recommendations. A customized NGS panel targeting 355 cancer related genes was used to sequence DNA of all samples. The probes for capturing exon regions in these genes were manufactured by Roche NimbleGen. SeqCap EZ Library SR User's Guide (Roche, Pleasanton, CA) was followed for library preparation and capture of targeted sequences. Paired-end sequencing of 2 x 150 bp was performed on an Illumina NextSeq500. The median coverage for these 64 samples was 324 X. Paired-end reads were aligned to the GRCh37 version of the

human genome using Burrows-Wheeler Aligner v0.7 to generate BAM files (8). After sorting the BAM files using samtools, PCR duplicates marked using Picard and realignment around putative gaps was performed using the Genome Analysis Toolkit (GATK) v3.2-2. Variant calling was performed with the GATK Haplotype caller. ANNOVAR (http://annovar.openbioinformatics.org/en/latest) and snpEff were used for annotating variants and for retrieving information on variants in the population-based studies such as the 1000 Genomes Project (www.1000genomes.org), NHLBI-ESP 6500 exomes, or ExAC (http://exac.broadinstitute.org/) or gnomAD(http:// http://gnomad.broadinstitute.org/), and clinical databases such as the Human Gene Mutation Database (HGMD) (9). CCLE database was downloaded from the Depmap Portal (https://depmap.org/portal/download/), focused on the CGA_WES/WGS column(10). Allele Fraction was calculated as the number of alternative reads over total reads (t_alt/t_depth).

Orthogonal copy number analysis was carried out by the HMMCopy Suite and R package (Ver. 3.14) (11). For copy number analysis to address cell line drift between passages, and to compare between CCLE and our Linked read data, read counts from every consecutive 50kb bins across the genome (Hg19) were generated by mapCounter, blacklist filtered, and median read counts across every 4 bins were calculated to account for potential outliers in high coverage reads. Pearson correlation coefficients between samples were calculated using the corrplot R package (12). For high resolution analysis of Linked-read data, 1kb bin was used for segmentation calls also by HMMCopy. For Gene level copy number, bedtools was used to intersect Hg19-gene list (UCSC) with GC and mappability normalized compy number in 50kb bins (median was taken for genes spanning multiple bins).

Sequencing and phasing performance for Linked-read sequencing of human PCa cell lines

For all cell lines, >80% of the extracted DNA had high molecular weight >60 kbp as measured by TapeStation Bioanalyzer (Table S2). Next generation sequencing of linked-read libraries generated from these samples resulted in 35X to 54X average genome coverage (Table S2). Initial quality analysis of the linked-read sequencing data revealed that on average, 53 to 97 linked reads could be phased to high molecular weight DNA molecules with a calculated average size of 59 kb to 144 kb (Table S2). These data could be used to assemble haplotype phased blocks typically ranging from 10s to >100 megabase (Mb), allowing phasing of >98% of all single nucleotide polymorphisms (SNPs) and genes for all cell lines (Table S2).

Identification of somatic mutations in the PCa cell lines

We assessed likely somatic mutations, defined as nucleotide variants (SNV) and small insertions and deletions, by filtering for mutations that exhibited very low population allele frequencies (<0.1%) in the Genome Aggregation Database (gnomAD)(26). Somatic mutations accounted for 0.88% of all variants detected.

Analysis for phased variants and SV

Only passed variants from phased_variants.vcf files produced by Long Ranger were annotated using the vcf2maf (13) pipeline and the Uniport consensus reference. To filter for likely somatic variants, only variants not described in the gnomAD database or those with gnomAD_AF of less than .1% were retained. To limit to variants with potential protein functional consequences, any variants annotated as "IGR" (integenic), "Silent", "Targeted_Region" under Variant classification were filtered out. Categories that started with "regulatory_region_variant", "TF_binding_site_variant", "splice_donor_variant" ,

"splice_region_variant", "coding_sequence_variant" under Consequence were also filtered out. For the list of genes frequently implicated in cancer, the list of 97 genes recurrently mutated in PCa as identified by Armenia et al. as well as all Cosmic Tier 1(19) 576 genes were used as a reference list of genes recurrently mutated in PCa and cancer in general. *Cosmic-Longtail* is the Gene-level haplotype was inferred through a custom R (version 4.0.0) script that used the haplotype (column GT) and phase set (column PS) information provided by Long Ranger in phase_variants.vcf files. In short, any gene with at least one mutation with GT="1|1" was categorized as "Hemi/homozygous", and any gene with only one heterozygous (GT="0|1" or "1|0") mutation was categorized as "Monoallelic". For any gene with multiple heterozygous mutations, if all mutations were on the same phase set and belonged to the same haplotype, it was categorized as "Multi-monoallelic". If all mutations were on the same phase set and there was at least one mutation belonging to the opposite allele, it was categorized as "Biallelic heterozygous". In cases where multiple heterozygous mutations did not belong to the same phase set or were unphased (GT="0/1"), the genes were categorized as "no info".

References

1. SOBEL,R.E. and SADAR,M.D. (2005) Cell lines used in prostate cancer research: A compendium of old and new lines—part 1. *The Journal of Urology,* 173, 342-359.
2. SOBEL,R.E. and SADAR,M.D. (2005) Cell lines used in prostate cancer research: A compendium of old and new lines—part 2. *The Journal of Urology,* 173, 360-372.
3. van Bokhoven,A., Varella-Garcia,M., Korch,C., Johannes,W.U., Smith,E.E., Miller,H.L., Nordeen,S.K., Miller,G.J. and Lucia,M.S. (2003) Molecular characterization of human prostate carcinoma cell lines. *The Prostate,* 57, 205-225.
4. Haffner,M.C., Bhamidipati,A., Tsai,H.K., Esopi,D.M., Vaghasia,A.M., Low,J.Y., Patel,R.A., Guner,G., Pham,M.T., Castagna,N., et al. (2021) Phenotypic characterization of two novel cell line models of castration-resistant prostate cancer. *Prostate,* 81, 1159-1171.
5. Culig,Z., Hoffmann,J., Erdel,M., Eder,I.E., Hobisch,A., Hittmair,A., Bartsch,G., Utermann,G., Schneider,M.R., Parczyk,K., et al. (1999) Switch from antagonist to agonist of the androgen receptor bicalutamide is associated with prostate tumour progression in a new model system. *Br. J. Cancer,* 81, 242-251.
6. Thalmann,G.N., Sikes,R.A., Wu,T.T., Degeorges,A., Chang,S.M., Ozen,M., Pathak,S. and Chung,L.W. (2000) LNCaP progression model of human prostate cancer: Androgen-independence and osseous metastasis. *Prostate,* 44, 91-103 Jul 1;44(2).
7. Capes-Davis,A., Reid,Y.A., Kline,M.C., Storts,D.R., Strauss,E., Dirks,W.G., Drexler,H.G., MacLeod,R.A.F., Sykes,G., Kohara,A., et al. (2013) Match criteria for human cell line authentication: Where do we draw the line? *Int. J. Cancer,* 132, 2510-2519.
8. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics,* 25, 1754-1760.
9. Stenson,P.D., Ball,E.V., Mort,M., Phillips,A.D., Shiel,J.A., Thomas,N.S., Abeysinghe,S., Krawczak,M. and Cooper,D.N. (2003) Human gene mutation database (HGMD): 2003 update. *Hum. Mutat.,* 21, 577-581.
10. Ghandi,M., Huang,F.W., Jané-Valbuena,J., Kryukov,G.V., Lo,C.C., McDonald,E.R., Barretina,J., Gelfand,E.T., Bielski,C.M., Li,H., et al. (2019) Next-generation characterization of the cancer cell line encyclopedia. *Nature,* 569, 503-508.
11. Daniel Lai, Gavin Ha, Sohrab Shah. HMMCopy. 1.36.0, .
12. Taiyun Wei,V.S. Corrplot.
13. Kandoth,C. (2020) Vcf2maf. 1.6.19, .
14. Armenia,J., Wankowicz,S.A.M., Liu,D., Gao,J., Kundra,R., Reznik,E., Chatila,W.K., Chakravarty,D., Han,G.C., Coleman,I., et al. (2018) The long tail of oncogenic drivers in prostate cancer. *Nat Genet,* 50, 645.
15. Sondka,Z., Bamford,S., Cole,C.G., Ward,S.A., Dunham,I. and Forbes,S.A. (2018) The COSMIC cancer gene census: Describing genetic dysfunction across all human cancers. *Nature Reviews Cancer,* 18, 696-705.
16. Waters,C.E., Saldivar,J.C., Hosseini,S.A. and Huebner,K. (2014) The FHIT gene product: Tumor suppressor and genome "caretaker". *Cellular and Molecular Life Sciences : CMLS,* 71, 4577-4587.
17. Tate,J.G., Bamford,S., Jubb,H.C., Sondka,Z., Beare,D.M., Bindal,N., Boutselakis,H., Cole,C.G., Creatore,C., Dawson,E., et al. (2018) COSMIC: The catalogue of somatic mutations in cancer. *Nucleic Acids Res.,* 47, D941-D947.

Supplementary Tables:

Table S1: Origins, characteristics, and authentication information of cell lines used in this study

Table S2: HMW DNA extraction QC, sequencing depth, and Long Ranger output metrics for each cell line

Table S3: Allele fraction analysis between linked-read data and other data sets

Table S4: Copy number concordance analysis between linked-read data and CCLE data or cell lines of other passages

Table S5: Microsatellite instability status, somatic mutation count and somatic mutation rate for each cell line.

Table S6: Mutations on Homology-directed repair genes in this panel of cell lines

Table S7: Somatic mutations in Cosmic-Longtail gene list found in this panel of 12 PCa cell lines, with putative driver mutations being deleterious (frameshift, splice site, or nonsense) or inframe indels and missense mutations found in Cancer Mutation Census (CMC) list (PMID: 32461654). Any missense or inframe indel mutation not found in CMC is potentially a passenger. Column headings are according to MAF column specification annotated by vcf2maf (2020, doi:10.5281).

Table S8: List of simple and complex SVs in each cell lines and the phase information used for complex SV calling by ChainLink method

Table S9: SVs involved in the TMPRSS2-ERG rearrangements.

Table S10: List of SVs and genes interrupted by SVs in each cell line

Table S11: List of CNV in each cell line as called by barcode coverage by Long Ranger

Table S12: List of full deletions (CNV=0) in each cell line and associated gene loss

Table S13: List of small deletions (> 40bp and < 30kb) detected in each cell line

Supplementary Figure Legends:

Figure S1: Workflow, cell lines provenance, and mutations profiles in common PCa cell lines **(A)** Schematic of workflow and analysis steps. Green boxes represent wet lab experiments; blue boxes represent computational analysis. **(B)** Pearson correlation coefficient between copy numbers measured by normalized read counts in different PCa cell lines from different passages and experiments. Heatmap colors indicate coefficient value. Same or related cell lines clustered together. Red labels indicate samples from the current linked-read experiment. Black labels indicates samples from past experiments. Black box enclosed LNCaP samples from different experiments and passages that clustered together. "Digit_X" next to sample names indicates mean sequencing coverage. **(C)** All somatic mutations in this set of PCa cell lines stratified by variant classification, with MSI status indicated below. **(D)** All somatic mutations in the Longtail gene list, stratified by variant classification, with MSI status indicated below.

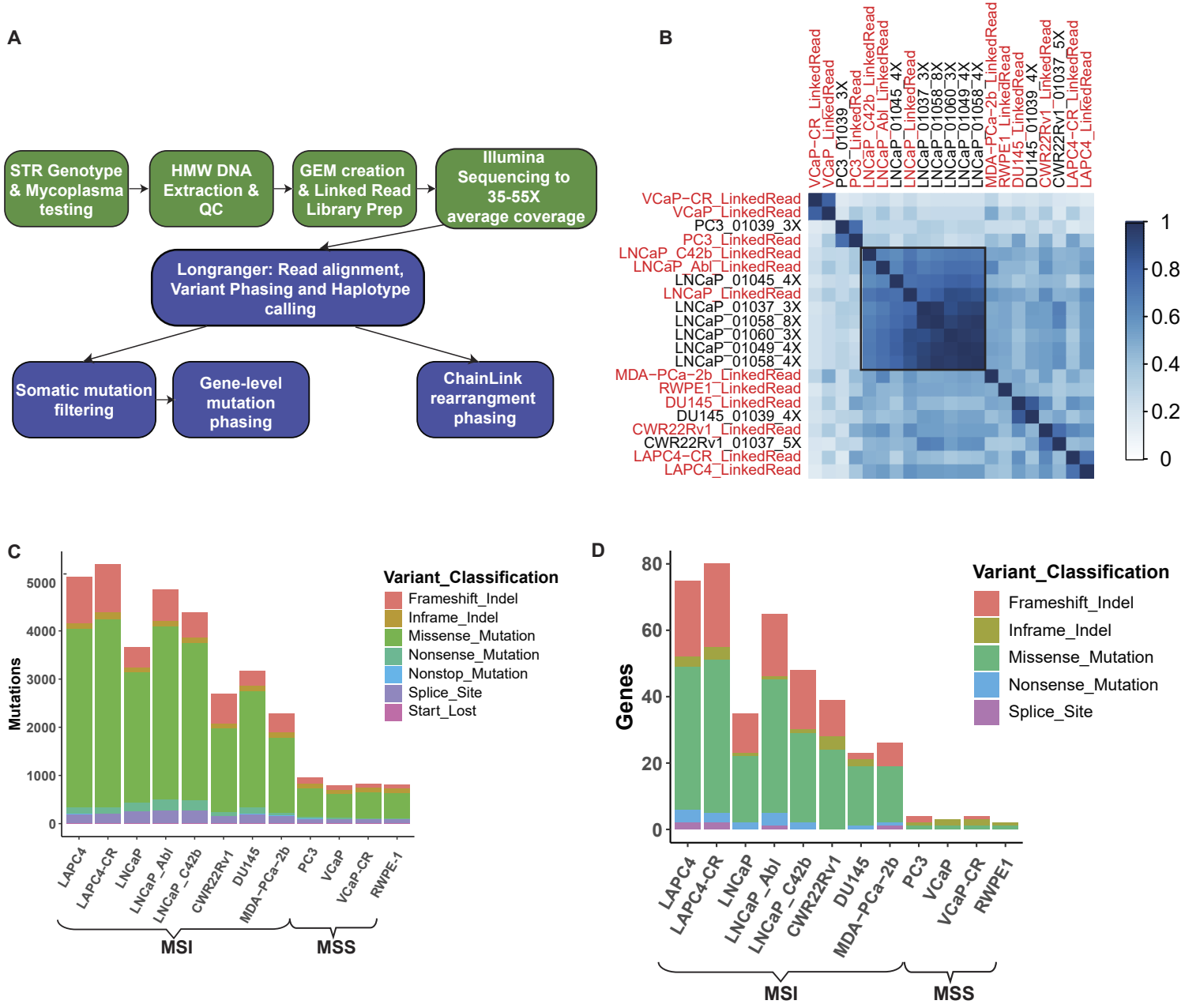
Figure S2: Number of SVs and CNVs present in each cell line and comparison between ChainLink and ChainFinder in chaining complex SVs (A) Number of SVs and CNVs found in each cell line stratified by type. DEL= deletion, DISTAL= distal, DUP= duplication, INV=inversion, UNK=unknown. **(B)** Number of chains and SVs per chain as identified by ChainFinder or Chainlink. **(C)** Complex chained SVs for representative cell lines MDA-PCa-2b and PC3 identified by ChainLink or Chainfinder. Heatmap track beneath chromosome band represents CNV with colors representing copy number. Innermost link track represents large SVs as defined by barcode overlap: complex SVs are colored by chain, with simple SVs identified in gray. **(D)** Rearrangements and copy number alterations involved in a chromothripsis event on chromosome 1, 10, 5 and 8 in PC3 or **(E)** on chromosomes 5 in VCaP. Heatmap tracks beneath chromosome bands represent CNV with colors representing copy number. Innermost link track joins the genomic breakpoints of large SVs. Color of SV links indicates orientation of breakpoints as deletion like (DEL like), duplication like (DUP like), tail to tail inversion (t2tINV), or head to head inversion (h2hINV).

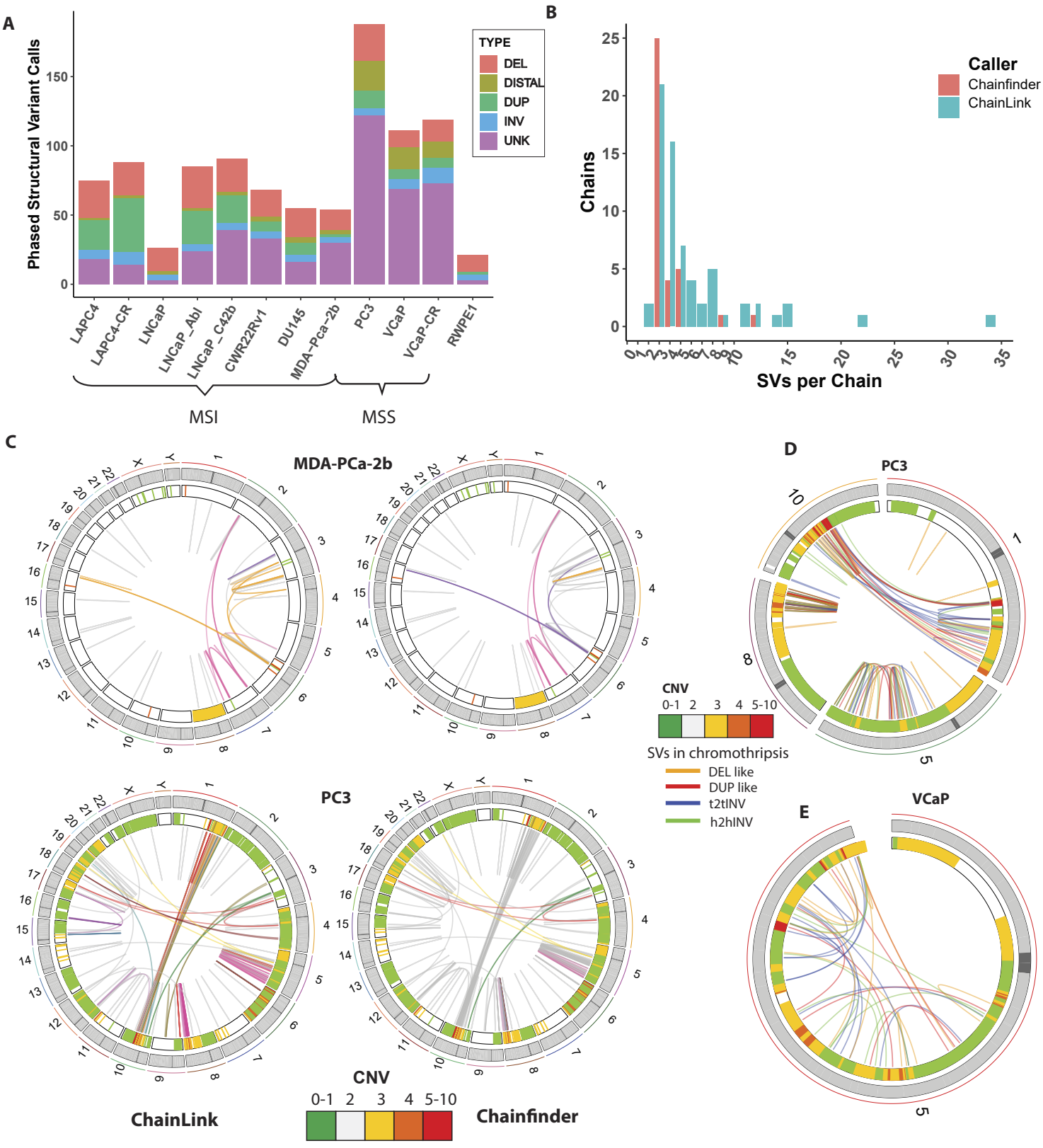
Figure S3: Evidence of a potential sub-clonal population in RWPE-1 and AR status in this panel of PCa. (A) Density plot showing number of 10kb bin over readcounts per bin, demonstrating a shoulder of copy number larger than 2 in RWPE-1. The gradual decline in bin count as CN increases indicates a sub-clonal population in RWPE-1, consistent with **(B)** Mutation count versus allele fraction of each mutation, showing a minor mode at allele fraction of 0.13, indicating sub-clonal divergence in RWPE-1. **(C)** Lollipop graphs showing Pfam domains and positions of mutations found in AR proteins in different PCa cell lines. Colors of the lollipop head indicate variant classification for each mutation **(D)** Copy number and segmentation analysis by HMMCopy showed an amplification event around exon 3 of AR in CWR22Rv1, as indicated by GC and mappability normalized copy number change on Y axis and genomic location on chromosome X on x-axis (Hg19).

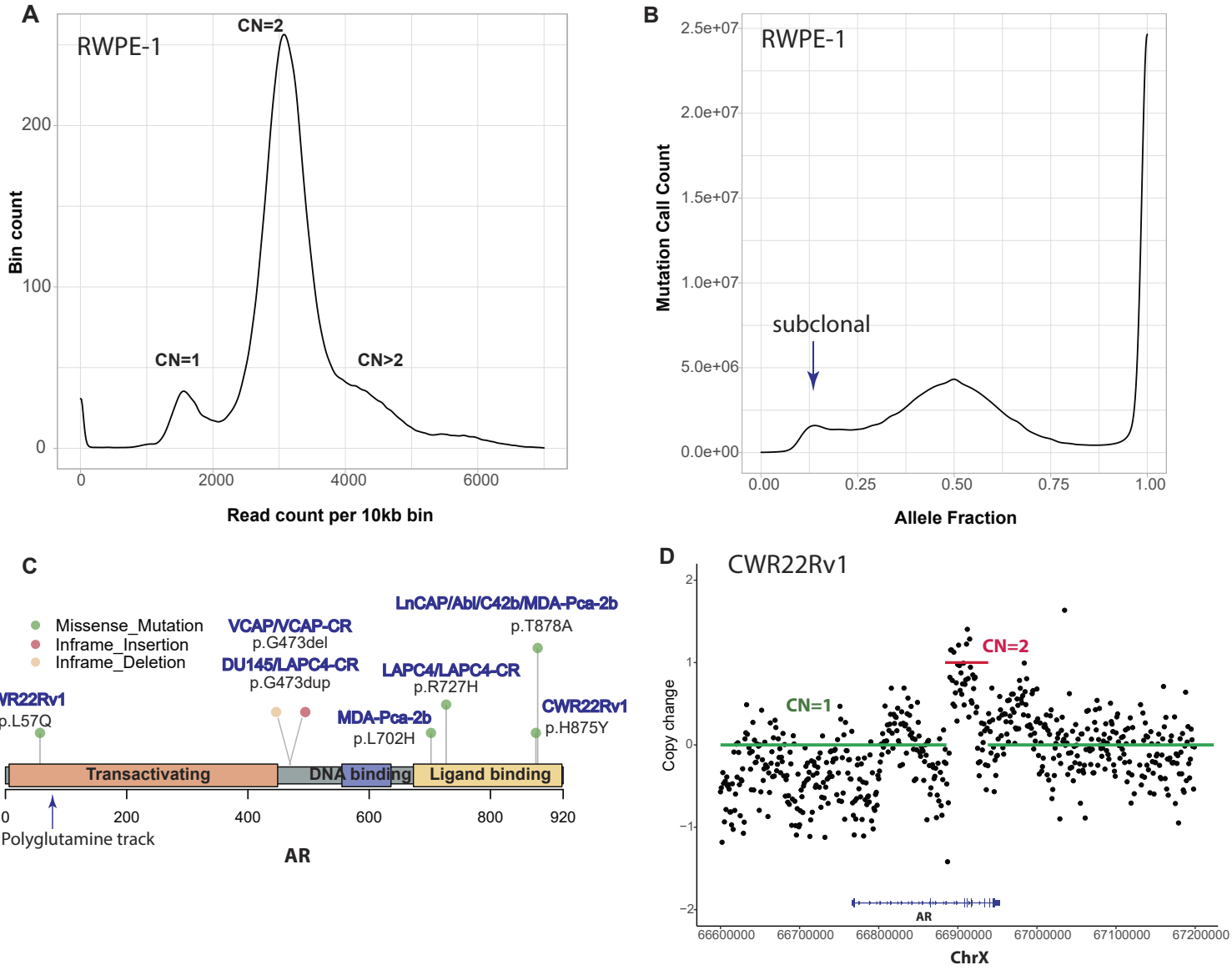
Figure S4: Mutations in SWI/SNF complexes and its association with castrate resistant phenotype. **(A)** Onco-print showing mutations in SWI/SNF subunits present in 11% of 1578 patients and 1626 samples from 5 Prostate Cancer studies. Top color bar represents study origins, colors represent alteration classification. Only samples with at least one SWI/SNF subunit mutated are shown (n=170). **(B)** Frequency of SWI/SNF subunit mutations found in each of the 5 studies queried, from left to right: Metastatic Castration Resistant Prostate Cancer (SU2C/PCF Dream Team, Cell 2015, n=150), Metastatic

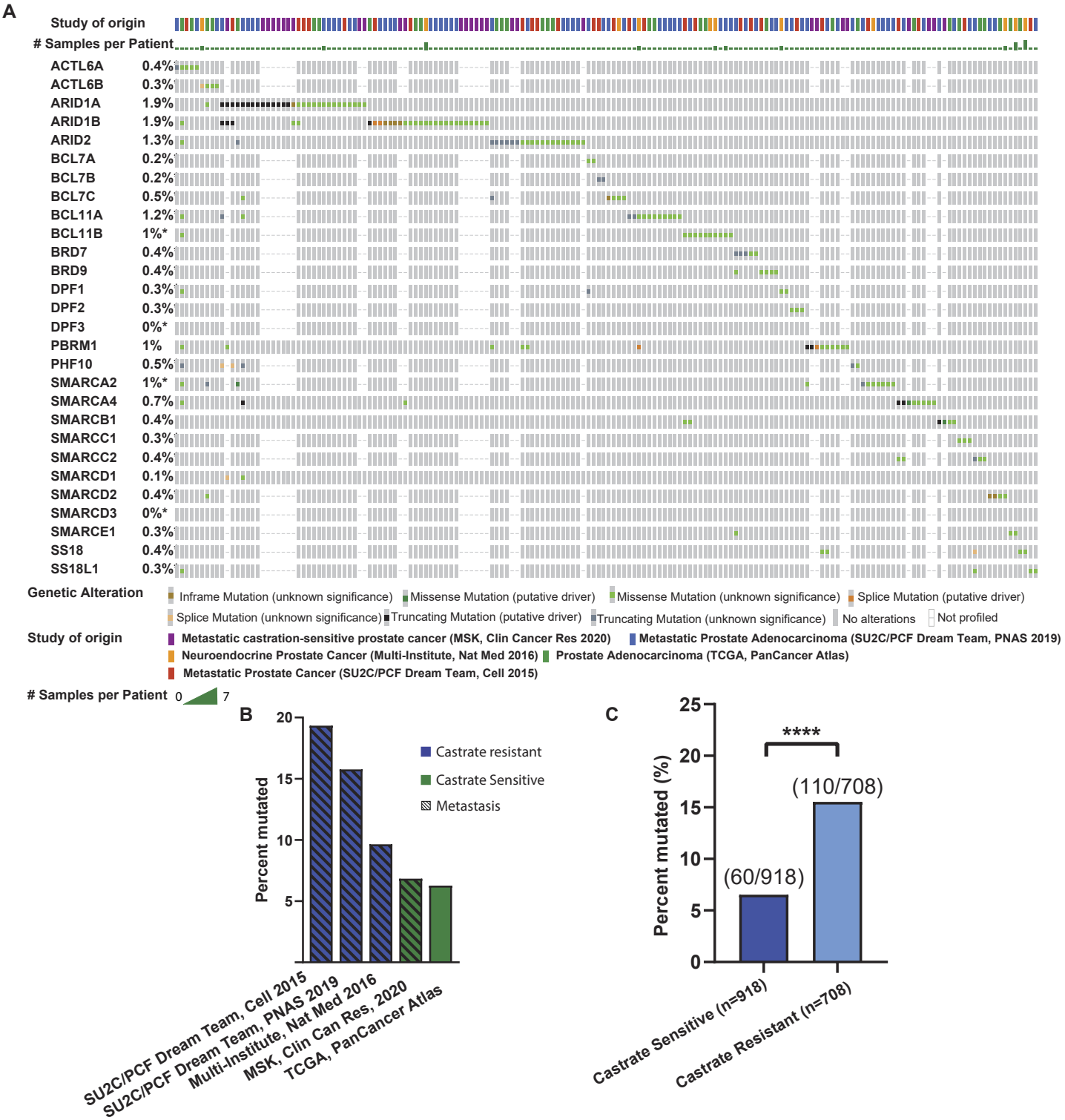
Castration Resistant Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019, n=444), Castrate Resistant Neuroendocrine Prostate Cancer (Multi-Institute, Nat Med 2016, n=114), Metastatic castration-sensitive prostate cancer (MSK, Clin Cancer Res 2020, n=424), Prostate Adenocarcinoma (TCGA, PanCancer Atlas, n=494), total n=1578 patients/1626 samples. **(C)** Percentage of samples with at least one SWI/SNF mutation in castrate resistant (n=708) and castrate sensitive (n=918) PCa patients. SWI/SNF mutations are enriched in Castration resistant PCa compared to Castration sensitive PCa (Chi-squared, $p<0.0001$).

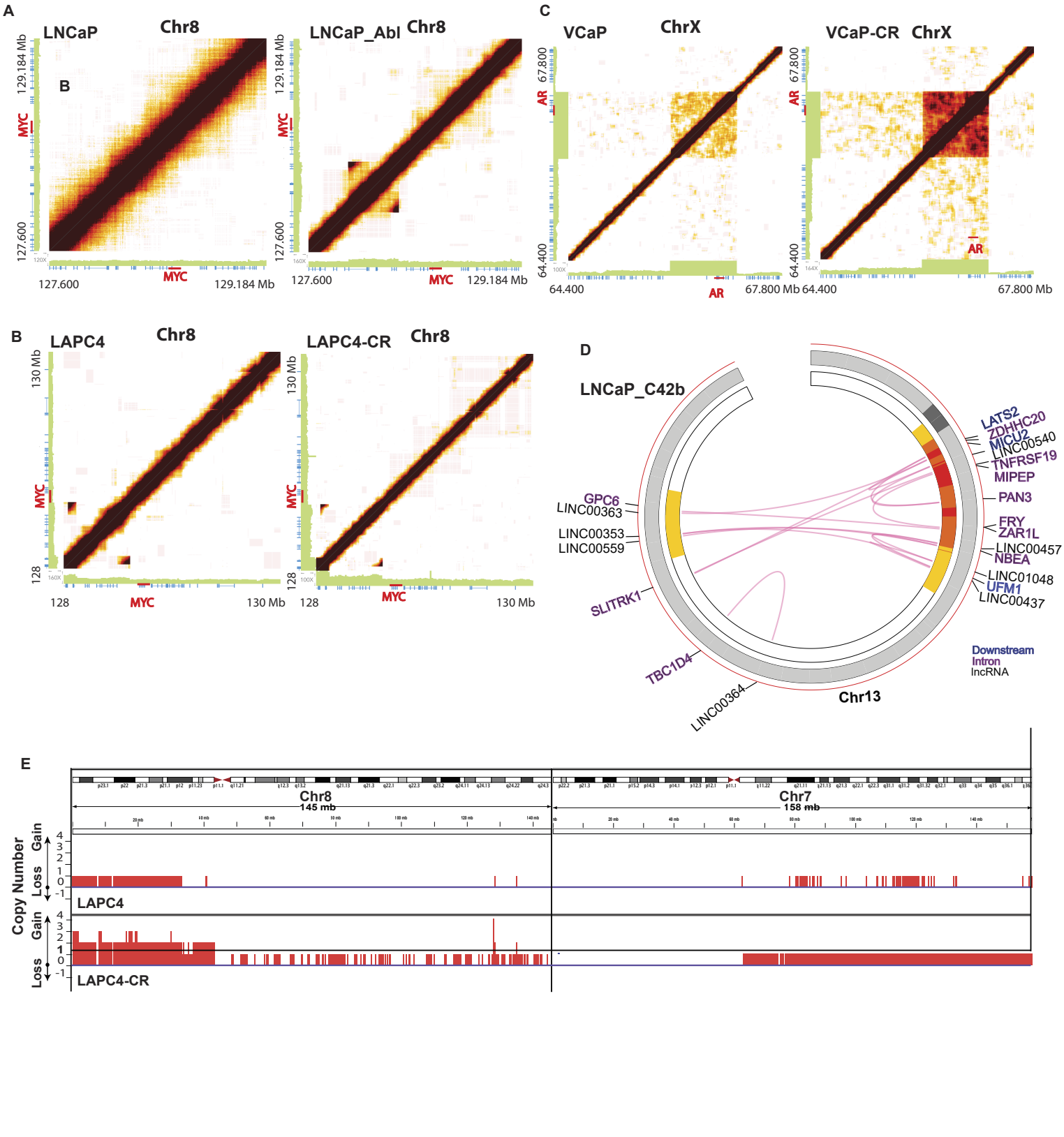
Figure S5: Structural variants and copy number variants found in CR clones.
(A) *MYC* enhancer amplification on chromosome 8 in LNCaP_Abl compared to LNCaP. **(B)** Tandem duplications around 8q24 in LAPC4-CR compared to LAPC4. **(C)** VCaP-CR showed copy number gain at *AR* enhancer and *AR* gene compared to VCaP. In A, B, C, x and y axis represents the same genomic coordinates on chromosome X or 8, with diagonal line representing self-interaction, and heatmap represents barcode overlap and genomic interactions.**(D)** Chromothripsis -like complex SV event on chromosome 13 disrupted many protein-coding introns and lncRNAs in LNCaP_C42b. **(E)** Tandem duplication observed in LAPC4 and LAPC4-CR, especially in chromosome 7 and 8. Purple line represents the state of copy number neutral. Red bars above the purple line represent copy number gains.


A
STR Genotype & Mycoplasma testing
HMW DNA Extraction & QC
GEM creation & Linked Read Library Prep
Illumina Sequencing to 35-55X average coverage
Longranger: Read alignment, Variant Phasing and Haplotype calling
Somatic mutation filtering
Gene-level mutation phasing
ChainLink rearrangment phasing
B
VCaP-CR_LinkedRead
VCaP_LinkedRead
PC3_01039_3X
PC3_LinkedRead
LNCaP_C42b_LinkedRead
LNCaP_Abl_LinkedRead
LNCaP_01045_4X
LNCaP_LinkedRead
LNCaP_01037_3X
LNCaP_01058_8X
LNCaP_01060_3X
LNCaP_01049_4X
LNCaP_01058_4X
MDA-PCa-2b_LinkedRead
RWPE1_LinkedRead
DU145_LinkedRead
DU145_01039_4X
CWR22Rv1_LinkedRead
CWR22Rv1_01037_5X
LAPC4-CR_LinkedRead
LAPC4_LinkedRead
1
0.8
0.6
0.4
0.2
0
C
Mutations
5000
4000
3000
2000
1000
0
Variant_Classification
Frameshift_Indel
Inframe_Indel
Missense_Mutation
Nonsense_Mutation
Nonstop_Mutation
Splice_Site
Start_Lost
LAPC4
LAPC4-CR
LNCaP
LNCaP_Abl
LNCaP_C42b
CWR22Rv1
DU145
MDA-PCa-2b
PC3
VCaP
VCaP-CR
RWPE-1
MSI
MSS
D
Genes
80
60
40
20
0
Variant_Classification
Frameshift_Indel
Inframe_Indel
Missense_Mutation
Nonsense_Mutation
Splice_Site
LAPC4
LAPC4-CR
LNCaP
LNCaP_Abl
LNCaP_C42b
CWR22Rv1
DU145
MDA-PCa-2b
PC3
VCaP
VCaP-CR
RWPE1
MSI
MSS


A
Phased Structural Variant Calls
TYPE
DEL
DISTAL
DUP
INV
UNK
LAPC4
LAPC4-CR
LNCaP
LNCaP_Abl
LNCaP_C42b
CWR22Rv1
DU145
MDA-Pca-2b
PC3
VCaP
VCaP-CR
RWPE1
MSI
MSS
B
Chains
SVs per Chain
Caller
Chainfinder
ChainLink
C
MDA-PCa-2b
PC3
ChainLink
Chainfinder
CNV
0-1 2 3 4 5-10
D
PC3
CNV
0-1 2 3 4 5-10
SVs in chromothripsis
DEL like
DUP like
t2tINV
h2hINV
E
VCaP


A
RWPE-1
CN=2
CN=1
CN>2
Bin count
Read count per 10kb bin
B
RWPE-1
subclonal
Mutation Call Count
Allele Fraction
C
Missense_Mutation
Inframe_Insertion
Inframe_Deletion
CWR22Rv1
p.L57Q
VCAP/VCAP-CR
p.G473del
DU145/LAPC4-CR
p.G473dup
LAPC4/LAPC4-CR
p.R727H
MDA-Pca-2b
p.L702H
LnCAP/Abl/C42b/MDA-Pca-2b
p.T878A
CWR22Rv1
p.H875Y
Transactivating
DNA binding
Ligand binding
Polyglutamine track
AR
D
CWR22Rv1
CN=2
CN=1
Copy change
AR
ChrX


A
Study of origin
Samples per Patient
ACTL6A 0.4%
ACTL6B 0.3%
ARID1A 1.9%
ARID1B 1.9%
ARID2 1.3%
BCL7A 0.2%
BCL7B 0.2%
BCL7C 0.5%
BCL11A 1.2%
BCL11B 1%*
BRD7 0.4%
BRD9 0.4%
DPF1 0.3%
DPF2 0.3%
DPF3 0%*
PBRM1 1%
PHF10 0.5%
SMARCA2 1%*
SMARCA4 0.7%
SMARCB1 0.4%
SMARCC1 0.3%
SMARCC2 0.4%
SMARCD1 0.1%
SMARCD2 0.4%
SMARCD3 0%*
SMARCE1 0.3%
SS18 0.4%
SS18L1 0.3%
Genetic Alteration
Inframe Mutation (unknown significance)
Missense Mutation (putative driver)
Missense Mutation (unknown significance)
Splice Mutation (putative driver)
Splice Mutation (unknown significance)
Truncating Mutation (putative driver)
Truncating Mutation (unknown significance)
No alterations
Not profiled
Study of origin
Metastatic castration-sensitive prostate cancer (MSK, Clin Cancer Res 2020)
Metastatic Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019)
Neuroendocrine Prostate Cancer (Multi-Institute, Nat Med 2016)
Prostate Adenocarcinoma (TCGA, PanCancer Atlas)
Metastatic Prostate Cancer (SU2C/PCF Dream Team, Cell 2015)
Samples per Patient 0 7
B
Percent mutated
20
15
10
5
Castrate resistant
Castrate Sensitive
Metastasis
SU2C/PCF Dream Team, Cell 2015
SU2C/PCF Dream Team, PNAS 2019
Multi-Institute, Nat Med 2016
MSK, Clin Can Res, 2020
TCGA, PanCancer Atlas
C
Percent mutated (%)
25
20
15
10
5
0
****
(110/708)
(60/918)
Castrate Sensitive (n=918)
Castrate Resistant (n=708)


A
LNCaP
Chr8
LNCaP_Abl
Chr8
B
MYC
129.184 Mb
127.600
C
VCaP
ChrX
VCaP-CR
ChrX
AR
67.800
64.400
67.800 Mb
B
LAPC4
Chr8
LAPC4-CR
Chr8
130 Mb
128
D
LNCaP_C42b
GPC6
LINC00363
LINC00353
LINC00559
SLITRK1
TBC1D4
LINC00364
Chr13
LATS2
ZDHHC20
MICU2
LINC00540
TNFRSF19
MIPEP
PAN3
FRY
ZAR1L
LINC00457
NBEA
LINC01048
UFM1
LINC00437
Downstream
Intron
lncRNA
E
Chr8
145 mb
Chr7
158 mb
Copy Number
Gain
Loss
LAPC4
LAPC4-CR