

Supplementary Methods

Here we have provided additional information on the methods used for this project developing on those already stated in the main manuscript.

2.1 The clinical relevance of PCa cell lines

A literature search was performed and detailed variables of PCa cell lines were collated. Sailer et.al's paper contained a list of 42 cell lines¹ and three additional cells lines were included from the prostate adenocarcinoma MSK study along with a new cell line derived from an Afro-Caribbean patient, selected for inclusion due to the lack of available models representing this cohort of distinct unmet clinical need². cBioPortal is an open access gold standard platform for comparison of large patient cohorts from multiple institutes^{3,4}. From cBioPortal, 24 PCa studies containing multi omics datasets from 10,239 patients were analysed, and an additional published study with 45 Nigerian patients was also included⁵. The collated clinical cohort and cell line data were compared available global data, where available to assess their broader relevance. Often this was in fact from single countries and used as proxies due to lack of available data.

Initial variables searched for are shown in Supplementary Table 1 but not all cell lines had available data for each variable (Supplementary Table 2). The number of citations for each cell line was recorded as the highest number of results from databases Ovid MEDLINE or Web of Science Core Collection using search strategy “((cell line name) OR (cell line name with no hyphen or spaces) AND prostate).” This was done to maximise the identification of relevant cell line literature despite differing notations, whilst yielding only relevant studies by including “prostate”.

25

26 To compare cell lines against clinical cohorts the clinical variables analysed on cBioPortal
27 included age at sample collection, geographical location, race, cancer type, Gleason grade
28 and metastatic stage. Clinical cohorts in scientific studies are often selected based on specific
29 clinical course or patient features, and thus we analysed the broader relevance of both cell
30 lines and the clinical cohort against PCa patients worldwide (Supplementary Table 3). Age at
31 sample collection was analysed using a systematic review on robotic radical prostatectomy
32 with a large patient cohort from 33 studies in 14 different countries⁶. There is often confusion
33 regarding the correct terminology for race, ethnicity or ancestry. We followed guidance from
34 epidemiologists at our centre and the Office for National Statistics (ONS) in the United
35 Kingdom who use the term ethnicity⁷. To analyse the ethnicity distribution of prostate cancer
36 incidence, a paper from the Journal of the American Medical Association (JAMA) was used
37 as a proxy to represent worldwide PCa ethnicity demographics⁸. Their large cohort of almost
38 230,000 patients were from the Surveillance, Epidemiology, and End Results (SEER)
39 database⁹. Currently there is no publicly available database encompassing race data in PCa as
40 opposed to geographical distribution. The SEER database encompasses 18 geographic areas
41 in the United States and despite the paper only using US data it was deemed suitable because
42 of the large patient cohort and because all PCa clinical cohorts on cBioPortal except the
43 Prostate Adenocarcinoma (CPC-GENE, Nature 2017) study¹⁰ had a minimum of one
44 American institution collaborator. The SEER database was also used to investigate the
45 proportion of patients with cancer stages one to four at diagnosis. Average worldwide
46 Gleason grade was calculated using a large Australian study documenting Gleason grade at
47 diagnosis. The figure for proportion of different PCa types is not definitive and so a figure
48 from a classical book on surgical pathology by Ackerman et.al with over 3500 citations was
49 used¹¹. Prior treatment history of cell lines and clinical cohorts was also investigated. No

worldwide source was used for comparison due to the vast heterogeneity of treatment in healthcare systems globally.

2.2 The biological relevance of PCa cell lines

To analyse the biological features of cell lines the KEGG prostate cancer gene set was used (systematic name M13191,¹²⁻¹⁴). The Kyoto Encyclopaedia of Genes and Genomes (KEGG), details the most important gene alterations in PCa by linking genomic and functional information with knowledge on cellular processes involved in oncogenesis. A further ten cell lines had limited biologic data available on the COSMIC (Catalogue of Somatic Mutations in Cancer) database (Supplementary Table 2) but often had few genes or only mutational data analysed, thus, were not analysed in depth in our study. The biological variables screened for in each cell line is shown in Supplementary Table 4.

The clinical cohorts on cBioPortal used to compare biological data in this study were selected on the basis of all having at least 200 patients and biological data of at least two alteration types. Additionally, a study of neuroendocrine PCa¹⁵ was included due to many 3D models taking on a neuroendocrine phenotype in vitro and ex vivo in laboratory settings and to allow comparison of different PCa types.

A heatmap identifying which cell lines are most suited to investigate specific molecular pathways was created and the KEGG PCa gene set was used to identify the top 50 pathways genes from the gene set were involved in (number of genes from the set involved in the pathway/ total number of genes involved in that pathway). The proportion related to the colour was the number of genes altered in each cell line/ total number of genes involved in that pathway.

2.3 Statistical Analysis

All raw data was recorded on Microsoft Excel. Tests for normality used in this study were the D'Agostino & Pearson, Anderson-Darling, Shapiro-Wilkins, and the Kolmogorov-Smirnov tests.

This study focused on the clinical and biological relevance of prostate cancer cell lines by thoroughly investigating various variables pertinent to translational research and clinical applicability. Age and ethnicity of cell lines have been explored to comprehend their impact on disease progression and treatment response, unveiling potential disparities in outcomes. Additionally, we examined Gleason grade and cancer type to ascertain the representativeness of cell line models across different disease stages and subtypes within prostate cancer. The assessment of treatment history aids in understanding drug resistance mechanisms and informs the development of novel therapeutic strategies for refractory disease. Our investigation encompassed multiomics variables including mutations, copy number alterations, structural variants, microsatellite instability, mRNA and protein expression, and tumour mutational burden in cell lines, facilitating a comprehensive analysis of molecular features reflective of patient tumours.

1. Sailer V, von Amsberg G, Duensing S, et al. Experimental in vitro, ex vivo and in vivo models in prostate cancer research. *Nature Reviews Urology*. 2023/03/01 2023;20(3):158-178. doi:10.1038/s41585-022-00677-z

2. Valentine H, Aiken W, Morrison B, et al. Expanding the Prostate Cancer Cell Line Repertoire with ACRJ-PC28, an AR-negative Neuroendocrine Cell Line Derived From an African-Caribbean Patient. *Cancer Research Communications*. 2022;2(11):1355-1371. doi:10.1158/2767-9764.Crc-22-0245

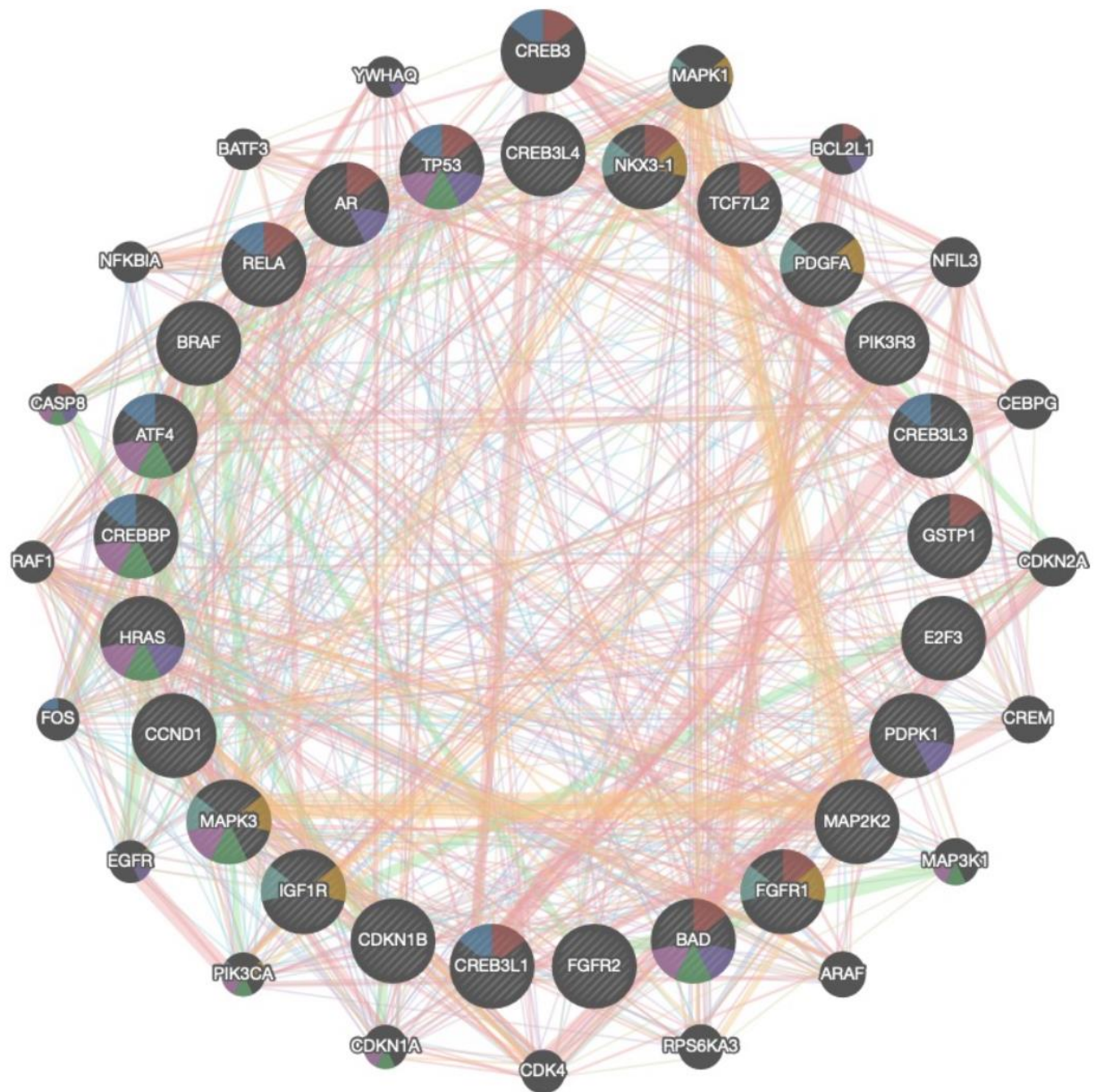
3. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. Apr 2 2013;6(269):pl1. doi:10.1126/scisignal.2004088

4. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. May 2012;2(5):401-4. doi:10.1158/2159-8290.Cd-12-0095

5. White JA, Kaninjing ET, Adeniji KA, et al. Whole-exome Sequencing of Nigerian Prostate Tumors from the Prostate Cancer Transatlantic Consortium (CaPTC) Reveals DNA Repair Genes Associated with African Ancestry. *Cancer Research Communications*. 2022;2(9):1005-1016. doi:10.1158/2767-9764.Crc-22-0136
6. Du Y, Long Q, Guan B, et al. Robot-Assisted Radical Prostatectomy Is More Beneficial for Prostate Cancer Patients: A System Review and Meta-Analysis. *Medical science monitor : international medical journal of experimental and clinical research*. 2018;24:272-287. doi:10.12659/msm.907092
7. Statistics OfN. Ethnic group, national identity and religion. Accessed 06 March, 2024. <https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion>
8. Fletcher SA, Marchese M, Cole AP, et al. Geographic Distribution of Racial Differences in Prostate Cancer Mortality. *JAMA Network Open*. 2020;3(3):e201839-e201839. doi:10.1001/jamanetworkopen.2020.1839
9. Institute. NC. Data from: SEER*Statdatabase: November 2016 submission . Deposited Released April 2017.
10. Fraser M, Sabelnykova VY, Yamaguchi TN, et al. Genomic hallmarks of localized, non-indolent prostate cancer. *Nature*. 2017/01/01 2017;541(7637):359-364. doi:10.1038/nature20788
11. Ackerman LA, Rosai J. Surgical pathology. *Surgical pathology*. 1974:1394-1394.
12. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. Jan 1 2000;28(1):27-30. doi:10.1093/nar/28.1.27
13. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. Nov 2019;28(11):1947-1951. doi:10.1002/pro.3715
14. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. Jan 6 2023;51(D1):D587-d592. doi:10.1093/nar/gkac963
15. Beltran H, Prandi D, Mosquera JM, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat Med*. Mar 2016;22(3):298-305. doi:10.1038/nm.4045

1
2

VCaP



2

Functions

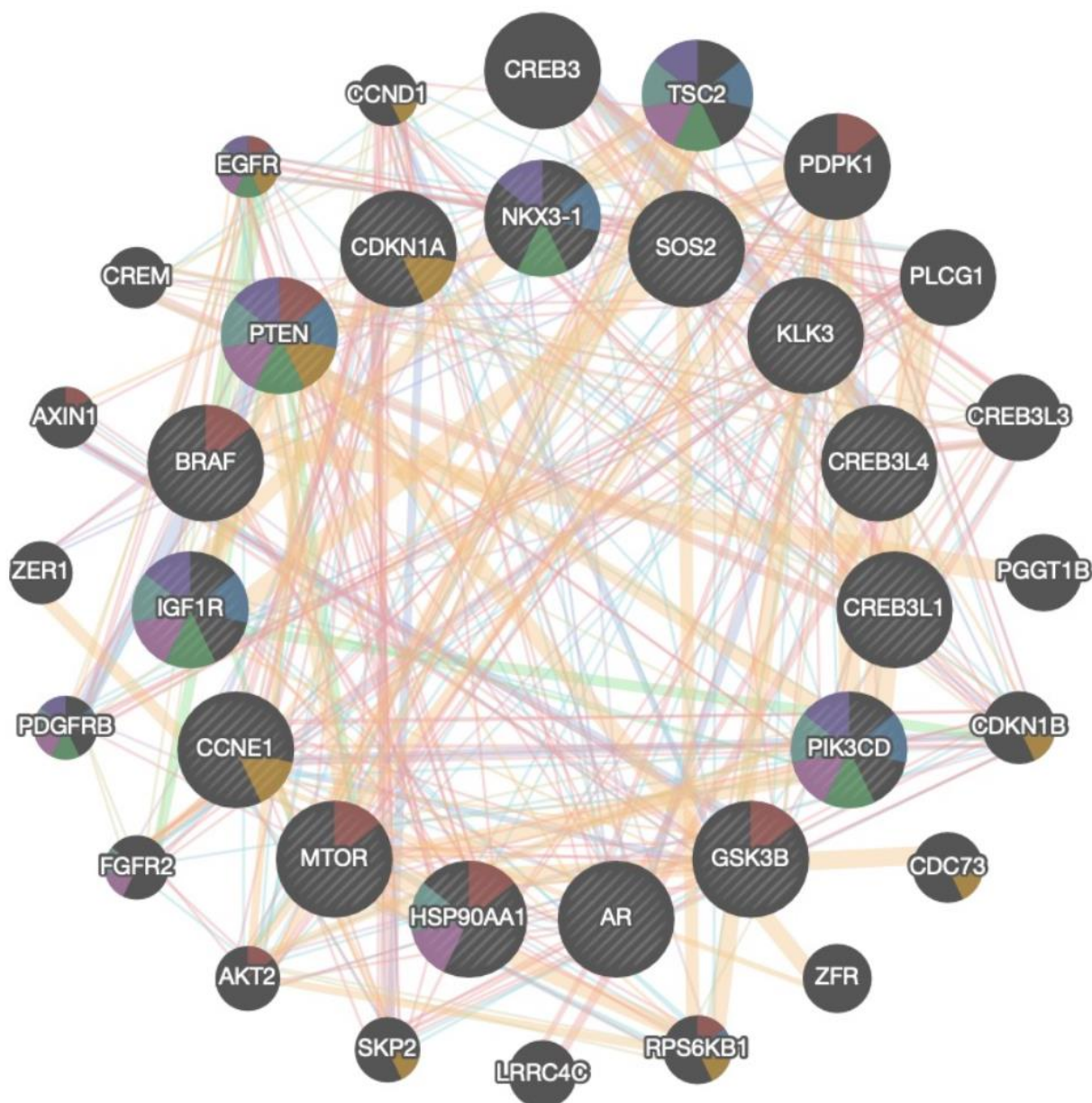
- regulation of apoptotic signaling pathway
- phosphatidylinositol 3-kinase signaling
- regulation of protein localization to membrane
- cellular response to abiotic stimulus
- cellular response to environmental stimulus
- phosphatidylinositol-mediated signaling
- positive regulation of transcription by RNA polymerase II

Networks

- Physical Interactions
- Co-expression
- Pathway
- Genetic Interactions
- Co-localization
- Predicted
- Shared protein domains

Supplementary Figure 1: VCaP has many genes altered from the KEGG prostate cancer gene set - Inner ring represents genes altered in this cell line and the outer ring represents predicted related genes. The top seven pathways are colour coded inside the gene bubbles and the different coloured lines represent networks.

MDA PCa 2b



Functions

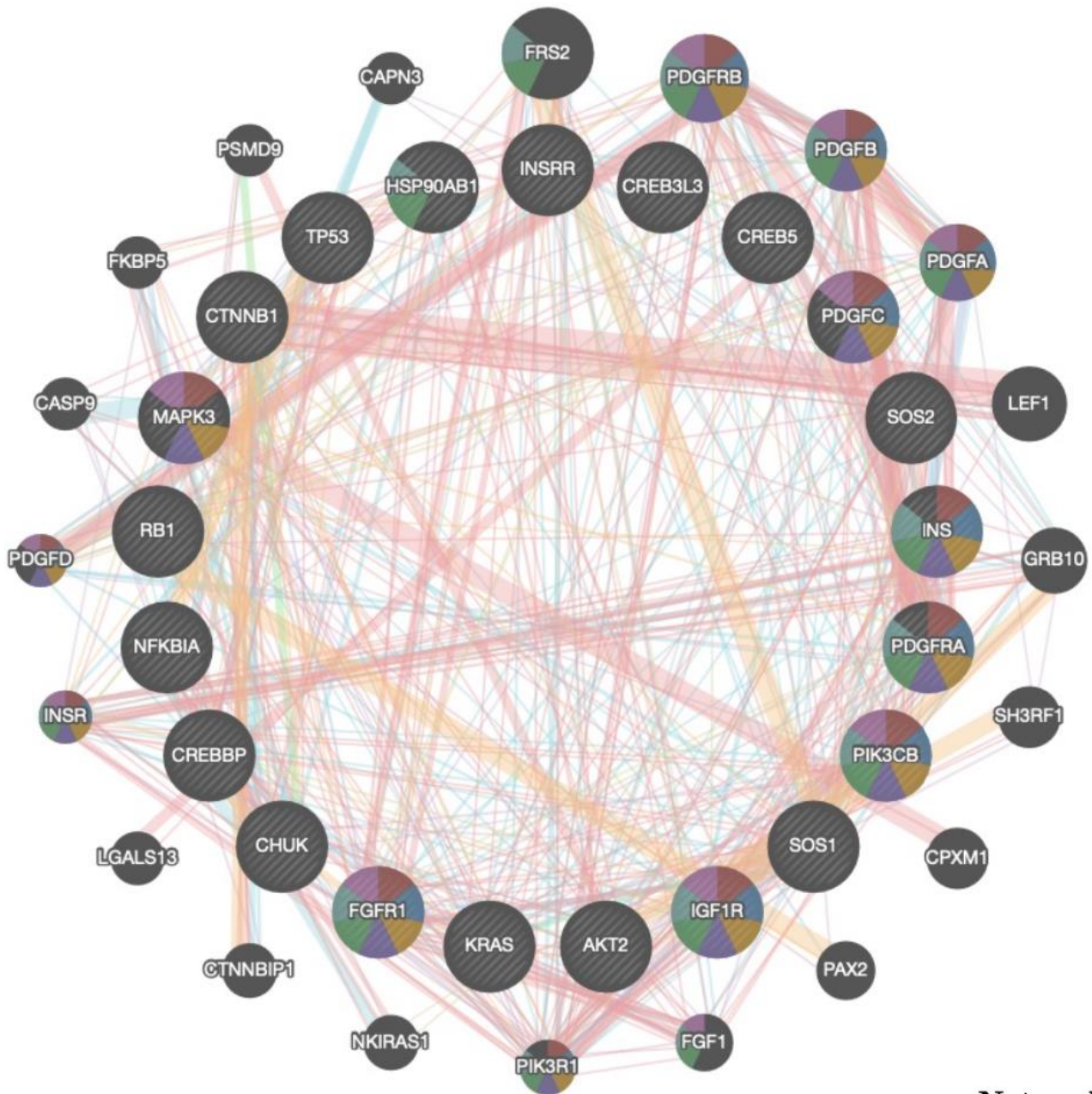
- peptidyl-serine modification
- inositol lipid-mediated signaling
- cell cycle G1/S phase transition
- phosphatidylinositol 3-kinase signaling
- regulation of protein kinase B signaling
- protein kinase B signaling
- phosphatidylinositol-mediated signaling

Networks

- Physical Interactions
- Co-expression
- Pathway
- Genetic Interactions
- Co-localization
- Predicted
- Shared protein domains

Supplementary Figure 2: MDA PCa 2b has 15 genes altered from the KEGG prostate cancer gene set, many of which are involved in peptidyl-serine modification - Inner ring represents genes altered in this cell line and the outer ring represents predicted related genes. The top seven pathways are colour coded inside the gene bubbles and the different coloured lines represent networks.

DU145



Functions

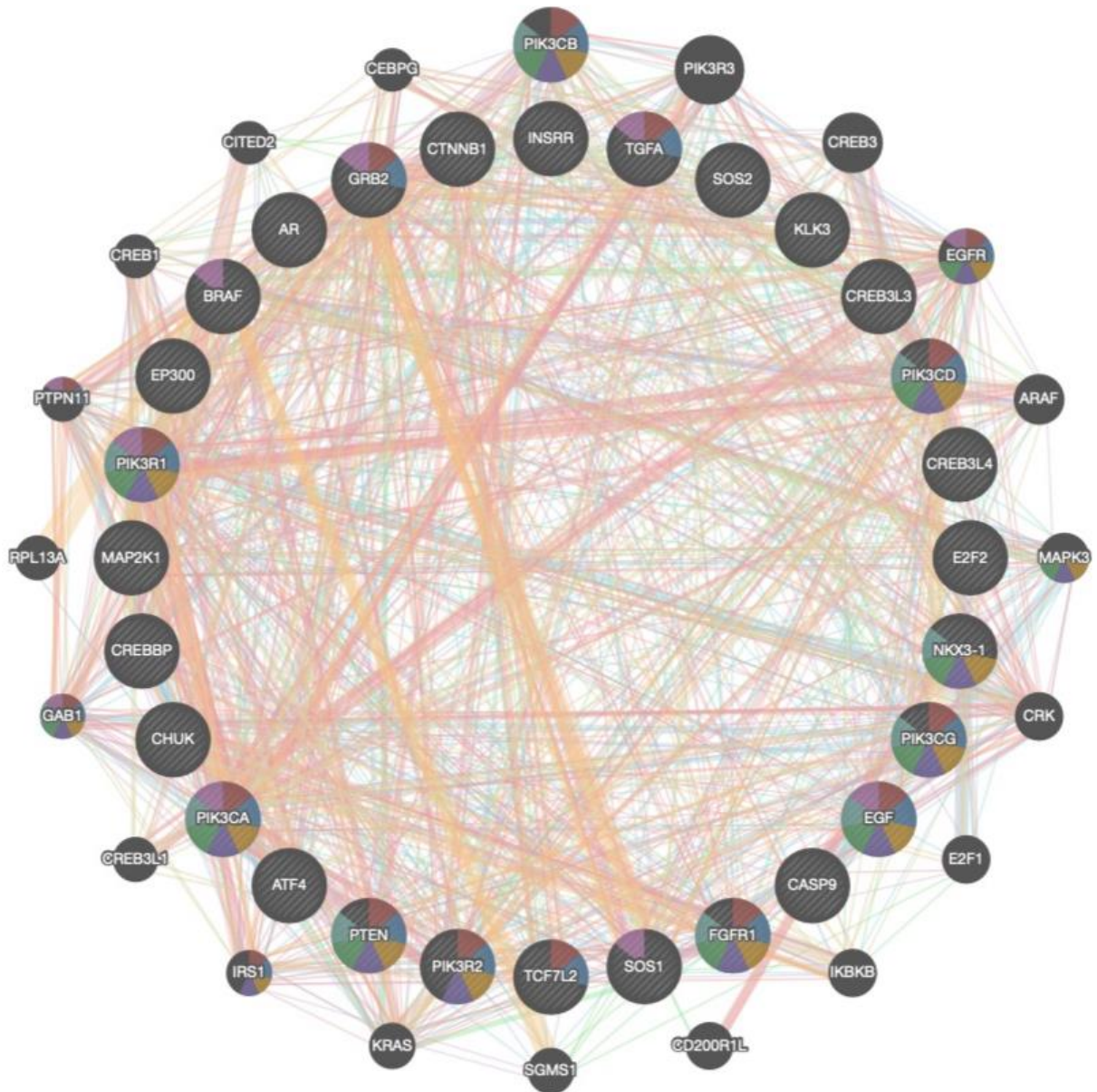
- phosphatidylinositol 3-kinase signaling
- regulation of phosphatidylinositol 3-kinase signaling
- phosphatidylinositol-mediated signaling
- inositol lipid-mediated signaling
- regulation of protein kinase B signaling
- protein kinase B signaling
- regulation of MAP kinase activity

Networks

- Physical Interactions
- Co-expression
- Pathway
- Genetic Interactions
- Co-localization
- Predicted
- Shared protein domains

Supplementary Figure 3: DU145 has 21 genes altered from the KEGG prostate cancer gene set, many of which are involved in phosphatidylinositol 3-kinase signalling - Inner ring represents genes altered in this cell line and the outer ring represents predicted related genes. The top seven pathways are colour coded inside the gene bubbles and the different coloured lines represent networks.

LNCaP-



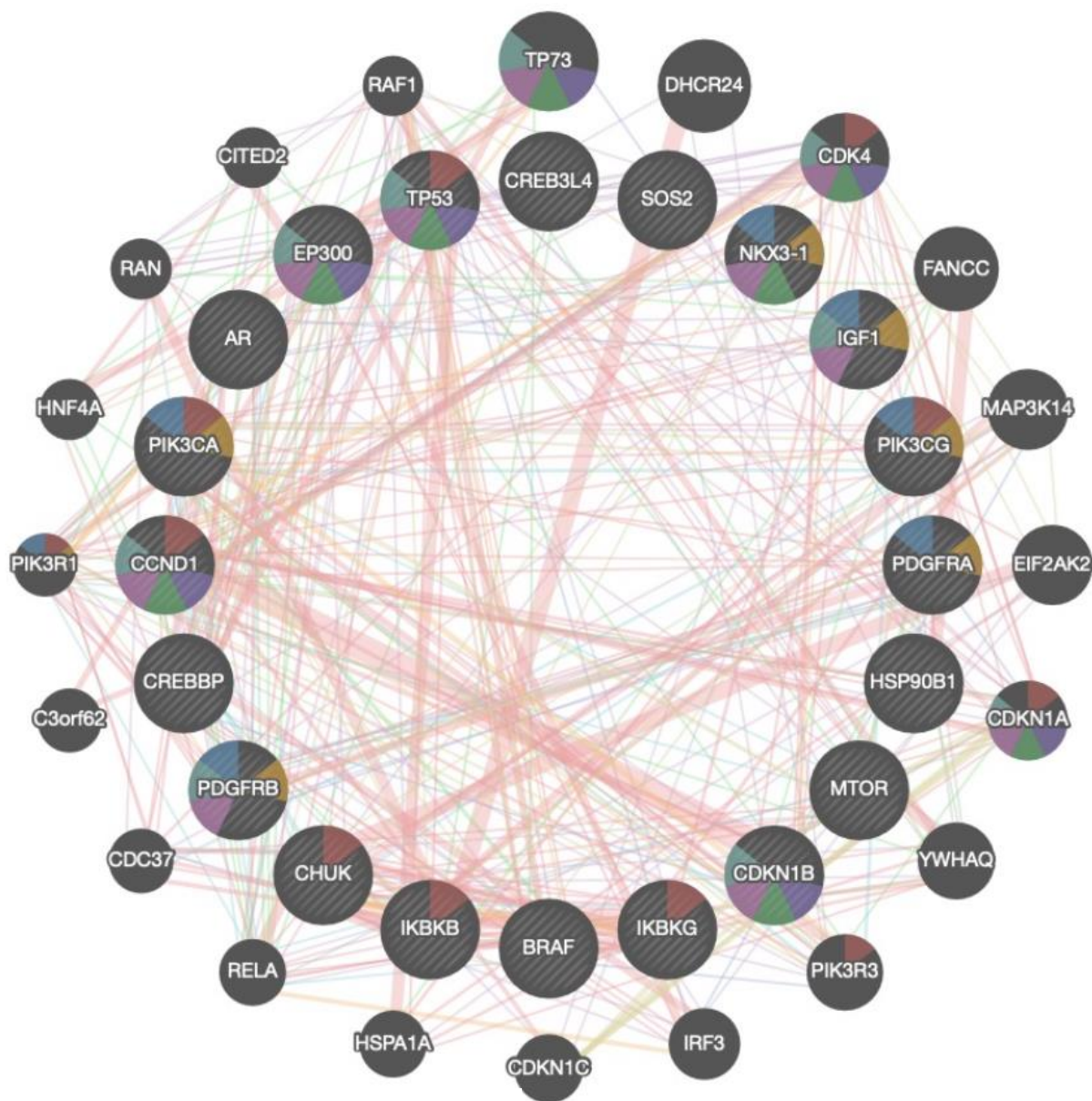
Functions

- regulation of protein kinase B signaling
- protein kinase B signaling
- phosphatidylinositol-mediated signaling
- inositol lipid-mediated signaling
- phosphatidylinositol 3-kinase signaling
- regulation of phosphatidylinositol 3-kinase signaling
- ERBB signaling pathway

Networks

- Physical Interactions
- Co-expression
- Pathway
- Genetic Interactions
- Co-localization
- Predicted
- Shared protein domains

Supplementary Figure 4: LNCaP-FGC has the greatest number of genes altered from the KEGG prostate cancer gene set - Inner ring represents genes altered in this cell line and the outer ring represents predicted related genes. The top seven pathways are colour coded inside the gene bubbles and the different coloured lines represent networks.



Functions

- transferase complex, transferring phosphorus-containing groups
- regulation of phosphatidylinositol 3-kinase signaling
- regulation of cell cycle arrest
- cell cycle arrest
- positive regulation of cell cycle
- positive regulation of cell cycle process
- phosphatidylinositol 3-kinase signaling

Networks

- Physical Interactions
- Co-expression
- Pathway
- Genetic Interactions
- Co-localization
- Predicted
- Shared protein domains

Supplementary Figure 5: 22Rv1 has 20 genes altered from the KEGG prostate cancer gene set, many of which are involved in the transferase complex, transferring phosphorous-containing groups - Inner ring represents genes altered in this cell line and the outer ring represents predicted related genes. The top seven pathways are colour coded inside the gene bubbles and the different coloured lines represent

Cell Line	ATCC and/or Source	Cell Culture Conditions (Medium, cell serum %, and plate)	References
LNCaP	Julius Horoszewicz (Roswell Park Memorial Institute, Buffalo, NY, USA)	RMPI-1640, 10% FBS, adherent tissue culture treated plastic	(1)
PC-3	ATCC (CRL-1435)	F-12K, 10% FBS, adherent tissue culture treated plastic	(2)
DU145	ATCC (HTB-81)	EMEM, 10% FBS, adherent tissue culture treated plastic	(3)
1013 L	Anita Bilström (Active Biotech Research AB, Lund, Sweden)		(4)
22Rv1	ATCC (CRL-2505) James Jacobberger (Case Western Reserve University, Cleveland, OH, USA)	RPMI-1640, 10% FBS, adherent tissue culture treated plastic	(5)
VCaP	ATCC (CRL-2876) Kenneth Pienta (University of Michigan, Ann Arbor, MI, USA)	DMEM, 10% FBS, adherent tissue culture treated plastic	(6)
MDA PCa 1	Nora Navone (University of Texas, Houston, TX, USA)		(7)
CWR22	NCI Division of Cancer Treatment and Diagnosis (DCTD) Tumour Repository		(8)
LAPC-4(LAPC-4-1)	Charles Sawyer (University of California, Los Angeles, CA, USA)	IMDM, 10% FBS, adherent tissue culture treated plastic	(9)
LNCaP-C4	ATCC (CRL-3313)	DMEM/F12, 10% FBS, adherent tissue culture treated plastic	(10)
MDA Pca 2b	ATCC (CRL-2422), Nora Navone (University of Texas, Houston, TX, USA)	F-12K, 20% FBS, adherent tissue culture treated plastic	(11)
PPC-1	No authentic stock is known		(12)
LNCaP-C4-2	ATCC (CRL-3315)	DMEM/F12, 10% FBS, adherent tissue culture treated plastic	(10)
ARCaP	Haiyen Zhau (University of Virginia, Charlottesville, VA, USA)		(7)
LNCaP-FGC	ATCC (CRL-1740)	RMPI-1640, 10% FBS, adherent tissue culture treated plastic	(13)
PZ-HPV-7	ATCC (CRL-2221)	Keratinocyte-SFM, 0% (defoned), adherent tissue culture treated plastic	(14)
PWR-1E	ATCC (CRL-11611)	Keratinocyte-SFM, 0% (defoned), adherent tissue culture treated plastic	(15)
DuCaP	Kenneth Pienta (University of Michigan, Ann Arbor, MI, USA)		(16)
CWR-R1	Christopher Gregory (University of North Carolina, Chapel Hill, NC, USA)		(17)
ALVA-31	No authentic stock is known		(18)
RWPE-2	ATCC (CRL-11610)	Keratinocyte-SFM, 0% (defoned), adherent tissue culture treated plastic	(19)
CA-HPV-10	ATCC (CRL-2220)	Keratinocyte-SFM, 0% (defoned), adherent tissue culture treated plastic	(14)
MDA Pca 2a	Nora Navone (University of Texas, Houston, TX, USA)		(11)
PSK-1	Chol Jang Kim (Shiga University of Medical Science, Otsu, Japan)		(20)
PC-93	Gert Jan van Steenbrugge (Erasmus University, Rotterdam, Netherlands)		(21)
ALVA-41	No authentic stock is known		(22)
PC-346C	Wyske van Weerden (Erasmus University, Rotterdam, Netherlands)		(23)
E006AA	Walter Rayford (Louisiana State University-Health Sciences Center, New Orleans, LA, USA)		(24)
NCI-H660	ATCC (CRL-5813)	RMPI-1640, 5% FBS, suspension	(25)
P69SV40T	No authentic stock is known		(26)
ALVA-101	Steven Loop (American Lake Veterans Administration Hospital, Tacoma, WA, USA)		(22)
DuPro-1	William Isaacs (Johns Hopkins University, Baltimore, MA, USA)		(27)
WPMY-1 normal prostatic myofibroblasts	ATCC (CRL-2854)	DMEM, 5% FBS, adherent tissue culture treated plastic	(28)
UM-SCP-1	Herb Barton Grossman (University of Texas, Houston, TX, USA)		(29)
LAPC-3	No authentic stock is known		(9)
RC-77T/E	Clayton Yates (Tuskegee University, Tuskegee, AL, USA)		(30)
ALVA-55	Steven Loop (American Lake Veterans Administration Hospital, Tacoma, WA, USA)		(31)
KuCaP13	No authentic stock is known		(32)
LASCPC-01	ATCC (CRL-3356) Owen N. Witte (University of California, Los Angeles, CA, USA)		(33)
hTERT EP156T	ATCC (CRL-3289)	Modified MCDB-153, 1% FBS, adherent tissue culture treated plastic	(34)
LNCaP-LN-3	Korean Cell Line Bank 80018		(35)
PNF-08	Gerhard Unteregger (University of Saarland Medical School, Germany)		(36)

LAPC9	Noah Craft (Department of Medicine, Molecular Biology Institute, University of California, Los Angeles, USA)		(37)
LNCaP104S	John Kokontis (Ben May Institute, University of Chicago, Illinois, USA)		(38)
LuCAP35	Eva Corey (Department of Urology, University of Washington, Seattle, Washington 98195, USA)		(39)
ACRJ-PC28	The University of the West Indies, Mona, Kingston, Jamaica, West Indies		(40)

Supplementary Table 1: Prostate cancer cell lines analysed in this study. Abbreviations : FBS – Foetal Bovine Serum

cBioPortal Name	Number of samples	Number of patients
Race Differences in Prostate Cancer (MSK, 2021) (41)	2448	2435
Prostate Cancer (MSK, JCO Precis Oncol 2017) (42)	504	401
Prostate Adenocarcinoma (TCGA, Firehose Legacy) (43)	501	500
Prostate Adenocarcinoma (MSK/DFCI, Nature Genetics 2018) (44)	1013	1013
Prostate Adenocarcinoma (MSK, Eur Urol 2020) (45)	1465	1465
Prostate Adenocarcinoma (MSK, Clin Cancer Res. 2022) (46)	1417	1417
Metastatic Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019) (47)	444	429
Neuroendocrine Prostate Cancer (Multi-Institute, Nat Med 2016)	114	81
Metastatic castration-sensitive prostate cancer (MSK, Clin Cancer Res 2020) (48)	424	424
Prostate Adenocarcinoma (MSK, Cancer Cell 2010) (49)	240	238
Prostate Cancer (DKFZ, Cancer Cell 2018) (50)	324	292

Supplementary Table 2: Clinical cohorts accessed on cBioPortal used in this study.

Variable	Source	Data available, n (%)
Number of citations	Ovid MEDLINE Web Of Science Core Collection	45 (100)
Number of patents	World Intellectual Property PATENTSCOPE database	45 (100)
Anatomical Location of sample collection	Original cell line papers (18, 19, 21-57), Cellosaurus website, ATCC website, COSMIC database	45 (100)
Patient age at sample collection	Original cell line papers (18, 19, 21-57), Cellosaurus website, ATCC website, COSMIC database	27 (60)
Geographical location of sample collection	Original cell line papers (18, 19, 21-57)	45 (100)

Supplementary Table 3: Baseline characteristics investigated for each cell line, COSMIC = Catalogue of Somatic Mutations in Cancer

Cell Line	Data Available	Data not available	Cell Line	Data Available	Data not available
All	Year of sample collection Number of citations Number of patents Anatomical location of sample collection Cancer Type		DuCaP	Age at sample collection COSMIC data Ethnicity M stage	CCLE data MSI Gleason grade
LNCaP	Age at sample collection COSMIC data MSI Ethnicity M stage	CCLE data Gleason grade	CWR-R1	Ethnicity M stage	Age at sample collection COSMIC data CCLE data MSI Gleason grade
PC-3	Age at sample collection COSMIC data CCLE data MSI Ethnicity M stage	Gleason grade	ALVA-31	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade
DU145	Age at sample collection COSMIC data CCLE data MSI Ethnicity M stage	Gleason grade	RWPE-2	Age at sample collection Ethnicity Gleason grade M stage	COSMIC data CCLE data MSI
1013 L	COSMIC data MSI Ethnicity	Age at sample collection CCLE data Gleason grade M stage	CA-HPV-10	Age at sample collection Ethnicity Gleason grade M stage	COSMIC data CCLE data MSI
22Rv1	COSMIC data CCLE data Ethnicity Gleason grade	Age at sample collection MSI M stage	MDA Pca 2a	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade
VCaP	Age at sample collection COSMIC data CCLE data MSI Ethnicity M stage	Gleason grade	PSK-1	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade
MDA PCa 1	Age at sample collection MSI Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade	PC-93	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade
CWR22	COSMIC data MSI Ethnicity Gleason grade M stage	Age at sample collection CCLE data	ALVA-41	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade
LAPC-4(LAPC-4-1)	COSMIC data M stage	Age at sample collection CCLE data MSI Ethnicity Gleason grade	PC-346C	Gleason grade M stage	Age at sample collection COSMIC data CCLE data MSI Ethnicity
LNCaP-C4	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade	E006AA	Age at sample collection Ethnicity Gleason grade M stage	COSMIC data CCLE data MSI
MDA Pca 2b	Age at sample collection COSMIC data CCLE data MSI Ethnicity M stage	Gleason grade	NCI-H660	Age at sample collection COSMIC data CCLE data MSI Ethnicity M stage	Gleason grade
PPC-1	Age at sample collection COSMIC data Ethnicity	CCLE data MSI Gleason grade M stage	P69SV40T	Age at sample collection COSMIC data Ethnicity Gleason grade M stage	CCLE data MSI
LNCaP-C4-2	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade	ALVA-101	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade
ARCaP	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade	DuPro-1	Age at sample collection Ethnicity Gleason grade	COSMIC data CCLE data MSI

				M stage	
LNCaP-FGC	Age at sample collection COSMIC data CCLE data MSI Ethnicity M stage	Gleason grade	WPMY-1 normal prostatic myofibroblasts	Age at sample collection Ethnicity Gleason grade M stage	COSMIC data CCLE data MSI
PZ-HPV-7	Age at sample collection Ethnicity Gleason grade M stage	COSMIC data CCLE data MSI	UM-SCP-1	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade
PWR-1E	Age at sample collection MSI Ethnicity Gleason grade M stage	COSMIC data CCLE data	PNF-08	Ethnicity Gleason grade M stage	Age at sample collection COSMIC data CCLE data
ALVA-55	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade	LAPC9	COSMIC data M stage	Age at sample collection CCLE data MSI Ethnicity Gleason grade
KuCaP13	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI Gleason grade	LNCaP104S	Age at sample collection COSMIC data Ethnicity M stage	CCLE data MSI Gleason grade
LASCPC-01	Age at sample collection MSI Ethnicity M stage	COSMIC data CCLE data Gleason grade	LuCAP35	Age at sample collection COSMIC data Ethnicity M stage	CCLE data MSI Gleason grade
hTERT EP156T	Age at sample collection Gleason grade M stage	COSMIC data CCLE data MSI Ethnicity	LAPC-3	COSMIC data M stage	Age at sample collection CCLE data MSI Ethnicity Gleason grade
LNCaP-LN-3	Age at sample collection Ethnicity M stage	COSMIC data CCLE data MSI	RC-77T/E	Age at sample collection Ethnicity Gleason grade M stage	COSMIC data CCLE data MSI
ACRJ-PC28	Age at sample collection Ethnicity M stage Gleason grade	CCLE data COSMIC data MSI			

Supplementary Table 4 : Data available for each cell line

Variable	Source
Patient age at sample collection (radical prostatectomy)	Systematic Review on robotic-assisted radical prostatectomy (51)
Geographical distribution of prostate cancer (incidence)	GCO, IARC, WHO Cancer Today Website (52, 53, 54)
Ethnicity of prostate cancer patients	JAMA worldwide racial demographics of prostate cancer study and SEER database (55)
Gleason grade of prostate cancer patients	Weerakoon et. al The current use of active surveillance in an Australian cohort of men (56)
Prostate cancer type	Ackerman et.al Prostate surgical pathology (57)

Supplementary Table 5: Sources of worldwide Prostate Cancer data. GCO = Global Cancer Observatory, IARC = International Agency for Research on Cancer, WHO = World Health Organization. JAMA = The Journal of the American Medical Association, SEER = The Surveillance, Epidemiology, and End Results Program.

Variable	Variable Type	Variable options detailed
Mutation	Categorical	Missense mutation Truncating mutation Inframe mutation Splice
Copy Number Alteration	Categorical	Amplification Deletion
Structural Variant	Categorical	Fusion
mRNA expression	Continuous	Low (Z score <2 relative to rest of cohort) High (Z score >2 relative to rest of cohort)
Protein expression	Continuous	Low (Z score <2 relative to rest of cohort) High (Z score >2 relative to rest of cohort)
Microsatellite Instability	Categorical	Stable Instable

Supplementary Table 6: Biological variables investigated for each cell line.

Cell Line	Patients with one alteration the same N (%)	Patients with 2-4 alterations the same N (%)	Patients with 5-6 alterations the same N (%)
NCIH660	1299 (90.21)	141 (9.79)	0 (0)
VCaP	1897 (74.71)	625 (24.62)	17 (0.67)
MDA PCA 2b	661 (82.94)	134 (16.81)	2 (0.25)
DU145	1408 (87.95)	193 (12.05)	0 (0)
LNCaP-FGC	1738 (82.02)	380 (19.93)	1 (0.05)
22Rv1	1462 (84.95)	256 (14.88)	3 (0.17)
PC-3	1352 (83.92)	258 (16.02)	1 (0.06)

Supplementary Table 7: Number of patients with alterations the same as cell line

	Cell Line p value	Clinical cohort p value	Worldwide p value
D'Agostino & Pearson Test	<0.0001	0.2983 (ns)	<0.0001
Anderson-Darling test	0.0015	0.0549 (ns)	<0.0001
Shapiro-Wilkins test	0.0008	0.0384	<0.0001
Kolmogorov-Smirnov test	0.0181	>0.100 (ns)	<0.0001

Supplementary Table 8: Results of normality distribution tests for distribution of age at sample collection.

Treatment	Clinical Cohort (%)	Cell Line (%)
No treatment	8.99	30.40
ADT	1.45	30.40
Chemotherapy	0.59	13.04
Unknown	88.97	28.26

Supplementary Table 9: Table demonstrating percentage of cell lines, or patients receiving different forms of treatment for prostate cancer prior to sampling

cBioPortal Name	mRNA Expression	Protein Expression	Mutations	Copy Number Alterations	Structural Variants	Microsatellite Instability
Race Differences in Prostate Cancer (MSK, 2021) (41)	✓		✓	✓	✓	
Prostate Cancer (MSK, JCO Precis Oncol 2017) (42)			✓	✓	✓	
Prostate Adenocarcinoma (TCGA, Firehose Legacy) (43)	✓	✓	✓	✓		
Prostate Adenocarcinoma (MSK/DFCI, Nature Genetics 2018) (44)	✓	✓	✓	✓		
Prostate Adenocarcinoma (MSK, Eur Urol 2020) (45)			✓	✓	✓	✓
Prostate Adenocarcinoma (MSK, Clin Cancer Res. 2022) (46)			✓	✓	✓	✓
Metastatic Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019) (47)	✓		✓	✓	✓	
Neuroendocrine Prostate Cancer (Multi-Institute, Nat Med 2016)	✓		✓			
Metastatic castration-sensitive prostate cancer (MSK, Clin Cancer Res 2020) (48)			✓	✓	✓	✓
Prostate Adenocarcinoma (MSK, Cancer Cell 2010) (49)	✓		✓	✓		
Prostate Cancer (DKFZ, Cancer Cell 2018) (50)	✓		✓			

Supplementary Table 10: Biological alterations screened for in clinical cohorts accessed on cBioPortal.

1. Horoszewicz JS, Leong SS, Chu TM, Wajsman ZL, Friedman M, Papsidero L, et al. The LNCaP cell line--a new model for studies on human prostatic carcinoma. *Prog Clin Biol Res.* 1980;37:115-32.
2. Kaighn ME, Lechner JF, Narayan KS, Jones LW. Prostate carcinoma: tissue culture cell lines. *Natl Cancer Inst Monogr.* 1978(49):17-21.
3. Stone KR, Mickey DD, Wunderli H, Mickey GH, Paulson DF. Isolation of a human prostate carcinoma cell line (DU 145). *Int J Cancer.* 1978;21(3):274-81.
4. Williams RD. Human urologic cancer cell lines. *Invest Urol.* 1980;17(5):359-63.
5. Sramkoski RM, Pretlow TG, 2nd, Giaconia JM, Pretlow TP, Schwartz S, Sy MS, et al. A new human prostate carcinoma cell line, 22Rv1. *In Vitro Cell Dev Biol Anim.* 1999;35(7):403-9.
6. Korenchuk S, Lehr JE, L MC, Lee YG, Whitney S, Vessella R, et al. VCaP, a cell-based model system of human prostate cancer. *In Vivo.* 2001;15(2):163-8.
7. Zhau HY, Chang SM, Chen BQ, Wang Y, Zhang H, Kao C, et al. Androgen-repressed phenotype in human prostate cancer. *Proc Natl Acad Sci U S A.* 1996;93(26):15152-7.
8. Pretlow TG, Wolman SR, Micale MA, Pelley RJ, Kursh ED, Resnick MI, et al. Xenografts of primary human prostatic carcinoma. *J Natl Cancer Inst.* 1993;85(5):394-8.
9. Klein KA, Reiter RE, Redula J, Moradi H, Zhu XL, Brothman AR, et al. Progression of metastatic human prostate cancer to androgen independence in immunodeficient SCID mice. *Nat Med.* 1997;3(4):402-8.
10. Wu HC, Hsieh JT, Gleave ME, Brown NM, Pathak S, Chung LW. Derivation of androgen-independent human LNCaP prostatic cancer cell sublines: role of bone stromal cells. *Int J Cancer.* 1994;57(3):406-12.
11. Navone NM, Olive M, Ozen M, Davis R, Troncoso P, Tu SM, et al. Establishment of two human prostate cancer cell lines derived from a single bone metastasis. *Clin Cancer Res.* 1997;3(12 Pt 1):2493-500.
12. Brothman AR, Lesho LJ, Somers KD, Wright GL, Jr., Merchant DJ. Phenotypic and cytogenetic characterization of a cell line derived from primary prostatic carcinoma. *Int J Cancer.* 1989;44(5):898-903.
13. Carroll AG, Voeller HJ, Sugars L, Gelmann EP. p53 oncogene mutations in three human prostate cancer cell lines. *Prostate.* 1993;23(2):123-34.
14. Weijerman PC, König JJ, Wong ST, Niesters HG, Peehl DM. Lipofection-mediated immortalization of human prostatic epithelial cells of normal and malignant origin using human papillomavirus type 18 DNA. *Cancer Res.* 1994;54(21):5579-83.
15. Webber MM, Bello D, Kleinman HK, Wartinger DD, Williams DE, Rhim JS. Prostate specific antigen and androgen receptor induction and characterization of an immortalized adult human prostatic epithelial cell line. *Carcinogenesis.* 1996;17(8):1641-6.
16. Lee YG, Korenchuk S, Lehr J, Whitney S, Vessella R, Pienta KJ. Establishment and characterization of a new human prostatic cancer cell line: DuCaP. *In Vivo.* 2001;15(2):157-62.
17. Gregory CW, Johnson RT, Jr., Mohler JL, French FS, Wilson EM. Androgen receptor stabilization in recurrent prostate cancer is associated with hypersensitivity to low androgen. *Cancer Res.* 2001;61(7):2892-8.
18. Loop SM, Rozanski TA, Ostenson RC. Human primary prostate tumor cell line, ALVA-31: a new model for studying the hormonal regulation of prostate tumor cell growth. *Prostate.* 1993;22(2):93-108.
19. Webber MM, Bello D, Quader S. Immortalized and tumorigenic adult human prostatic epithelial cell lines: characteristics and applications Part 2. Tumorigenic cell lines. *Prostate.* 1997;30(1):58-64.

20. Kim CJ, Kushima R, Okada Y, Seto A. Establishment and characterization of a prostatic small-cell carcinoma cell line (PSK-1) derived from a patient with Klinefelter syndrome. *Prostate*. 2000;42(4):287-94.
21. Claas FH, van Steenbrugge GJ. Expression of HLA-like structures on a permanent human tumor line PC-93. *Tissue Antigens*. 1983;21(3):227-32.
22. Plymate SR, Loop SM, Hoop RC, Wiren KM, Ostenson R, Hryb DJ, et al. Effects of sex hormone binding globulin (SHBG) on human prostatic carcinoma. *J Steroid Biochem Mol Biol*. 1991;40(4-6):833-9.
23. Romijn J.C. E-SS, van Weerden W.M., Teubel W., Trapman J., Schroder F.H. Development and characterization of a new in vitro model to study the mechanisms of prostate cancer progression. *Urol Res*. 1995;23:P81-265.
24. Koochekpour S, Maresh GA, Katner A, Parker-Johnson K, Lee TJ, Hebert FE, et al. Establishment and characterization of a primary androgen-responsive African-American prostate cancer cell line, E006AA. *Prostate*. 2004;60(2):141-52.
25. Carney DN, Gazdar AF, Bepler G, Guccion JG, Marangos PJ, Moody TW, et al. Establishment and identification of small cell lung cancer cell lines having classic and variant features. *Cancer Res*. 1985;45(6):2913-23.
26. Bae VL, Jackson-Cook CK, Brothman AR, Maygarden SJ, Ware JL. Tumorigenicity of SV40 T antigen immortalized human prostate epithelial cells: association with decreased epidermal growth factor receptor (EGFR) expression. *Int J Cancer*. 1994;58(5):721-9.
27. Gingrich JR, Tucker JA, Walther PJ, Day JW, Poulton SH, Webb KS. Establishment and characterization of a new human prostatic carcinoma cell line (DuPro-1). *J Urol*. 1991;146(3):915-9.
28. Webber MM, Trakul N, Thraves PS, Bello-DeOcampo D, Chu WW, Storto PD, et al. A human prostatic stromal myofibroblast cell line WPMY-1: a model for stromal-epithelial interactions in prostatic neoplasia. *Carcinogenesis*. 1999;20(7):1185-92.
29. Grossman HB, Wedemeyer G, Ren L, Carey TE. UM-SCP-1, a new human cell line derived from a prostatic squamous cell carcinoma. *Cancer Res*. 1984;44(9):4111-7.
30. Theodore S, Sharp S, Zhou J, Turner T, Li H, Miki J, et al. Establishment and characterization of a pair of non-malignant and malignant tumor derived cell lines from an African American prostate cancer patient. *Int J Oncol*. 2010;37(6):1477-82.
31. Mehta PP, Lokeshwar BL, Schiller PC, Bendix MV, Ostenson RC, Howard GA, et al. Gap-junctional communication in normal and neoplastic prostate epithelial cells and its regulation by cAMP. *Mol Carcinog*. 1996;15(1):18-32.
32. Okasho K, Mizuno K, Fukui T, Lin YY, Kamiyama Y, Sunada T, et al. Establishment and characterization of a novel treatment-related neuroendocrine prostate cancer cell line KUCaP13. *Cancer Sci*. 2021;112(7):2781-91.
33. Lee JK, Phillips JW, Smith BA, Park JW, Stoyanova T, McCaffrey EF, et al. N-Myc Drives Neuroendocrine Prostate Cancer Initiated from Human Prostate Epithelial Cells. *Cancer Cell*. 2016;29(4):536-47.
34. Kogan I, Goldfinger N, Milyavsky M, Cohen M, Shats I, Dobler G, et al. hTERT-immortalized prostate epithelial and stromal-derived cells: an authentic in vitro model for differentiation and carcinogenesis. *Cancer Res*. 2006;66(7):3531-40.
35. Pettaway CA, Pathak S, Greene G, Ramirez E, Wilson MR, Killion JJ, et al. Selection of highly metastatic variants of different human prostatic carcinomas using orthotopic implantation in nude mice. *Clin Cancer Res*. 1996;2(9):1627-36.
36. Szczyrba J, Löprich E, Wach S, Jung V, Unteregger G, Barth S, et al. The microRNA profile of prostate carcinoma obtained by deep sequencing. *Mol Cancer Res*. 2010;8(4):529-38.

37. Craft N, Chhor C, Tran C, Beldegrun A, DeKernion J, Witte ON, et al. Evidence for clonal outgrowth of androgen-independent prostate cancer cells from androgen-dependent tumors through a two-step process. *Cancer Res.* 1999;59(19):5030-6.
38. Kokontis J, Takakura K, Hay N, Liao S. Increased androgen receptor activity and altered c-myc expression in prostate cancer cells after long-term androgen deprivation. *Cancer Res.* 1994;54(6):1566-73.
39. Corey E, Quinn JE, Buhler KR, Nelson PS, Macoska JA, True LD, et al. LuCaP 35: a new model of prostate cancer progression to androgen independence. *Prostate.* 2003;55(4):239-46.
40. Valentine H, Aiken W, Morrison B, Zhao Z, Fowle H, Wasserman JS, et al. Expanding the Prostate Cancer Cell Line Repertoire with ACRJ-PC28, an AR-negative Neuroendocrine Cell Line Derived From an African-Caribbean Patient. *Cancer Research Communications.* 2022;2(11):1355-71.
41. Stopsack KH, Nandakumar S, Arora K, Nguyen B, Vasselman SE, Nweji B, et al. Differences in Prostate Cancer Genomes by Self-reported Race: Contributions of Genetic Ancestry, Modifiable Cancer Risk Factors, and Clinical Factors. *Clinical Cancer Research.* 2022;28(2):318-26.
42. Abida W, Armenia J, Gopalan A, Brennan R, Walsh M, Barron D, et al. Prospective Genomic Profiling of Prostate Cancer Across Disease States Reveals Germline and Somatic Alterations That May Affect Clinical Decision Making. *JCO Precision Oncology.* 2017(1):1-16.
43. Abeshouse A, Ahn J, Akbani R, Ally A, Amin S, Andry Christopher D, et al. The Molecular Taxonomy of Primary Prostate Cancer. *Cell.* 2015;163(4):1011-25.
44. Armenia J, Wankowicz SAM, Liu D, Gao J, Kundra R, Reznik E, et al. The long tail of oncogenic drivers in prostate cancer. *Nature Genetics.* 2018;50(5):645-51.
45. Nguyen B, Mota JM, Nandakumar S, Stopsack KH, Weg E, Rathkopf D, et al. Pan-cancer Analysis of CDK12 Alterations Identifies a Subset of Prostate Cancers with Distinct Genomic and Clinical Characteristics. *European Urology.* 2020;78(5):671-9.
46. Chakraborty G, Nandakumar S, Hirani R, Nguyen B, Stopsack KH, Kreitzer C, et al. The Impact of PIK3R1 Mutations and Insulin-PI3K-Glycolytic Pathway Regulation in Prostate Cancer. *Clinical Cancer Research.* 2022;28(16):3603-17.
47. Abida W, Cyrta J, Heller G, Prandi D, Armenia J, Coleman I, et al. Genomic correlates of clinical outcome in advanced prostate cancer. *Proceedings of the National Academy of Sciences.* 2019;116(23):11428-36.
48. Stopsack KH, Nandakumar S, Wibmer AG, Haywood S, Weg ES, Barnett ES, et al. Oncogenic Genomic Alterations, Clinical Phenotypes, and Outcomes in Metastatic Castration-Sensitive Prostate Cancer. *Clinical Cancer Research.* 2020;26(13):3230-8.
49. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative Genomic Profiling of Human Prostate Cancer. *Cancer Cell.* 2010;18(1):11-22.
50. Gerhauser C, Favero F, Risch T, Simon R, Feuerbach L, Assenov Y, et al. Molecular Evolution of Early-Onset Prostate Cancer Identifies Molecular Risk Markers and Clinical Trajectories. *Cancer Cell.* 2018;34(6):996-1011.e8.
51. Du Y, Long Q, Guan B, Mu L, Tian J, Jiang Y, et al. Robot-Assisted Radical Prostatectomy Is More Beneficial for Prostate Cancer Patients: A System Review and Meta-Analysis. *Medical science monitor : international medical journal of experimental and clinical research.* 2018;24:272-87.
52. Ferlay J EM, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F Global Cancer Observatory: Cancer Today. : Lyon, France: International Agency for Research on Cancer; 2020 [Available from: <https://gco.iarc.fr/today>,.

53. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*. 2021;71(3):209-49.
54. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. 2019;144(8):1941-53.
55. Fletcher SA, Marchese M, Cole AP, Mahal BA, Friedlander DF, Krimphove M, et al. Geographic Distribution of Racial Differences in Prostate Cancer Mortality. *JAMA Network Open*. 2020;3(3):e201839-e.
56. Weerakoon M, Papa N, Lawrentschuk N, Evans S, Millar J, Frydenberg M, et al. The current use of active surveillance in an Australian cohort of men: a pattern of care analysis from the Victorian Prostate Cancer Registry. *BJU International*. 2015;115(S5):50-6.
57. Ackerman LA, Rosai J. Surgical pathology. *Surgical pathology*1974. p. 1394-.