Supplementary Material

Supplementary Tables

Supplementary Table 1: Patient characteristics distributed by study, with The Cancer Genome Atlas (TCGA) set appended.

Study	Clinical variable	Class	# patients
Australia			
	Relapsed	0	54
		1	79
	Gleason Score	<4+3	23
		≥4+3	110
	PSA	median	8.90
	Pathological T	T stage < 3	37
		T stage ≥ 3	96
	Age at diagnosis	median	66
		< 65	61
		≥ 65	72
	Follow-up (days)	median	259
Canada			
	Relapsed	0	196
		1	92
	Gleason Score	<4+3	225
		≥4+3	63
	PSA	median	6.90

	Pathological T	T stage < 3	187
		T stage ≥ 3	101
	Age at diagnosis	median	64
		< 65	151
		≥ 65	137
	Follow-up (days)	median	229
France			
	Relapsed	0	5
		1	10
	Gleason Score	<4+3	0
		≥4+3	15
	PSA	median	7.90
	Pathological T	T stage < 3	1
		T stage ≥ 3	14
	Age at diagnosis	median	64
		< 65	8
		≥ 65	7
	Follow-up (days)	median	210
Germany			
	Relapsed	0	162
		1	68
	Gleason Score	<4+3	145
		≥4+3	85
	PSA	median	9.20
	Pathological T	T stage < 3	136

		T stage ≥ 3	94
	Age at diagnosis	median	48
		< 65	204
		≥ 65	256
	Follow-up (days)	median	753
UK			
	Relapsed	0	148
		1	36
	Gleason Score	<4+3	121
		≥4+3	63
	PSA	median	8.00
	Pathological T	T stage < 3	61
		T stage ≥ 3	123
	Age at diagnosis	median	62
		< 65	113
		≥ 65	71
	Follow-up (days)	median	991
TCGA			
	Relapsed	0	339
		1	44
	Gleason Score	<4+3	150
		≥4+3	233
	Pathological T	T stage < 3	142
		T stage ≥ 3	241
	Follow-up (days)	median	417

Supplementary Table 2: Patient data.

Separate File

Supplementary Table 3: Multifactor *Cox* model results for predicted-deleterious mutations in 778 out of 850 germline samples (excluding patients treated with radiotherapy), grouped into 52 gene-sets. Shown are *p*-values and hazard ratios of LASSO-selected gene-sets as well as clinical variables reported at time of biochemical recurrence (BCR) or last check-up, impacting the predicted time until BCR.

	HR (95% CI)	***p*-value**
Gleason (≥4+3 : <4+3)	2.38 (1.72 – 3.31)	**2.13e-7**
Stage (T3-T4 : T1-T2)	1.99 (1.46 – 2.70)	**1.18e-5**
PI3K/AKT/mTOR signalling	1.60 (1.09 – 2.36)	**0.0177**
Inflammatory response	1.41 (1.03 – 1.93)	**0.0321**
KRAS signalling (up)	1.39 (1.02 – 1.88)	**0.0346**
Myc targets v2	1.28 (0.86 – 1.91)	0.22
p53 pathway	1.26 (0.91 – 1.76)	0.166
DRG	1.26 (0.95 – 1.67)	0.115
Age	1.22 (0.93 – 1.60)	0.152
Fatty acid metabolism	1.22 (0.90 – 1.65)	0.21
G2-M checkpoint	1.18 (0.86 – 1.61)	0.30

IL-6/JAK/STAT3 signalling	1.11 (0.71 – 1.73)	0.66
Mitotic spindle	1.10 (0.83 – 1.45)	0.51
Preop_PSA	1.03 (1.01 – 1.06)	**0.0117**
UV response (dn)	0.71 (0.50 – 1.02)	0.064
Cholesterol homeostasis	0.50 (0.28 – 0.92)	**0.0244**

Supplementary Table 4: Gene-sets used in study: 50 Hallmark sets from Gene-set Enrichment Analysis Molecular Signatures Database (GSEA MsigDB); BROCA extended panel and DRG panel.

Separate File

Supplementary Table 5: Multifactor *Cox* model results for clinical variables in 850 germline samples, impacting the predicted time until biochemical recurrence. Gleason and T-stage were reported at time of biochemical recurrence or last follow-up, while age and PSA were reported at time of surgery.

	HR (95% CI)	***p*-value**
Gleason (≥4+3 : <4+3)	1.99 (1.49 – 2.66)	**2.81×10^{-6}**
Stage (T3-T4 : T1-T2)	1.71 (1.31 – 2.24)	**7.70×10^{-5}**
Age	1.47 (1.15 – 1.87)	**2.01×10^{-3}**
Preop_PSA	1.04 (1.01 – 1.06)	**7.81×10^{-3}**

Supplementary Table 6: Univariate Cox model results for predicted-deleterious mutations in 850 germline samples, grouped into 52 gene-sets. Shown are hazard ratios, *p* and adjusted *p*-values of the most significant (*p*-value threshold < 0.1) gene-sets in terms of predicting time until biochemical recurrence.

	HR (95% CI)	***p*-value**	***q*-value**
PI3K/AKT/mTOR signalling	1.44 (0.99 – 2.07)	0.0536	0.143
G2-M checkpoint	1.33 (1.02 – 1.74)	0.0361	0.29
KRAS signalling (up)	1.32 (1.00 – 1.75)	0.0508	0.20
TNFA signalling via NFKB	1.31 (0.97 – 1.77)	0.0786	0.157
Inflammatory response	1.29 (0.97 – 1.72)	0.0823	0.110
DRG	1.24 (0.96 – 1.59)	0.0955	0.0955
Mitotic spindle	1.23 (0.98 – 1.58)	0.0910	0.104
Cholesterol homeostasis	0.63 (0.37 – 1.06)	0.0801	0.128

Supplementary Table 7: Multifactor *Cox* model results for predicted-deleterious mutations in 383 The Cancer Genome Atlas (TCGA) germline samples, stratified by location and grouped into 52 gene-sets. Shown are *p*-values and hazard ratios of the same predictors identified by the Pan Prostate Cancer Group (PPCG) *Cox* model

(cholesterol homeostasis was removed as samples have no mutations in this gene-set, which caused convergence errors).

	HR (95% CI)	*p*-value	Bootstrap HR (95% CI)	Bootstrap *p*-value
Myc targets v2	4.46 (1.73 – 11.5)	**$1.99x10^{-3}$**	6.43 (6.17 – 6.72)	$6.00x10^{-3}$
Coagulation	3.49 (1.47 – 8.30)	**$4.64x10^{-3}$**	5.42 (5.15 – 5.72)	0.0110
Gleason (≥4+3 : <4+3)	2.98 (1.06 – 8.33)	**0.0377**	$1.07x10^{6}$ ($4.91x10^{5}$ – $2.08x10^{6}$)	$6.00x10^{-3}$
Stage (T3-T4 : T1-T2)	2.89 (1.01 – 8.28)	**0.0484**	$7.22x10^{12}$ ($1.41x10^{6}$ – $4.33x10^{13}$)	0.0400
G2-M checkpoint	2.19 (0.91 – 5.25)	0.0805	3.08 (2.94 – 3.26)	0.0540
Inflammatory response	1.71 (0.62 – 4.77)	0.30	2.12 (2.03 – 2.24)	0.21
Fatty acid metabolism	1.44 (0.50 – 4.18)	0.50	1.77 (1.69 – 1.84)	0.27
KRAS signalling (up)	1.16 (0.53 – 2.54)	0.71	1.29 (1.25 – 1.34)	0.40
p53 pathway	0.83 (0.31 – 2.26)	0.72	0.91 (0.88 – 0.95)	0.36
DRG	0.81 (0.41 – 1.62)	0.56	0.88 (0.85 – 0.91)	0.30
Mitotic spindle	0.73 (0.30 – 1.78)	0.49	0.76 (0.74 – 0.79)	0.22
PI3K/AKT/mTOR signalling	0.70 (0.09 – 5.49)	0.74	0.92 (0.85 – 1.01)	0.35
IL-2/STAT5 signalling	0.65 (0.26 – 1.62)	0.36	0.81 (0.77 – 0.845)	0.25
UV response (dn)	0.39 (0.10 – 1.47)	0.165	0.52 (0.49 – 0.56)	0.131
Glycolysis	0.37 (0.10 – 1.29)	0.117	0.46 (0.44 – 0.49)	0.0870

Supplementary Table 8: Odds Ratio results for the event of biochemical recurrence given predicted-deleterious mutations in 850 germline samples.

Separate File

Supplementary Table 9: Odds Ratio results for the event of biochemical recurrence given predicted-deleterious mutations in 850 germline samples. Results are filtered to include only genes with OR > 2 and a difference between *Has Mutation + Has BCR* vs *Has Mutation + No BCR* of at least two within the significant all sample gene-sets: PI3K/AKT/mTOR signalling, KRAS signalling (up) and Inflammatory response, and high-Gleason gene-sets: Hypoxia, PI3K/AKT/mTOR signalling, TNFA signalling via NFKB and KRAS signalling (up). Pancreas-beta cells is a significant high-Gleason gene-set, but has no genes with OR > 2.

Gene-set	Gene	Has Mutation (Has BCR)	No Mutation (Has BCR)	*p*
Hypoxia	**GAPDHS**	8 (6)	842 (279)	0.0198
	GRHPR	2 (2)	848 (283)	0.112
	PGM1	2 (2)	848 (283)	0.112
	SELENBP1	2 (2)	848 (283)	0.112
	NAGK	2 (2)	848 (283)	0.112
	SLC6A6	2 (2)	848 (283)	0.112
PI3K/AKT/mTOR signalling	**PIKFYVE**	6 (5)	844 (280)	0.0180
	MYD88	2 (2)	848 (283)	0.112
	CAB39	2 (2)	848 (283)	0.112
	RPS6KA1	2 (2)	848 (283)	0.112
TNFA signalling via NFKB	**DDX58**	3 (3)	847 (282)	0.0374
	KYNU	2 (2)	848 (283)	0.112
	NR4A1	2 (2)	848 (283)	0.112
	DENND5A	2 (2)	848 (283)	0.112
KRAS signalling (up)	**MMP10**	7 (5)	843 (280)	0.0457
	HKDC1	6 (4)	844 (281)	0.0938
	RBM4	2 (2)	848 (283)	0.112

Inflammatory response	**IRAK2**	4(4)	846 (281)	0.0125
	IL2RB	3(3)	847 (282)	0.0374
	MSR1	2(2)	848 (283)	0.112
	ITGB8	2(2)	848 (283)	0.112
	PIK3R5	2(2)	848 (283)	0.112

Supplementary Figure Legends

Supplementary Figure 1: For Supplementary Table 5: Horizontal box plot of the coefficient / log hazard ratios with lower and upper 95% confidence intervals.

Supplementary Figure 2: Oncoplot of 22 genes from Supplementary Table 9 altered in 211 of 850 samples. Variants are unfiltered. Right chart shows mutation distribution per gene. Variants annotated as Multi_Hit are those genes which are mutated more than once in the same sample.

Supplementary Figure 3: Oncoplot of 22 genes from Supplementary Table 9 altered in 107 of 285 samples with biochemical recurrence. Variants are unfiltered. Right chart shows mutation distribution per gene. Variants annotated as Multi_Hit are those genes which are mutated more than once in the same sample.

Supplementary Figure 4: Oncoplot of 22 genes from Supplementary Table 9 altered in 102 of 565 samples without biochemical recurrence. Variants are unfiltered. Right chart shows mutation distribution per gene. Variants annotated as Multi_Hit are those genes which are mutated more than once in the same sample.

Supplementary Methods

Supplementary Method 1

Unless otherwise stated, all patients underwent radical prostatectomy (RP), and biochemical recurrence (BCR) was defined as two consecutive post-RP PSA measurements of more than 0.2 ng/ml (backdated to the date of the first increase). If a patient had successful salvage radiation therapy, this was not considered BCR. If PSA continued to rise after radiation therapy, BCR was backdated to first PSA>0.2 ng/ml. If a patient received other salvage treatment (such as hormones or chemotherapy), this was considered BCR.

Melbourne, Australian research group

All patients were hormone-naïve at the time of treatment. Patients were retrospectively selected from our tissue biorepositry enriching for patients with high grade disease.

DNA and RNA were simultaneously extracted using the Allprep Micro Kit (Qiagen, CA) following manufacturer instructions and including on column DNAse digestion of the RNA. Genomic DNA was extracted from fresh frozen samples of whole blood with the DNeasy Blood & Tissue Kit (Qiagen, Maryland) following manufacturer instructions.

Canadian Prostate Cancer Genome Network

All patients underwent either image-guided radiotherapy (IGRT) or radical prostatectomy (RP), with curative intent, for pathologically confirmed prostate

cancer. All patients were hormone-naïve at the time of definitive local therapy. In the IGRT cohort, a single ultrasound-guided needle biopsy was obtained before the start of therapy. Fresh-frozen RP specimens were obtained from the University Health Network (UHN) Pathology BioBank or from the Genito-Urinary BioBank of the Centre Hospitalier Universitaire de Québec (CHUQ).

For IGRT patients, BCR was defined as a rise in PSA concentration of more than 2.0 ng/ml above the nadir (after radiotherapy, PSA levels drop and stabilize at the nadir).

Whole blood was collected and informed consent, consistent with local Research Ethics Board (REB) and International Cancer Genome Consortium (ICGC) guidelines, was obtained at the time of clinical follow-up. All patients were N0M0 as an entry criterion for the study.

Fraser M, Sabelnykova VY, Yamaguchi TN, et al. Genomic hallmarks of localized, non-indolent prostate cancer. *Nature*. 2017;**541**:359-64. https://doi.org/10.1038/nature20788

French ICGC Prostate Cancer Group

The French cohort is comprised of Caucasian patients with aggressive prostate cancer characterized by a clinical-pathological aggressive pattern (D'Amico 3 with primary Gleason grade 4). All patients were treatment-naïve at the time of surgery.

They provided written informed consent, consistent with local Research Ethics Board (REB) and the International Cancer Genome Consortium (ICGC) guidelines. For

germline DNA extraction, saliva was collected using the Oragene DNA collection kit (DNA Genotek Inc) at the time of consent.

Germany ICGC Prostate Cancer Group – Early Onset (EO)

The EO cohort is composed of patients diagnosed with PC <= 55 years of age. Except for two patients (PCA125 and PCA176) who received pre-operation hormone therapy with LH-RH, the patients did not receive any neo-adjuvant radiotherapy, androgen deprivation therapy, or chemotherapy prior to the surgical removal of tumour tissue.

DNA and RNA were extracted as described previously:

Weischenfeldt J, Simon R, Feuerbach L, et al. Integrative genomic analyses reveal an androgen-driven somatic alteration landscape in early-onset prostate cancer. *Cancer Cell*. 2013;**23**:159-70. https://doi.org/10.1016/j.ccr.2013.01.002

CRUK-ICGC Prostate Group, UK

Fresh frozen tumour and matching whole blood samples were collected from radical prostatectomy patients treated at The Royal Marsden NHS Foundation Trust, London, at the Addenbrooke's Hospital, Cambridge, or at Oxford University Hospitals NHS Trust. Consequently those samples with >40% tumour content and their matching blood samples were whole genome sequenced. All patients were treatment naïve at the time of surgery.

This data was collected as part of the CRUK-ICGC prostate project within the framework of ICGC and more information can be found in previous publications:

Cooper CS, Eeles R, Wedge DC, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat Genet*. 2015;**47**:367-72. https://doi.org/10.1038/ng.3221

Wedge DC, Gundem G, Mitchell T, et al. Sequencing of prostate cancers identifies new cancer genes, routes of progression and drug targets. *Nat Genet*. 2018;**50**:682-92. https://doi.org/10.1038/s41588-018-0086-z

Supplementary Method 2

In brief, after read alignment and duplicate removal, Base Quality Score Recalibration (BSQR) was performed to detect errors introduced by the sequencer and correct the quality scores assigned to each base call. Variants were called using GATK HaplotypeCaller via local de-novo assembly of haplotypes in a region, producing one gvcf file per sample. Joint-genotyping was performed on the whole cohort, producing one multi-sample VCF file. Variant Quality Score Recalibration (VQSR) was performed to remove false positive variants by comparing them against a high quality set. Genotype posteriors were calculated using 1000 Genomes phase 3 VCF. Indels were left-aligned, and multi-allelic variants were decomposed into bi-allelic components.

Supplementary Method 3

Single sample- and paired-sample calling modes were used for discovery of SNVs, multi nucleotide variants, and indels <50bp. Raw variant predictions were further filtered for quality (QUAL>20, QUAL/AO>2), strand bias artefacts (SAF>1, SAR>1), read position artefacts (RPR>1, RPL>1), and normalized for consistent representation across patients with *vt* v0.5.

Supplementary Author list:

Benedicte Ulhøi[a], Jørgen Bjerggaard Jensen[b,c], Michael Borre[b,d], Leila Farajzadeh[b,e], Philippe Lamy[b], Justin Peters[f], Anne Warren[g], Peter Van Loo[h,i,j], Tom Lesluyes[k], Nana Mensah[k], Geoff Macintyre[l], Lars Feuerbach[m], Clarissa Gerhäuser[n], Christoph Plass[n], Holger Sültmann[o,p], Sarah Minner[q], Ronald Simon[q], William B. Isaacs[r], Paul Boutros[s], Takafumi Yamaguchi[s], Maria Jakobsdottir[t], Atef Sahli[t], Markus Graefen[u], Peter Georgeson[v,w,x], Chol-Hee Jung[v], Edmund Lau[v,y], Daniel Park[v,z], Sebastian Uhrig[aa,ab], Abraham Gihawi[ac], Valeriia Haberland[ac], Jingjing Zhang[ac,ad], Ian Mills[ae], Andrew Erickson[ae], Alastair D. Lamb[ae,af], Clare Verrill[ae], Kevin Cheng[ag], Diogo Pellegrina[ag], Housheng Hansen He[ah], G. Steven Bova[ai], Jocelyn Penington[aj], Katherine O'Neill[ak], Bethany Campbell[al], Ken Chow[al], Michael J. Clarkson[al], Marek Cmero[al,v], Anthony Costello[al,f], Matthew K H. Hong[al], Michael Kerger[al,am], Natalie Kurganovs[al], Patrick McCoy[al], Anne Nguyen[al], Ryan Stuchbery[al], Andrew Ryan[an], Alejandro Berlin[ao], Theo Van Der Kwast[ap], Andreas J. Gruber[aq], Weerachai Jaratlerdsiri[ar], Pamela Soh[ar], Justin Bedő[as,aj], Ruining Dong[as,aj], Breon Feran[as], Stefano Mangiola[as,at], Ramyar Molania[as,au], Anthony Papenfuss[as,av,au], Alan Rubin[as,aj], Ismael Vergara[as], Matthew Wakefield[as,aj], David Brundage[aw], Ryan Carelli[aw], Eddie L. Imada[aw], Francesca Khani[aw], Luigi Marchionni[aw], Brian D. Robinson[aw], Claudio Zanettini[aw], Yaobo Xu[ax]

[a]Aarhus University Hospital, Department of Pathology, Palle Juul-Jensens Boulevard 99, DK-8200 Aarhus N, Denmark
[b]Aarhus University, Department of Clinical Medicine, Palle Juul-Jensens Boulevard

99, DK-8200 Aarhus N, Denmark
[c]Regional Hospital West Jutland, Palle Juul-Jensens Boulevard 99, DK-7500 Holstebro, Denmark
[d]Aarhus University Hospital, Department of Urology, Palle Juul-Jensens Boulevard 99, DK-8200 Aarhus N, Denmark
[e]Aarhus University Hospital, Department of Molecular Medicine, Palle Juul-Jensens Boulevard 99, DK-8200 Aarhus N, Denmark
[f]Australian Prostate Cancer Centre, Melbourne, Victoria, Australia
[g]Cambridge University Hospitals NHS Foundation Trust, Department of Histpathology, Hill Road, Cambridge, CB2 0QQ, UK
[h]Cancer Genomics Group, The Francis Crick Institute, London, NW1 1AT, UK
[i]The University of Texas MD Anderson Cancer Center, Department of Genetics, Houston, TX 77030, USA
[j]The University of Texas MD Anderson Cancer Center, Department of Genomic Medicine, Houston, TX 77030, USA
[k]Cancer Genomics Group, The Francis Crick Institute, London, UK
[l]Fundación del Sector Público Estatal Centro Nacional de Investigaciones Oncológicas Carlos III (F.S.P. CNIO), Spain
[m]German Cancer Research Center (DKFZ), Divison Applied Bioinformatics, 69120 Heidelberg, Germany
[n]German Cancer Research Center (DKFZ), Division Cancer Epigenomics, 69120 Heidelberg, Germany
[o]German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), Division Cancer Genome Research, 69120 Heidelberg, Germany
[p]National Center for Tumor Diseases (NCT), 69120 Heidelberg, Germany
[q]Institute of Pathology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[r]John Hopkins School of Medicine, Baltimore, MD 21205, USA
[s]Jonsson Comprehensive Cancer Center, University of California - Los Angeles, 10833 Le Conte Ave., Los Angeles, California 90024, USA
[t]Manchester Cancer Research Centre, University of Manchester, Division of Cancer Sciences, 555 Wilmslow Road, Manchester, M20 4GJ, UK
[u]Martini-Klinik, University Medical Centre Hamburg - Eppendorf, Martinistrasse 52, Hamburg 20246, Germany
[v]Melbourne Bioinformatics, The University of Melbourne, Gratton Street, Parkville, Victoria 3010, Australia
[w]The University of Melbourne, Department of Clinical Pathology, Gratton Street, Melbourne, Victoria, Australia
[x]University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Melbourne, Victoria, Australia
[y]School of Mathematics and Statistics, The University of Melbourne, Gratton Street, Parkville, Victoria 3010, Australia
[z]The University of Melbourne, Department of Biochemistry and Pharmacology, Gratton Street, Parkville, Victoria 3010, Australia

[aa]National Center for Tumor Diseases (NCT), Computational Oncology, Molecular Precision Oncology Program, 69120 Heidelberg, Germany
[ab]German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
[ac]Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, NR4 7JT, UK
[ad]The Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ, UK
[ae]Nuffield Department of Surgical Sciences, University of Oxford, Headington, Oxford, OX3 9DQ, UK
[af]Oxford University Hospitals NHS Foundation Trust, Department of Urology, Oxford, UK
[ag]Ontario Institute for Cancer Research, MaRS Centre, 661 University Avenue, Suite 510, Toronto, Ontario, M5G 0A3, Canada
[ah]Princess Margaret Cancer Centre, University Health Network, University of Toronto, Department of Medical Biophysics, Toronto, Ontario, M5G 1L7, Canada
[ai]Prostate Cancer Research Center, Faculty of Medicine and Health Technology, Tampere University, FI-33014 Tampere, Finland
[aj]School of Computing and Information Systems, The University of Melbourne, Parkville, Victoria 3010, Australia
[ak]School of Medicine, University of St Andrews, North Haugh, St Andrews, KY16 9TF, UK
[al]The University of Melbourne, Department of Surgery, Gratton Street, Parkville, Victoria 3010, Australia
[am]Australian Prostate Cancer Centre, Melbourne, Victoria, Australia
[an]TissuPath, Mount Waverly, Victoria, Australia
[ao]University Health Network, Toronto, Canada
[ap]University Health Network, Laboratory Medicine Program, Toronto, Canada
[aq]University of Konstanz, Universitaetsstrasse 10, D-78464 Konstanz, Germany
[ar]University of Sydney, Sydney, Australia
[as]Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia
[at]The University of Melbourne, Department of Medical Biology, Melbourne, Victoria 3010, Australia
[au]The University of Melbourne, Department of Medical Biology, Melbourne, Victoria 3010, Australia
[av]Peter MacCallum Cancer Centre, Melbourne, Victoria 3000, Australia
[aw]Weill Cornell Medicine, Department of Pathology & Laboratory Medicine, 1300 York Avenue, New York, NY 10065, USA
[ax]Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK