Genomics of lethal prostate cancer at diagnosis and castration-resistance


Published in:
The Journal of clinical investigation

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© 2020 American Society for Clinical Investigation. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Genomics of lethal prostate cancer at diagnosis and castration-resistance

Joaquin Mateo, …, Suzanne Carreira, Johann S. de Bono

*J Clin Invest.* 2019. [https://doi.org/10.1172/JCI132031.](https://doi.org/10.1172/JCI132031)

Genomics of primary prostate cancer differs from that of metastatic castration-resistant prostate cancer (mCRPC). We studied genomic aberrations in primary prostate cancer biopsies from patients who developed mCRPC, also studying matching, same patient, diagnostic and mCRPC biopsies following treatment. We profiled 470 treatment-naïve, prostate cancer diagnostic biopsies and for 61 cases, mCRPC biopsies using targeted and low-pass whole genome sequencing (*n* = 52). Descriptive statistics were used to summarize mutation and copy number profile. Prevalence was compared using Fisher's exact test. Survival correlations were studied using log-rank test. *TP53* (27%) and *PTEN* (12%) and DDR gene defects (*BRCA2* 7%; *CDK12* 5%; *ATM* 4%) were commonly detected. *TP53, BRCA2,* and *CDK12* mutations were significantly commoner than described in the TCGA cohort. Patients with *RB1* loss in the primary tumour had a worse prognosis. Among 61 men with matched hormone-naïve and mCRPC biopsies, differences were identified in *AR, TP53, RB1,* and PI3K/AKT mutational status between same-patient samples. In conclusion, the genomics of diagnostic prostatic biopsies acquired from men who develop mCRPC differs to that of the primary prostatic cancers. RB1/TP53/AR aberrations are enriched in later stages, but the prevalence of DDR defects in diagnostic samples is similar to mCRPC.

Find the latest version:

[https://jci.me/132031/pdf](https://jci.me/132031/pdf)
Title: **Genomics of lethal prostate cancer at diagnosis and castration-resistance**

**Authors:** Joaquin Mateo$^{1,2,3,*}$, George Seed$^{2,*}$, Claudia Bertan$^{2}$, Pasquale Rescigno$^{2,3}$, David Dolling$^{2}$, Ines Figueiredo$^{2}$, Susana Miranda$^{2}$, Daniel Nava Rodrigues$^{2}$, Bora Gurel$^{2}$, Matthew Clarke$^{2}$, Mark Atkin$^{2}$, Rob Chandler$^{2,3}$, Carlo Messina$^{2,3}$, Semini Sumanasuriya$^{2,3}$, Diletta Bianchini$^{2,3}$, Maialen Barrero$^{2,3}$, Antonella Petermolo$^{2,3}$, Zafeiris Zafeirou$^{2,3}$, Mariane Fontes$^{2,3,4}$, Raquel Perez-Lopez$^{1,2,3}$, Nina Tunariu$^{2,3}$, Ben Fulton$^{5}$, Robert Jones$^{5}$, Ursula McGovern$^{6}$, Christy Ralph$^{7}$, Mohini Varughese$^{8}$, Omi Parikh$^{9}$, Sunil Jain$^{10}$, Tony Elliott$^{11}$, Shahneen Sandhu$^{12}$, Nuria Porta$^{2}$, Emma Hall$^{2}$, Wei Yuan$^{2}$, Suzanne Carreira$^{2,*}$, Johann S. de Bono$^{2,3,*}$.

**Affiliations:**

1. Vall d’Hebron Institute of Oncology (VHIO) and Vall d’Hebron University Hospital, Barcelona, Spain
2. The Institute of Cancer Research, London, UK
3. The Royal Marsden NHS Foundation Trust, London, UK
4. Instituto Oncoclinicas - Grupo Oncoclinicas, Rio de Janeiro, Brazil
5. The Beatson West of Scotland Cancer Centre, Glasgow, UK
6. University College Hospital, London, UK
7. St James’s University Hospital, Leeds, UK
8. Musgrove Park Hospital, Taunton, UK
9. Royal Blackburn Hospital, Blackburn, UK
10. Belfast City Hospital, Belfast, UK
11. The Christie Hospital, Manchester, UK
12. Peter McCallum Cancer Center, Melbourne, Australia

*J. Mateo and G. Seed contributed equally to this work.*

*S. Carreira and J.S. de Bono are co-senior authors.*
CORRESPONDING AUTHORS:
Joaquin Mateo, MD PhD
Clinical Research Program, Vall Hebron Institute of Oncology
Medical Oncology Department, Vall Hebron University Hospital
Natzaret 115, 08035 Barcelona
Spain
Tel: +349 32543450, ext 8690
Email: jmateo@vhio.net

and

Professor Johann S. de Bono, MB ChB, MSc, FRCP, PhD, FMedSci
Regius Professor of Experimental Cancer Medicine
Division of Clinical Studies, The Institute of Cancer Research
Drug Development Unit, The Royal Marsden NHS Foundation Trust
Downs Rd, Sutton, Surrey SM2 5PT
United Kingdom
Telephone: +44 (0)2087224028
Fax: +44 (0)2086427979
Email: johann.de-bono@icr.ac.uk

Conflicts of Interest:
JM has served as a consultant for AstraZeneca, Roche, Janssen, Clovis and Amgen. TE
has received educational grants from Janssen. RJ reports fees/advisory role from
Astellas, AstraZeneca, BMS, Bayer, Exelixis, Janssen, Ipsen, Merck Serono, Novartis, Pfizer, Roche, Sanofi, EUSA. CR has received research grants from Oncolytics and Tusk Therapeutics, honoraria from BMS and support for travel from Roche, GSK, Viralytics, Janssen, Novartis, Pfizer, Ipsen. SJ reports fees/advisory role from Astellas, Janssen, Bayer, Boston Scientific, Almac Diagnostics. JDB has served as a consultant for Astellas, AstraZeneca, Bayer, Daiichi, Genentech, GSK, Janssen, Merck Serono, MSD, Orion, Pfizer Oncology, Sanofi-Aventis, Silicon Biosystems and Taiho. No relevant conflicts of interest were disclosed by other authors.

Presented in part at the 2018 ASCO Annual Meeting

Keywords: prostate cancer, genomics, targeted NGS, tumour evolution, castration-resistance.
Genomics of primary prostate cancer differs from that of metastatic castration-resistant prostate cancer (mCRPC). We studied genomic aberrations in primary prostate cancer biopsies from patients who developed mCRPC, also studying matching, same patient, diagnostic and mCRPC biopsies following treatment.

We profiled 470 treatment-naïve, prostate cancer diagnostic biopsies and for 61 cases, mCRPC biopsies using targeted and low-pass whole genome sequencing (n=52).

Descriptive statistics were used to summarize mutation and copy number profile. Prevalence was compared using Fisher’s exact test. Survival correlations were studied using log-rank test.

TP53 (27%) and PTEN (12%) and DDR gene defects (BRCA2 7%; CDK12 5%; ATM 4%) were commonly detected. TP53, BRCA2 and CDK12 mutations were significantly commoner than described in the TCGA cohort. Patients with RB1 loss in the primary tumour had a worse prognosis. Among 61 men with matched hormone-naïve and mCRPC biopsies, differences were identified in AR, TP53, RB1 and PI3K/AKT mutational status between same-patient samples.

In conclusion, the genomics of diagnostic prostatic biopsies acquired from men who develop mCRPC differs to that of the non-lethal primary prostatic cancers. RB1/TP53/AR aberrations are enriched in later stages, but the prevalence of DDR defects in diagnostic samples is similar to mCRPC.
INTRODUCTION

Inter-patient genomic heterogeneity in prostate cancer is well-recognized (1). However, molecular stratification of prostate cancer to guide treatment selection based on predictive genomic biomarkers remains an unmet clinical need. Recent genomic studies have elucidated this inter-patient heterogeneity, identifying multiple potentially actionable alterations which are now being evaluated in clinical trials. These studies have also described differences in the genomic landscape of the different clinical states of the disease (localized vs metastatic)(1, 2). Alterations in the AR gene (mutations, amplifications and structural variants) are increased the prevalence in mCRPC, and associated with the development of castration-resistance, as well as resistance to abiraterone acetate and enzalutamide (3, 4). Moreover, loss-of-function events in TP53, RB1, PTEN and DNA damage repair (DDR) genes are more common in mCRPC compared to non-metastatic, prostate cancer cohorts. It remains unclear whether these differences are the result of evolutionary processes in response to therapy exposure, or whether these reflect different disease sub-types with differing outcomes.

An ultimate aim of understanding the genomic landscape of cancer is the implementation of more precise therapeutic strategies, but metastatic biopsy acquisition is a key obstacle for implementing genomic stratification in clinical practice. Liquid biopsies can partially overcome this limitation, but these assays are not yet validated to replace tumour biopsy testing, at least for prostate cancer(5, 6). Understanding if primary tumour biopsies can be used for molecular stratification to guide the treatment of advanced mCRPC years later remains a key question.
This study aims to describe the genomic profile of primary tumour biopsies from lethal prostate cancers, either presenting as metastatic hormone treatment-naïve prostate cancers, or locoregional tumours that later evolve to metastatic disease; we hypothesized that these primary tumours would be enriched for alterations previously associated with mCRPC, and would be different to those primary prostate tumours that do not recur. Additionally, we assessed a cohort of same-patient, matched, treatment-naïve and mCRPC biopsies to determine if these genomic defects change during treatment with tumour evolution.
RESULTS

Patient and sample disposition

Between March 2015 and December 2017, 652 primary tumor samples from consenting patients were received; 87 cases (13%) were discarded due to either low DNA yield or excessive DNA degradation. Hence, targeted NGS was successfully performed on 565 prostate cancer diagnostic biopsies. Fifty-four cases were excluded due to either: 1) the biopsy not being collected prior to ADT; or 2) diagnosis being based on a metastatic biopsy (Supplementary Figures 1 and 2 in the Appendix). Next generation sequencing of 511 samples was analysed; of those, 41 (8%) cases did not meet quality control criteria for copy-number calling (7) and were discarded, so the final analysis evaluated 470 cases. Two cohorts were defined for the planned analyses based on disease extent at the time of original diagnosis: Cohort 1 was composed of 175 cases with locoregional prostate cancer at diagnosis (69.5% confined to the prostate, 30.5% with pelvic nodal extension); Cohort 2 included 292 primary tumours from patients with metastatic disease at diagnosis. The clinical records of 3 subjects were unobtainable (Table 1).

Genomic profile of lethal primary prostate tumours

Recurrent aberrations in genes and pathways related to lethal prostate cancer were identified, the commonest being mutations and homozygous loss of TP53, (27%) (Figure 1 and Appendix). Deleterious mutations and/or homozygous deletions in genes involved in DNA damage repair pathways were identified in 23% of primary tumours. BRCA2 was the DDR gene most commonly altered (7%). Alterations in mismatch repair genes were detected in 11/470 (2%) cases.
Activating mutations in PIK3CA and AKT1 were detected in 5%, with PTEN loss-of-function mutations or deep deletions in 12%. Deep deletions of RB1 were uncommon in the primary tumours (5%), although shallow deletions in RB1 were frequent. Genes in the WNT pathway (loss of APC or activating mutations in CTNNB1) were altered in 7% of cases (8, 9). SPOP mutations were identified in 7% cases (10, 11).

Surprisingly, low-allele frequency AR T878A or R630Q mutations (always with low MAF, ranging 0.06 to 0.18) were detected in 1% of treatment-naïve samples (12).

Our Cohort 1 of primary tumours, without detectable metastases at diagnosis, was enriched for alterations in TP53 (25 vs 8%; p<0.001), BRCA2 (8 vs 3%; p=0.015) and CDK12 (6 vs 2%; p=0.04) when compared with the TCGA series (Table 2). Conversely, SPOP mutations were less common in our population than in the better prognosis TCGA series (3% vs 11%; p=0.001). No relevant differences in prevalence of other mutations were observed when comparing Cohort 1 and Cohort 2. After adjusting for Gleason score, CDK12 mutations were enriched in Gleason 8 or higher cases (1/105 cases in Gleason 6-7 vs 21/353 in Gleason ≥8) (Appendix)

Clinical outcome based on primary tumour genomics.

Median time to ADT progression and start of first mCRPC therapy was 1.17 years (95%CI: 1.08-1.26 years) among the subset (n=210) of patients with clinical data available. Median overall survival from first evidence of metastatic disease was 4.28 years (95%CI: 3.72-4.84 years).
None of the gene alterations were associated with a significantly different time to ADT progression; patients with germline or somatic \textit{BRCA2} alterations had the lowest median time to ADT progression among the subgroups but the differences were not significant (median 0.92 years, 95\%CI 0.5-1.17, \(p=0.39\)). (Table 3)

Patients with \textit{RB1} alterations in the primary tumour had a significantly shorter overall survival (median OS from metastatic disease 2.32 years, 95\%CI 1.82-3.84; \(p=0.006\)). (Table 3 and Appendix)

\textit{Changes when assessing clinically actionable genomic alterations in patient-matched treatment-naïve and castration-resistant.}

We pursued NGS of mCRPC biopsies acquired from 61 patients participating in this study to further investigate if certain gene aberrations were detected more often in biopsies after progression on ADT and subsequent lines of therapy. Overall, we performed targeted NGS on 61 mCRPC biopsies (using the same panel as for the primary treatment-naïve samples) and copy-number profiles for both primary and mCRPC samples were compared using low-pass WGS in 52 cases with sufficient DNA in both samples. Copy number estimation was adjusted for ploidy, and tumour purity, since mCRPC biopsies overall had higher tumour content than the primary prostate biopsies (Appendix).

The median time between the two same-patient biopsies was 45.2 months (range 12 to 211 months). All mCRPC samples were obtained after progression on ADT, and in 50/61 (82\%) cases after progression on at least 2 further lines of therapy for mCRPC.
(80% after at least one taxane and 90% after abiraterone acetate and/or enzalutamide) (Table 4).

The commonest finding, when comparing same-patient primary treatment-naïve and mCRPC samples, was an increase in AR mutations and amplification. Other than AR, the main differences between the two same-patient biopsies were increased TP53, RB1 and PI3K/AKT pathway alterations in mCRPC (Figure 2 and Appendix) suggesting that these may emerge with treatment selection pressures.

In several cases, mutations in TP53 (n=4) and RB1 (n=4), detected in mCRPC samples, were not detected in the same patient’s, matched, treatment-naïve and diagnostic primary tumour biopsies. Overall, there was a decrease in copy-number for both TP53 and RB1 in mCRPC, even after adjusting for tumour purity based on low-pass WGS. More deep deletions in PTEN were also detected in the mCRPC cohort. Mutations in the WNT pathway genes CTNNB1 and APC, as well as MYC amplification, were also more common in mCRPC.

Conversely, aberrations in DNA damage repair pathway genes were relatively unchanged from diagnosis to mCRPC. Eleven truncating mutations in BRCA2, CDK12, ATM, MSH6 and PALB2 were identified in the mCRPC biopsies of 9/61 patients (one patient had both CDK12 and PALB2 mutations; one patient CDK12 and MSH6 mutations). Two patients had pathogenic germline BRCA2 mutations; in both of these cases, both the primary untreated tumour and the mCRPC biopsy presented loss of heterozygosity resulting in biallelic BRCA2 loss. The other 8 deleterious mutations (4 in CDK12, 2 BRCA2, 1 ATM, 1 PALB2, 1 MSH6) were only detected in somatic DNA; all
8/8 were also detected in the patient-matched, metachronous, diagnostic, treatment-naïve, biopsies. In 3 of 4 cases with CDK12 truncating mutations, there was a second missense mutation in CDK12; again, these second events were also detected in both the diagnostic patient-matched biopsies. However, 2 truncating mutations in ATRX and FANCM were detected only in the mCRPC samples.

With regards to copy number aberrations in DNA repair genes, we identified a trend for lower tumour suppressor gene copy number in mCRPC samples, only partially explained by the higher tumour purity of mCRPC biopsies. No deep deletions in BRCA1/BRCA2/ATM were identified, although changes indicating single copy loss with disease evolution to mCRPC were detected.

Generally, the number of private events was small. An outlier case was P001, a patient with a MMR-defective prostate cancer who had the highest mutation burden, including several shared mutations between primary and mCRPC (APC, CDK12, MSH6, ERBB4, PTEN and TP53), several private mutations only detected in mCRPC (including missense, non-truncating, mutations in APC, ATM, EZH2, JAK1) and several private mutations of the primary tumour not detected in the later mCRPC biopsy (CTNNB1, PRKDC, ERCC3 and ERRC6), suggesting the presence of different clones coming from a shared origin.
Molecular stratification of prostate cancer promises to impact patient care and deliver more precise treatments, but several challenges remain to be addressed including the elucidation of the genomic profiles of distinct clinical states and understanding the impact of drug resistance and tumour evolution (13, 14). Here, we show that the primary prostatic biopsies of patients who develop metastatic prostate cancer are enriched for genomic aberrations typically found in mCRPC, even prior to exposure to androgen deprivation. These data may help define a subset of patients with locoregional disease at diagnosis with higher risk of lethal disease; clinical trials should test if these patients may benefit from more intense therapeutic approaches. Furthermore, our data support the use of primary prostate biopsies to characterize metastatic hormone-naïve prostate cancers, which may facilitate the implementation of genomic testing into clinical practice.

Defects in some DNA damage repair genes have been identified as promising predictive biomarkers for PARP inhibitors or platinum chemotherapy (15-18). The prevalence of mutations and deletions in DNA repair genes in our cohorts of patients with only locoregional disease detected at diagnosis or metastatic, hormone-naïve prostate cancer was similar to what has been previously described for mCRPC. In a recent study, Marshall et al found an increased prevalence of these mutations in higher-Gleason score primary tumours, which also indirectly supports the association of these mutations with more aggressive primary tumours (19). These data in a cohort of 470 primary tumours suggest that lethal prostate cancer is enriched for DNA repair defects from diagnosis, prior to developing castration-resistance. However, the limited number of cases with DDR gene alterations in the cohort of matched primary-metastatic biopsies, including
only 4 cases with BRCA2 mutations, prevents us from making broad conclusions with regards to the genomic evolution of these tumours. Indeed, we and others have reported sub-clonal homozygous deletions of DDR genes (20, 21). Detecting these subclonal deletions is technically challenging with targeted NGS assays used for patient stratification in clinical practice or in clinical trials, particularly when studying primary tumour samples with low tumour content and degraded DNA.

Alterations in TP53 were common in diagnostic biopsies in this cohort. Moreover, several loss-of-function alterations of TP53, RB1 and PTEN were detected in mCRPC biopsies but not in patient-matched, treatment-naïve, primary tumours. Concurrent loss of RB1 and TP53 function has been postulated to drive a phenotypic change associated with resistance to endocrine therapies (22, 23); additionally, TP53 mutations have been associated with more aggressive disease (24-26), which may in part explain why we are observing TP53 mutations more often than expected in primary prostate cancer in this cohort of patients who all had lethal forms of the disease, even if many presented as localized tumours.

As precision medicine strategies are developed for prostate cancer patients, our findings become clinically-relevant. Firstly, our analyses indicate that RB1 loss in the primary tumour associates with poor prognosis; these data confirm recently published results from two independent studies looking at genomics-clinical outcome correlations in metastatic samples (27, 28). In our series, DDR defects and particularly BRCA2 mutations did not associate with shorter survival; however, most of these patients were enrolled into PARPi clinical trials; data from randomized trials has confirmed the improved outcome of patients with DDR defects receiving PARPi; this needs to be
taken into consideration when interpreting our results. Secondly, these data are critically important for designing precision medicine strategies: if DNA repair defects are already detectable in the primary tumour, there is a rational for testing synthetic lethal strategies with PARP inhibitors or platinum, in metastatic hormone-naïve prostate cancer, where the magnitude of benefit for patients could be larger. These data also support the use of diagnostic prostate cancer biopsies for the patient stratification based on DNA repair gene defects in trials of men with mCRPC, as the prevalence of these alterations in primary tumours from patients with lethal prostate cancer was similar to what has been reported for metastatic disease, and in the small number of same-patient sample pairs available, DDR mutational status was concordant (29). Conversely, trials investigating novel therapeutic approaches in the TP53/RB1-deficient phenotype should take into account that a proportion of genomic aberrations in TP53 and RB1 are not detected when assessing diagnostic treatment-naïve primary tumour specimens.

The main limitation of our study comes from having only one biopsy core available per time point and patient; we therefore could not assess spatial tumour heterogeneity. Primary prostate cancers can be multifocal, and previous studies have reported on inter-foci genomic heterogeneity(30, 31). We cannot rule out that in some cases the primary tumour sample may not represent the dominant tumour clone in the primary biopsy; hence, it is possible that some of the differences we observe in paired mCRPC biopsies may have not resulted from treatment-selective pressure but been in other areas of these primary tumours. However, genomic testing in clinical practice is largely based on the analyses of single biopsy cores. With the advent of novel imaging modalities, genomic stratification of prostate cancer could be improved by better identifying aggressive areas of prostate cancer in clinical diagnostic pathways (32, 33). Another key limitation is the
inability to pursue subclonality assessments using our clinically-oriented targeted sequencing assay. Hence, we cannot prove if some of the gene aberrations detected in the mCRPC biopsies, but not in the treatment-naïve samples, were already present at a subclonal level at the time of diagnosis. Regardless of whether these events emerge de novo or as a result of expansion of a subclone, the observed enrichment for certain alterations (such as TP53 or RB1) in the post-treatment resistance samples supports the clinical relevance of such alterations.

In conclusion, this study describes the genomic landscape of primary prostate tumours that will evolve to lethal prostate cancer across a cohort of 470 cases, with this being characterized by higher frequencies of TP53 and DNA repair gene aberrations. Significant differences in the detection of AR, TP53, RB1 and PTEN alterations, but not of DNA repair genes, was observed when comparing same patient mCRPC and treatment-naïve biopsies. These data are important for the genomic stratification of primary prostate cancer to identify higher risk cases, support the use of primary prostate tumour biopsies for molecular stratification of metastatic hormone-naïve prostate cancer and provide a rational for the study of DNA repair-targeting therapies, including PARP inhibitors, in earlier stages of the disease.
METHODS

Study design

This analysis included all consecutive patients consented between March 2015 and December 2017 for molecular characterization of prostate cancer biopsies at The Institute of Cancer Research (London, UK). These studies involved either prostate tumour samples and/or newly acquired metastatic biopsies. We report here on all patients for whom a treatment-naïve primary prostate tumour sample was successfully sequenced. Primary tumour samples were retrieved from referring hospitals; in most cases, only one sample was made available for the study; if more than one sample from the primary tumour was available, the highest Gleason lesion was used. Additionally, metastatic biopsies in castrate-resistant conditions were pursued in consenting patients.

Sample acquisition and processing

All prostate cancer treatment-naïve and metastatic biopsy samples were centrally reviewed by a pathologist (D.N.R). DNA was extracted from formalin-fixed and paraffin embedded (FFPE) tumour blocks (average, 6 sections of 10mic each per sample) using the FFPE Tissue DNA kit (Qiagen). DNA was quantified with the Quant-iT high-sensitivity PicoGreen double-stranded DNA Assay Kit (Invitrogen). The Illumina FFPE QC kit (WG-321-1001) was used for DNA quality control tests according to the manufacturer’s protocol as previously described (34). In brief, quantitative polymerase chain reaction (qPCR) was performed using 4ng of sample or control DNA and the average Cq (quantification cycle) was determined. The average Cq value for the control DNA was subtracted from the average Cq value of the samples to obtain a ΔCq. DNA samples with a ΔCq<4 were selected for sequencing; double amount of DNA was used for cases with ΔCq between 2-4.
Sequencing and bioinformatic analyses

Libraries for next-generation targeted sequencing were constructed using a customized panel (Generead DNaseq Mix-n-Match Panel v2; Qiagen) covering 6025 amplicons (398702 bp) across 113 genes used in (35) (Appendix). Libraries were run using the MiSeq Sequencer (Illumina). FASTQ files were generated using the Illumina MiSeq Reporter v2.5.1.3. Sequence alignment and mutation calling were performed using the Qiagen GeneRead Targeted Exon Enrichment Panel Data Analysis Portal (https://ngsdataanalysis.qiagen.com). Mutation calls were reviewed manually in IGV according to the standard operating procedure for somatic variant refinement of tumour sequencing data, following the principles described in (36). This manual review included assessing read strand quality, base quality, read balance and sequencing artefacts (high discrepancy regions, adjacent indels, multiple mismatches, start or end of amplicons. Mutation annotation was based on data from publically available databases (ClinVar, COSMIC, Human Genome Mutation Database, IARC TP53 Database), published literature and in silico prediction tools, and only deleterious mutations were included in the analysis.

Copy number variations (CNV) in prostatic biopsies were assessed using CNVkit (v0.3.5, https://github.com/etal/cnvkit(37)), which we previously validated in an independent cohort of prostate cancer samples(7). The read depths of tumour samples were normalized and individually compared to a reference consisting of non-matched male germline DNA; the circular binary segmentation (CBS) algorithm was used to infer copy number segments. Quality estimation of the CNV was based on distribution of bin-level copy ratios within segments. Cases were excluded from the analysis if any of the following criteria were met: IQR>1, total reads<500000, <99.9% of reads on
target, <95% paired reads or single reads>0. Manual review of copy number calls for selected oncogenes and tumour suppressors was pursued, accounting for tumour content. Oncoprints and heatmaps representing mutations and copy number calls were generated using R and cBioportal OncoPrinter (38-40).

Low-Pass Whole Genome Sequencing was performed on the mCRPC, and same patient, treatment-naïve, diagnostic, paired samples for copy-number profiling. Libraries where constructed using the NEBNext Ultra FS II DNA kit (NEB) according to the manufacturer’s protocol. Samples where pooled and run on the NextSeq (Illumina) at 0.5X mean coverage, using the 300 cycles High Output V2.5 kit. BCL files were converted to FASTQ files using BCL2FASTQ v2.17. Sequence alignments were performed using Burrows-Wheeler Aligner (BWA mem v0.7.12) to the hg19 human genome build. Copy number analysis was performed using IchorCNA(41). In short, hg19 genomes (filtered centromeres) were divided into 500kb non-overlapping bins, and the abundance of the mapped reads was counted by HMMcopy Suite in each bin and predicted segments of CNAs. GC and mappability bias were corrected by loess regression and based on a panel of germline DNA sequencing from healthy donors. The maximum CNA detection was set to 20 copies.

Raw sequencing data has been deposited at the European Nucleotide Archive with Accession number PRJEB32038. VCF files with mutation calls and CN values for the targeted sequencing data are available in the appendix.
Statistical considerations

Descriptive statistics were used to summarize patient, and sample, characteristics data as well as mutation frequency. The prevalence of mutations was compared between cohorts using Fisher’s exact test. The statistical analysis plan and the gene list to be analysed was designed prior to data collection. A Bonferroni correction was applied; p-values of <0.01 were considered statistically significant and all tests were two-sided unless otherwise specified.

Additionally, exploratory associations between the pre-selected list of gene alterations and patient outcomes were tested in a subset of the study population (n=210) with available consent for clinical data collection (all at The Royal Marsden). Clinical data was captured retrospectively from electronic patient records. Time to ADT progression was defined from the date of starting ADT to start of first mCRPC therapy. Overall survival was defined as time from the date of diagnosis, date of metastatic disease and the date of CRPC to the date of death or last follow up. To account for variability between patients who were diagnosed with de-novo metastatic vs localized disease, survival data is presented from the first evidence of metastatic disease. Patients alive at the time of last follow up were censored. Association of genomic aberrations with survival are presented using Kaplan-Meier curves and log-rank test. All calculations were performed using STATA v15.1(Stata Corp,TX).

Study Approval

The study included all patients with mCRPC who, between March 2015 and December 2017 provided written consent to participate in one of two IRB-approved molecular characterization programs for prostate cancer: 1) an internal molecular characterization
study at The Royal Marsden Hospital (London, UK) and/or 2) a tumour next-generation sequencing (NGS) pre-screening study at 17 hospitals (Appendix) for the TOPARP-B study, an investigator-initiated clinical trial of the PARP inhibitor olaparib in mCPRC (42) (TOPARP, CR-UK 11/029, NCT 01682772).
Author contributions

JM, SC, JSDB designed the study. JM, DD, NP, EH, JSDB created the study methodology. JM, PR, RC, CM, SS, DB, MB, AP, ZZ, MF, RPL, NT, BF, RJ, UM, CR, MV, OP, SJ, TE, SS consented patients, acquired samples and collected clinical data. JM, CB, IF, SM, DNR, BG, MA, SC processed samples and generated experimental data. GS, WY, SC planned and conducted bioinformatics analysis. DD, NP designed and conducted the statistical analysis plan. JM, GS, WY, SC, NP, DD, JSDB analysed and interpreted data. JM, GS, SC, JSDB wrote the manuscript. EH, JSDB obtained funding. SC and JSDB supervised the study. All authors reviewed and approved the manuscript. Order of joint first authors was determined based on their role in data interpretation and manuscript preparation.

Acknowledgements:

We would like to acknowledge funding support from the Prostate Cancer Foundation; Prostate Cancer UK; Movember; Cancer Research UK (Centre Programme grant); Experimental Cancer Medicine Centre grant funding from Cancer Research UK and the Department of Health; and Biomedical Research Centre funding to the Royal Marsden. TOPARP is an investigator-initiated study supported by Cancer Research UK and conducted with support from the Investigator-Sponsored Study Collaboration between AstraZeneca and the National Institute for Health Research Cancer Research Network. J. Mateo was supported by a Prostate Cancer Foundation Young Investigator Award and has acknowledges research funding from Fundacio LaCaixa, FERO Foundation, Cellex Foundation and Instituto de Salud Carlos III. G. Seed was supported by a Prostate Cancer UK PhD Studentship. We acknowledge contribution to patient recruitment from all investigators involved in the TOPARP trial and the staff at the ICR.
Clinical Trials and Statistics Unit; full list of clinical investigators involved is presented in the Appendix.
Figure 1. Oncoprint of genomic aberrations (non-sense, indels, splice site mutations, relevant missense mutations and copy number changes) for 470 untreated primary prostate cancer biopsies from patients who later developed metastatic castration-resistant disease.
Figure 2. Differences in genomic profiles between same patient, matched, primary untreated and mCRPC biopsies. A) Mutation calls in genes of interest for the mCRPC biopsies which were not present in the treatment-naïve primary tumour for the same patient (61 pairs, full gene set in Suppl Fig 6); B) Overall copy number profiles based on low-pass WGS (52 pairs); C) amplifications and deep deletions detected in the mCRPC biopsies and not present in the treatment-naïve primary tumours for the same patient (based on low-pass WGS, after adjusting for tumour purity and ploidy, and validated by SNP data from targeted panel sequencing).
Table 1. Population characteristics and sample disposition for the overall study population (n=470)

<table>
<thead>
<tr>
<th>Metastatic disease at original diagnosis of prostate cancer</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No (Cohort 1)</td>
<td>175</td>
<td>37.5%</td>
</tr>
<tr>
<td>Yes (Cohort 2)</td>
<td>292</td>
<td>62.5%</td>
</tr>
<tr>
<td>Not recorded</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gleason score primary tumour (overall population)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;7</td>
<td>15</td>
<td>3.3%</td>
</tr>
<tr>
<td>7</td>
<td>90</td>
<td>19.7%</td>
</tr>
<tr>
<td>8</td>
<td>85</td>
<td>18.6%</td>
</tr>
<tr>
<td>9</td>
<td>245</td>
<td>53.5%</td>
</tr>
<tr>
<td>10</td>
<td>23</td>
<td>5.0%</td>
</tr>
<tr>
<td>Gleason not recorded</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian</td>
<td>431</td>
<td>96.9%</td>
</tr>
<tr>
<td>African or african-american</td>
<td>7</td>
<td>1.6%</td>
</tr>
<tr>
<td>asian</td>
<td>4</td>
<td>0.9%</td>
</tr>
<tr>
<td>Caribbean</td>
<td>4</td>
<td>0.9%</td>
</tr>
<tr>
<td>Not recorded</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Staging of patients in Cohort 1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>6</td>
<td>3.7%</td>
</tr>
<tr>
<td>T2</td>
<td>20</td>
<td>12.2%</td>
</tr>
<tr>
<td>T3</td>
<td>131</td>
<td>79.9%</td>
</tr>
<tr>
<td>T4</td>
<td>7</td>
<td>4.3%</td>
</tr>
<tr>
<td>N0</td>
<td>114</td>
<td>69.5%</td>
</tr>
<tr>
<td>N1</td>
<td>50</td>
<td>30.5%</td>
</tr>
<tr>
<td>T-N not recorded</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gleason score in Cohort 1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;7</td>
<td>11</td>
<td>6.5%</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>29.6%</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
<td>16.6%</td>
</tr>
<tr>
<td>9</td>
<td>76</td>
<td>45.0%</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>2.4%</td>
</tr>
<tr>
<td>Not recorded</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Comparison of cohort 1 in this study (patients with primary, non-metastatic at diagnosis, prostate cancer) and the TCGA series for primary prostate cancers (distribution of genomic events per Gleason score group are available in the Appendix).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Events considered</th>
<th>TCGA (N=333)</th>
<th>Cohort 1 (N=175)</th>
<th>p-value (Fisher exact test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td>AKT1</td>
<td>Activating mutations</td>
<td>3 (0.9%)</td>
<td>0 (0%)</td>
<td>0.56</td>
</tr>
<tr>
<td>ATM</td>
<td>Loss-of-function mutations and deep deletions</td>
<td>20 (6%)</td>
<td>10 (6%)</td>
<td>1.00</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Loss-of-function mutations and deep deletions</td>
<td>3 (1%)</td>
<td>3 (2%)</td>
<td>0.42</td>
</tr>
<tr>
<td>BRCA2</td>
<td>Loss-of-function mutations and deep deletions</td>
<td>10 (3%)</td>
<td>14 (8%)</td>
<td>0.015</td>
</tr>
<tr>
<td>CDK12</td>
<td>Loss-of-function mutations and deep deletions</td>
<td>7 (2%)</td>
<td>10 (6%)</td>
<td>0.04</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>Activating mutations</td>
<td>7 (2%)</td>
<td>3 (2%)</td>
<td>1.00</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>Activating mutations and copy number gains</td>
<td>7 (2%)</td>
<td>7 (4%)</td>
<td>0.26</td>
</tr>
<tr>
<td>PTEN</td>
<td>Loss-of-function mutations and deep deletions</td>
<td>57 (17%)</td>
<td>20 (11%)</td>
<td>0.09</td>
</tr>
<tr>
<td>RB1</td>
<td>Loss-of-function mutations and deep deletions</td>
<td>3 (1%)</td>
<td>6 (3%)</td>
<td>0.07</td>
</tr>
<tr>
<td>SPOP</td>
<td>Hotspot mutations</td>
<td>37 (11%)</td>
<td>5 (3%)</td>
<td>0.001</td>
</tr>
<tr>
<td>TP53</td>
<td>Loss-of-function mutations and deep deletions</td>
<td>27 (8%)</td>
<td>44 (25%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Table 3. Association of gene defects with clinical outcome. Long-rank p-values are presented unadjusted and adjusted for both Gleason score (≤7 vs ≥8) and presence/absence of metastatic disease at initial diagnosis.

<table>
<thead>
<tr>
<th>Gene alteration</th>
<th>Time to ADT progression</th>
<th>Overall Survival (from metastatic disease)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Median (Years)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall population</td>
<td>202</td>
<td>1.17 (95%CI: 1.08-1.27)</td>
</tr>
<tr>
<td><strong>TP53</strong></td>
<td>47</td>
<td>1.19 (95%CI: 1.00-1.67)</td>
</tr>
<tr>
<td><strong>PTEN</strong></td>
<td>23</td>
<td>1.58 (95%CI: 0.83-2.15)</td>
</tr>
<tr>
<td><strong>RB1</strong></td>
<td>13</td>
<td>1.17 (95%CI: 0.56-2.33)</td>
</tr>
<tr>
<td><strong>SPOP</strong></td>
<td>9</td>
<td>1.25 (95%CI: 0.50-2.23)</td>
</tr>
<tr>
<td><strong>BRCA2</strong></td>
<td>15</td>
<td>0.92 (95%CI: 0.50-1.17)</td>
</tr>
<tr>
<td><strong>CDK12</strong></td>
<td>12</td>
<td>1.20 (95%CI: 0.58-2.82)</td>
</tr>
<tr>
<td><strong>ATM</strong></td>
<td>11</td>
<td>1.07 (95%CI: 0.42-2.33)</td>
</tr>
<tr>
<td><strong>PIK3CA</strong></td>
<td>7</td>
<td>1.62 (95%CI: 0.58-2.41)</td>
</tr>
<tr>
<td><strong>CTNNB1</strong></td>
<td>7</td>
<td>1.42 (95%CI: 0.50-2.00)</td>
</tr>
<tr>
<td><strong>AKT1</strong></td>
<td>2</td>
<td>1.58 (95%CI: NA)</td>
</tr>
<tr>
<td><strong>BRCA1</strong></td>
<td>3</td>
<td>1.08 (95%CI: 0.42-NA)</td>
</tr>
<tr>
<td><strong>BRCA1/2 / ATM</strong></td>
<td>28</td>
<td>1.07 (95%CI: 0.83-1.21)</td>
</tr>
<tr>
<td><strong>PIK3CA/ AKT1/PTEN</strong></td>
<td>32</td>
<td>1.59 (95%CI: 1.00-2.15)</td>
</tr>
</tbody>
</table>
Table 4. Sample disposition for the patient-matched primary untreated and mCRPC biopsies (n=61 cases with paired samples). Median time between the two same-patient samples were taken was 45.2 months (range: 12 to 211 months)

| Location Hormone-Naive Sample | Prostate | 61 | 100%
|------------------------------|---------|----|----|
| Location CRPC Sample         | Bone    | 24 | 39.4%
|                              | Lymph Node | 22 | 36.17%
|                              | Liver   | 4  | 6.6%
|                              | Other   | 11 | 18.0%
| Metastatic status at original diagnosis | M0 | 25 | 41.7%
|                              | M1      | 35 | 58.3%
| Treatments received between the two samples acquisition | Prostatectomy | 10 | 16.4%
|                              | Pelvic radiotherapy | 27 | 44.3%
|                              | Androgen deprivation therapy | 61 | 100%
|                              | First gen antiandrogen | 41 | 67.2%
|                              | Abiraterone acetate | 34 | 55.7%
|                              | Enzalutamide | 33 | 54.1%
|                              | Abiraterone and/or enzalutamide | 55 | 90.2%
|                              | Docetaxel | 49 | 80.3%
|                              | Cabazitaxel | 20 | 32.8%
|                              | Radium-223 | 4  | 6.5%
|                              | Investigational agents | 14 | 22.9%
| Lines of therapy for CRPC before mCRPC biopsy | 0 | 2  | 3.2%
|                              | 1        | 9  | 14.7%
|                              | 2        | 21 | 34.4%
|                              | 3 or more | 29 | 47.5%
REFERENCES


