

Title: RB1 Heterogeneity in Advanced Metastatic Castration Resistant Prostate Cancer

Running Title: RB1 heterogeneity in mCRPC

Authors: Daniel Nava Rodrigues^{1,2,*}, Nicola Casiraghi^{3,*}, Alessandro Romanel⁴, Mateus Crespo², Susana Miranda², Pasquale Rescigno^{1,2}, Ines Figueiredo², Ruth Riisnaes², Suzanne Carreira², Semini Sumanasuriya^{1,2}, Paola Gasperini³, Adam Sharp^{1,2}, Joaquin Mateo⁶, Alan Makay², Christopher McNair⁵, Matthew Schiewer⁵, Karen Knudsen⁵, Gunther Boysen², Francesca Demichelis^{3,5,§} and Johann S. de Bono^{1,2,§}

Keywords: Rb1, prostate cancer, castration-resistant, heterogeneity,

Affiliations:

(1) The Royal Marsden NHS Foundation Trust, Downs Road, London, UK

(2) The Institute of Cancer Research, 15 Cotswold Road, London, UK

(3) Centre for Integrative Biology (CIBIO), University of Trento, Laboratory of Functional and Computational Oncology, Via Sommarive 9, Trento 38123, Italy

(4) Centre for Integrative Biology (CIBIO), University of Trento, Laboratory of Bioinformatics and Computational Genomics, Via Sommarive 9, Trento 38123, Italy

(5) Caryl and Israel Englander Institute for Precision Medicine, New York Presbyterian Hospital–Weill Cornell Medicine, 413 East 69th Street, New York, NY, 10021, USA

(6) Vall Hebron Institute of Oncology (VHIO), Barcelona

* These authors contributed equally to this work.

§ These authors jointly directed this work and are corresponding authors.

Words: 5145

Corresponding Authors:

Johann de Bono

Prostate Cancer Targeted Therapy Group

The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust

15 Cotswold Road, London SM2 5NG, United Kingdom

Tel: +44 2087224029

E-mail: Johann.de-Bono@icr.ac.uk

Francesca Demichelis

Centre for Integrative Biology (CIBIO),

University of Trento,

Laboratory of Functional and Computational Oncology,

Via Sommarive 9, Trento 38123, Italy

E-mail: f.demichelis@unitn.it

Conflicts of Interest: Dr. de Bono reports personal fees from Astellas Pharma, personal fees from Astra Zeneca, personal fees from Bayer, personal fees from GenMab, personal fees from Genentech, personal fees from GlaxoSmithKline, personal fees from Janssen, personal fees from Medivation, personal fees from Orion Pharma, personal fees from Pfizer, personal fees from Sanofi, during the conduct of the study. Other authors have no conflicts of interest do declare.

Statement of Translational Relevance:

Intra-patient molecular heterogeneity in advanced metastatic castration resistant has been underexplored. Using whole-genome sequencing of 21 tumors from 10 patients, we show, in agreement with previous studies, an overall limited genomic heterogeneity in putative cancer drivers. However, private aberrations in putative drivers were identified. We also uncovered intra-patient heterogeneity in aberrations involving the *RB1* gene and proceeded to study this in an independent cohort of metastatic castration-resistant prostate cancers (mCRPCs). We show *RB1* copy number losses are common (~56%) in tumors with previous exposure to taxanes and anti-androgens. We also identify genomic rearrangements as a common putative second hit. Finally, RB1 protein expression was heterogeneous in ~28% of tumors, which has implications for ongoing trials of CDK4/6 inhibitors, as absence of functional RB1 is a negative predictive biomarker of response to these agents.

ABSTRACT (229 words)

Purpose: Metastatic castration resistant prostate cancer (mCRPC) is a lethal but clinically heterogeneous disease, with patients having variable benefit from endocrine and cytotoxic treatments. Intra-patient genomic heterogeneity could be a contributing factor to this clinical heterogeneity. Here we used whole-genome sequencing (WGS) to investigate genomic heterogeneity in 21 previously treated CRPC metastases from 10 patients to investigate intra-patient molecular heterogeneity (IPMH).

Experimental Design: WGS was performed on topographically separate metastases from patients with advanced metastatic prostate cancer (PCa). IPMH of the *RB1* gene was identified and further evaluated by fluorescent *in situ* (FISH) and immunohistochemistry (IHC) assays.

Results: WGS identified limited IPMH for putative driver events. However, heterogeneous genomic aberrations of *RB1* were detected. We confirmed the presence of these *RB1* somatic copy number aberrations (SCNA), initially identified by WG, with FISH, and identified novel structural variants (SV) involving *RB1* in 6 samples from three of these ten patients (30%; 3/10). WGS uncovered a novel deleterious *RB1* structural lesion constituted of an intra-genic tandem duplication involving multiple exons and associating with protein loss. Using *RB1* IHC in a large series of mCRPC biopsies, we identified heterogeneous expression in ~28% of mCRPCs.

Conclusion: mCRPCs have a high prevalence of *RB1* genomic aberrations, with structural variants, including rearrangements, being common. Intra-patient genomic and expression heterogeneity favor *RB1* aberrations as late, sub-clonal events that increase in prevalence due to treatment selective pressures.

Introduction

Prostate cancer (PCa) is among the most frequently diagnosed malignancies worldwide (1). However, only about a quarter of all new cases will lead to cancer-specific death, highlighting the clinical heterogeneity of these tumors (1-3). Several large scale genomic studies of cohorts of primary, hormone-naïve and metastatic, castration resistant prostate cancers (mCRPC) have been performed suggesting both intra- and inter-patient molecular heterogeneity (4-11). Although multi-focality and genomic heterogeneity are well-known features of primary PCa (12-14), the level of intra patient molecular heterogeneity (IPMH) in mCRPCs is less well established. Samples of late stage disease are rarely acquired, with autopsy studies being expensive and logistically difficult, and few reports on IPMH in mCRPCs are available.

Two studies have investigated the extent of IPMH in mCRPC patients. In a comprehensive molecular study using array comparative genomic hybridisation (aCGH), whole exome sequencing, and transcriptional profiling, Kumar *et al* reported that most presumptive driving genomic events and actionable targets are either present as common roots or “result from convergent evolution conferred by therapeutic pressures” (15). Further supporting this phenomenon of divergent genomics with convergent phenotypes, a WGS study of multiple metastases from 10 autopsied patients identified multiple separate endocrine treatment resistance mechanisms arising independently and subclonally within individual patients (16). Crucially, although both studies report some heterogeneity for known oncogenic drivers, neither followed up in dissecting gene-specific intra-patient heterogeneity using independent cohorts. Such efforts could shed light on the hierarchy of genomic events in PCa progression.

The goal of this work was to identify recurrent intra-patient molecular diversity. We conducted WGS at >100X median coverage of 21 mCRPCs from 10 consecutive patients with >1 biopsiable non-bone metastases between 01/01/2013 to 31/12/2014. Our analyses, focused on 1,392 cancer-related genes, revealed intra-patient heterogeneity for the Retinoblastoma-Associated Protein (RB1), a tumor suppressor that inhibits pro-oncogenic E2F1 mediated transcription in its hypophosphorylated state (17). In a study of 500

metastatic samples of 20 different tumor types, including 93 mCRPCs, RB1 was shown to be genomically aberrant in 13.6% of the cases, this being among the most common deleteriously aberrant genes in metastatic disease overall (18). Critically, aberrations of RB1 are commoner in mCRPC, indicating that these are either present at onset in clinically aggressive primary tumors or are acquired late in the natural history of prostatic malignancies.

We then employed fluorescent *in situ* hybridisation (FISH) and immunohistochemistry (IHC) to further study RB1 loss and show that *RB1* genomic aberrations are common in clinically aggressive primary PCa, becoming commoner in mCRPC regardless of histology. We demonstrate that structural variants involving *RB1* are frequent in mCRPCs and may be a common putative second hit. Finally, we show that heterogeneous RB1 protein loss is common, favoring RB1 depletion as a late event in the natural history of PCa.

Methods

Population and Tissue Samples

Patients gave their written informed consent and were enrolled to protocols approved by the Royal Marsden NHS Foundation Trust Hospital (London, UK) ethics review committee (Reference no. 04/Q0801/60). Demographics and clinical data were retrospectively collected from electronic hospital records. WGS data was generated from 21 samples from 10 consecutive patients with >1 accessible non-bone metastases acquired between 01/01/2013 and 31/12/2014. Additionally, FISH and IHC data were generated from randomly selected metastatic samples collected between 01/01/2013 and 31/12/2016. In this second cohort, metastatic sites included lymph node, bone, viscera, and soft tissues. Minimum tissue quality criteria required the presence of at least 50 well-preserved cancer cells. Patient matched hormone-sensitive tumors for a subset of cases were studied.

Whole Genome Sequencing

For each of the 21 metastases, tumor DNA was extracted from seven 10 μ m thick frozen sections after confirmation of tumor cell presence on a 4 μ m thick H&E stained slide. Germline DNA was extracted from saliva. The QIAmp[®] DNA Mini Kit (#51304; Qiagen, Hilden, Germany) was used for DNA isolation from both fresh/frozen sections and buccal swabs. Library preparation was performed using the TruSeq[®] Nano DNA Library Prep kit. Sequencing was performed using the Illumina HiSeq[®] 2000 System. Tumor and matched germline DNA samples were profiled with a WGS protocol. Paired-end sequencing reads were aligned to the human reference genome (GRCh37/hg19) using Illumina Isaac Genome Alignment Software (19) embedded in Illumina pipeline version v2.0.1. Genome Analysis Toolkit (GATK v2.6) (20) best practices for variant calling that included marking of duplicate reads, recalibration of base quality scores and local realignment were adopted for all aligned samples. Germline-tumor pair consistency was tested using the SNP panel identification assay (SPIA v1.1.0) (21). To identify and characterize somatic single-nucleotide variants (SNVs) in exons, the MuTect v1.1.6 (22) tool was used. To reduce false positives, calls were refined and confirmed using an ad-hoc pileup approach using ASEQ v.1.1.8 (23) and adopting stringent filtering quality criteria. Finally, each retained SNV was annotated with genomic features and effect/impact prediction using the SnpEff v4.1a tool (24). For each tumor and matched normal WGS sample, the algorithm BreakDancer v1.2 (25) was run with default parameters to detect genomic SVs including intra- and inter-chromosomal translocations. The raw output of BreakDancer was filtered using stringent criteria based on quality score and minimum number of reads supporting the SV. Additionally, an ad-hoc WES based computational strategy was developed to support evidence of significant enrichment in the coverage of exons involved in tandem duplications detected in the *RB1* gene. Briefly, the mean coverage in each *RB1* exon was computed separately in tumor and normal samples and normalized to the total coverage of all *RB1* exons (N=27). Then, the ratio between tumor and matched normal normalized mean coverages was computed for each *RB1* exon. Finally, the method computed the median ratio grouping exons 1 to 6 (R_{upstream}), 7 to 17 ($R_{\text{tandem duplication}}$), and 8 to 27 ($R_{\text{downstream}}$). Tandem duplication events were compatible with high ratios $R_{\text{tandem duplication}} / R_{\text{upstream}}$ and $R_{\text{tandem duplication}} / R_{\text{downstream}}$. Estimation of genomic segments and SCNA was obtained using BICseq v1.1.11 algorithm (26). Segmented

data were used in combination with CLONET v2.0.0 (27) to estimate ploidy and purity for each tumor sample, and further to determine allele-specific copy number.

Cell lines

22Rv1 (ATCC®CRL-2505), a PCa- and MD-MBA-468 (ATCC®HTB-132), a triple-negative breast cancer- cell lines were purchased from ATCC and cultured according to the manufacturer's protocol. Cell pellets were fixed in 10% neutral buffered formalin overnight at 4oC. After dehydration in ethanol, at 70% (10 minutes), 80% (10 minutes), 90% (10 minutes), and 100% (3x10 minutes), lipid removal and diaphonisation were performed with xylene (3x15 minutes).

Antibody Validation, Immunohistochemistry, and Scoring Methods

Antibody target specificity validation was confirmed by performing Western Blots of cell lysates as previously described (28). For immunohistochemistry (IHC), FFPE samples were cut at 4µm thick sections onto glass slides and dewaxed with xylene. Antigen unmasking was performed by heating slides in a pH6 citrate buffer solution for 18-minutes using a microwave at 900W. Staining was performed with conventional diaminobenzidine method using the i6000 autostainer (Biogenex; Fremont – CA, USA). The Dako-Envision kit (Agilent-Dako, Santa Clara – CA, USA) was used for reaction visualisation. A mouse monoclonal anti-RB1 antibody targeting amino acids 332-344 (Clone G3-245, BD Biosciences, Franklin Lakes – NJ, USA) diluted 1:50 was used. Endogenous peroxide was blocked using a 3% H₂O₂ solution; non-specific staining was blocked using Dako protein block serum-free X0909, and the diluted primary antibody was incubated on tissue samples for 1 hour. We used MDA-MB-468, a triple-negative breast cancer cell line with biallelic deletion of the RB1 locus (c.265_2787del2523) as negative control, and 22RV1, an AR positive PCa cell line with genomically intact RB1 as a positive control. Absence of IHC AR staining (Clone AR441, Agilent-Dako, Santa Clara – CA, USA) was used as a surrogate marker of neuroendocrine prostate cancer (NEPC). This validated anti-AR antibody assay was performed as previously described (28). An H-Score determined by the formula: (% of negative)x0 + (% of weak positivity)x1 + (% of moderate positivity)x2 + (% of strong positivity)x3, yielding a result

between 0 and 300, was used for RB1 protein expression semi-quantitative analyses. Samples with any proportion of tumor cells lacking RB1 staining with good internal controls, i.e. RB1 positive endothelial or stromal cells, were interpreted as heterogeneous, and positivity was considered of weak intensity to reflect an H-Score of <100. AR was dichotomously scored as either positive or negative, with the cut-off for positivity being expression >0.

Fluorescent *in situ* Hybridisation (FISH)

Fluorescent *in situ* hybridisation (FISH) was performed as previously described (29). We optimised a dual colour FISH assay using two commercially available FISH probes: 1) a ~202 Kb probe directly labelled with Spectrum Orange targeting the 13q14 locus spanning RB1, and 2) a ~612 Kb probe directly labelled with Spectrum Green targeting the subtelomeric 13q34 locus (Abbot Laboratories, Lake Bluff, IL-USA). A sub-telomeric control probe was selected because homologies in centromeric alpha satellites between chromosomes 13 and 17 could compromise gene copy number enumeration (30). We defined copy number gains (CNG) as a ratio of RB1 signals/nuclei ratio of >2; copy number neutral (CNN) as a ratio of RB1 signals/nuclei ratio of <2 and ≥ 1.7 ; shallow deletions (SD), i.e. mostly heterozygously deleted, as a ratio of RB1 signals/nuclei <1.7 and ≥ 1 , and deep deletions (DD), i.e. at least focally homozygously deleted, as a ratio of <1. At least 50 intact non-overlapping nuclei were counted per sample and the number of cells with >2, 2, 1, or 0 signals was recorded for both probes. The controls used for RB1 IHC were the same as those used for FISH, namely MDA-MB-468 as negative and 22RV1 as positive controls respectively.

Figures and Statistical Analysis

Figures were generated using R version 3.3.2 and GraphPad Prism v7. Statistical tests were performed using GraphPad Prism v7 and STATA v15.

Ethics Statement

The work herein presented was conducted to Good Clinical Practice standards and in accordance to the Declaration of Helsinki.

Results

Clinical features of the patients selected for WGS

Patients in this study had clinically aggressive disease, the majority being stage IV at diagnosis; their clinical data are summarised in **Table 1**. Median age of the WGS cohort was 58.8 years (IQR: 50.84-61.4). Gleason scores were available for 9 patients, 4 having scores of ≤ 7 and 5 of >7 . Two patients had local treatment with radical prostatectomy followed by salvage radiotherapy. Another two patients had radical radiotherapy. Eight of the ten patients had locally advanced disease, i.e. $\geq T3$, with six having M1 disease, at diagnosis. Median time on androgen deprivation therapy was 14.99 months (IQR: 6.95-49.15 months). A summary of all patient treatments and best clinical responses are available in **Supplementary Figure 1**. Twenty-one (21) mCRPC biopsies were collected from these 10 patients; 2 samples per patient except one subject from whom we had three separate nodules from a penectomy specimen. In 7 patients, these mCRPC samples came from enlarged, topographically distant, lymph nodes; in two men, liver mCRPC biopsies and lymph-node biopsies were available; in one subject liver and muscle mCRPC biopsies were acquired. No complications occurred during tumor sampling. All biopsies included a fresh frozen sample for genomic sequencing studies and a formalin-fixed, paraffin embedded, sample to confirm tumor histology and for IHC studies. Fresh-frozen samples were cut at $4\mu\text{m}$ and rapidly stained with H&E to confirm tumor content. Nineteen of 21 samples had $>50\%$ tumor content, with the remaining two cases having $\sim 10\%$ and $\sim 25\%$.

Bulk-tissue WGS shows limited Intra-patient genomic heterogeneity

We obtained WGS data from these 21 tumor samples from 10 patients; 21 tumor and 10 matched germline DNA samples were profiled at a median sequencing coverage depth of 109X and 37X, respectively. Genomic profiles of tumor samples were defined from WGS data by applying an *ad hoc* computational workflow. CLONET estimated tumor ploidy, purity and corrected SCNA in all but 2 tumor samples that showed low quality data. Overall, tumor cellularity was high (median=74%, min=38%, max=94%) and polyploidy (genomes with more than 2 paired sets of chromosomes) was detected in all tumor samples.

Our initial approach focused on genes of taxonomical relevance, i.e. clonal, mutually exclusive events that define genomic subtypes (31) and previously reported by the SU2C

International team CRPC study (32). We identified ETS rearrangements in 13 samples from 6 patients; in 11 of these samples from 5 patients, *ERG* was the ETS partner involved. WGS segmented data indicated the presence of an interstitial deletion spanning approximately 3Mb between *ERG* and *TMPRSS2* in 7 samples from 3 patients (V4074, V5162 and V4038). Breakpoints of this deletion were consistent among samples within the same patient. In 2 samples from one patient (V4002), we identified a previously unreported fusion between *ETV1* and *MIOS*. ETS rearrangement status were identical in different samples from an individual patient. The remaining 4 ETS rearrangement negative cases did not have clonal somatic mutations in *SPOP*, *IDH1*, or *FOXA1*, subgroup-defining genomic lesions described by the TCGA consortium (4).

The landscape of identified coding variants of interest through this WGS are summarised in **Figure 1**. First, we quantified the total number of somatic single nucleotide variants (SNVs) in coding regions for each tumor sample and found a median number of 173 (min=127, max=494) variants. When the analysis was stratified by functional annotation, the median number of non-synonymous SNVs was 43 (min=27, max=127). The two tumors harbouring the highest number of SNVs, with 494 and 456 bases mutated (127/494 and 123/456 being non-synonymous SNVs), were from the same patient and presented a deleterious mismatch repair gene aberration (i.e. missense variant p.Ile185Val in *MLH3*). Second, we estimated for each sample the fraction of the genome that was impacted by a SCNA, here defined as copy number altered fraction (CNAF). The median CNAF across tumor samples was 0.83 (min=0.35, max=0.94). These fractions were significantly higher than those reported by Hieronymus *et al.* in aCGH and low-pass WGS studies in hormone sensitive primary and metastatic prostate cancers, with CNAF of 0.04-0.05 and 0.32 respectively (33). Our data indicate that CNAF increases significantly in later stage mCRPC, suggesting that increased aneuploidy favours cell fitness and treatment resistance. To quantify the somatic CN similarities among metastatic sites from the same patient, we measured the correlation between log₂ ratios in a set of 1,392 cancer related genes. Results based on log₂ ratios of these genes indicated overall strong correlations with a median of 0.85 (min=0.53, max=0.95). We also focused on a subset of genes (N=32), reported in **Figure 1** and selected based on their relevance to prostate carcinogenesis and key roles in important molecular pathways. The fraction of these genes with concordant copy number status across metastases from the same patient was on average 0.93 (median=0.95, min=0.75,

max=1). Overall, our somatic SNVs and CN analyses to assess intra-patient heterogeneity, confirmed low level diversity at a bulk tumor level as previously reported (15).

However, genomic heterogeneity was identified in some patients. Aberrations in the WNT signalling pathway were seen as private events in two metastases from two distinct patients; the first showing a missense point mutation causing the protein change Ser680Arg in *CTNNB1*, with the latter showing a missense variant in *APC* (protein change Gly108Val). Amplification of *AURKA*, a serine-threonine kinase with key roles in centrosome maturation and spindle assembly (34), was seen in one of two samples from a single patient (V4038). Finally, in two patients (20%; 2/10), heterogeneous somatic genomic aberrations of *RB1* were identified between different metastases (**Figure 1**). In the first patient (V5128), one metastatic site had copy number neutral loss-of-heterozygosity with the other site having heterozygous loss. A second patient (V5033) had heterozygous loss in both metastases, but an additional single base substitution causing premature protein truncation at Glu204 in only one of them. Importantly, CLONET based tumor purity estimations of these samples were high (94%, 74% and 85%, 84%).

***RB1* Copy Number Enumeration by FISH and Protein Expression in mCRPC**

Given the observed intra-patient heterogeneity of *RB1* loss in mCRPC, and the fact that these aberrations were more common than anticipated in mCRPC samples, we decided to further investigate *RB1* loss by orthogonal assays pursuing *RB1* FISH and *RB1* IHC. We hypothesised that the presence of *RB1* deletions in mCRPC from patients with matched, previously copy number normal, HSPC samples, as well as the presence of heterogeneous IHC staining in mCRPC biopsies would both favour disruptions of *RB1* as late, sub-clonal, molecular events. These would have therapeutic relevance since *RB1* loss-of-function is a putative predictive biomarker of resistance to CDK4/6 inhibitors (35), which have anti-tumor activity in ER+/HER2- breast cancer (36, 37), and are being evaluated in clinical trials of mCRPC (NCT02905318). The validation of *RB1* FISH and IHC assays is summarised in **Supplementary Figure 2**. The MDA-MB-468 cell line had complete absence of IHC staining with rare cells having a single *RB1* FISH signal, resulting in an *RB1* signal to nuclei ratio (*RB1*s/Nuc) of 0.43 in 100 nuclei counted. The 22RV1 cell line had diffuse *RB1* positivity on IHC and an *RB1*s/Nuc ratio of 2.03 in 100 nuclei counted. To validate copy number analyses

determined by sequencing and to query RB1 heterogeneity at the protein level, we analysed the 21 samples from 10 patients in the WGS cohort using RB1 IHC and FISH. FISH results were successfully acquired from 20 of 21 samples, with 19 of these also having RB1 copy number determined by WGS; FISH copy number analyses were consistent with WGS estimates in 16 samples (84%). All seven cases identified as hemizygotously deleted for *RB1* by WGS also had *RB1* signal/Nucleus ratios of <1.7 by FISH enumeration; which validated our cut-off to define *RB1* loss. Interestingly, one case was considered *RB1* wild-type (WT) by WGS but had an *RB1*/Nucleus FISH ratio of 1.01, whereas two cases had ratios >2 but were reported WT by WGS. These discrepancies may have resulted from stromal contamination or pseudo-normalisation of genomic data in polyploid tumors. A comparison between *RB1* copy number evaluation by WGS and FISH, *RB1* IHC and histological tumor type is presented in **Supplementary Table 1**.

Structural variants of *RB1*: a putative mechanism of *RB1* inactivation

The identification of four mCRPCs with complete absence of *RB1* expression without a significant proportion of cells showing genomic copy loss led us to hypothesise that other genomic events could account for a presumptive inactivation mechanism. Exploiting the WGS data we identified four distinct putative SVs involving *RB1* in 6 metastases (6/21;28.57%) from 3 patients (3/10; 30%); 5 out of these 6 mCRPC biopsies (83.33%) had complete protein loss by IHC. Three of the four SVs were inter-chromosomal translocations and one was an intra-chromosomal tandem duplication. In two of these patients, these SVs were shared by multiple metastatic samples (S115, S116, S117 in patient V4074; and S363, S364 in patient V5191; all of these samples were IHC negative). In the third patient (V4038), a single sample presented with two distinct *RB1* SVs but preserved IHC expression.

In three samples with ADC histology from patient V4074, our computational analysis revealed *RB1* copy number neutral LOH profile. Concomitantly, an increment in the number of sequencing reads spanning exons from 7 to 17 suggested an intragenic tandem duplication of the genomic segment containing exons 7-17. Using matched RNAseq data, we were able to validate the abnormal transcript generated by this intra-genic tandem duplication that would impair protein translation. Next, we first queried whole exome sequencing data from the same patient and verified the capability to detect the genomic

event by such a strategy, and next queried a large set of individuals (11) for the same event (N=149) (**Figure 2**). Although this specific *RB1* structural event associating with full protein loss does not appear to be frequent, it evidences the importance of appropriate assay design to accurately query *RB1* status for patient management.

***RB1* deletions are frequent in clinically aggressive HSPC and more common in mCRPC**

To further evaluate *RB1* loss heterogeneity in lethal PCa we generated FISH data for *RB1* copy number analysis from 70 samples from 41 patients (including the 19 samples with WGS copy number data available). Importantly, the set included 20 HSPC and 50 CRPC biopsy samples, including 20 matched, same-patient HSPC-CRPC pairs (**Figure 3A**). Of the 20 HSPC samples, 35% (7/20) had at least shallow *RB1* deletions, i.e. a gene signal/nuclei ratio <1.7, compared with 65% (13/20) of the matched, same patient, CRPC samples. Of the 13 HSPC samples with ratios ≥ 1.7 prior to any systemic therapy, 7 had matching CRPC samples with ratios <1.7, whereas 6 retained a ratio ≥ 1.7 . Of the 7 HSPC samples with ratios <1.7 at diagnosis, only one matched CRPC sample had a ratio ≥ 1.7 . Our data indicate an enrichment for *RB1* deletions in clinically aggressive HSPCs, in line with a previous study on non-indolent PCa (38). More importantly, however, we also show *RB1* deletions evolving in matched mCRPC from previously copy number neutral HSPCs.

Deletions of *RB1* in mCRPCs

Overall, of the 50 mCRPC samples with FISH data, 43 were adenocarcinomas, 6 were NEPCs (cohort was enriched for NEPC), and 1 had mixed histology with predominantly (~95%) NEPC phenotype. An XY plot showing ratio between *RB1*/nuclei versus IHC H-Score and histological type notation is seen in **Figure 3B**. Copy number estimation of *RB1* by FISH revealed a high degree of heterogeneity, with most tumors showing a subset of *RB1* deleted cells larger than could be attributed by artefactual nuclear truncation, i.e. tangential sectioning of nuclei. Only twelve (24%; 12/50) mCRPCs were copy number neutral for both *RB1* and reference probe. *RB1* deletions, shallow or deep, were evident by FISH in 56% (28/50) of the mCRPC samples; shallow deletions were noted in 10 samples and deep deletions in 18 samples. The reference probe was not helpful for *RB1* copy number enumeration as it was detected in various combinations being detected to be lost, neutral,

or gained independently of *RB1* status, indicating frequent structural aberrations involving the 13q34 locus.

Relationship of *RB1* FISH enumeration and *RB1* protein expression by IHC

Twenty-two of 43 ADCs (22/43; 51.16%) had diffusely *RB1* IHC positivity (IHC^{Pos}) (**Figure 3B**); *RB1*/nuclei FISH ratios for these 22 tumors was >2 in 5, <2 and ≥1.7 in 8, and <1.7 in 9. Critically none of the diffusely positive cases had an *RB1*/nuclei ratio of <1. In the 16 ADCs with IHC heterogeneity (IHC^{Het}), *RB1*/nuclei FISH ratios were >2 in 1, <2 and ≥1.7 in 2, and <1.7 in 13, of which 5 were <1. Of the 5 ADC with *RB1* IHC negative staining (IHC^{Neg}), FISH profiles showed 4 cases with an *RB1*/nuclei ratio <2 and ≥1.7, and one with a ratio of 1.16. Of the seven NEPC samples, one had a ratio >2, one <2 and ≥1.7, and 5 had ratios <1.7, of which 3 were <1. All seven NEPCs with available FISH data were *RB1* IHC^{Neg}.

Kendall's Correlation Coefficient was used to determine the association between *RB1* signal/nuclei ratio and H-score. Tumors with an *RB1*s/nuclei ratio <1.7 were defined as having loss; ROC curves with bootstrapped confidence intervals were used to determine an optimal cut-off for *RB1* H-score to detect cases with a FISH ratio of ≥1.7. The area under the ROC curve was 0.55 (95% CI: 0.39-0.74) with an optimal cut-point for H-score of 105 (Sensitivity 0.77, Specificity 0.25) (**Supplementary Figure 3**). Overall, these data indicated that *RB1* copy number losses are common in mCRPCs but did not correlate well with *RB1* IHC protein loss of expression. As demonstrated in the WGS cohort, this was likely due to undetected complex structural rearrangements as putative second hits to *RB1*.

To better understand the relationship between *RB1* genomics and protein, we turned to the IHC and WGS data of the 21 tumor samples to pursue a detailed characterization of *RB1* loss. For each sample we annotated genomic aberrations with potentially damaging functional impact, such as copy number loss, somatic missense SNVs and structural rearrangements (tandem duplication and chromosomal translocations) impacting the coding region of *RB1*. These studies showed that *RB1* protein expression by IHC correlated well with the total number of genomic aberrations detected; the higher the number of genomic aberrations at the *RB1* locus, the lower the mean *RB1* protein level. Specifically, tumors carrying no (N=7), one (N=8) or two (N=6) genomic aberrations in *RB1* had median

IHC scores of 120 (mean=120, SD=35.11), 65 (mean=61.25, SD=57.92) and 0 (mean=0, SD=0), respectively (**Figure 4**).

RB1 expression can be highly heterogeneous in advanced prostate cancer

Different RB1 IHC staining patterns were observed in mCRPC biopsies (**Table 2**). Overall, homogenous RB1 IHC staining (positive or negative) was seen in 71.7% of the samples (76/106). Of these, 56.6% (60/106) of mCRPC biopsies were positive and 15.1% (16/106) negative (**Figures 5A₁₋₄ and 5B₁₋₄**). When categorizing RB1 IHC expression by histological phenotype, 61.4% (59/96) of ADCs and one NE cancer (12.5%; 1/8) showed homogeneous positivity (**Figure 5A₄**). Complete loss of RB1 expression was seen in 8.3% (8/96) of ADCs (**Figure 5B₁₋₂**) and in 87.5% (7/8) of NE tumors (**Figure 5B₃₋₄**). Two cases showed mixed histology with co-existence of adenocarcinoma and AR-low NEPC. One of these cases with had homogenous loss of RB1 IHC expression in both components, whereas the other showed positivity in the ADC component and negativity in the NE component in a single bone metastases biopsy (**Supplementary Figures 5 and 6**). Heterogeneous RB1 staining patterns were noted in 28% (30/106) of mCRPC samples (**see Figure 5C₁₋₄**), with most of these showing an ADC phenotype (96.6%; 29/30) apart from one sample which had mixed histology (**Figure 5C₄; Supplementary Figure 6**). Our data indicate that heterogeneous/focal RB1 protein loss of expression is common in mCRPC corroborating reports of loss of RB1 as a late event and associated with treatment selective pressures.

Discussion

CRPC is invariably lethal. However, substantial inter-patient genomic heterogeneity and variable sensitivity to established treatments, including endocrine agents and taxane chemotherapy lead to variable clinical courses. Robinson *et al.* suggest that “nearly 90% of cases have potentially actionable somatic or germline events” (11). Molecular stratification of mCRPC promises to improve the treatment of these diseases (39, 40) but these strategies are challenged by intra-patient heterogeneity, which can be difficult to identify from bulk exome sequencing. Intra-patient heterogeneity results in mixed responses to anticancer drugs; imaging studies, for example, have shown divergent inter-lesion responses to the PARP inhibitor olaparib in individual patients with advanced solid tumors (41). Identifying

significant intra-patient genomic heterogeneity in mCRPC is thus key to validating biomarkers and deliver molecularly stratified treatments.

Herein we obtained whole-genome sequencing of mCRPC tumor biopsies at 100X mean coverage and performed focused analyses on somatic variants involving 1,392 cancer related genes. In agreement with the data reported by Kumar *et al.* (15), despite significant inter-patient heterogeneity, the majority of somatic genomic aberrations were shared between different biopsies of each individual patient at the bulk level. Events involving taxonomy defining genes, e.g. *ETS* fusions, and certain recurrently aberrant tumor suppressors such as *TP53* mutations and deletions of *PTEN*, were consistently shared between samples from the same patient, in keeping with these being earlier, i.e. prior to CRPC state, events (31, 42, 43). Certain cancer-related genes previously reported in prostate cancer, however, were identified as private events, favouring later, likely subclonal origins. These included: *AURKA* amplification, mutations in *WNT*-signalling genes, and *RB1* aberrations. Since TCGA and SU2C/PCF sequencing data suggested an increase in *RB1* loss in mCRPC compared to primary disease and because of the established role for *RB1* loss as a predictive biomarker of CDK4/6 inhibitor resistance (35), we pursued a deeper study of *RB1* in a wider cohort of mCRPC biopsies. We found that intra-patient heterogeneity for *RB1* loss is common in mCRPC both at the genomic and protein levels. Highly relevant for the current treatment of mCRPC, we uncover novel complex genomic *RB1* structural variants that impact *RB1* protein expression and that would have gone undetected by exon focused sequencing assays. This discovery supports the use of immunohistochemistry, and perhaps multiple orthogonal assays, for determining evidence of *RB1* loss.

In mCRPC, *RB1* inactivation can promote survival benefit via distinct mechanisms including an increased output of AR directed transcription (44) or, in combination with inactivation of *TP53*, by inducing a state of plasticity which allows the emergence of AR-low NEPCs (45-47). Using WGS, Fraser *et al.* identified *RB1* losses, mono- or biallelic, in 35% of non-indolent HSPCs (38), a similar frequency to what we observed in the HSPC samples from our mCRPC matched, same-patient, cohort. Furthermore, in agreement with previous studies (5, 9, 11), we identify a higher frequency of *RB1* loss in mCRPC (56%) than in HSPC (35%). This is further evidence that *RB1* aberrations, including monoallelic deletions, confer fitness

advantages to prostate cancer cells. Critically, in agreement with our data, a study evaluating somatic copy number aberrations from WGS data from HSPCs and mCRPC, found that partial losses of *RB1* are significantly more common than homozygous deletions (48). We identified intra-sample heterogeneous loss of RB1 expression by IHC in 28% of our mCRPC samples, implicating RB1 depletion as a late and sub-clonal event. We also observed an erratic relationship between RB1 IHC protein loss and copy number status as determined by FISH. Samples with *RB1* mono-allelic deletions showed both homogeneous and heterogeneous expression patterns by IHC, and some copy number neutral cases had complete absence of protein expression. These data indicate that CN status alone, assessed by either massive-parallel sequencing of bulk biopsies or FISH, provides insufficient information on RB1 protein expression and function. Even when protein expression is present, however, determining RB1 functionality is challenging. Using mRNA derived signatures from bulk sequencing has been a recurrent approach (15, 44), but focal RB1 loss-of-expression/loss-of-function may become diluted in tumors with large proportions of *RB1* functional components.

Our study has limitations. Firstly, our patient cohorts were retrospectively collected and heterogeneous in terms of prior treatments with different timing of biopsies in relation to treatment lines. This impairs our ability to draw meaningful clinical conclusions regarding the prognostic value of the biomarkers we studied. Secondly, we were unable to obtain satisfactory RB1 staining from diagnostic HSPC samples. This could be due to different factors including tissue block age (several blocks were over a decade old) and the fact that these samples were acquired from multiple pathology laboratories with different tissue processing protocols. In contrast, our mCRPC samples were, at maximum, up to three years old, and were uniformly processed to rigorous protocols. Finally, we did not perform assays that could inform on epigenetic silencing of *RB1* and thus are unable to assess the contribution of methylation events to *RB1* inactivation.

In conclusion, previous studies indicate that the proportion of advanced PCas showing 13q14 (locus of *RB1*) mono- or bi-allelic deletions ranges from as low as 10% (11) to as high as >90% (8) depending on assay used for gene copy number estimation. Herein, we report a frequency of *RB1* loss by FISH of 56% and intra-sample focal IHC losses of RB1 expression in

28% of mCRPCs. Our findings support *RB1* aberrations in mCRPCs as resulting from treatment selection pressures in most cases. These data are critically important to the study of CDK4/6 inhibitors in mCRPC, with such trials now ongoing (e.g. NCT02905318). Overall, our data indicate that a significant number of mCRPC have heterogeneous loss of RB1 protein suggesting CDK4/6 inhibitors may be less efficacious in later stage mCRPC with efficacy being limited to patients without RB1 depletion.

Acknowledgements: We would like to acknowledge funding and support from the European Research Council Consolidator Grant 648670 (F.D.), Movember, Prostate Cancer UK, Prostate Cancer Foundation/SU2C to the International Prostate Cancer Dream team, Cancer Research UK, Experimental Cancer Medicine Centre grant funding from Cancer Research UK and the Department of Health, and Biomedical Research Centre funding to the Royal Marsden. We would also like to thank Dr. Scott Tomlins for his advice on antibody selection and assay development. Research Supported by a Stand Up To Cancer-Prostate Cancer Dream Team Translational Research Grant (SU2C-AACR-DT0712). Stand Up to Cancer is a division of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of SU2C.

Author's Contributions: D.N.R. and N.C. conceived and wrote this manuscript; Study design and manuscript structure were developed by G.B., J.D.B. and F.D.; Bioinformatic analyses were performed by N.C., A.M., A.R., P.G., and F.D.; Histological experiments were performed by M.C., S.M., I.F., and R.R.; Histopathological analyses were performed by D.N.R.; P.R., S.S., A.S., and J.M. curated clinical data; S.C., C.M., M.S., and K.K. reviewed and contributed to manuscript's final format.

References:

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA: a cancer journal for clinicians*. 2011;61(2):69-90.
2. Attard G, Parker C, Eeles RA, Schroder F, Tomlins SA, Tannock I, et al. Prostate cancer. *Lancet*. 2016;387(10013):70-82.
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA: a cancer journal for clinicians*. 2017;67(1):7-30.

4. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell*. 2015;163(4):1011-25.
5. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nature genetics*. 2012;44(6):685-9.
6. 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.
7. Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, et al. Punctuated evolution of prostate cancer genomes. *Cell*. 2013;153(3):666-77.
8. Williams JL, Greer PA, Squire JA. Recurrent copy number alterations in prostate cancer: an in silico meta-analysis of publicly available genomic data. *Cancer genetics*. 2014;207(10-12):474-88.
9. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature*. 2012;487(7406):239-43.
10. Grasso CS, Cani AK, Hovelson DH, Quist MJ, Douville NJ, Yadati V, et al. Integrative molecular profiling of routine clinical prostate cancer specimens. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 2015.
11. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, et al. Integrative clinical genomics of advanced prostate cancer. *Cell*. 2015;161(5):1215-28.
12. Barry M, Perner S, Demichelis F, Rubin MA. TMPRSS2-ERG fusion heterogeneity in multifocal prostate cancer: clinical and biologic implications. *Urology*. 2007;70(4):630-3.
13. Arora R, Koch MO, Eble JN, Ulbright TM, Li L, Cheng L. Heterogeneity of Gleason grade in multifocal adenocarcinoma of the prostate. *Cancer*. 2004;100(11):2362-6.
14. Cooper CS, Eeles R, Wedge DC, Van Loo P, Gundem G, Alexandrov LB, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nature genetics*. 2015;47(4):367-72.
15. Kumar A, Coleman I, Morrissey C, Zhang X, True LD, Gulati R, et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nature medicine*. 2016;22(4):369-78.
16. Gundem G, Van Loo P, Kremeyer B, Alexandrov LB, Tubio JM, Papaemmanuil E, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature*. 2015;520(7547):353-7.
17. Dyson NJ. RB1: a prototype tumor suppressor and an enigma. *Genes & development*. 2016;30(13):1492-502.
18. Robinson DR, Wu YM, Lonigro RJ, Vats P, Cobain E, Everett J, et al. Integrative clinical genomics of metastatic cancer. *Nature*. 2017;548(7667):297-303.
19. Raczy C, Petrovski R, Saunders CT, Chorny I, Kruglyak S, Margulies EH, et al. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics*. 2013;29(16):2041-3.
20. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-303.
21. Demichelis F, Greulich H, Macoska JA, Beroukhim R, Sellers WR, Garraway L, et al. SNP panel identification assay (SPIA): a genetic-based assay for the identification of cell lines. *Nucleic Acids Res*. 2008;36(7):2446-56.

22. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology*. 2013;31(3):213-9.
23. Romanel A, Lago S, Prandi D, Sboner A, Demichelis F. ASEQ: fast allele-specific studies from next-generation sequencing data. *BMC Med Genomics*. 2015;8:9.
24. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6(2):80-92.
25. Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods*. 2009;6(9):677-81.
26. Xi R, Hadjipanayis AG, Luquette LJ, Kim TM, Lee E, Zhang J, et al. Copy number variation detection in whole-genome sequencing data using the Bayesian information criterion. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(46):E1128-36.
27. Prandi D, Baca SC, Romanel A, Barbieri CE, Mosquera JM, Fontugne J, et al. Unraveling the clonal hierarchy of somatic genomic aberrations. *Genome Biol*. 2014;15(8):439.
28. Welti J, Rodrigues DN, Sharp A, Sun S, Lorente D, Riisnaes R, et al. Analytical Validation and Clinical Qualification of a New Immunohistochemical Assay for Androgen Receptor Splice Variant-7 Protein Expression in Metastatic Castration-resistant Prostate Cancer. *European urology*. 2016.
29. Reid AH, Attard G, Ambroisine L, Fisher G, Kovacs G, Brewer D, et al. Molecular characterisation of ERG, ETV1 and PTEN gene loci identifies patients at low and high risk of death from prostate cancer. *British journal of cancer*. 2010;102(4):678-84.
30. Verlinsky Y, Ginsberg N, Chmura M, Freidine M, White M, Strom C, et al. Cross-hybridization of the chromosome 13/21 alpha satellite DNA probe to chromosome 22 in the prenatal screening of common chromosomal aneuploidies by FISH. *Prenat Diagn*. 1995;15(9):831-4.
31. Cancer Genome Atlas Research Network. Electronic address scmo, Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell*. 2015;163(4):1011-25.
32. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, et al. Integrative clinical genomics of advanced prostate cancer. *Cell*. 2015;161(5):1215-28.
33. Hieronymus H, Schultz N, Gopalan A, Carver BS, Chang MT, Xiao Y, et al. Copy number alteration burden predicts prostate cancer relapse. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(30):11139-44.
34. Nikonova AS, Astsaturov I, Serebriiskii IG, Dunbrack RL, Jr., Golemis EA. Aurora A kinase (AURKA) in normal and pathological cell division. *Cell Mol Life Sci*. 2013;70(4):661-87.
35. Malorni L, Piazza S, Ciani Y, Guarducci C, Bonechi M, Biagioni C, et al. A gene expression signature of retinoblastoma loss-of-function is a predictive biomarker of resistance to palbociclib in breast cancer cell lines and is prognostic in patients with ER positive early breast cancer. *Oncotarget*. 2016;7(42):68012-22.
36. Turner NC, Huang Bartlett C, Cristofanilli M. Palbociclib in Hormone-Receptor-Positive Advanced Breast Cancer. *The New England journal of medicine*. 2015;373(17):1672-3.

37. Finn RS, Crown JP, Lang I, Boer K, Bondarenko IM, Kulyk SO, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. *The Lancet Oncology*. 2015;16(1):25-35.
38. Fraser M, Sabelnykova VY, Yamaguchi TN, Heisler LE, Livingstone J, Huang V, et al. Genomic hallmarks of localized, non-indolent prostate cancer. *Nature*. 2017;541(7637):359-64.
39. Mateo J, Carreira S, Sandhu S, Miranda S, Mossop H, Perez-Lopez R, et al. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. *The New England journal of medicine*. 2015;373(18):1697-708.
40. de Bono JDG, U.; Massard, C.; Bracarda, S.; Kocak, I.; Font, A.; Arranz Arijia, J.A.; Shih, K.; Radavoi, G.; Yu, W.; Chan, W.; Huang, J.; Musib, L.; Gendreau, S.; Meng, R.; Patel, P.; Maslyar, D.; Jinga, V. Randomized phase II study of AKT blockade with ipatasertib (GDC-0068) and abiraterone (Abi) vs. Abi alone in patients with metastatic castration-resistant prostate cancer (mCRPC) after docetaxel chemotherapy (A. MARTIN Study). ASCO; Chicago, IL2016. p. 5017-.
41. Perez-Lopez R, Roda D, Jimenez B, Brown J, Mateo J, Carreira S, et al. High frequency of radiological differential responses with poly(ADP-Ribose) polymerase (PARP) inhibitor therapy. *Oncotarget*. 2017;8(61):104430-43.
42. Ferraldeschi R, Nava Rodrigues D, Riisnaes R, Miranda S, Figueiredo I, Rescigno P, et al. PTEN protein loss and clinical outcome from castration-resistant prostate cancer treated with abiraterone acetate. *European urology*. 2015;67(4):795-802.
43. 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 Precis Oncol*. 2017;2017.
44. Sharma A, Yeow WS, Ertel A, Coleman I, Clegg N, Thangavel C, et al. The retinoblastoma tumor suppressor controls androgen signaling and human prostate cancer progression. *The Journal of clinical investigation*. 2010;120(12):4478-92.
45. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nature medicine*. 2016.
46. Ku SY, Rosario S, Wang Y, Mu P, Seshadri M, Goodrich ZW, et al. Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science*. 2017;355(6320):78-83.
47. Mu P, Zhang Z, Benelli M, Karthaus WR, Hoover E, Chen CC, et al. SOX2 promotes lineage plasticity and antiandrogen resistance in TP53- and RB1-deficient prostate cancer. *Science*. 2017;355(6320):84-8.
48. Camacho N, Van Loo P, Edwards S, Kay JD, Matthews L, Haase K, et al. Appraising the relevance of DNA copy number loss and gain in prostate cancer using whole genome DNA sequence data. *PLoS Genet*. 2017;13(9):e1007001.

Figure and Table legends

Figure 1: Landscape of genomic aberrations. Each row represents a gene and each column a tumor sample. Specifically, the genomic status (wild type or mutated) is established based on the presence or absence of at least one missense somatic point mutation. Grey bars at top indicate the fraction of the genome altered by a SCNA event (Copy Number Altered Fraction, CNAF) and to the total number of non-synonymous SNVs. Grey bars on the left indicate, for each gene, the fraction of samples affected by somatic missense SNV (violet), copy number loss (blue) or focal amplification (red). Overall, the total number of non-synonymous SNVs and CNAFs indicate low intra-patient tumor heterogeneity. Analysis focused on a reduced set of known cancer-associated genes revealed heterogeneous genomic status for *RB1* in patients V5128 and V5033.

Figure 2: Alternative molecular mechanism causing RB1 inactivation in patient V4074. A) Paired-end whole genome sequencing data mapped on *RB1* locus for 3 metastatic sites of patients V4074. Histograms at the top of each alignment track show coverage profiles. Grey sequencing reads indicate expected (corrected orientation and insert size of paired reads) mapping. Vertical black lines indicate break points detected by BreakDancer algorithm. Green highlights pairs of reads with anomalous orientation compatible with a tandem duplication event. B) Paired-end RNA sequencing reads from V4074 supporting (green) the tandem duplication event involving exons 7 to 17. C) Schematic representation of detected tandem duplication. At the top is shown *RB1* genomic architecture when tandem duplication involving exons 7 to 17 occurs (black boxes). Read pairs with coordinates or insert size spanning break points (junctions between yellow and black boxes) will correctly map (grey reads) in the reference genome (bottom), while read pairs with insert size spanning the region between the repeated region (junctions between consecutive black boxes) will map with anomalous orientation (green reads). D) Left, immunohistochemistry of metastatic samples for RB show absence of expression. Right, FISH data for *RB1* (target probes in red and reference probes in green) support concordantly hemizygous loss across all sites. E) Analysis extended to PCF SU2C cohort (N=149) confirms that the event is detectable in WES data by ad hoc coverage bases computation. In red, the median tumor/normal normalized coverage log-ratio profile (y-axis) computed across *RB1*

exons (grouped as indicated on x-axis) for a metastatic site of patient V4074 affected by the tandem duplication. Grey lines represent profiles of all other PCF SU2C patients.

Figure 3. RB1 deletion studied by FISH and IHC. A) 20 HSPC-mCRPC paired samples (same patient); 35% percent (7/20; 35%) of HSPCs presented at least shallow deletion by our FISH definition. In comparison, sixty-five percent (13/20; 65%) of mCRPCs presented at least a shallow deletion. Of the 7 patients with HSPCs showing copy number losses by FISH, 6 presented with copy number lost mCRPC and one case relapsed with a copy number neutral mCRPC. Our data shows an enrichment for RB1 deletion in clinically aggressive primary prostate cancer. B) X-Y plot of IHC H-Score x RB1 Signal/Nuclei ratio in mCRPCs. A general positive correlation is observed between the two variables but several discrepancies exist. Our data shows that gene copy number assessed by FISH alone provides insufficient data regarding RB1 IHC expression.

Figure 4. Genomic data explains Retinoblastoma protein levels. A) IHC Retinoblastoma protein levels against the number of genomic aberrations at RB1 genomic locus with potential damaging functional impact. B) Pie chart summarizing RB1 genomic data: the external level indicates the number of samples (N=21) carrying none, one or two (yellow, orange and red sections, respectively) putative deleterious genomic hits in RB1. Inner level specifies the type of RB1 aberrations (light blue, copy number loss; pink, missense SNVs; green, structural variant) affecting samples in each previous section.

Figure 5: Micrographs of RB1 IHC staining (200x magnification). Row A) shows examples homogeneously positive cases. A range of staining intensities can be appreciated in diffusely positive cases. A4) accounts for an AR negative tumor with typical neuroendocrine morphology and diffuse positivity for RB1. Row B) shows homogeneously negative cases. B1) and B2) are have a typical ADC morphology and were AR positive on IHC. B3) and B4) are examples of NEs with confirmed absence of AR staining on IHC. Row C) shows examples of RB1 IHC heterogeneous cases. C1), C2), and C3) have an ADC phenotype (confirmed by AR IHC). C4) shows one of two cases with RB1 heterogeneous staining pattern and mixed histology.

Table 1: Summary of clinical data.

		WGS data + RB1 FISH/IHC	RB1 FISH/IHC only	RB1 IHC only	Total
	n. Patients	10	30	55	95
Age	Median	58.8	61.7	62.99	61.87
	IQR	50.84-61.4	59.04-66.86	57.61-67.39	57.09-66.58
Staging at Diagnosis	I	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	II	0 (0%)	0 (0%)	3 (5.45%)	4 (4.2%)
	III	4 (40%)	4 (13.33%)	10 (18.18%)	18 (18.94%)
	IV	6 (60%)	22 (73.33%)	33 (60%)	61 (64.21%)
	NA	0 (0%)	4 (13.33%)	9 (16.36%)	13 (13.68%)
Gleason Score	GS ≤ 7	4 (40%)	5 (16.66%)	15 (27.27%)	19 (20%)
	GS > 7	5 (50%)	17 (56.66%)	32 (58.18%)	52 (54.73%)
	NA	1 (10%)	8 (26.66%)	8 (14.54%)	17 (17.89%)
Exposure prior to biopsy	Docetaxel	7 (70%)	20 (66.66%)	36 (65.45%)	63 (66.31%)
	Cabazitaxel	5 (50%)	9 (30%)	14 (25.45%)	28 (29.47%)
	Abiraterone	5 (50%)	23 (76.66%)	38 (69.09%)	66 (69.47%)
	Enzalutamide	4 (40%)	8 (26.67%)	8 (14.54%)	20 (21.05%)
	Carboplatinum	2 (20%)	-	-	-

Table 2: Summary of samples, tissue sites of origin, and histology.

		n. of mCRPC samples	RB1 IHC ^{pos}	RB1 IHC ^{het}	RB1 IHC ^{neg}
		106 (100%)	60 (56.6%)	30 (28.3%)	16 (15.09%)
Site of metastases	Lymph Node	44 (41.51%)	29 (27.3%)	10 (9.4%)	5 (4.71%)
	Bone	39 (36.79%)	22 (20.75%)	15 (14.15%)	2 (1.88%)
	Visceral	12 (11.32%)	6 (5.66%)	3 (2.83%)	3 (2.8%)
	Soft Tissue	11 (10.38%)	4 (3.77%)	1 (0.0094%)	6 (5.66%)
Histology	ADC	96 (90.57%)	59 (55.66%)	29 (27.35%)	8 (7.54%)
	NE	8 (7.55%)	1 (0.0094%)	0 (0%)	7 (6.6%)
	Mixed	2 (1.89%)	0 (0%)	1 (0.0094%)	1 (0.0094%)

Figure 1.

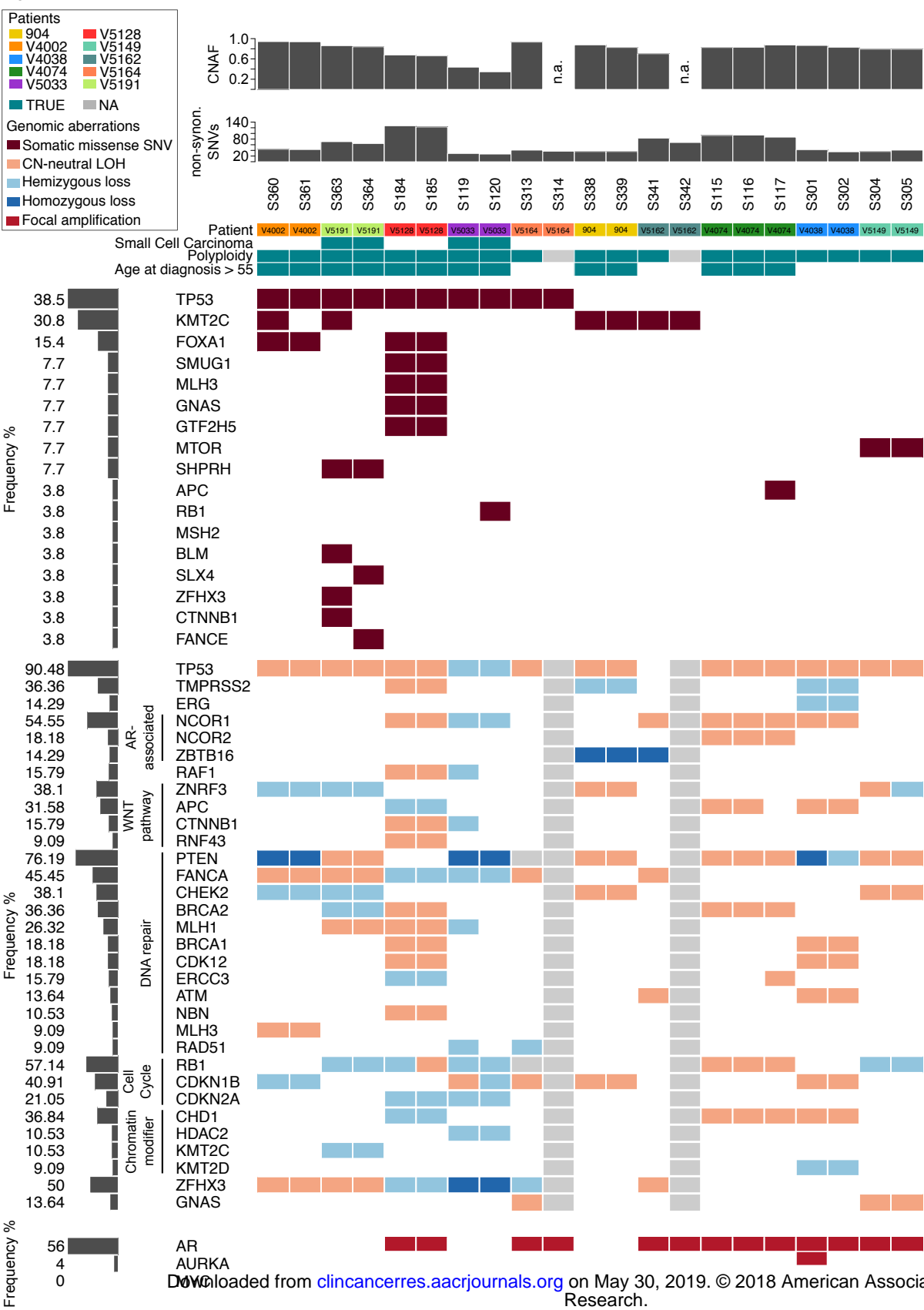


Figure 2.

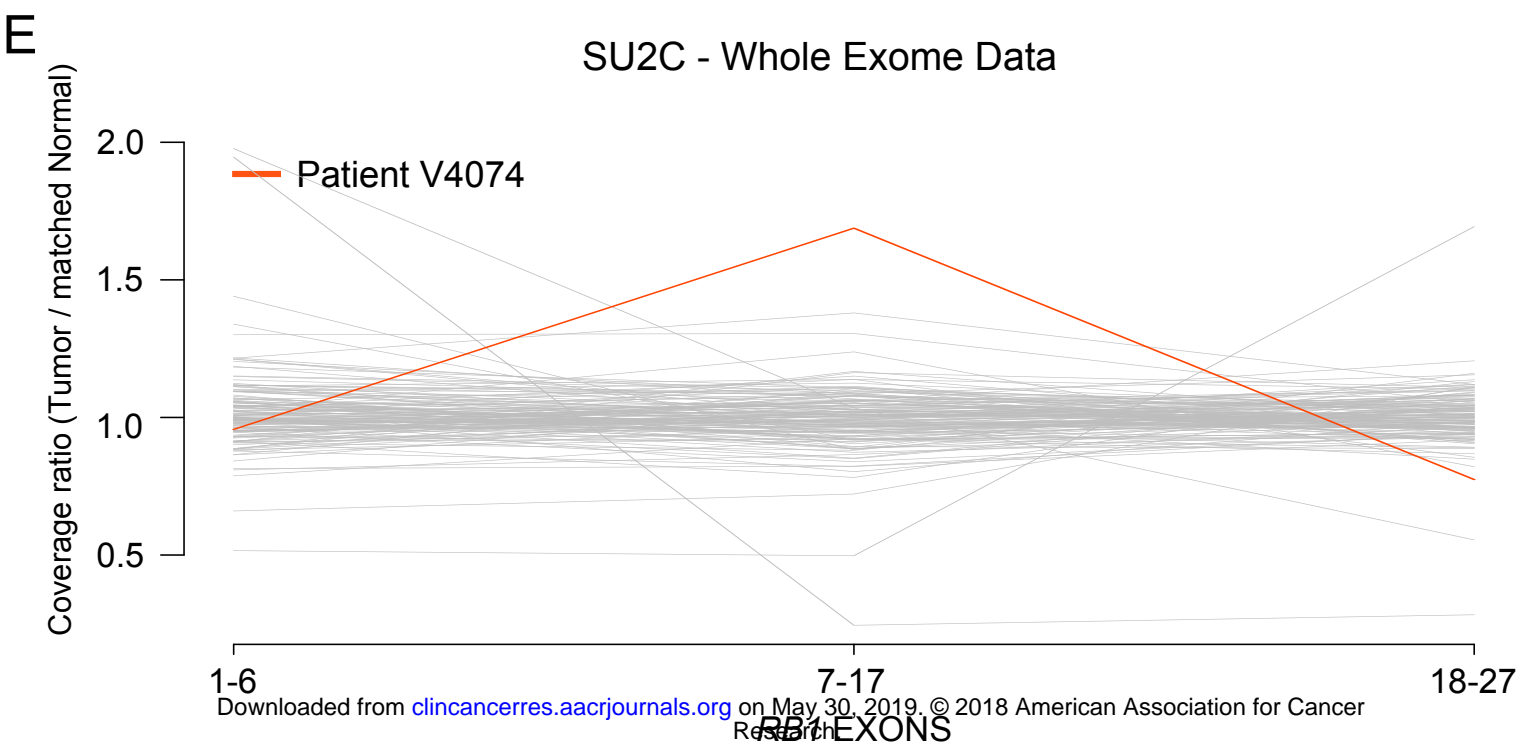
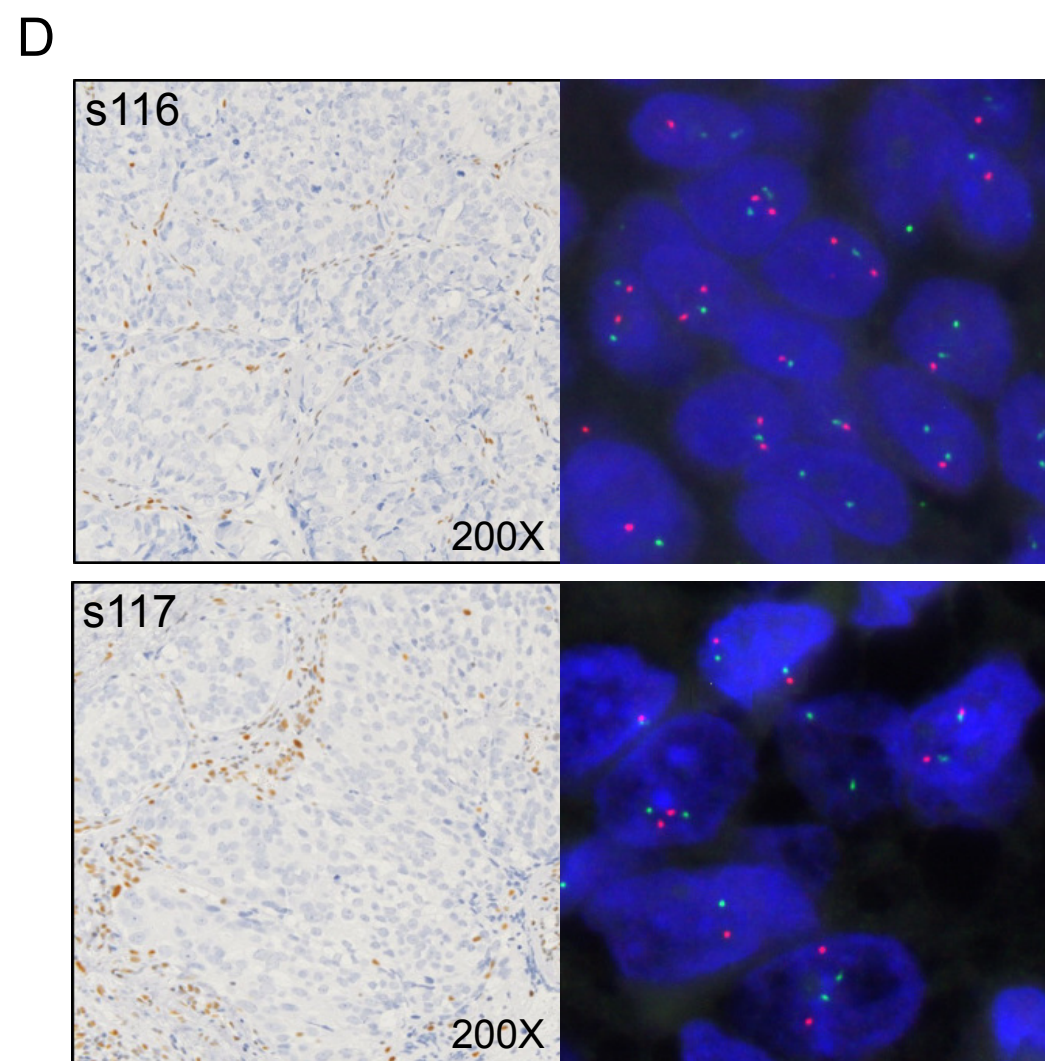
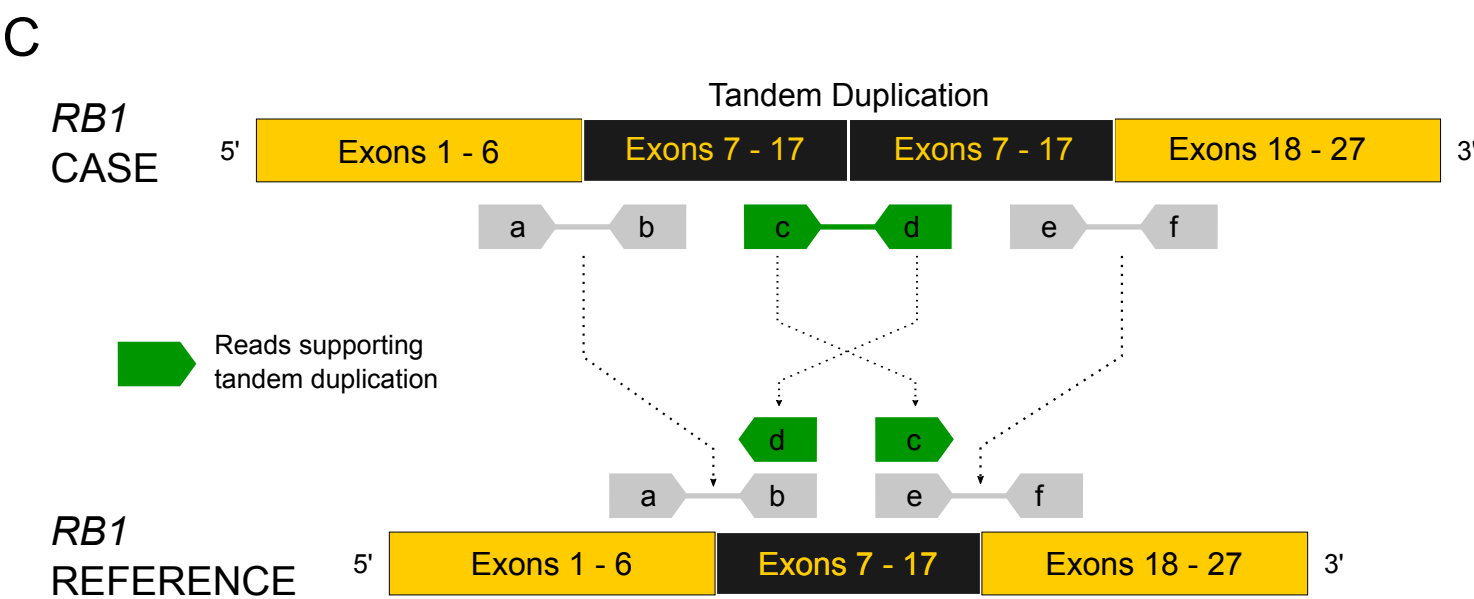
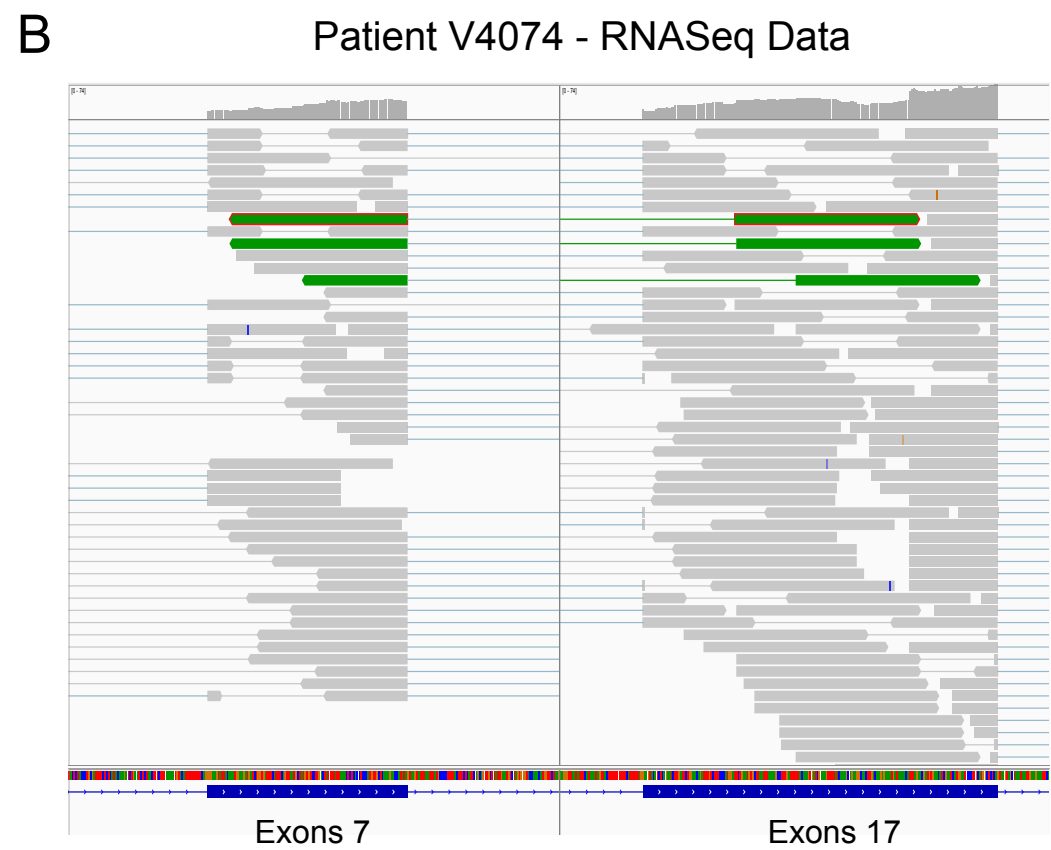
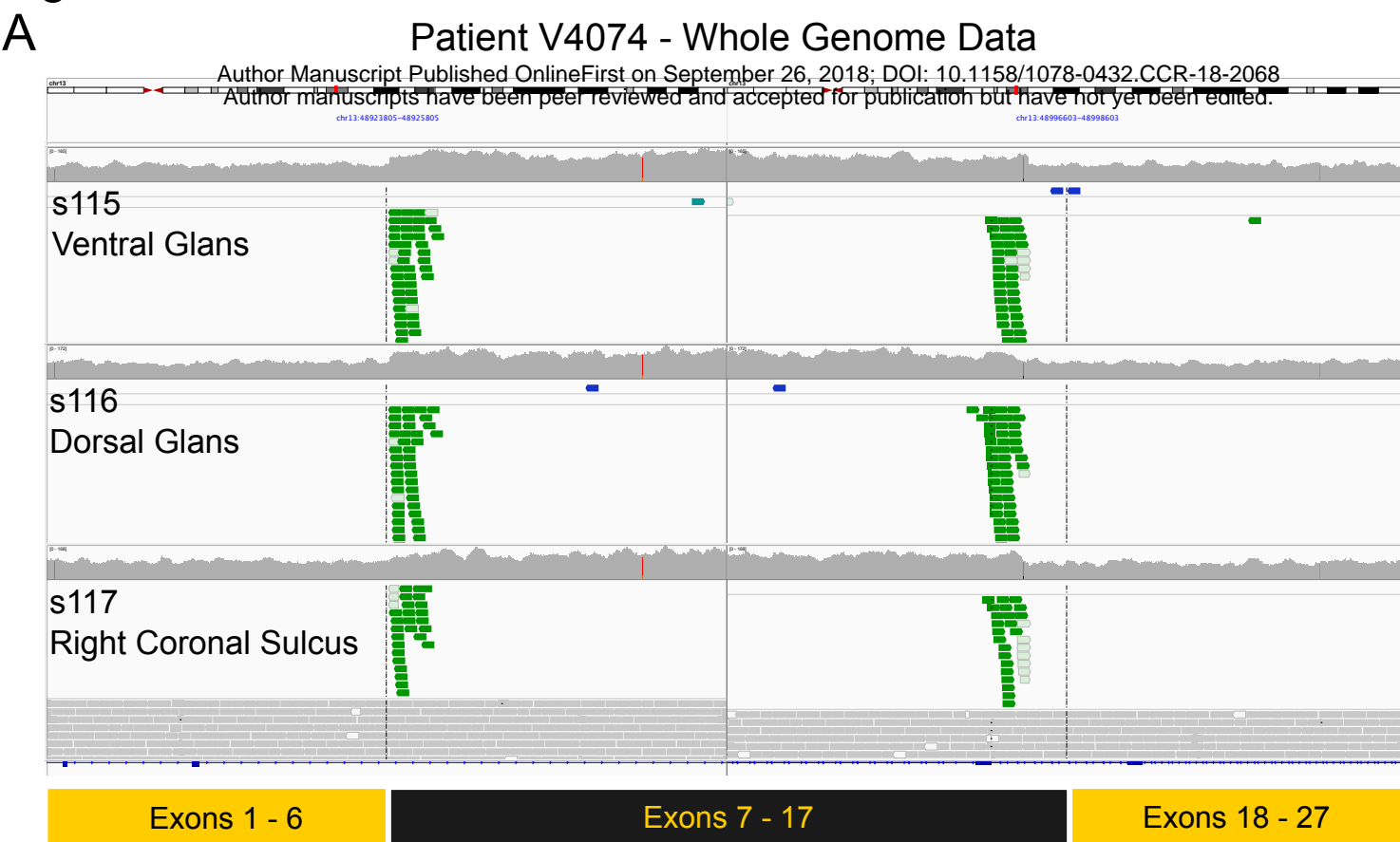


Figure 3.

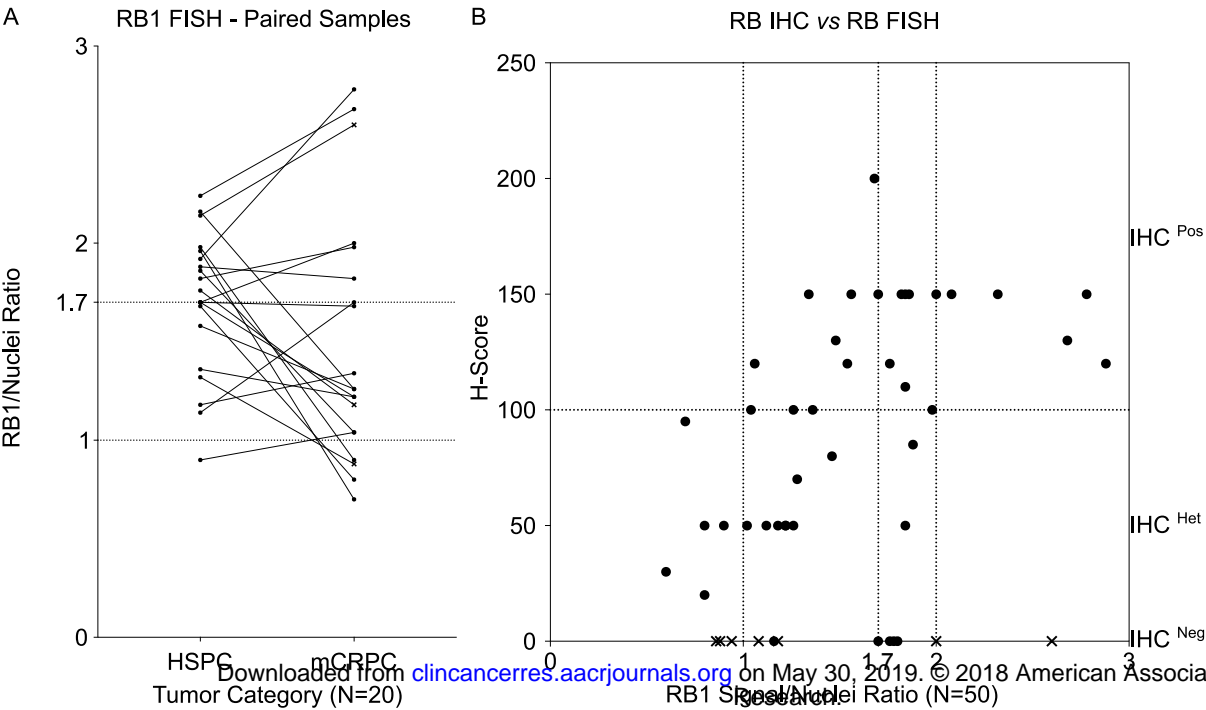
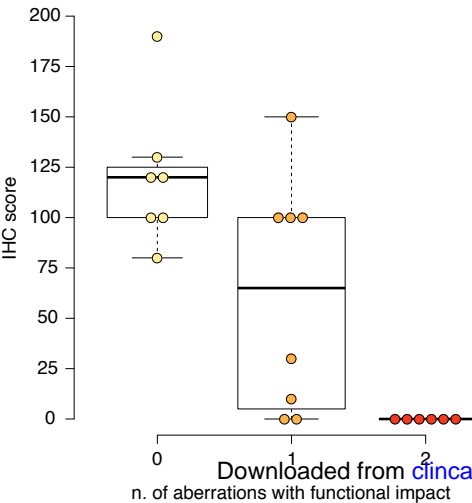


Figure 4.

A



B

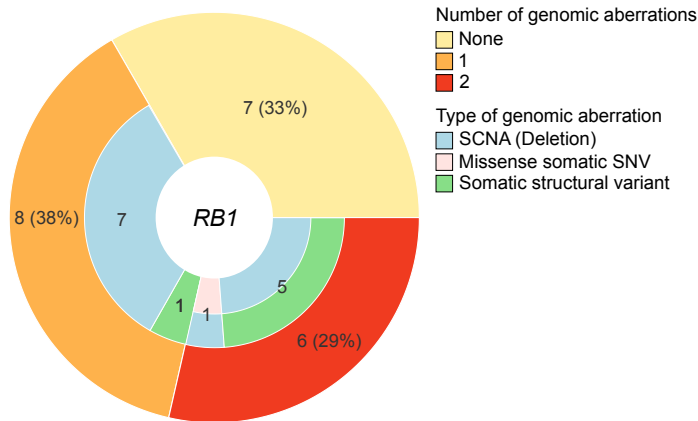
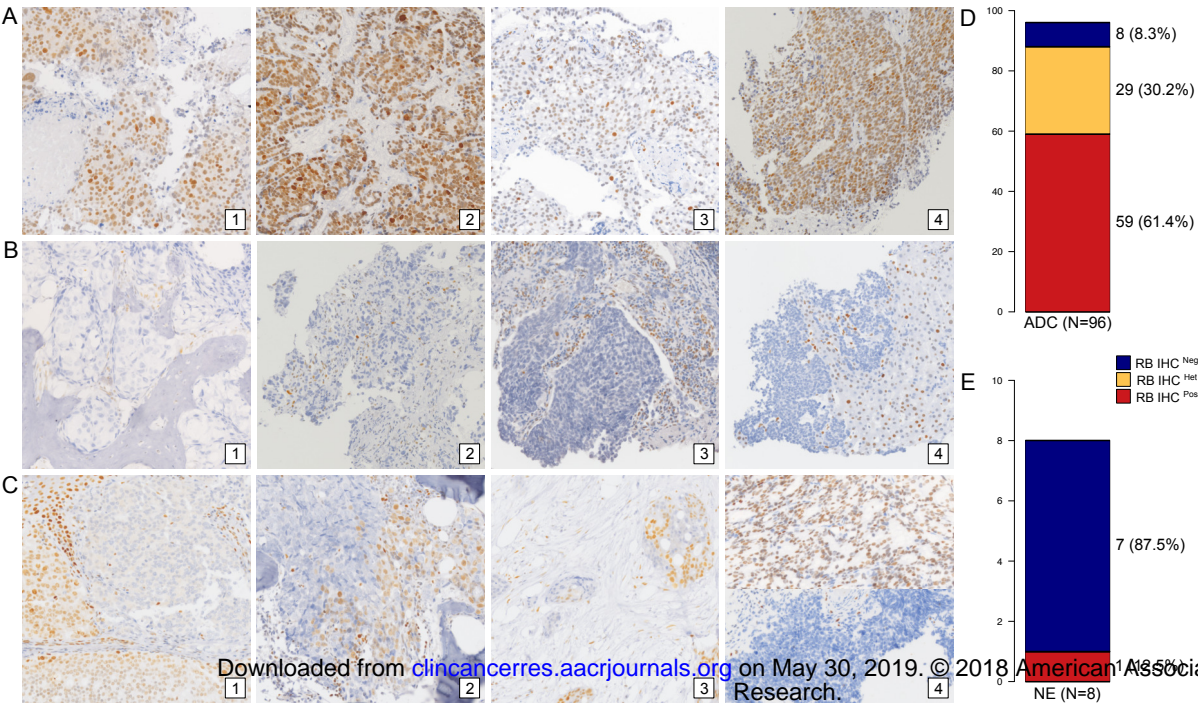


Figure 5.



Clinical Cancer Research

RB1 Heterogeneity in Advanced Metastatic Castration Resistant Prostate Cancer

Daniel Nava Rodrigues, Nicola Casiraghi, Alessandro Romanel, et al.

Clin Cancer Res Published OnlineFirst September 26, 2018.

Updated version	Access the most recent version of this article at: doi: 10.1158/1078-0432.CCR-18-2068
Supplementary Material	Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2018/09/26/1078-0432.CCR-18-2068.DC1
Author Manuscript	Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/early/2018/09/26/1078-0432.CCR-18-2068 . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.