medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Increased frequency of CHD1 deletions in prostate cancers of African American men is associated with distinct homologous recombination deficiency associated DNA aberration profiles

Miklos Diossy[1*], Viktoria Tisza[2,3*], Hua Li[4,5*], Jia Zhou[6], Zsofia Sztupinszki[1], Denise Young[4,5], Darryl Nousome[4,5], Claire Kuo[4,5], Yongmei Chen[4,5], Reinhard Ebner[7], Isabell A. Sesterhenn[8], Joel T. Moncur[8], Gregory T. Chesnut[4], Gyorgy Petrovics[4,5], Gregory T. Klus[2,9], Sandor Spisak[3,10], Gabor Valcz[11], Pier Vitale Nuzzo[3,10], Dezso Ribli[12], Aimilia Schina[1,13], Judit Börcsök[1], Aurel Prosz[1], Marcin Krzystanek[1], Thomas Ried[9], David Szuts[14], Salma Kaochar[15], Shailja Pathania[16,17], Alan D. D'Andrea[6,18], Istvan Csabai[12], Shiv Srivastava[4,19], Albert Dobi[3,4], Matthew L Freedman[3,10,20], Zoltan Szallasi[1,2,21]

[1] Danish Cancer Society Research Center, Copenhagen, Denmark

[2] Computational Health Informatics Program, Boston Children's Hospital, USA, Harvard Medical School, Boston, USA

[3] Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts

[4] Center for Prostate Disease Research, Murtha Cancer Center / Research Program, Department of Surgery, Uniformed Services University of the Health Sciences, Bethesda, MD

[5] The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc, Bethesda, MD

[6] Department of Radiation Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA.

[7] CytoTest Inc., Rockville, Maryland, USA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[8] Joint Pathology Center, Silver Spring, Maryland, USA

[9] Genetics Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA.

[10] Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA

[11] MTA-SE Molecular Medicine Research Group, Hungarian Academy of Sciences, 1051 Budapest, Hungary

[12] Department of Physics of Complex Systems, Eötvös Loránd University, Budapest, Hungary

[13] National Center for Cancer Immune Therapy, Copenhagen, Denmark

[14] Institute of Enzymology, Research Centre for Natural Sciences, Budapest, Hungary

[15] Department of Medicine, Baylor College of Medicine, Houston, USA

[16] Center for Personalized Cancer Therapy, University of Massachusetts, Boston, MA

[17] Department of Biology, University of Massachusetts, Boston, MA

[18] Center for DNA Damage and Repair, Dana-Farber Cancer Institute, Boston, MA 02215

[19] Department of Biochemistry and Molecular & Cell Biology, Georgetown University School of Medicine, Washington DC

[20] The Eli and Edythe L. Broad Institute, Cambridge, MA, USA

[21] 2nd Department of Pathology, SE NAP, Brain Metastasis Research Group, Hungarian Academy of Sciences, Semmelweis University, Budapest, Hungary

* These authors contributed equally

Correspondence should be addressed to: Zoltan.Szallasi@childrens.harvard.edu

Word counts

abstract: 139

Article: 3442 without the methods section, 5673 including the methods section

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background

African American (AA) men have significantly higher mortality rates from prostate cancer (PC) than individuals of European ancestry (EA). Therapeutically targetable molecular differences may hold the potential to reduce this disparity.

Objective

To investigate chromodomain helicase DNA-binding protein 1 (*CHD1*) deletion both as a cause of aggressive disease and therapeutic vulnerability in the prostate cancer of AA men.

Design, setting, and participants

91 AA and 109 EA prostate cancer cases were analyzed by fluorescence in situ hybridization (FISH) for the deletion of CHD1. Whole exome and whole genome sequencing data from prostate adenocarcinoma cases were analyzed for mutational signatures from AA and EA individuals.

Outcome measurements and statistical analysis

Associations with biochemical recurrence were evaluated using Cox proportional hazard regression models. Association between mutational signatures and CHD1 deletion were assessed by Wilcoxon ranked sum tests.

Results and limitations

Subclonal deletion of *CHD1* is nearly three times as frequent in prostate tumors of men than in EA men. *CHD1* deletion is associated with some of the homologous recombination deficiency associated mutational signatures in prostate cancer. In a cell line model *CHD1* deletion induced 1-10 kb deletions resembling those induced by *BRCA2* deficiency. *CHD1* deficient cells showed markedly increased sensitivity to both talazoparib and the radiomimetic bleomycin.

Conclusions

CHD1 is more frequently deleted in the prostate cancer of AA men. This deletion is both associated with and induces mutational signatures characteristic of *BRCA2* deficiency. *CHD1* deficient prostate cancer is more sensitive to talazoparib or bleomycin treatment.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Patient summary

Subclonal deletion of *CHD1* is more frequent in the prostate cancer of AA men and this could be one of the reasons behind more aggressive disease. *CHD1* deletion, however, also constitutes a therapeutic vulnerability to the PARP inhibitor talazoparib. This treatment may significantly improve the outcome of disease in AA men.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

African American (AA) men have significantly higher incidence and mortality rates from prostate cancer (PC) compared to individuals of European ancestry (EA). Recent studies demonstrated that AA men are at higher risk of progression after radical prostatectomy, even in equal access settings and when accounting for socioeconomic status[1,2]. While the reasons underlying these disparities are multifactorial, these data strongly argue that germline and/or somatic genetic differences between AA and EA men may in part explain these differences.

Comparative analysis of AA and EA prostate tumors have identified several genetic differences. *PTEN* deletions, *ERG* translocations and/or *ERG* over-expression are more frequent in PCs of EA men[3–5]. In contrast, *LSAMP* and *ETV3* deletions, *ZFHX3* mutations, *MYC* and *CCND1* amplifications and *KMT2D* truncations are more frequent in PCs of AA men[6–8]. *ERF*, an ETS transcriptional repressor, also showed an increased mutational frequency in AA prostate cancer cases with probable functional consequences such as increased anchorage independent growth[9], and *SPINK1* expression is also enriched in African American PC[10].

Chromodomain helicase DNA-binding protein 1 (*CHD1*) deletion is frequently present in prostate cancer. Deletions are associated with increased Gleason score and faster biochemical recurrence[11], activation of transcriptional programs that drive prostate tumorigenesis[12] and enzalutamide resistance[13]. Mechanistically, *CHD1* loss influences prostate cancer biology in at least two ways. *CHD1*, an ATPase-dependent chromatin remodeler, contributes to a specific distribution of androgen receptor (AR) binding in the genome of prostate tissue. When lost, the AR cistrome

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

redistributes to HOXB13 enriched sites and thus alters the transcriptional program of prostate cancer cells [12]. *CHD1* also contributes to genome integrity. It is required for the recruitment of CtIP, an exonuclease, to DNA double strand breaks (DSB) to initiate end resection. Upon *CHD1* loss this important step in DSB repair is impaired leading to homologous recombination deficiency[14,15]. The functional impact of *CHD1* loss is further influenced by the presence of SPOP mutations, which were reported to be associated with the suppression of DNA repair[16].

CHD1 loss is frequently subclonal[17], which makes its detection by next generation sequencing more challenging[18] and it may go undetected depending on the subclonal fraction of cells harboring this aberration. Therefore, the true proportion of PC cases with *CHD1* may be underestimated. Thus, we decided to investigate the frequency of *CHD1* loss in EA and AA PC by methods more sensitive to detecting subclonal deletions including evaluations of multiple tumor foci present in each prostatectomy specimen.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

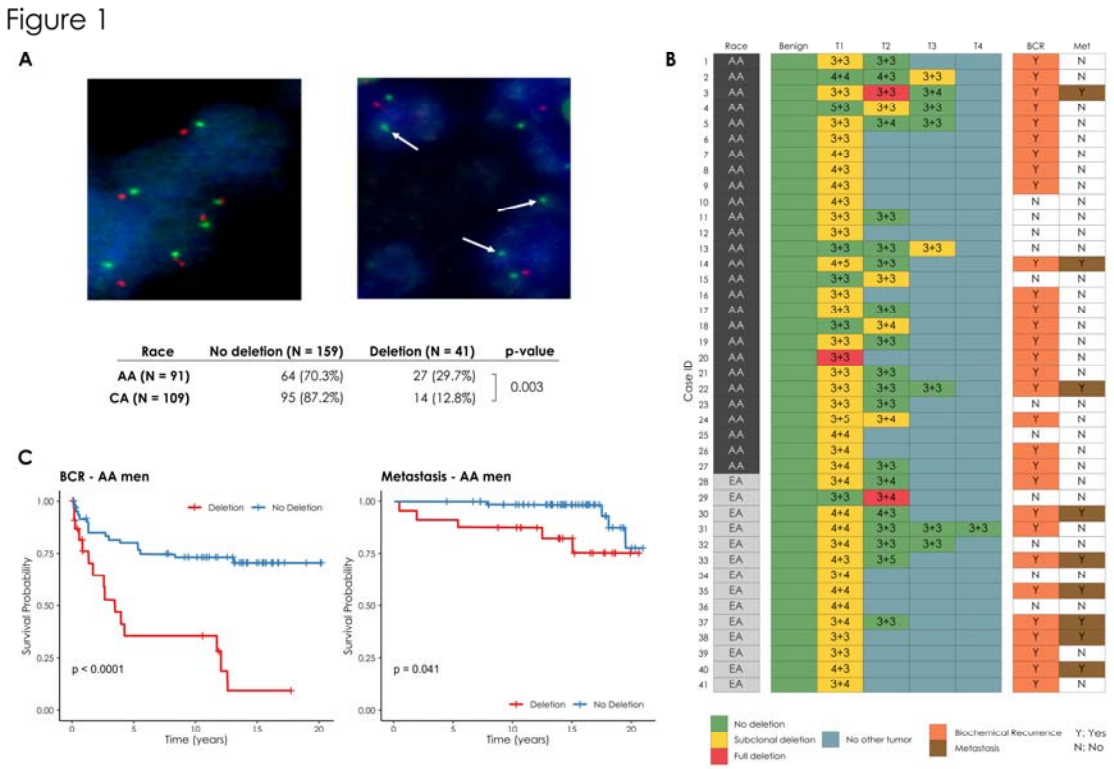

FIGURE 1: *CHD1* copy number by FISH in tissue microarrays. (A) Prostate cancer cells with wild type (diploid) *CHD1* (upper left) vs. prostate cancer cells harboring mono-allelic deletion for *CHD1* (upper right) are visualized by FISH assay. Orange signal: *CHD1* probe; green signal: human chromosome 5 short arm probe; blue color: DAPI nuclear stain. Arrows are representing the lack of CHD1. Representative view fields capture 3-3 cell nuclei at 60X magnification. Inset table summarizes the higher frequency of *CHD1* deletion in prostatic carcinoma of AA vs. EA patients. **(B)** *CHD1* deletion is a subclonal event in prostate cancer. Multiple tumor samples from 200 patients were assessed by FISH assay that identified 41 patients with CHD1 deletion (left panel). The heatmap depicts the sampled largest tumor 1 (T1), second largest tumor (T2), and so on. Numbers denote pathological Gleason grade for each tumor. BCR: biochemical recurrence (grey); Met: metastasis (brown). **(C)** Deletion of CHD1 (clonal or subclonal in any of the nodes) is strongly associated with disease progression in AA prostate cancer patients (N=91). BCR: univariable Kaplan-Meier curve; Metastasis: univariable Kaplan-Meier curve.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RESULTS

Subclonal *CHD1* deletion is more frequent in African American prostate cancers and associated with worse clinical outcome.

CHD1 is frequently subclonally deleted in prostate cancer [17]. Our initial analysis on the SNP array data from TCGA comparing AA and EA PC cases suggested that the subclonal loss of *CHD1* may be a more frequent event in AA men (Suppl. Figures 1 and 2). To independently validate this observation, we assessed *CHD1* copy number by FISH in tissue microarrays (TMAs) constructed from multiple tumor foci per prostatectomy specimen in a matched cohort of 91 AA and 109 EA patients from the equal-access military healthcare system (Figure 1A, Suppl. Figure 3). Key clinico-pathological features including diagnosis, age, diagnosis PSA levels, pathological stages, Gleason sums, Grade groups, margin status, adjuvant therapy, biochemical recurrence (BCR) and metastasis did not differ between AA and EA cases (Suppl. Table 1A). Consistent with the long-term follow up (median: 16 years) of the cohort, we observed a 40% biochemical recurrence (BCR) and 16% metastasis rate[19]. For each clinical case up to four different cancerous areas were analyzed comprising 4-10 different tissue cores (for details see methods and Suppl. Table 1).

We detected subclonal *CHD1* loss in 27 out of 91 AA cases (29.7%), and 14 out of 109 (11%) EA cases indicating that *CHD1* deletion is about three times more frequent in prostate tumors of AA men. Our FISH data confirmed the subclonal nature of *CHD1* deletion in prostate cancer cells (Figure 1B). In most cases *CHD1* deletion was present in only a subset of tumor cores (see Suppl. Table 1 for details).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

As a control, we performed FISH staining of *PTEN* and immunohistochemistry (IHC) staining of ERG in a subset of the cohort (42 AA and 59 EA prostate cancer cases) confirming previously described differences[3,4] (Suppl. Table 1).

Further analyses revealed a significant association between *CHD1* deletion and pathologic stages and Gleason sum. Higher frequency of *CHD1* deletion was detected in T3-4 pathological stage compared to T2 stage (P=0.043, Suppl. Table 1). Prostate cancer cases with higher Gleason sum scores (3+4, 4+3, 8-10) were seen more frequently in the *CHD1* deletion group than in the non-deletion group (P<0.001). In contrast, lower Gleason sum score (3+3) was more often seen in non-deletion cases (P<0.001, table 1c). Notably, *CHD1* deletion was strongly associated with rapid biochemical recurrence (Figure 1C) in AA cases (P<0.0001). The multivariate Cox model analysis showed that *CHD1* deletion was an independent predictor of BCR in the entire cohort (P=0.0006) after adjusting for age at diagnosis, PSA at diagnosis, race, pathological tumor stage, grade group, and surgical margins. Moreover, a significant correlation between *CHD1* deletion and metastasis was also detected in AA patients with Kaplan-Meier analysis (P=0.041). Following adjustment for age at diagnosis, PSA at diagnosis, race, pathological tumor stage, grade group, and surgical margins in the Cox proportional hazards model, *CHD1* deletion was significantly associated with metastasis (P=0.047, Suppl. Figure 4). Taken together, our data strongly support the association of *CHD1* deletions with aggressive prostate cancer and worse clinical outcomes in AA PC.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Estimating the frequency of subclonal *CHD1* loss in next generation sequencing data of AA and EA prostate cancer.

Previous publications characterizing the genome of AA prostate cancer cases [9,20] did not report an increased frequency of *CHD1* loss as we observed in the FISH-based analysis presented above. Methods to detect copy number variations from WGS or WES data have at least two major limitations. First, subclonal copy number variations (sCNV) can be missed if they are present in fewer than 30%, of the cells[18]. Second, copy number loss can be underestimated with smaller deletions (e.g., <10 kb). Although various tools are available for inferring sCNVs from WES, WGS or SNP array data, such as TITAN[18], THetA[21] , and Sclust[22], they are designed to work on the entire genome, and likely miss small (~1-10kb) CNVs during the data segmentation process. In order to maximize the accuracy of our analysis we performed a gene focused analysis of the copy number loss in *CHD1*. We considered several factors such as the change in the normalized coverage in the tumors relative to their normal pairs', the cellularity of the tumor genome, and the approximate proportion of tumor cells exhibiting the loss. We also evaluated whether the deletion was heterozygous or homozygous using a statistical method designed for calling subclonal loss of heterozygosity (LOH) events within a confined genomic region (details are available in the Materials and Methods section, and in the Supplementary Material).

Using this approach in a large cohort (N=530 cases; 59 AA WES, 18AA WGS, 408 EA WES and 45 EA WGS, for details see supplementary material and Suppl. Figures 5-28), we observed that *CHD1* is more frequently deleted in AA tumors (N=20; 26%)

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

than in EA tumors (N=73 EA; 16%). Taken together, when next generation sequencing based copy number variations were analyzed with a more sensitive method, *CHD1* loss was detected more frequently in the AA cases than in the EA cases (p=0.029, Fisher exact test), which is consistent with our observations with FISH method in the TMA cohort.

CHD1 loss is associated with genomic signatures frequently observed in _BRCA2_ deficient prostate cancers.

CHD1 loss was shown to reduce HR competence in cell line model systems[14,23]. Detecting and quantifying HR deficiency in tumor biopsies is currently best achieved by analyzing whole genome sequencing data for specific HR deficiency associated mutational signatures. Those include: 1) A single nucleotide variation based mutational signature ("COSMIC signatures 3[24] and SBS3[25]); 2) a short insertions/deletions based mutational profile, often dominated by deletions with microhomology, a sign of alternative repair mechanisms joining double-strand breaks in the absence of HR, which is also captured by COSMIC indel signatures ID6 and ID8[25] ; 3) large scale rearrangements such as non-clustered tandem duplications in the size range of 1-100kb (mainly associated with *BRCA1* loss of function) or deletions in the range of 1kb-1Mb (mainly associated with *BRCA2* loss of function)[26]. All of these signatures can be efficiently induced by the inactivation of *BRCA1*, *BRCA2* or several other key downstream HR genes (Suppl. Figures 31-48) [27] .

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

HR deficiency is also assessed in the clinical setting by a large scale genomic aberration based signature, namely the HRD score[28], which is also approved as

Figure 2

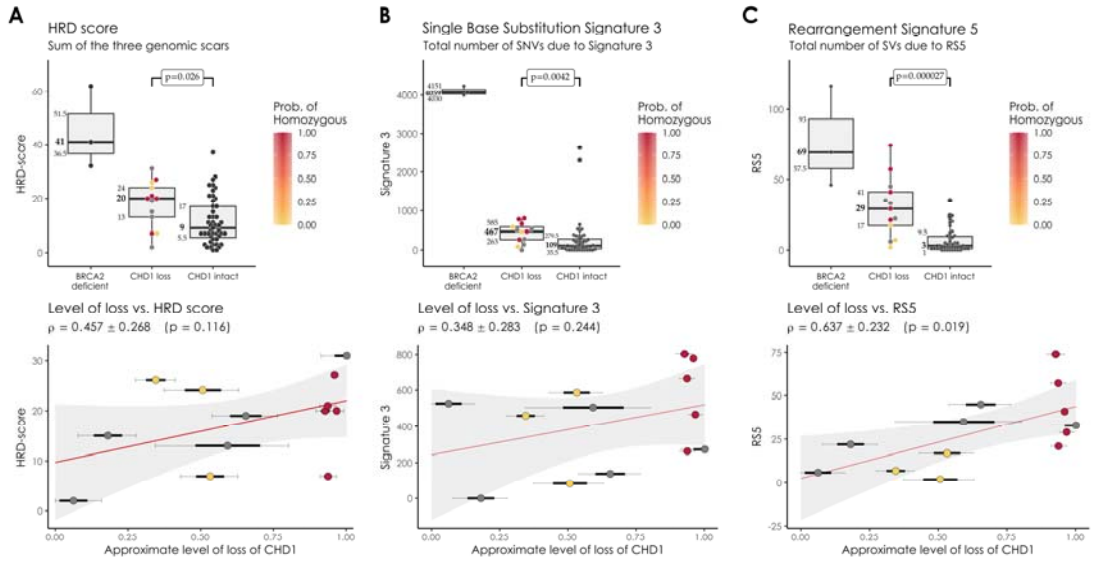

FIGURE 2: ***HRD markers in the PRAD WGS cohorts.*** **(A)** HRD-score, the sum of the three genomic scars, HRD-LOH, LST, and ntAI, **(B)** n*umber of somatic mutations due to single-base substitution signature 3,* ***(C)*** number of structural variants due to rearrangement signature 5.

The significance of the difference between the means of the "CHD1 loss" and "control" groups were assessed with Wilcoxon ranked sum tests. Below the box plots are the correlations between the approximate levels of loss in CHD1 and the HRD measures are visualized. The standard errors and the corresponding p-values of the correlation coefficients (Pearson) are also indicated. Horizontal lines indicate the uncertainty in the level of loss in each sample. Thick black lines correspond to the 66%, thin black error-bars to the 95% percentile intervals.

companion diagnostic for PARP inhibitor therapy. Recently a composite mutational signature, HRDetect[29], combining several of the mutational features listed above was evaluated as an alternative method to detect HR deficiency in prostate adenocarcinoma[30]. In order to further strengthen the link between *CHD1* loss, HR deficiency and potentially increased PARP inhibitor sensitivity we performed a

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

detailed analysis on the mutational signature profiles of *CHD1* deficient prostate cancer.

We analyzed whole exome and whole genome sequencing data of several prostate adenocarcinoma cohorts (For the detailed results see the Supplementary methods) containing samples both from AA (52 WES and 18 WGS cases) and EA (387 WES and 45 WGS cases) individuals in order to determine whether *CHD1* loss is associated with the HRD mutational signatures.

We divided the cohorts into three groups: 1) *BRCA2* deficient cases that served as positive controls for HR deficiency, 2) *CHD1* deleted cases without mutations in HR genes (Supplementary Figures 22,25 and 29), and 3) cases without *BRCA* gene aberration or *CHD1* deletion (for details see suppl material).

In the WGS cohorts *CHD1* deficient cases showed increased HRD score relative to the control cases but lower than the *BRCA2* deficient cases (Figure 2A). It is important to emphasize, however, that the HRD score was positively correlated with the estimated fraction of the subclonal loss of *CHD1* (Figure 2A, Suppl. Figure 30-31), suggesting that the signal of HRD score was "diluted" likely due to subclonality. The most characteristic HRD associated single nucleotide variation signature (signature 3), was significantly increased in the *BRCA2* deficient cases and slightly increased in the CHD1 deficient cases (Figure 2B).

The increase of the relative contribution of short indel signatures ID6 and ID8 to the total number of indels characteristic of loss of function on BRCA2 biallelic mutants was not observed in the CHD1 loss cases (Suppl. Fig. 36-39). This suggests,

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

that the alternative end-joining repair pathways do not dominate the repair of DSBs in those cases.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3

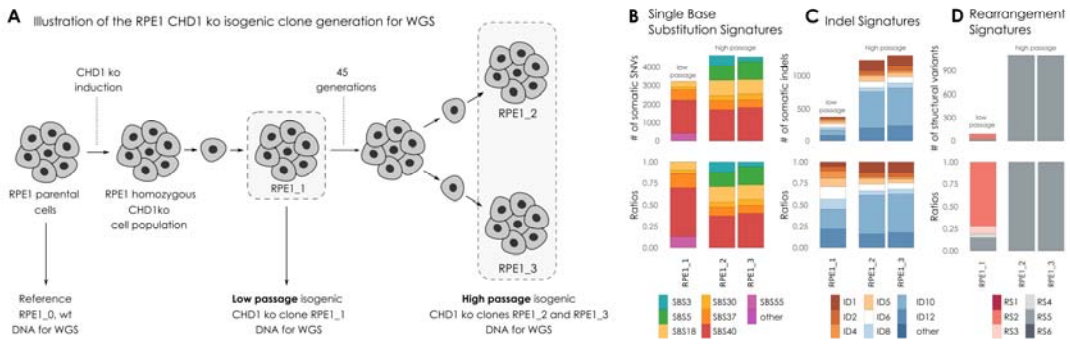

FIGURE 3: *RPE1 CHD1* ko cell line experiment and somatic signature extraction. (A) Illustration of the RPE1 CHD1 knock out (ko) isogenic clone generation for whole genome sequencing (WGS). DNA was extracted from RPE1 parental cell line (wild type, wt) and used as a reference genome. CHD1 ko was induced in parental RPE1 cell line. RPE1 CHD1 ko cell population was single cell cloned. Isogenic cell lines displaying homozygous CHD1 ko were identified. DNA was extracted directly from the regenerated population (RPE1_1, low passage stage). Cells were further propagated through 45 generations, then high passage cell population was single cell cloned. DNA was extracted from two isogenic CHD1 ko clones (RPE1_2 and RPE1_3, high passage stage) after propagation. **(B)** Single Nucleotide Substitution (SBS) signatures, **(C)** Indel signatures, **(D)** Rearrangement signatures. The number of mutations indicated originate from the reconstructed mutational spectra.

In the WGS cohort we also determined the number of structural variants as previously defined (Suppl.Fig. 40)[26]. As expected, RS5 was significantly increased in the *BRCA2* mutant cases since this signature (an increase in the number of non-clustered 1kb-1Mb deletions) was identified as a specific feature of such tumors. *CHD1* deficient cases also displayed a significant increase in RS5 structural variations but the signal showed a strong subclonal dilution (Figure 2C) suggesting that the number of RS5 aberrations may be similar in the *BRCA2* and *CHD1* deficient cases. Finally, the *BRCA2* deficient cases showed high HRDetect scores (Suppl.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figures 41-42). The HRDetect scores were also elevated relative to the controls but significantly lower than the previously published threshold for HR deficiency. However, since the HRDetect scores arise from a logistic regression, which involves the non-linear transformation of the weighted sum of its attributes, even slightly lower linear sums in the *CHD1* loss cases compared to the *BRCA2* mutant cases can result in substantially lower HRDetect scores (Suppl. Figure 43).

We have previously processed WES prostate adenocarcinoma data for the various HR deficiency associated mutational signatures[30]. When the *CHD1* deficient cases were compared to the *BRCA1/2* deficient and *BRCA1/2* intact cases we obtained results that were consistent with the WGS based results outlined above (Suppl. Figures 44-49).

Deleting *CHD1* in cell lines induce some aspects of homologous recombination deficiency-associated mutational signatures

In order to investigate the functional impact of the biallelic loss of CHD1 we created several CRISPR-Cas9 edited cell lines. DNA repair pathway aberration induced mutational signatures can be detected in cell lines by whole genome sequencing[27,31]. This analysis is more efficient if the starting cell line has a diploid genome, such as in the RPE1 (retinal pigment epithelium) cells in which we previously deleted *CHD1* using CRISPR-Cas9 editing (Suppl. Figure 50) [14]. We grew single cell clones from these cell lines for 45 generations to accumulate the genomic aberrations induced by CHD1 loss (Suppl. Figures 51-64). Two of such late passage

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

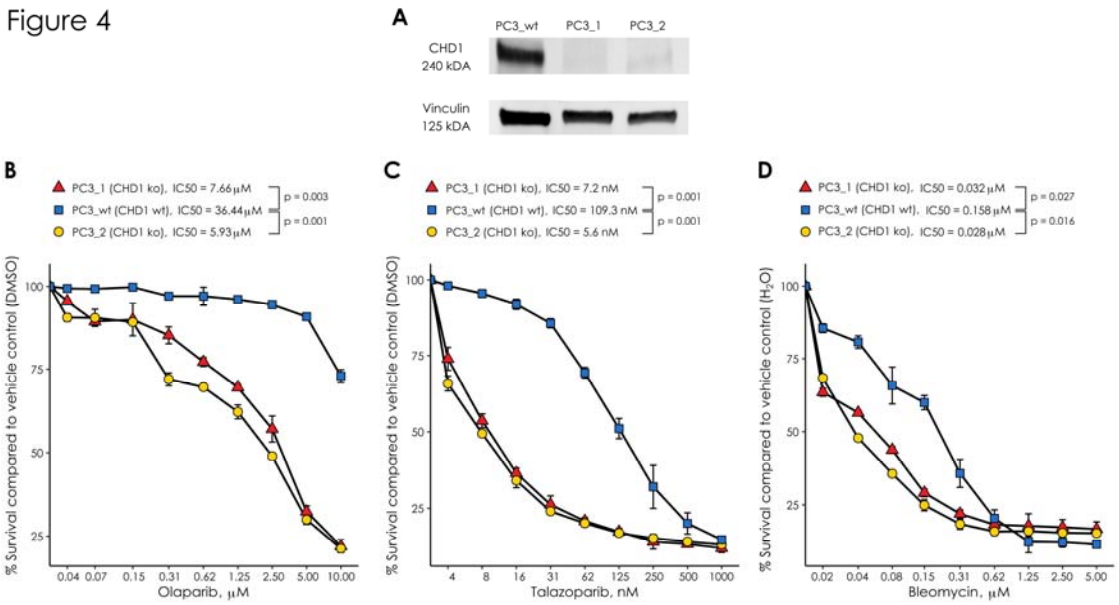

FIGURE 4: ***CHD1 loss PC-3 prostate cells show significant response to HR deficiency directed therapy.*** **(A)** Immunoblot shows that CHD1 was successfully knocked out in PC-3 cells. Sensitivity assays of parental wt and chd1 ko clones to PARP inhibitor Olaparib **(B)**, Talazoparib **(C)**, and the radiomimetics bleomycin sulfate **(D)**. Cells viability was measured using PresoBlue™ reagent. SD of triplicates are shown, p-values were calculated using student's t-test. p-values <0.05 were considered statistically significant.

clones and an early passage clone were subjected to WGS analysis (Figure 3A). All the clones retained the *BRCA2* wild type background of their parental clone.

CHD1 elimination induced some increase in SBS3 (HR deficiency associated) but more significant increase in SBS5, SBS18 (Figure 3B). Short indels flanked by microhomology (ID6 signature) constitute a robust sign of HR deficiency as a result of microhomology mediated-end joining (MMEJ) repair of DSBs in the absence of HR[32]. *CHD1* loss did not induce a significant increase of this mutational signature. Instead, the most significant increase was observed in ID10 (Figure 3C). Large genomic rearrangement signatures showed a significant increase in RS5, the

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

mutational signature strongly associated with *BRCA2* loss, which was detected in prostate cancer cases with *CHD1* deletion as shown in the previous section (Figure 3D).

Taken together, *CHD1* loss in cell line model systems mainly induced deletions of 1-10 kb, but only modestly induced the other types of mutational signatures that are associated with the loss of key members of the HR machinery.

CHD1 deficient cell lines show increased sensitivity to talazoparib and the radiomimetic agent bleomycin.

CHD1 deficient cancer cells have an increased sensitivity to the PARP inhibitor olaparib[14]. While this synthetic-lethal relationship is worth investigating, the olaparib sensitivity of *CHD1*-deficient is relatively mild. PARP inhibitors were initially thought to exert their therapeutic activity by inhibiting the enzymatic activity of PARP, but it was later revealed that trapped PARP on DNA may have a more significant contribution to cytotoxicity (reviewed in[33]). Therefore, we tested the efficacy of the strong PARP trapping agent talazoparib in the prostate cancer cell line PC-3 with or without CRISPR-Cas9-mediated *CHD1* deletion. *CHD1* knock out clones were identified by immunoblotting (Figure 4A). Consistent with previous reports, deleting CHD1 induced an approximately 5-fold increase in olaparib sensitivity (Figure 4B)[14]. In contrast, the sensitivity to talazoparib increased by about 15-20-fold in the same *CHD1* deficient PC-3 cells (Figure 4C). These data suggest that trapped PARP may have a more toxic effect in cells with *CHD1* deficiency.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Consistent with the significant functional evidence linking *CHD1* deletion and HR repair of DSBs, *CHD1* deficient cells also showed increased sensitivity to

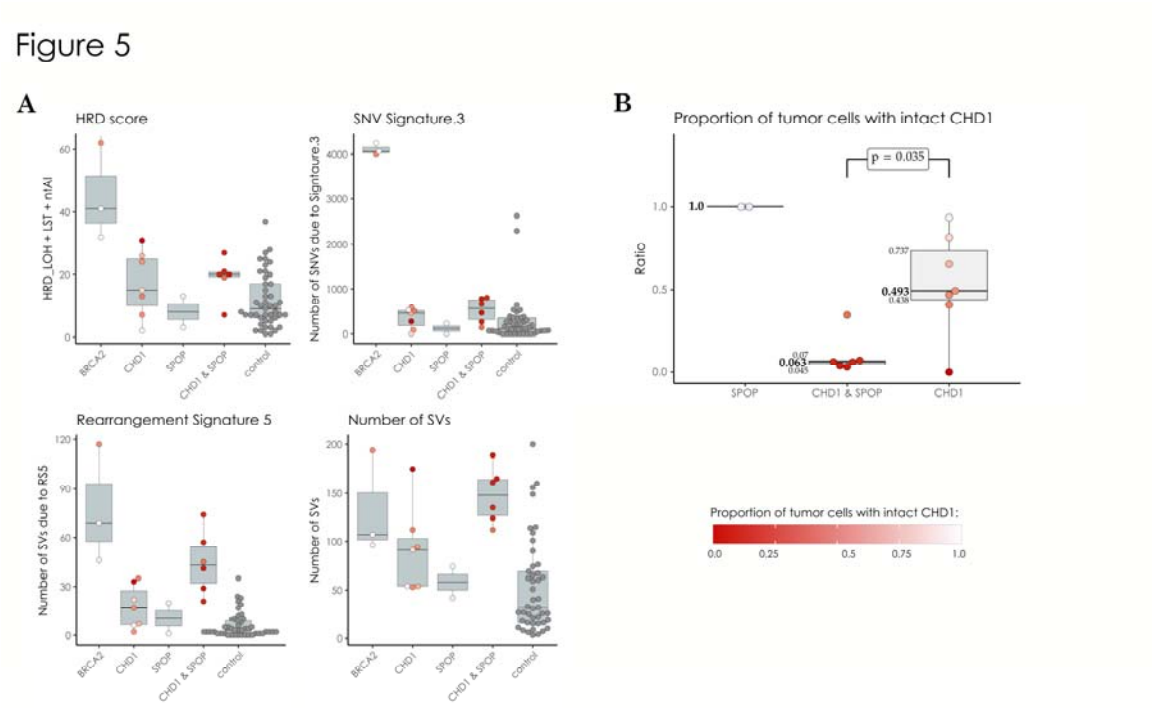

FIGURE 5: ***CHD1 loss and SPOP mutation in the WGS cohorts.*** **(A)** HRD-related markers and total number of structural variants in samples with mutations in SPOP, BRCA2 and loss in CHD1 versus the controls. Samples that simultaneously harbor mutations in SPOP and a loss in CHD1 tend to have higher markers. **(B)** Proportion of cells with intact CHD1 in SPOP mutants and samples identified with CHD1 loss. While the deletion in CHD1 in SPOP mutants is mostly clonal, in samples with wild type SPOP background it is mostly subclonal. The color-code for points in both panels A and B is illustrated in the bottom right corner of the figure.

irradiation[14]. We investigated, whether this increased sensitivity also applies to chemotherapy agents that induce DSBs, such as radiomimetic drugs. As shown in Figure 4D, *CHD1* deficient cells show significantly increased (5-fold) sensitivity to bleomycin, the most frequently used radiomimetic therapeutic agent.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The impact of SPOP mutations on the clonality of *CHD1* deletions and HR deficiency associated mutational signatures.

SPOP mutations and *CHD1* deletions show a strong tendency to co-exist in prostate cancer[34] and SPOP mutations have been shown to suppress key HR genes[16]. Therefore, we investigated whether the presence of *SPOP* mutation in a *CHD1* deficient prostate cancer is associated with a further increase of HR deficiency associated mutational signatures. We identified cases with *SPOP* mutations or *CHD1* deletions only, cases with both *SPOP* mutations and *CHD1* deletions and cases without either of those aberrations (Figure 5A). Cases with both mutations showed significantly higher levels of signature SBS3, RS5 and the total number of large-scale structural rearrangements relative to cases with either mutation alone. It should be noted, however, that the proportion of cells with *CHD1* deletions tended to be significantly higher in *SPOP* mutant cases than those with *CHD1* deletions without *SPOP* mutations. Thus, considering the previously demonstrated impact of *CHD1* subclonality on the intensity of HR deficiency associated mutational signatures (Figure 2), it is possible that the presence of *SPOP* will intensify HR deficiency associated mutational signatures by enhancing the proportion of *CHD1* deficient cells in a tumor (Figure 5B).

DISCUSSION

The presence of functionally relevant subclonal mutations in various solid tumor types is well documented[35,36]. Deletions present only in a minority of tumor cells

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

are difficult to detect unless more targeted analytical approaches are applied. Here we present one example of such detection bias with significant functional relevance. We used a FISH based approach to detect *CHD1* deletion in PC. Consistent with the previously described subclonal nature of *CHD1* loss, we found that while this gene is often deleted in prostate cancer, it is rarely deleted in every tumor cell. When we took the subclonal nature of *CHD1* loss into consideration a significant racial disparity emerged, with an approximately 3-fold increase in the frequency of CHD1 deletion in AA PC patients. This loss was also significantly associated with early biochemical recurrence. Since *CHD1* loss is associated with a more malignant phenotype, the significantly higher frequency of *CHD1* loss in AA PC may account for the diverging clinical course observed in PC between men of African and European Ancestry. It is possible that *CHD1* loss is in fact more frequent in EA PC as well but with a lower focal density than in AA cases. This is certainly a limitation of our study, but with the sensitivity thresholds we established the difference between AA and EA are significant.

Several studies pointed out a potentially intimate link between *CHD1* loss an homologous recombination deficiency[14,15,23]. Interestingly, *CHD1* null cells showed only a modest (3-fold) increase in sensitivity to PARP inhibitor or platinum-based therapy[14,15,23]. This suggested that *CHD1* loss may not lead to the same level or the same completeness of HR deficiency as that detected upon loss of function of *BRCA1* or *BRCA2*. The loss of function of those key HR genes usually leads to various DNA repair deficiencies such as stalled fork destabilization or reduced

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

capacity of DSB repair. The presence of those DNA repair deficiencies can often be detected by different types of DNA aberration profiles and they can be associated with an up to 1000-fold increase in PARP inhibitor sensitivity. The modest increase in PARP inhibitor sensitivity suggests that *CHD1* loss may lead to some but not all DNA repair aberrations usually associated with loss of function of *BRCA1/2*. Indeed, *CHD1* deficient tumors and cell line models displayed strong signals of the *BRCA2* deficiency associated structural variation signature (SV5), but only modest or no increase of the single nucleotide variation or short indel based signatures. This suggests, that *CHD1* loss "mimics" some but not all of the consequences of *BRCA2* deficiency. The precise mechanistic nature of this similarity needs further biochemical studies.

Identification of synthetic lethal agents with *CHD1* deficiency is expected to benefit those prostate cancer cases that harbor this aberration. In early clinical studies, patients with *CHD1* deficient prostate cancer responded to PARP inhibitor and platinum-based therapy[23]. However, the subclonal nature of *CHD1* loss we have highlighted here may have considerable clinical consequences. Tumors with *CHD1* loss in a significant subset of the cells may show significant response to HR deficiency directed therapy. HR deficiency associated mutational signatures are used to prioritize ovarian cancer patients for PARP inhibitor therapy and a similar strategy may be considered for prostate cancer as well. However, as we showed here, the HR deficiency associated mutational signatures are "diluted out" proportionally to the subclonality of *CHD1* loss. Therefore, diagnostic cut-off values

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

may need to be readjusted for *CHD1* deficient cases. If only a smaller subset of tumor cells harbor *CHD1* deficiency, then a synthetic lethal agent may have only a modest benefit in terms of tumor shrinkage. However, as it was suggested recently, *CHD1* loss may play a key role developing enzalutamide resistance[13]. Therefore, it is possible that eliminating the subset of *CHD1* deleted tumor cells, even if a minority, will significantly delay antiandrogen therapy resistance. Moreover, the majority of specimens analyzed for *CHD1* deletions in our study represent treatment naïve primary prostate tumor specimens. It would be reasonable to expect clonal expansion of *CHD1* deleted tumor cells in more advanced heavily treated metastatic and castration resistant prostate cancers (CRPC), which calls for further analysis. Therefore, agents with higher specificity for *CHD1* deficiency, such as talazoparib or bleomycin, may be used as agents to stave off resistance to antiandrogen therapy. Considering the higher frequency of *CHD1* loss in AA PC, such *CHD1* directed therapy may stop the early development of more malignant clones and may reduce the racial differences in the overall outcome of prostate cancer.

Materials and Methods:

Cohort selection and Tissue Microarray (TMA) generation: The aggregate cohort was composed of 2 independently selected cohort samples from Bio-specimen bank of Center for Prostate Disease Research and the Joint Pathology Center. Wholemount prostates were collected from 1996 to 2008 with minimal follow-up time of 10 years. The first cohort of 42 AA and 59 EA cases was described

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

before[6,37]. Similarly, the second cohort of 50 AA and 50 EA cases was selected based on the tissue availability (>1.0 cm tumor tissue) and tissue differentiation status (1/3 well differentiated, 1/3 moderately differentiated and 1/3 poorly differentiated). All the selected cases had the signed patient consent forms for tissue research applications. Patients who have donated tissue for this study also contributed to the long term follow-up data (mean 14.5 years). Our study was reviewed and approved by institutional review board (IRB) of WRNMMC and Uniformed Services University of the Health Sciences, Bethesda, MD. TMA block was assigned as 10 cases each slide and each case with 2 benign tissue cores, 2 Prostatic intraepithelial neoplasia (PIN) cores if available and 4-10 tumor cores covering the index and non-index tumors from formalin fixed paraffin embedded (FFPE) wholemount blocks. The description of numbers of patients, tumors and tumor cores of combined cohort was in Supplementary table 1d. All the blocks were sectioned into 8 μM tissue slides for FISH staining.

Fluorescence in situ hybridization (FISH) assay: A gene-specific FISH probe for CHD1 was generated by selecting a combination of bacterial artificial chromosome (BAC) clones (Thermo Fisher Scientific, Waltham, MA) within the region of observed deletions near 5q15-q21.1, resulting in a probe matching ca. 430 kbp covering the CHD1 gene as well as some upstream and downstream adjacent genomic sequences including the complete repulsive guidance molecule B (RGMB) gene. Due to the high degree of homology of chromosome 5-specific alpha satellite centromeric DNA to the centromere repeat sequences on other chromosomes, and the resulting potential for cross-hybridization to other centromere sequences,

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

particularly on human chromosomes 1 and 19, a control probe matching a stable genomic region on the short arm of chromosome 5 – instead of a centromere 5 probe - was used for chromosome 5 counting (supplementary figure 1e). The FISH assay of CHD1 was performed on TMA as previously described[6]. The green signal was from probe detecting control chromosome 5 short arm and the red signal was from probe detecting CHD1 gene copy. The FISH stained TMA slides were scanned with Leica Aperio VERSA digital pathology scanner for further evaluation. The criteria for CHD1 deletion was that in over 50% of counted cancer cells (with at least 2 copies of chromosome 5 short arm detected in one tumor cell) more than one copy of CHD1 gene had to be undetected. Examining tumor cores, deletions were called when more than 75% of evaluable tumor cells showed loss of allele. Focal deletions were called when more than 25% of evaluable tumor cells showed loss of allele or when more than 50% evaluable tumor cells in each gland of a cluster of two or three tumor glands showed loss of allele. Benign prostatic glands and stroma served as built-in control.

The sub-clonality of CHD1 deletion was presented with a heatmap showing CHD1 deletion status in all the given tumors sampled from whole-mount sections of each patient. The color designations were denoted as: red color (full deletion) meaning all the tumor cores carrying CHD1 deletion within a given tumor, yellow color (sub-clonal deletion) meaning only partial tumor cores carrying CHD1 deletion within a given tumor and green color (no deletion) meaning no tumor core carry CHD1 deletion (supplementary table 1b).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Statistics Analysis: The correlations of CHD1 deletion and clinic-pathological features, including pathological stages, Gleason score sums, Grade groups, margin status, and therapy status were calculated using an unpaired t-test or chi-square test. Gleason Grade Groups were derived from the Gleason patterns for cohort from Grade group 1 to Grade group 5. Due to the small sample sizes within each Grade group, Grade group 1 through Grade group 3 were categorized as one level as well as Grade group 4 through Grade group 5. A BCR was defined as either two successive post-RP PSAs of ≥0.2 ng/mL or the initiation of salvage therapy after a rising PSA of ≥0.1 ng/mL. A metastatic event was defined by a review of each patient's radiographic scan history with a positive metastatic event defined as the date of a positive CT scan, bone scan, or MRI in their record. The associations of CHD1 deletion and clinical outcomes with time to event outcomes, including BCR and metastasis, were analyzed by a Kaplan–Meier survival curves and tested using a log-rank test. Multivariable Cox proportional hazards models were used to estimated hazard ratios (HR) and 95% confidence intervals (Cis) to adjust for age at diagnosis, PSA at diagnosis, race, pathological tumor stage, grade group, and surgical margins. We checked the proportional hazards assumption by plotting the log-log survival curves. A P-value < 0.05 was considered statistically significant. Analyses were performed in R version 4.0.2.

Immunohistochemistry for ERG: ERG immunohistochemistry was performed as previously described[38]. Briefly, four μm TMA sections were dehydrated and blocked in 0.6% hydrogen peroxide in methanol for 20 min. and were processed for

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

antigen retrieval in EDTA (pH 9.0) for 30 min in a microwave followed by 30 min of cooling in EDTA buffer. Sections were then blocked in 1% horse serum for 40 min and were incubated with the ERG-MAb mouse monoclonal antibody developed at CPDR (9FY, Biocare Medical Inc.) at a dilution of 1:1280 for 60 min at room temperature. Sections were incubated with the biotinylated horse anti-mouse antibody at a dilution of 1:200 (Vector Laboratories) for 30 min followed by treatment with the ABC Kit (Vector Laboratories) for 30 min. The color was developed by VIP (Vector Laboratories,) treatment for 5 minutes, and the sections were counter stained by hematoxylin. ERG expression was reported as positive or negative. ERG protein expression was correlated with clinico-pathologic features.

Prostate cancer patients and specimens in the in-silico study cohorts

Evaluation of the self-declared ancestries: Since the available ancestry data were based on the self-assessment of the patients, and it was a crucial part of our study to identify the samples accurately, we have interrogated the genotypes of 3000 SNPs that are specific to one of the greater Caucasian, African and Asian ancestries, in each of the germline samples [39]. The data was collected into a single genotype matrix, the first two principal components of which was used to train a non-naïve Bayes classifier to differentiate between the three ancestries (details are available in the supplementary material, Supp. Figures 5-21).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Identification of local subclonal loss of CHD1 in prostate adenocarcinoma: The paired germline and tumor binary alignment (bam) files were analyzed using bedtools genomcecov (v2.28.0)[40], and their mean sequencing depths were determined. The coverage above and within the direct vicinity of CHD1 (*chr5:98,853,485-98,930,272 in grch38 and chr5:98,190,408-98,262,740 in grch37*) was collected in 50 bp wide bins into d-dimensional vectors (d_grch37 = 1447, d_grch38 = 1536) using an in-house tool and samtools (v1.6)[41], and were normalized using their corresponding mean sequencing depths. The linear relationship between the paired germline-tumor coverages were determined in the following form:

$$c_n = \alpha + \beta_0 c_t,$$

where $c_n$is the normalized coverage of the germline sample and $c_t$is the normalized coverage of its corresponding tumor pair. The intercept (α) was used to ensure that the data was free of outliers, and the slope (β_0) was used as a raw measure of the observable loss in the tumor. Similar slopes were calculated for 14 housekeeping genes in each of the sample-pairs, which were used to assess the significance of the loss (Supplementary Material).

The cellularity (c) of the tumors were estimated using sequenza[42] after the rigorous selection of the most reliable cellularity-ploidy pair offered by the tool as alternative solutions. In order to account for the uncertainty of the reported cellularity values, a beta distribution was fitted on the grid-approximated marginal posterior densities of c. These were used to simulate random variables to determine

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the proportion of the approximate loss of CHD1 in the tumors, by the following formula:

$$\beta_t = \frac{\beta - 1 + c}{c}$$

Here, $\beta \sim \text{Normal}(\beta_0, \sigma)$, where σ is the standard error of β_0, $c \sim \text{Beta}(s_1, s_2)$, where s_1 and s_2 are the fitted shape-parameters of the cellularity, and β_t is the cellularity-adjusted slopes of the curve. The approximate level of loss in CHD1 is distributed as 1-β_t (Further details are available in the supplementary materials, Suppl. Figures 22-27).

Local subclonal LOH-calling: The SNP variant allele frequencies (VAF) in the close vicinity of *CHD1* in the tumor were collected with GATK HaplotypeCaller (v4.1.0) [43]. The coverage and VAF data were carefully analyzed in order to ensure that we are strictly focusing on regions that have suffered the most serious loss (e.g., if only a part of the gene were lost, the unaffected region was excluded from the analysis). By using the tumor cellularity (c) and the estimated level of loss in the tumor (β_t), we assessed whether a heterozygous or a homozygous subclonal deletion is more likely to result in the observed frequency pattern (A detailed explanation is available in the supplementary notes, Suppl. Figure 30, Suppl. Tables 2-3).

Cell culture models. PC-3 prostate cell line was purchased from ATCC® and grown in RPMI 1640 (Gibco) supplemented with 10% FBS (Gibco). RPE1 WT and CHD1 knock out cells were provided by Jia Zhou as previously reported[14]. RPE1 cells

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

were incubated in DMEM with 10% FBS at 37°C in 5% CO_2, and regularly tested negative for Mycoplasma spp. contamination.

Stable CRISPR-Cas9 expressing isogenic PC-3 cell line generation. Full length SpCas9 ORF was introduced in PC-3 cell population by Lentiviral transduction using lentiCas9-Blast (Addgene #52962) construction. After antibiotics (blasticidin) selection, survival populations were single cell cloned, isogenic cell lines were generated and tested for Cas9 activity by cleavage assay.

Gene knock out induction. Generation of RPE1-*CHD1*$^{-/-}$ cells were reported previously[14]. CHD1 was targeted in CRISPR-Cas9 expressing PC-3 cell line using guide RNA CHD1_ex2_g1 (gCTGACTGCCTGATTCAGATC), resulting in PC-3 chd1 ko 1, and chd1 ko 2 homozygous loss cell lines.

Transfection. Cells were transiently transfected by Nucleofector® 4D device (Lonza) by using supplemented, Nucleofector® SF solution and 20 µl Nucleocuvette® strips following the manufacturer's instructions. Following transfection, cells were resuspended in 100 µl culturing media and plated in 1.5 ml pre-warmed culturing media in a 24 well tissue culture plate. Cells were subjected to further assays 72 h post transfection.

In vitro T7 EndonucleaseI (T7E1) Assay. Templates used for T7E1 were amplified by PCR using CGTCAACGATGTCACTAGGC forward and ATGATTTGGGGCTTTCTGCT reverse oligos generating a 946 bp amplicon. 500 ng PCR products were denatured

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

and reannealed in 1x NEBuffer 2.1 (New England Biolabs) using the following protocol: 95°C, 5 min; 95-85°C at -2°C/sec; 85-25°C at -0.1°C/sec; hold at 4°C. Hybridized PCR products were then treated with 10 U of T7E1 enzyme (New England Biolabs) for 30 min in a reaction volume of 30 µl. Reactions were stopped by adding 2 µl 0.5 M EDTA, fragments were visualized by agarose gel electrophoresis.

Immunoblot Analysis. Freshly harvested cells were lysed in RIPA buffer. Protein concentrations were determined by Pierce BCA™ Protein Assay Kit (Pierce). Proteins were separated via Mini Protean TGX stain free gel 4-15% (BioRad) and transferred to polyvinilydene difluoride membrane by using iBlot 2 PVDF Regular Stacks (Invitrogene) and iBlot system transfer system (LifeTechnologies). Membranes were blocked in 5% BSA solution (Sigma). Primary antibodies were diluted following the manufacturer's instructions: anti-Vinculin antibody (Cell Signaling) (1:1000) and antiCHD1 (Novus Biologicals) (1:2000).
Signals were developed by using Clarity Western ECL Substrate (BioRad) and Image Quant LAS4000 System (GEHealthCare).

Sample preparation for Whole Genome Sequencing (WGS).

RPE1 DNA was extracted from WT and *CHD1* knock out isogenic cell lines at low passage number of the cells. Following 45 passages, CHD1 knock out isogenic cell line was single cell cloned, and two colonies were propagated for DNA isolation.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DNA was extracted by using QIAamp DNA Mini Kit (QIAGENE). Whole Genome Sequencing of the DNA samples was carried out at Novogene and BGI service companies.

Viability cell proliferation assays.

Exponentially growing PC-3 cell lines WT, *CHD1* ko1, *CHD1* ko2 were seeded in 96-well plates (1000 cells/well) and incubated for 36 hrs to allow cell attachment. Identical cell numbers of seeded paralel isogenic lines were verified by the Celigo Imaging Cytometer after attachment. Cells were exposed to Talazoparib (Selleckchem), Olaparib (MedChemExpress) and Bleomycin sulfate (Fisher Scientific) for 24 hrs, then kept in drug-free fresh media for 5 days until cell growth was determined by the addition of PrestoBlue (Invitrogen) and incubated for 2.5 hrs. Cell viability was determined by using the BioTek plate reader system. Fluorescence was recorded at 560 nm/590 nm, and values were calculated based on the fluorescence intensity. IC50 values were determined by using the AAT Bioquest IC50 calculator tool. P-values were calculated using student's t-test. P-values <0.05 were considered statistically significant.

NGS analysis of the RPE1 whole genomes sequences

The reads of the four RPE1 WGS (1 parental and 3 CHD1 ko) were aligned to the grch37 reference genome using the bwa-mem[44] aligner. The resulting bam files were post-processed according to the GATK best-practices guidelines. Novel variants were called using Mutect2 (v4.1.0) by using the parental clone as "normal" and the *CHD1* ko clones as "tumor" specimens[43]. These vcfs were converted into

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

tab-delimited files and further analyzed in R. Annotation was performed via Intervar[45] .

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Acknowledgement

The authors thank Zita Bratu for technical assistance, Alimamy Bundu and Treissy Soares for FISH probe preparation and testing, Dr. Hua Zou, Audrey Flores and Safaa Khairi for valuable experimental support, and Orsolya Pipek for the technical support.

This work was supported by the Research and Technology Innovation Fund (KTIA_NAP_13-2014-0021 and NAP2-2017-1.2.1-NKP-0002); Breast Cancer Research Foundation (BCRF-17-156 to Z. Szallasi) and the Novo Nordisk Foundation Interdisciplinary Synergy Program Grant (NNF15OC0016584), Det Fri Forskningsrad (award number #7016-00345B; to Z. Szallasi); Department of Defense through the Prostate Cancer Research Program (award number is W81XWH-18-2-0056; to Z. Szallasi, A. Dobi and M.L. Freedman). Z. Szallasi, Z. Sztupinszki and J. Borcsok were supported by Velux Foundation 00018310 grant. S.K. is supported by the Prostate Cancer Foundation (18YOUN09 and 19CHAL07).

Disclaimer

The contents of this publication are the sole responsibility of the author(s) and do not necessarily reflect the views, opinions or policies opinions of Uniformed Services University of the Health Sciences (USUHS), the Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., the Department of Defense (DoD) or the Departments of the Army, Navy, or Air Force. Mention of trade

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

names, commercial products, or organizations does not imply endorsement by the U.S. Government.

References

[1] Gaines AR, Turner EL, Moorman PG, Freedland SJ, Keto CJ, McPhail ME, et al. The association between race and prostate cancer risk on initial biopsy in an equal access, multiethnic cohort. Cancer Causes Control 2014;25:1029–35. https://doi.org/10.1007/s10552-014-0402-6.

[2] Chu DI, Moreira DM, Gerber L, Presti JC, Aronson WJ, Terris MK, et al. Effect of race and socioeconomic status on surgical margins and biochemical outcomes in an equal-access health care setting: results from the Shared Equal Access Regional Cancer Hospital (SEARCH) database. Cancer 2012;118:4999–5007. https://doi.org/10.1002/cncr.27456.

[3] Khani F, Mosquera JM, Park K, Blattner M, O'Reilly C, MacDonald TY, et al. Evidence for molecular differences in prostate cancer between African American and Caucasian men. Clin Cancer Res 2014;20:4925–34. https://doi.org/10.1158/1078-0432.CCR-13-2265.

[4] Rosen P, Pfister D, Young D, Petrovics G, Chen Y, Cullen J, et al. Differences in frequency of ERG oncoprotein expression between index tumors of Caucasian and African American patients with prostate cancer. Urology 2012;80:749–53. https://doi.org/10.1016/j.urology.2012.07.001.

[5] Sedarsky J, Degon M, Srivastava S, Dobi A. Ethnicity and ERG frequency in prostate cancer. Nat Rev Urol 2018;15:125–31. https://doi.org/10.1038/nrurol.2017.140.

[6] Petrovics G, Li H, Stümpel T, Tan S-H, Young D, Katta S, et al. A novel genomic alteration of LSAMP associates with aggressive prostate cancer in African American men. EBioMedicine 2015;2:1957–64. https://doi.org/10.1016/j.ebiom.2015.10.028.

[7] Koga Y, Song H, Chalmers ZR, Newberg J, Kim E, Carrot-Zhang J, et al. Genomic Profiling of Prostate Cancers from Men with African and European Ancestry. Clin Cancer Res 2020;26:4651–60. https://doi.org/10.1158/1078-0432.CCR-19-4112.

[8] Mahal BA, Alshalalfa M, Kensler KH, Chowdhury-Paulino I, Kantoff P, Mucci LA, et al. Racial Differences in Genomic Profiling of Prostate Cancer. N Engl J Med 2020;383:1083–5. https://doi.org/10.1056/NEJMc2000069.

[9] Huang FW, Mosquera JM, Garofalo A, Oh C, Baco M, Amin-Mansour A, et al. Exome Sequencing of African-American Prostate Cancer Reveals Loss-of-Function ERF Mutations. Cancer Discov 2017;7:973–83. https://doi.org/10.1158/2159-8290.CD-16-0960.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[10] Faisal FA, Kaur HB, Tosoian JJ, Tomlins SA, Schaeffer EM, Lotan TL. SPINK1 expression is enriched in African American prostate cancer but is not associated with altered immune infiltration or oncologic outcomes post-prostatectomy. Prostate Cancer Prostatic Dis 2019;22:552–9. https://doi.org/10.1038/s41391-019-0139-0.

[11] Burkhardt L, Fuchs S, Krohn A, Masser S, Mader M, Kluth M, et al. CHD1 is a 5q21 tumor suppressor required for ERG rearrangement in prostate cancer. Cancer Res 2013;73:2795–805. https://doi.org/10.1158/0008-5472.CAN-12-1342.

[12] Augello MA, Liu D, Deonarine LD, Robinson BD, Huang D, Stelloo S, et al. CHD1 Loss Alters AR Binding at Lineage-Specific Enhancers and Modulates Distinct Transcriptional Programs to Drive Prostate Tumorigenesis. Cancer Cell 2019;35:817–9. https://doi.org/10.1016/j.ccell.2019.04.012.

[13] Zhang Z, Zhou C, Li X, Barnes SD, Deng S, Hoover E, et al. Loss of CHD1 Promotes Heterogeneous Mechanisms of Resistance to AR-Targeted Therapy via Chromatin Dysregulation. Cancer Cell 2020;37:584-598.e11. https://doi.org/10.1016/j.ccell.2020.03.001.

[14] Zhou J, Li J, Serafim RB, Ketchum S, Ferreira CG, Liu JC, et al. Human CHD1 is required for early DNA-damage signaling and is uniquely regulated by its N terminus. Nucleic Acids Res 2018;46:3891–905. https://doi.org/10.1093/nar/gky128.

[15] Kari V, Mansour WY, Raul SK, Baumgart SJ, Mund A, Grade M, et al. Loss of CHD1 causes DNA repair defects and enhances prostate cancer therapeutic responsiveness. EMBO Rep 2016;17:1609–23. https://doi.org/10.15252/embr.201642352.

[16] Hjorth-Jensen K, Maya-Mendoza A, Dalgaard N, Sigurðsson JO, Bartek J, Iglesias-Gato D, et al. SPOP promotes transcriptional expression of DNA repair and replication factors to prevent replication stress and genomic instability. Nucleic Acids Res 2018;46:9484–95. https://doi.org/10.1093/nar/gky719.

[17] Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, et al. Punctuated evolution of prostate cancer genomes. Cell 2013;153:666–77. https://doi.org/10.1016/j.cell.2013.03.021.

[18] Ha G, Roth A, Khattra J, Ho J, Yap D, Prentice LM, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. Genome Res 2014;24:1881–93. https://doi.org/10.1101/gr.180281.114.

[19] Han M, Partin AW, Zahurak M, Piantadosi S, Epstein JI, Walsh PC. Biochemical (prostate specific antigen) recurrence probability following radical prostatectomy for clinically localized prostate cancer. J Urol 2003;169:517–23. https://doi.org/10.1097/01.ju.0000045749.90353.c7.

[20] Yuan J, Kensler KH, Hu Z, Zhang Y, Zhang T, Jiang J, et al. Integrative comparison of the genomic and transcriptomic landscape between prostate cancer patients of predominantly African or European genetic ancestry. PLoS Genet 2020;16:e1008641. https://doi.org/10.1371/journal.pgen.1008641.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[21] Oesper L, Mahmoody A, Raphael BJ. THetA: inferring intra-tumor heterogeneity from high-throughput DNA sequencing data. Genome Biol 2013;14:R80. https://doi.org/10.1186/gb-2013-14-7-r80.
[22] Cun Y, Yang T-P, Achter V, Lang U, Peifer M. Copy-number analysis and inference of subclonal populations in cancer genomes using Sclust. Nat Protoc 2018;13:1488–501. https://doi.org/10.1038/nprot.2018.033.
[23] Shenoy TR, Boysen G, Wang MY, Xu QZ, Guo W, Koh FM, et al. CHD1 loss sensitizes prostate cancer to DNA damaging therapy by promoting error-prone double-strand break repair. Ann Oncol 2017;28:1495–507. https://doi.org/10.1093/annonc/mdx165.
[24] Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21. https://doi.org/10.1038/nature12477.
[25] Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. Nature 2020;578:94–101. https://doi.org/10.1038/s41586-020-1943-3.
[26] Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47–54. https://doi.org/10.1038/nature17676.
[27] Póti Á, Gyergyák H, Németh E, Rusz O, Tóth S, Kovácsházi C, et al. Correlation of homologous recombination deficiency induced mutational signatures with sensitivity to PARP inhibitors and cytotoxic agents. Genome Biol 2019;20:240. https://doi.org/10.1186/s13059-019-1867-0.
[28] Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, et al. Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. Clin Cancer Res 2016;22:3764–73. https://doi.org/10.1158/1078-0432.CCR-15-2477.
[29] Davies H, Glodzik D, Morganella S, Yates LR, Staaf J, Zou X, et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nat Med 2017;23:517–25. https://doi.org/10.1038/nm.4292.
[30] Sztupinszki Z, Diossy M, Krzystanek M, Borcsok J, Pomerantz MM, Tisza V, et al. Detection of Molecular Signatures of Homologous Recombination Deficiency in Prostate Cancer with or without BRCA1/2 Mutations. Clin Cancer Res 2020;26:2673–80. https://doi.org/10.1158/1078-0432.CCR-19-2135.
[31] Zámborszky J, Szikriszt B, Gervai JZ, Pipek O, Póti Á, Krzystanek M, et al. Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions. Oncogene 2017;36:746–55. https://doi.org/10.1038/onc.2016.243.
[32] Feng W, Simpson DA, Carvajal-Garcia J, Price BA, Kumar RJ, Mose LE, et al. Genetic determinants of cellular addiction to DNA polymerase theta. Nat Commun 2019;10:4286. https://doi.org/10.1038/s41467-019-12234-1.
[33] Murai J, Pommier Y. PARP Trapping Beyond Homologous Recombination and Platinum Sensitivity in Cancers. Annu Rev Cancer Biol 2019;3:131–50. https://doi.org/10.1146/annurev-cancerbio-030518-055914.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21251199; this version posted February 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[34] Barbieri CE, Bangma CH, Bjartell A, Catto JWF, Culig Z, Grönberg H, et al. The mutational landscape of prostate cancer. Eur Urol 2013;64:567–76. https://doi.org/10.1016/j.eururo.2013.05.029.
[35] Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med 2015;21:751–9. https://doi.org/10.1038/nm.3886.
[36] Linch M, Goh G, Hiley C, Shanmugabavan Y, McGranahan N, Rowan A, et al. Intratumoural evolutionary landscape of high-risk prostate cancer: the PROGENY study of genomic and immune parameters. Ann Oncol 2017;28:2472–80. https://doi.org/10.1093/annonc/mdx355.
[37] Merseburger AS, Kuczyk MA, Serth J, Bokemeyer C, Young DY, Sun L, et al. Limitations of tissue microarrays in the evaluation of focal alterations of bcl-2 and p53 in whole mount derived prostate tissues. Oncol Rep 2003;10:223–8.
[38] Furusato B, Tan S-H, Young D, Dobi A, Sun C, Mohamed AA, et al. ERG oncoprotein expression in prostate cancer: clonal progression of ERG-positive tumor cells and potential for ERG-based stratification. Prostate Cancer Prostatic Dis 2010;13:228–37. https://doi.org/10.1038/pcan.2010.23.
[39] Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res 2017;45:D840–5. https://doi.org/10.1093/nar/gkw971.
[40] Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841–2. https://doi.org/10.1093/bioinformatics/btq033.
[41] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–9. https://doi.org/10.1093/bioinformatics/btp352.
[42] Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann Oncol 2015;26:64–70. https://doi.org/10.1093/annonc/mdu479.
[43] 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;20:1297–303. https://doi.org/10.1101/gr.107524.110.
[44] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–60. https://doi.org/10.1093/bioinformatics/btp324.
[45] Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet 2017;100:267–80. https://doi.org/10.1016/j.ajhg.2017.01.004.