AWARD NUMBER: W81XWH-18-2-0056

TITLE: Frequent Loss of CHD1 in the Prostate Cancer of African Americans and Its Potential Role in Increased Sensitivity to Platinum or PARP Inhibitor-Based Therapy

PRINCIPAL INVESTIGATOR: Dr. Zoltan Szallasi,

CONTRACTING ORGANIZATION: Children's Hospital, Boston

REPORT DATE: OCTOBER 2021

TYPE OF REPORT: ANNUAL

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
OCTOBER 2021	Annual	30SEPT2020 - 29SEPT2021

4. TITLE AND SUBTITLE

Frequent Loss of CHD1 in the Prostate Cancer of African Americans and Its Potential Role in Increased Sensitivity to Platinum or PARP Inhibitor-Based Therapy

5a. CONTRACT NUMBER
W81XWH-18-2-0056

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Zoltan Szallasi MD

E-Mail:Zoltan.szallasi@childrens.harvard.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Children's Hospital, Boston
Massachusetts, 02215

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT African American individuals with prostate cancer have significantly worse clinical outcome than European Americans with the same disease. In our preliminary analysis, we identified a specific molecular aberration that is present in African American prostate cancer cases two to three times more frequently than in European Americans with the same disease. This aberration, the loss of the CHD1 gene, is thought to lead to impaired homologous recombination and increased genomic instability. Homologous recombination (HR) deficiency can be therapeutically exploited by either platinum or PARP inhibitor-based therapy. Therefore, we are performing a comprehensive analysis of the clinical consequences of the increased prevalence of CHD1 loss in the prostate cancer of African American individuals. We are investigating the relationship between CHD1 loss and the HRD mutational signatures.A robust association between CHD1 loss and HR deficiency may be the mechanistic basis for effective, platinum or PARP inhibitor-based personalized therapy for African American prostate cancer cases.

15. SUBJECT TERMS:
African American men, prostate cancer, CHD1 gene, PARP inhibitor, platinum-based therapy

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRDC
Unclassified	Unclassified	Unclassified	Unclassified	66	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

Introduction:

African American individuals with prostate cancer have significantly worse clinical outcome than European Americans with the same disease. In our preliminary analysis, we identified a specific molecular aberration that is present in up to 40% of African American prostate cancer cases, two to three times more frequently than in European Americans with the same disease. This aberration, the loss of the CHD1 gene, is thought to lead to impaired homologous recombination and increased genomic instability. Homologous recombination (HR) deficiency can be therapeutically exploited by either platinum or PARP inhibitor-based therapy. The main aim of this proposal is to establish a firm link between CHD1 loss, increased homologous recombination deficiency and increased sensitivity to targeted therapy. This could lead to improved therapy for African American individuals with advanced prostate cancer.

Keywords:

African American men, prostate cancer, CHD1 gene, PARP inhibitor, platinum-based therapy

Accomplishments:

Major developments mainly associated with Children's Hospital, Boston as outlined in Specific Aim 2: Investigating the relationship between CHD1 loss and the HRD mutational signatures

Subtask 3: Download and process whole genome sequencing data from all available prostate cancer cases. Extract the various homologous recombination deficiency (HRD) measures from the sequencing data and correlate those with CHD1 loss.

Subtask 4: Catalogue all mutational signatures in the PRAD WGS data sets and identify those cases whose HRD signatures cannot be associated with a loss of function mutation, deletion in an HRD associated gene. Identify other possible causes of HR deficiency by RNASeq analysis and other strategies.

1) **Major activity:** We have downloaded and processed raw whole genome sequencing data (BAM files) from 311 prostate cancer samples. We have called germline mutations with HaplotypeCaller, we determined somatic point-mutations and short indels using Mutect2 (GATK 3.8). The high fidelity of the reported variants was ensured by the application of additional hard filters on top of the tools' default ones. Allele-specific copy number profiles had been estimated by using Sequenza. Structural Variants were called using BRASS (v6.0.0 - https://github.com/cancerit/BRASS).
The mutations were annotated using InterVar. Variants predicted as pathogenic or likely pathogenic were considered deleterious, while variants with unknown significance were marked differently. Copy number status of BRCA1/2 were based on Sequenza results.

Somatic point-mutational signatures were determined with the deconstructSigs R package, by using the cosmic signatures as a mutational-process matrix. The extraction of rearrangement signatures was executed as described in (Nik-Zainal S et al. Nature 2016;534:47–54.)
The calculation of the genomics scar scores (loss-of-heterozygosity: LOH, large scale transitions: LST and number of telomeric allelic imbalances: ntAI.) were determined using the scarHRD R package.
Due to the lack of sufficient numbers of bona fide HR-deficient cases within the prostate cancer cases, a prostate-specific HRDetect model could not be created at this time. Instead, the weights of the original, breast cancer-specific, whole genome-based HRDetect model (Davies H, Nat Med 2017;23:517–25) were used to calculate the HRDetect scores of the WGS prostate samples.
It should be noted that the processing of a single WGS case is about 2-3 weeks (depending on coverage). Of course, we are parallel processing the cohorts. Nevertheless, the computational investment should be acknowledged.

2) Specific objectives: The main objective was to set up robust methods to detect and quantify homologous recombination deficiency in prostate cancer. HR deficiency was mainly studied so far through mutational signatures in breast and ovarian cancer and we had to validate and benchmark all the previously described methods for the analysis of prostate cancer derived next generation sequencing data (both in EA and AA patients.) Such analysis would then allow us to correlate CHD1 loss with the downstream mutational signatures of HR deficiency.

3) Significant results:

CHD1 loss recapitulates genomic signatures frequently observed in BRCA2 deficient prostate cancer.

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 signature 3” or “BRCA signature” or SBS3); 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 homologous recombination, which is also captured by COSMIC indel signatures ID6 and ID8 ; 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).
HR deficiency is also assessed in the clinical setting by a large-scale genomic aberration-based signature, the HRD score, which is also approved as companion diagnostic for PARP inhibitor therapy. Recently a composite mutational signature, HRDetect, combining several of the mutational features listed above was evaluated as an alternative method to detect HR deficiency in prostate adenocarcinoma.

We analyzed whole exome and whole genome sequencing data of several prostate adenocarcinoma cohorts containing samples both from AA (52 WES and 18 WGS cases) and CA (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, and 3) cases without BRCA gene aberration or CHD1 deletion.

In the WGS cohorts CHD1 deficient cases showed increased HRD score relative to the control cases but lower than the BRCA2 deficient cases. (Figure 1)

Figure 1

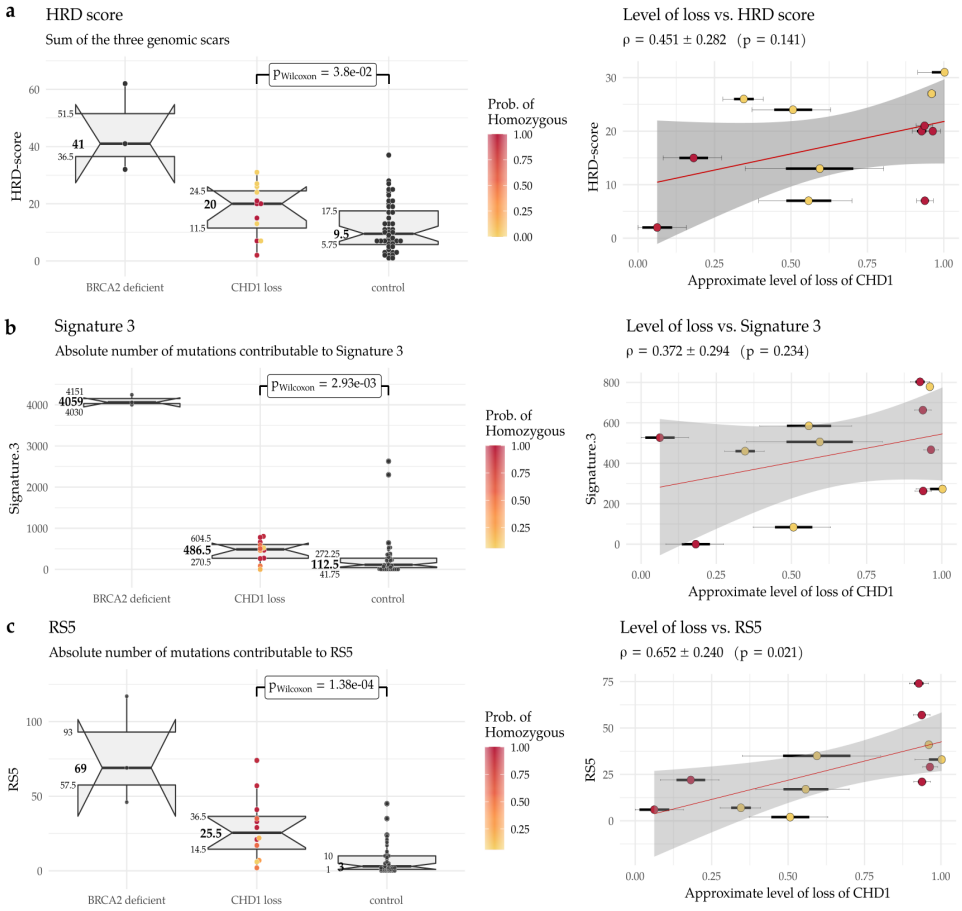


Figure 1. ***HRD markers in the PRAD WGS***
a) HRD-score, the sum of the three genomic scars, HRD-LOH, LST, and ntAI. **b)** number of somatic mutations contributable to single-base substitution signature 3, **c)** number of structural variants contributable to rearrangement signature 5.
The significance of the difference between the means of the "CHD1 loss" and "control" groups were assessed using Wilcoxon ranked sum tests. Next to the box plots are the correlations between the approximate levels of loss in CHD1 and the HRD measures are displayed. 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.

It is important to emphasize, however, that the HRD score was positively correlated with the estimated fraction of the sublonal loss of CHD1, suggesting that the signal of HRD score was "diluted" due to subclonality. The most characteristic HRD associated single nucleotide variation signature (signature 3), was significantly increased in the BRCA2 deficient cases and somewhat increased in the CHD1 deficient cases (figure 1b). Signature 5 was also slightly elevated both in the BRCA2 and CHD1 deficient cases. Similar results were obtained with version 3 signatures. In the WGS cohort we also determined the number of structural variants. 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 (Fig 1c) 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. 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.

We processed WES prostate adenocarcinoma data for the various HR deficiency associated mutational signatures. 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.

Tasks specific for the Subaward at Henry M. Jackson Foundation for the Advancement of Military Medicine:

Task #1: Detection of prevalence of CHD1 deletion and its correlation to more aggressive prostate cancer subtype in cohort of Center for Prostate Disease Research.

- Obtain IRB approval (mths 1-4)
- Construct tumor tissue microarrays (TMA), TMA-1: 42AA and 59CA and TMA-2 100AA and 100CA prostate cancers (mths 5-10)
- Assess CHD1 deletion frequency by Fluorescence In Situ Hybridization (FISH), compare to other cancer gene defects (ERG, PTEN, LSAMP) (mths 11-16)
- Perform association analyses with clinical-pathological features (mths 17-21)

Task #2: Detection of prevalence of CHD1 deletion in carboplatin treated cohort of castration resistant PCa cases from Dana Farber Cancer Institute.

- Prepare TMA from prostate specimens of patients treated with carboplatin at DFCI and assess by FISH and IHC (mths 6-7)

Task #3 (Subaward PI): Whole Genome Sequencing, data analysis of 35 CPDR AA PCa patients.

- Obtain IRB approval (mths 1-4)
- Complete Whole Genome Sequencing of prostate tumors of 35 AA patients (mths 5-16)

Task #4 (PI and Subaward PI): Assess patient derived AA patient derived cell lines with CHD1 loss for PARP inhibitor and platinum sensitivity.

- Characterize nine AA patient derived CPDR cell lines for CHD1 and for other cancer genes by FISH and IHC (mths 19-30)

1) Major activity: to obtain IRB approvals to perform Tasks (mths 1-9) followed by the initial and secondary reviews and approval by the USAMRMC, Office of Research Protections, Human Research Protection Office (mths 9-12) and perform FISH analysis.

Specific objectives: to obtain IRB approval from the Walter Reed National Military Medical Center IRB for: 1) generating TMA from prostate cancer tissue specimens archived in the CPDR Biobank; 2) IRB approval for using patient derived immortalized prostate cancer cell lines from the CPDR Biobank; 3) approval for utilizing whole genome sequences that have been generated under a CPDR data banking protocol. For the use of follow up data up to 20 years, we also obtained approval from the Uniform Services University of Health Sciences.

2) Significant results

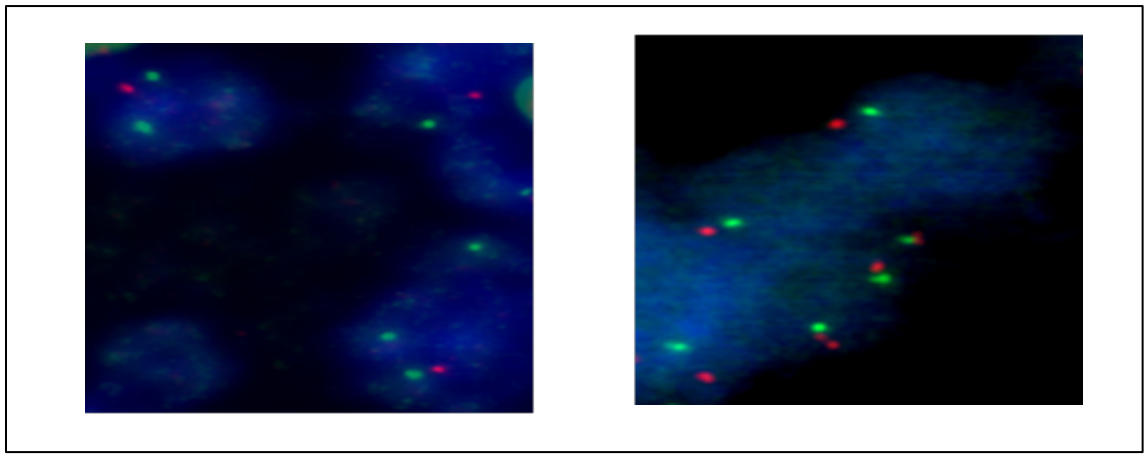
Subclonal CHD1 deletion is more frequent in African American prostate cancer cases and associated with worse clinical outcome.

CHD1 is frequently subclonally deleted in prostate cancer. Our preliminary analysis on the SNP array data from TCGA comparing African American and Caucasian American prostate cancer cases suggested that the subclonal loss of CHD1 may be a more frequent event in AA men. To independently validate this observation, we assessed CHD1 copy number by FISH in tissue microarrays (TMAs) in a cohort of 91 AA and 109 CA patients from the equal-access military healthcare system to minimize ascertainment criteria between the groups. 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 CA cases. Consistent with the long-term follow up, we observed a 40% biochemical recurrence (BCR) and 16% metastasis rate14. For each clinical case up to four different cancerous areas were analyzed comprising 4-10 different cores.

Our FISH data confirmed the subclonal nature of CHD1 deletion in prostate cancer cells. In most cases CHD1 deletion was present in only a subset of tumor cores. We detected subclonal CHD1 loss in 27 out of 91 AA cases (29.7%), and 14 out of 109 (11%) CA cases indicating that CHD1 deletion is about 3 times

more common in AA men. 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 CA prostate cancer cases) confirming previously described differences

.



Race	Deletion (N=41)	No deletion (N=159)	P value
AA (N=91)	27 (29.7%)	64 (70.3%)	0.003
EA (N=109)	14 (12.8%)	95 (87.2%)	

Figure 2a. Prostate cancer cells harboring mono-allelic deletion for *CHD1* (upper left) vs. prostate cancer cells with wild type (diploid) *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. 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.

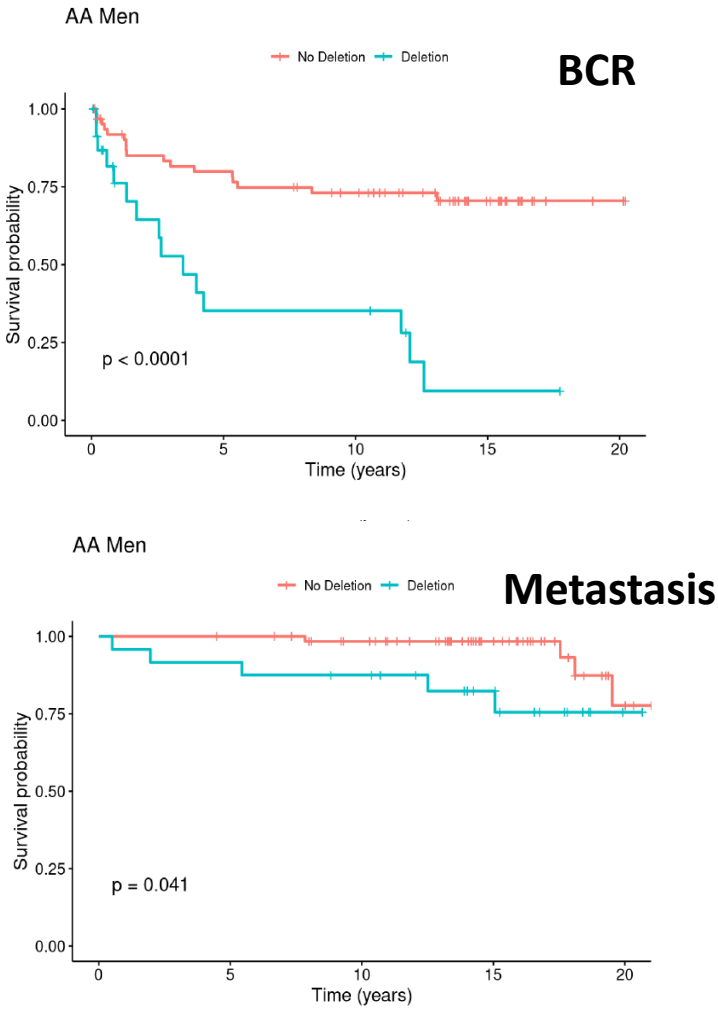


Figure 2b. Deletion of *CHD1* is strongly associated with disease progression in AA prostate cancer patients (N=91). BCR: univariable Kaplan-Meier curve; Metastasis: univariable Kaplan-Meier curve.

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. Prostate cancer cases with higher Gleason sum scores (3+4, 4+3, 8-10) were found more frequent in deletion group than non-deletion group. In contrast, lower Gleason sum score (3+3) was more often detected in non-deletion cases. Notably, CHD1 deletion was strongly associated with significantly earlier biochemical recurrence (figure 2b) with univariable Kaplan-Meier curve in AA cases (P<0.0001). The multivariable Cox model analysis showed that CHD1 deletion was an independent predictor of BCR in AA prostate cancer patients (P=0.0006)). Moreover, a significant correlation between CHD1 deletion and metastasis was also detected in the AA patients with univariable Kaplan-Meier curve model (P=0.041) and multivariable logistic regression predicting model. Taken together, our data indicate strong correlations between CHD1 deletion and aggressive prostate cancer and worse clinical outcomes in AA PC.

Other achievements:

Identified the AA patient derived CPDR RC165 cell line for CHD1 deletion: Encouraged by the high quality of CHD1 detection we have performed FISH assay for CHD1 on selected AA and CA patient-derived cell lines. Reviewing our previous data on the transcriptome of the AA prostate derived CPDR RC165 hTERT immortalized cell line (Kim, Dobi et al., PCPD 2007) we have noted gene expression signatures of prostate cancer progression resembling the consequences of homologous recombination defects. Thus, we evaluated the AA patient derived RC165 cell line for CHD1 deletion by FISH assay (Figure 2.). The experiment result indicates the mono-allelic deletion of CHD1 gene in this cell line.

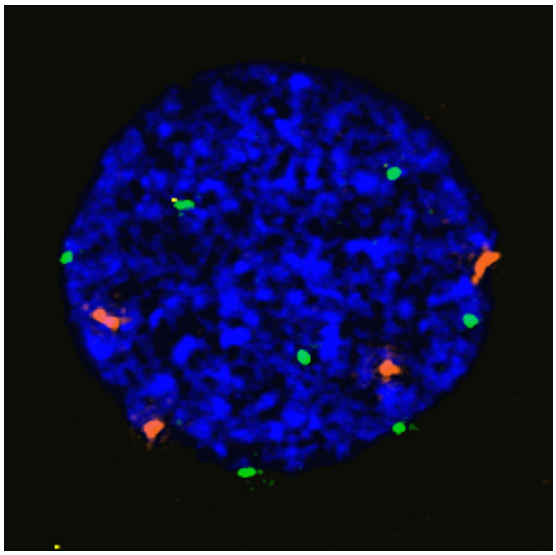


Figure 2. Allelic loss of CHD1 in the African American patient-derived CPDR cell line, RC165. The RC165 cell line is aneuploid for chromosome 5 (seven centromeric signals (green), whereas, only four CHD1 copy was detected in the cell nucleus (orange probe). Representative view of cell nucleus shown at 60X magnification.

Major developments mainly associated with the Dana Farer Cancer Institute site:

Specific Aim 3: Functional evaluation whether HRD mutational signatures can be directly induced in prostate cancer cell lines**.**

Major Task 1: To functionally evaluate if the HRD mutational signature can be induced in prostate cancer cell lines by deleting the CHD1 gene using clustered regularly interspaced short palindromic repeats (CRISPR) technology**.**

Subtask 1: In order to increase CRISPR based cleavage efficiency, we will create stable Cas9 protein expressing cell lines the PC-3 cell line.

Subtask 2: the CHD1 gene will be deleted using CRISPR technology both in a heterozygous and homozygous fashion. Single cell colonies will be isolated, validated and then grown for 50-100 generations to induce sufficient number of genomic aberrations to be detected by whole genome sequencing

Major activity

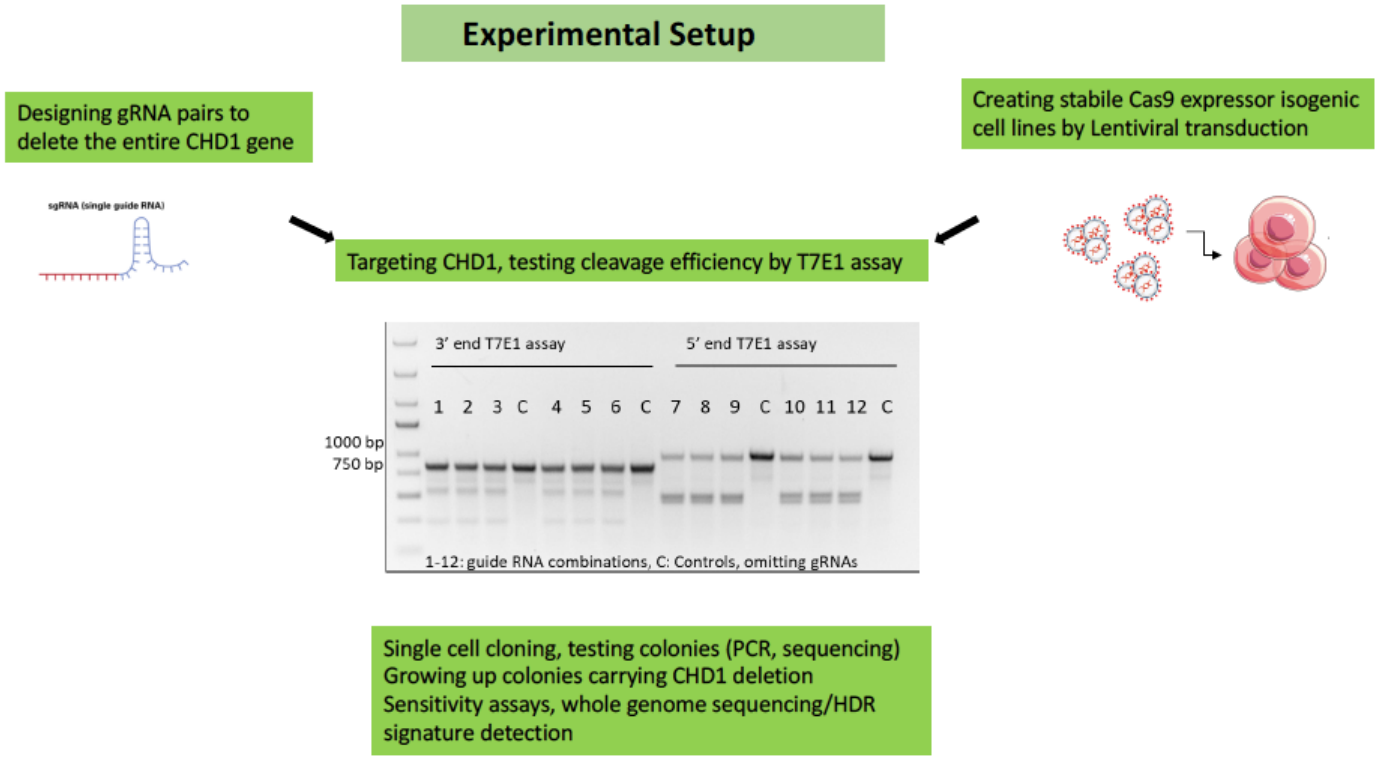
A stable Cas9 expressing version of the PC3 cell line was created by Lentiviral transduction.

After sufficient propagation, the stable Cas9 cell population was single cell cloned, and isogenic cell lines tested for Cas9 activity.

CHD1 deletion was induced by transfection using guide RNA pairs targeting the entire CHD1 gene. T7E1 assay protocol was optimized for testing Cas9 cleavage efficiency.

Single cell cloning: Cells were filtrated and plated 3 days after transfection into 20% FBS containing media with 1000, or 2000 cells per 10 cm dish (Corning). After 14-28 days, the formed colonies were picked up and growing in 96 well tissue culture plate (Corning).

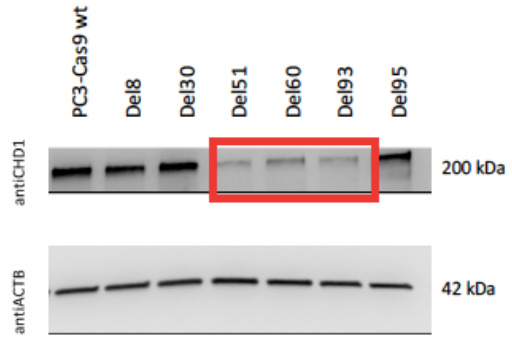
The experimental system is outlined on the figure below.



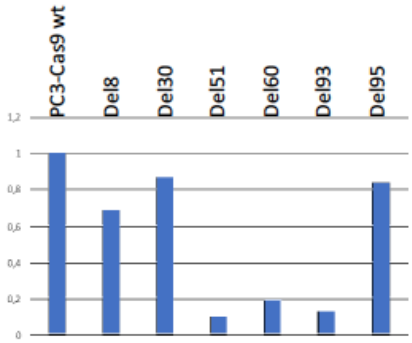
The figure below shows the significantly diminished expression of CHD1 in the heterozygous deleted mutant cell lines.

Induction of chd1 heterozygous mutation

Western hybridization results of parental PC3-Cas9 wt and CHD1 deleted clones

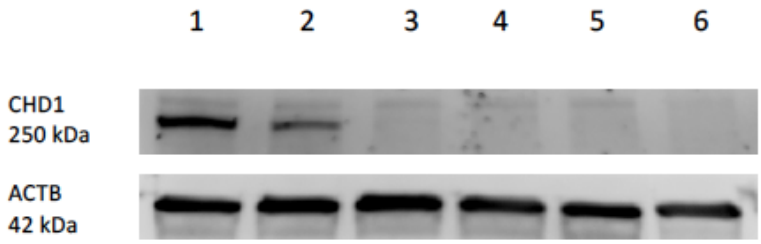


CHD1 intensities normalized to parental PC3-Cas9 wt cell line.



Induction chd1 null mutation in Del51 chd1 heterozygous isogenic cell line

CHD1 WB in parental and CHD1 ko cell lines in PC3



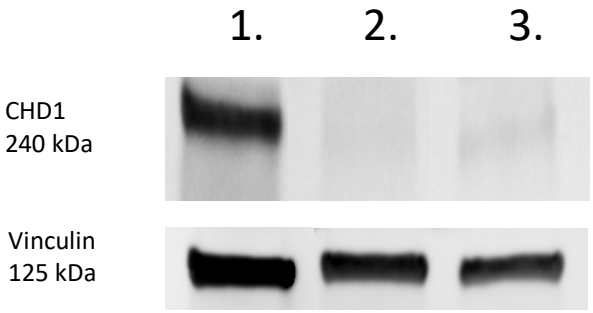
1: PC3/C9 Clone1 wt parental isogenic line P20
2: PC3/C9 del51_chd1 heterozygous mut isogenic line P20
3: PC3/C9 del6_chd1_null_mut P10
4: PC3/C9 del35_chd1_null_mut P12
5: PC3/C9 del62_chd1_null_mut P12
6: PC3/C9 del71_chd1_null_mut P12

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

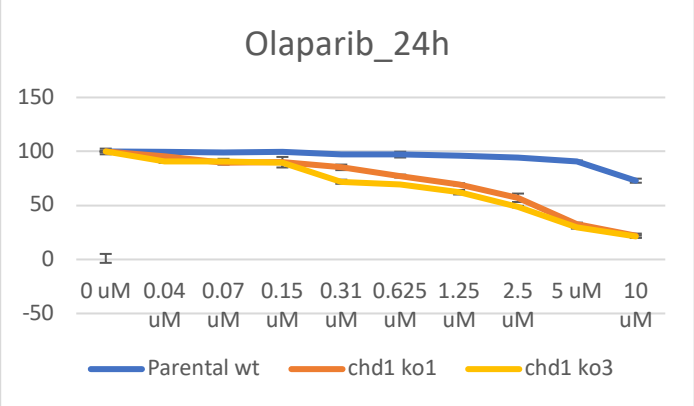
Considering the putative mechanistic link between CHD1 loss and HR deficiency it was previously shown that CHD1 deficient cancer cells show an increased sensitivity the PARP inhibitor olaparib. PARP inhibitors were initially thought to exert their therapeutic activity by inhibiting the enzymatic activity of PARP but it was later revealed that the toxic effect of trapped PARP on DNA may have a more significant contribution. Since trapped PARP is thought to be removed before end resection initiated by CtIP we were wondering whether a PARP inhibitor with a strong trapping ability would show preferential toxicity in cells where CtIP recruitment is impaired. Therefore, we tested the efficacy of both olaparib and the strong PARP trapping agent talazoparib in the RPE1 and PC3 cells with and without CHD1 deletion. Consistent with previous reports, deleting CHD1 induced an approximately 5-fold increase in olaparib sensitivity (Figure 3A). In contrast, the sensitivity to talazoparib increased by about 15-20 fold in the same CHD1 deficient PC3 cells. This would suggest that trapped PARP may have a more toxic effect in cells with impaired CtIP mediated end resection.

Consistent with the significant functional evidence linking CHD1 deletion to HR repair of double strand DNA breaks it was demonstrated that CHD1 deficient cells show increased sensitivity to irradiation. We investigated, whether this increased sensitivity also exists for chemotherapy agents that induce double strand breaks, often termed radiomimetics. As figure 3B demonstrates, CHD1 deficient cells show significantly increased (5-fold) sensitivity to bleomycin, the most frequently used radiomimetic therapeutic agent.

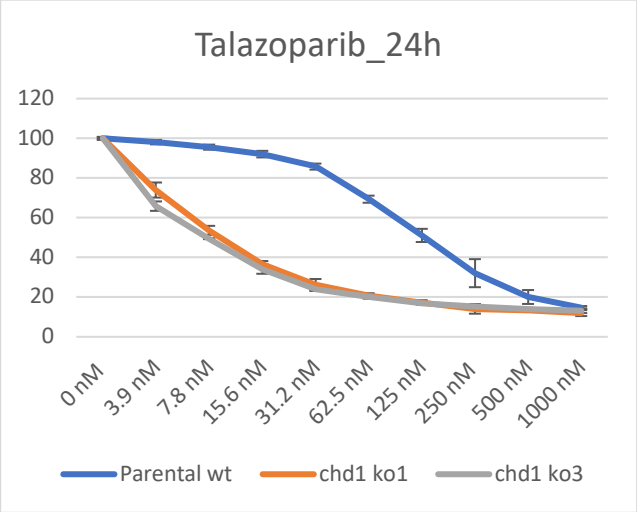
Figure 3



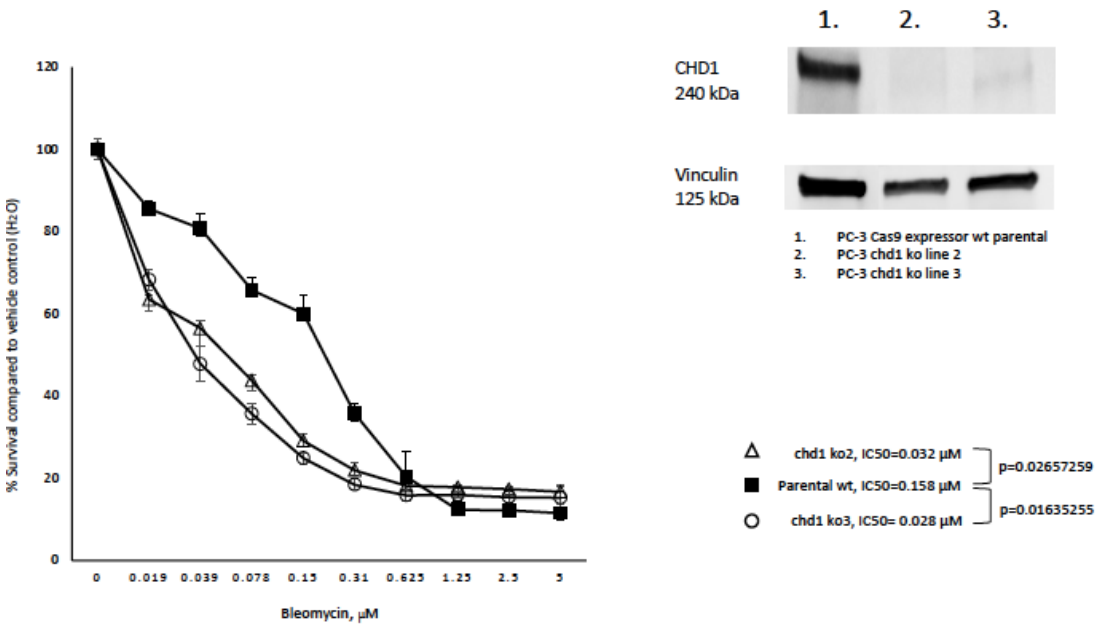
1. PC-3 Cas9 expressor wt parental
2. PC-3 chd1 ko line 1
3. PC-3 chd1 ko line 3



	Parental wt	chd1 ko1	chd1 ko3
IC50 (uM)	36.44	7.66	5.93



	Parental wt	chd1 ko1	chd1 ko3
IC50 (nM)	109.3	7.2	5.6

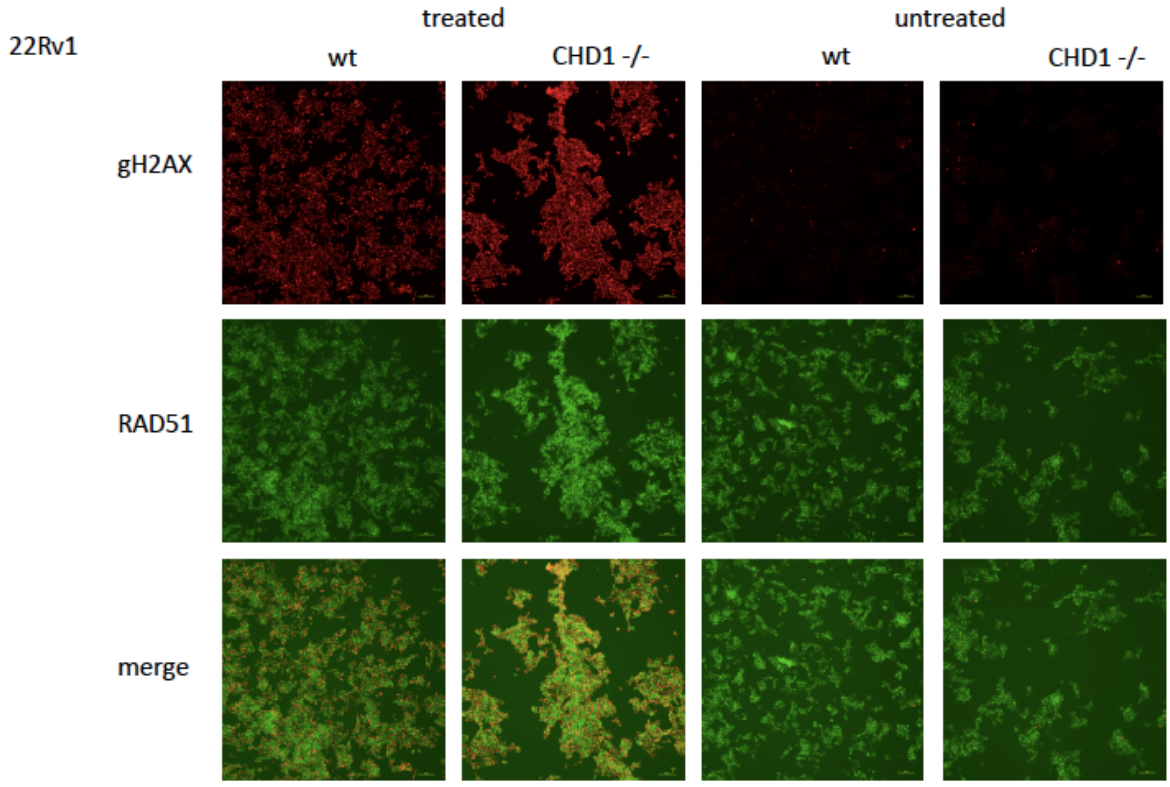


For a comprehensive overview of the results listed above, please consult the attached manuscript currently in revision at Nature Communications.

Further results:

Our next generations sequencing based mutational signature analysis showed that some homologous recombination deficiency induced events are induced in CHD1 deficient prostate cancer cells and some are not. Therefore, we decided to test HR deficiency directly in our cellular model systems by the RAD51 foci assay induced DNA double strand breaks.

As the figure below shows, CHD1 deficient cells did not lose RAD51 foci formation, suggesting that CHD1 loss does not inhibit this step of homologous recombination deficiency, which is consistent with our mutational signature analysis.



What opportunities for training and professional development has the project provided? Postdoctoral fellows were trained in histology, bioinformatics and cell biology.

We also initiated a collaboration within the framework of a U54 project grant between the University of Massachusetts Boston & Dana-Farber/Harvard Cancer Center as a Comprehensive Partnership to Advance Cancer Health Equity

How were the results disseminated to communities of interest?

Publication is in revision at Nature Communications, Diossy et al. : " Increased frequency of CHD1 deletions in prostate cancers of African American men is associated with distinct homologous recombination deficiency associated DNA aberration profiles"

What do you plan to do during the next reporting period to accomplish the goals?

Children's Hospital, Boston:

We have completed our part of the project.

Henry M. Jackson Foundation for the Advancement of Military Medicine (CPDR):

Due to the COVID pandemic our efforts of Task #2 have been significantly delayed.

Task #2: Detection of prevalence of CHD1 deletion in carboplatin treated cohort of castration resistant PCa cases from Dana Farber Cancer Institute.

- Prepare TMA from prostate specimens of patients treated with carboplatin at DFCI and assess by FISH and IHC (mths 6-7)

These efforts were restarted late summer as labs returned to normal functioning. (this effort involved several other personnel beyond our groups, identifying the blocks were considered less urgent relative to pathology services.)

This work will be completed during the next few months.

Dana Farer Cancer Institute:

We have completed our part of the project.

What was the impact on the development of the principal discipline(s) of the project?

In our benchmarking work for the HR deficiency induced mutational signature analysis we found that the number of HR deficiency cases (by robust display of HR deficiency induced mutational signatures) is about twice as high (5-8%) than expected by BRCA1/2 mutational analysis. This strongly suggests that proportion of patients significantly benefiting from PARP inhibitor/platinum treatment is higher than initially thought and careful analysis of HR deficiency in prostate cancer will likely become and important diagnostic decision point for more personalized therapy.

The design, selection and careful quality control of CHD1 FISH probe and the establishment of hybridization condition in prostate tissues will assure the rapid adaptation of this prognostic tool from research to clinical settings. The rapid adaptation for standard pathology has been a key consideration through our project recognizing the urgent need for prognostic tools and predictors of therapeutic response for African American prostate cancer patients.

All these results are important milestones towards a more targeted therapy of African American prostate cancer. We are, in fact, discussing this specific demographic aspect with planned clinical trials for talazoparib.

What was the impact on other disciplines? Nothing to report.

What was the impact on technology transfer? Nothing to report.

What was the impact on society beyond science and technology?

Our results constitute one of the few instances when a robust biological difference can be detected underlying the well-known disparities in cancer outcome. By the successful completion of the proposal we will have a significant advancement in determining which African American prostate cancer patient will benefit from platinum-based or PARP inhibitor treatment. Precise detection of CHD1 deletion will likely become a clinical tool aiding therapeutic stratification for therapies targeting homologous recombination.

CHANGES/PROBLEMS:

Changes in approach and reasons for change: Within the reporting period there was a change of Sub-award PI for the Henry M Jackson Foundation site that has been approved by the awarding agency. The project or its direction was not affected by the PI change.

Actual or anticipated problems or delays and actions or plans to resolve them:

An unexpected challenge was to obtain IRB approvals, originally planned for months 1-4. After 9 months in review, our protocol was determined as “research not involving human subjects as defined by 32 CFR 219.102(e) because the research involves the use of de-identified specimens and data not collected specifically for this study”. The initial and secondary reviews and approval by the USAMRMC, Office of Research Protections, Human Research Protection Office took an additional 3 month.

An anticipated challenge is the analyses of Whole Genome Sequences (WGS). One genome set of WGS at 110X depth is 3 Terabyte of uncompressed data. For 35 patients and their matched germ line DNA (total 70 WGS) it will be 210 Terabyte. The current secured file system for data transfer allows only 2 Gigabyte (that is 100,000 times smaller that the transfer capacity we need) while the use of portable data storage devises that could solve this issue is prohibited.

A potential solution for this problem is to obtain authorization/clearance for the DFCI PI to perform the bioinformatic analyses within our system, or to obtain a special permission for data transfer on portable device.

Changes that had a significant impact on expenditures: Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents: Nothing to report.

PRODUCTS:

Journal publications: Sztupinszki Z, Diossy M, Krzystanek M, Börcsök J, Pomerantz MM, Tisza V, Spisák S, Rusz O, Csabai I, Freedman ML, Szallasi Z. Detection of molecular signatures of homologous recombination deficiency in prostate cancer with or without BRCA1/2 mutations. Clin Cancer Res. 2020 Feb 18. pii: clincanres.2135.2019. doi: 10.1158/1078-0432.CCR-19-2135. [Epub ahead of print] PMID: 32071115

Books or other non-periodical, one-time publications: Nothing to report

Other publications, conference papers and presentations: Nothing to report

Website(s) or other Internet site(s): Nothing to report

Technologies or techniques: Nothing to report

Inventions, patent applications, and/or licenses: Nothing to report

PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

Name: Zoltan Szallasi, MD

Project Role: PI

Researcher Identifier (e.g. ORCID ID): 0000-0001-5395-7509

Nearest person month worked: 1month

Contribution to Project: Dr. Szallasi provides overall supervision for the project and supervises and performs some of the bioinformatic analysis.

Funding Support:

Name: Viktoria Tisza, PhD

Project Role: Research Associate

Researcher Identifier (e.g. ORCID ID):

Nearest person month worked: 5 months

Contribution to Project: Dr. Tisza works on preparing the CRISPR edited cell lines, she propagates the clones and prepares those for next generation sequencing.

Funding Support:

Name: Albert Dobi, PhD

Project Role: Subaward PI

Researcher Identifier (e.g. ORCID ID):

Nearest person month worked: 1

Contribution to Project: Dr. Dobi provided directions under this subaward, wrote the IRB protocol, corresponded with the IRB, coordinated and oversaw the generation of TMA, directed the selection and quality control of CHD1 FISH probe.
Funding Support:

Name: Hua Li, MD, PhD
Project Role: Co-Investigator
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: 2
Contribution to Project: Dr. Li contributed to the IRB protocol writing, generated the new TMA, performed FISH assay for CHD1, performed pathology and reading of FISH results.
Funding Support:

Name: Matthew Freedman, MD, PhD
Project Role: Subaward PI
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: 1
Contribution to Project: Dr. Freedman is supervising most of the CRISPR editing of the cell lines.
Funding Support:

Name: Ji-Heui Seo PhD
Project Role: research fellow
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: 1
Contribution to Project: Dr. Spisak is supporting the CRISPR editing of the cell lines as a technical expert in the field.
Funding Support:

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

[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

Abstract

Chromodomain helicase DNA-binding protein 1 (*CHD1*) is frequently deleted in a subset of prostate cancers. We show here that subclonal deletion of *CHD1* is nearly three times as frequent in prostate tumors of African American men than in men of European ancestry. We further show that *CHD1* deletion is associated with some of the homologous recombination deficiency associated mutational signatures in prostate cancer. In a cell line model *CHD1* deletion induces 1-10 kb deletions resembling those induced by *BRCA2* deficiency. This observation further validates a functional link between *CHD1* loss and homologous recombination deficiency in prostate cancer. *CHD1* deficient cells showed markedly increased sensitivity to both talazoparib and the radiomimetic bleomycin. These agents may be used in the personalized medicine setting to target *CHD1* deficient foci that may be a source of antiandrogen therapy resistance in AA prostate cancer.

Introduction

African American (AA) men have significantly higher incidence and mortality rates from prostate cancer (PC) compared to individuals of European ancestry (EA)[1]. 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[2,3]. 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[4–6]. In contrast, *LSAMP* and *ETV3* deletions, *ZFHX3* mutations, *MYC* and *CCND1* amplifications and *KMT2D* truncations are more frequent in PCs of AA men[7–9]. *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[10], and *SPINK1* expression is also enriched in African American PC[11].

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[12], activation of transcriptional programs that drive prostate tumorigenesis[13] and enzalutamide resistance[14]. 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

redistributes to HOXB13 enriched sites and thus alters the transcriptional program of prostate cancer cells [13]. *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[15,16]. 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[17].

CHD1 loss is frequently subclonal[18], which makes its detection by next generation sequencing more challenging[19] 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.

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 [18]. 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[20]. 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). 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[4,5] (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.

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 [10,21] 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[19]. 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

164 array data, such as TITAN[19], THetA[22] , and Sclust[23], they are designed to work on the
165 entire genome, and likely miss small (~1-10kb) CNVs during the data segmentation
166 process. In order to maximize the accuracy of our analysis we performed a gene
167 focused analysis of the copy number loss in *CHD1*. We considered several factors
168 such as the change in the normalized coverage in the tumors relative to their normal
169 pairs', the cellularity of the tumor genome, and the approximate proportion of
170 tumor cells exhibiting the loss. We also evaluated whether the deletion was
171 heterozygous or homozygous using a statistical method designed for calling
172 subclonal loss of heterozygosity (LOH) events within a confined genomic region
173 (details are available in the Materials and Methods section, and in the
174 Supplementary Material).
175 Using this approach in a large cohort (N=530 cases; 59 AA WES, 18AA WGS, 408 EA
176 WES and 45 EA WGS, for details see supplementary material and Suppl. Figures 5-
177 28), we observed that *CHD1* is more frequently deleted in AA tumors (N=20; 26%)
178 than in EA tumors (N=73 EA; 16%). Taken together, when next generation
179 sequencing based copy number variations were analyzed with a more sensitive
180 method, *CHD1* loss was detected more frequently in the AA cases than in the EA
181 cases (p=0.029, Fisher exact test), which is consistent with our observations with
182 FISH method in the TMA cohort.
183
184 **_CHD1_ loss is associated with genomic signatures frequently observed in _BRCA2_**
185 **deficient prostate cancers.**

186 *CHD1* loss was shown to reduce HR competence in cell line model systems[15,24].
187 Detecting and quantifying HR deficiency in tumor biopsies is currently best achieved
188 by analyzing whole genome sequencing data for specific HR deficiency associated
189 mutational signatures. Those include: 1) A single nucleotide variation based
190 mutational signature ("COSMIC signatures 3[25] and SBS3[26]); 2) a short
191 insertions/deletions based mutational profile, often dominated by deletions with
192 microhomology, a sign of alternative repair mechanisms joining double-strand
193 breaks in the absence of HR, which is also captured by COSMIC indel signatures ID6
194 and ID8[26] ; 3) large scale rearrangements such as non-clustered tandem
195 duplications in the size range of 1-100kb (mainly associated with *BRCA1* loss of
196 function) or deletions in the range of 1kb-1Mb (mainly associated with *BRCA2* loss
197 of function)[27]. All of these signatures can be efficiently induced by the inactivation of
198 *BRCA1*, *BRCA2* or several other key downstream HR genes (Suppl. Figures 31-48) [28] .
199 HR deficiency is also assessed in the clinical setting by a large scale genomic
200 aberration based signature, namely the HRD score[29], which is also approved as
201 companion diagnostic for PARP inhibitor therapy. Recently a composite mutational
202 signature, HRDetect[30], combining several of the mutational features listed above
203 was evaluated as an alternative method to detect HR deficiency in prostate
204 adenocarcinoma[31]. In order to further strengthen the link between *CHD1* loss, HR
205 deficiency and potentially increased PARP inhibitor sensitivity we performed a
206 detailed analysis on the mutational signature profiles of *CHD1* deficient prostate
207 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, that the alternative end-joining repair pathways do not dominate the repair of DSBs in those cases.

In the WGS cohort we also determined the number of structural variants as previously defined (Suppl.Fig. 40)[27]. 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. 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[31]. 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[28,32]. 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) [15]. 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 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[33]. *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 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[15]. 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[34]). 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)[15]. 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.

Consistent with the significant functional evidence linking *CHD1* deletion and HR repair of DSBs, *CHD1* deficient cells also showed increased sensitivity to irradiation[15]. 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.

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[35] and SPOP mutations have been shown to suppress key HR genes[17]. 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[36,37]. Deletions present only in a minority of tumor cells 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[15,16,24]. Interestingly, *CHD1* null cells showed only a modest (3-fold) increase in sensitivity to PARP inhibitor or platinum-based therapy[15,16,24]. 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 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[24]. 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 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[14]. 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 before[7,38]. 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, 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[7]. 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).

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[39]. Briefly, four □m TMA sections were dehydrated and blocked in 0.6% hydrogen peroxide in methanol for 20 min. and were processed for 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 [40]. 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).

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)[41], 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)[42], 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[43] 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 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) [44]. 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[15]. RPE1 cells 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[15]. 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 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. 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[45] aligner. The resulting bam files were post-processed according to the GATK best-practices guidelines. Novel

597 variants were called using Mutect2 (v4.1.0) by using the parental clone as "normal"
598 and the *CHD1* ko clones as "tumor" specimens[44]. These vcfs were converted into tab-
599 delimited files and further analyzed in R. Annotation was performed via Intervar[46] .
600
601
602

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

Author Contributions

Conception and design: Miklos Diossy, V. Tisza, H. Li, J. Zhou, Zs. Sztupinszki, S. Spisak, G. Valcz, P. V. Nuzzo, D. Ribli, T. Ried, S. Kaochar, S. Pathania, A. D'Andrea, I. Csabai, S. Srivastava, A. Dobi, M. L. Freedman, Z. Szallasi

Development of methodology: M. Diossy, V. Tisza, H. Li, J. Zhou, Zs. Sztupinszki, M. Krzystanek, A. Dobi, Z. Szallasi

TMA analysis: H. Li, D. Young, D. Nousome, C. Kuo, Y. Chen, R. Ebner, I. A. Sesterhenn, Gy. Petrovics, G. Valcz

Acquisition of data: M. Diossy, V. Tisza, H. Li, J. Zhou, Zs. Sztupinszki, D. Young, D. Nousome, C. Kuo, G. Valcz, D. Ribli

Cell line experiments: M. Diossy, V. Tisza, J. Zhou, G. T. Klus, S. Spisak, T. Ried, Z. Szallasi

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M. Diossy, V. Tisza, H. Li, Zs. Sztupinszki, D. Nousome, A. Schina, J. Börcsök, A. Prosz, I. Csabai, S. Srivastava, M. L. Freedman, Z. Szallasi

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M. Diossy, V. Tisza, H. Li, J. T. Moncur, G. T. Chesnut, S. Srivastava, A. Dobi, Z. Szallasi

Study supervision: A. Dobi, Z. Szallasi

All authors were involved in the preparation of the manuscript and the supplementary materials.

Data Availability:

Whole exome and whole genome TCGA data presented in this study are available from the GDC (https://portal.gdc.cancer.gov/) and ICGC (https://dcc.icgc.org/) data portals respectively. The whole genomes from the Mayo clinic are available from dbGap (phs001105.v1.p1), while whole genomes from DFCI and CPDR are available upon request.

Figure captions:

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.

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.

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.

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.

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.

References

1. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer Statistics, 2017. *CA Cancer J Clin* **67**, 7–30 (2017).
2. Gaines, A. R. *et al.* The association between race and prostate cancer risk on initial biopsy in an equal access, multiethnic cohort. *Cancer Causes Control* **25**, 1029–1035 (2014).
3. Chu, D. I. *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* **118**, 4999–5007 (2012).
4. Khani, F. *et al.* Evidence for molecular differences in prostate cancer between African American and Caucasian men. *Clin. Cancer Res.* **20**, 4925–4934 (2014).
5. Rosen, P. *et al.* Differences in frequency of ERG oncoprotein expression between index tumors of Caucasian and African American patients with prostate cancer. *Urology* **80**, 749–753 (2012).
6. Sedarsky, J., Degon, M., Srivastava, S. & Dobi, A. Ethnicity and ERG frequency in prostate cancer. *Nat Rev Urol* **15**, 125–131 (2018).
7. Petrovics, G. *et al.* A novel genomic alteration of LSAMP associates with aggressive prostate cancer in African American men. *EBioMedicine* **2**, 1957–1964 (2015).
8. Koga, Y. *et al.* Genomic Profiling of Prostate Cancers from Men with African and European Ancestry. *Clin Cancer Res* **26**, 4651–4660 (2020).

9. Mahal, B. A. *et al.* Racial Differences in Genomic Profiling of Prostate Cancer. *N Engl J Med* **383**, 1083–1085 (2020).
10. Huang, F. W. *et al.* Exome Sequencing of African-American Prostate Cancer Reveals Loss-of-Function ERF Mutations. *Cancer Discov* **7**, 973–983 (2017).
11. Faisal, F. A. *et al.* 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* **22**, 552–559 (2019).
12. Burkhardt, L. *et al.* CHD1 is a 5q21 tumor suppressor required for ERG rearrangement in prostate cancer. *Cancer Res.* **73**, 2795–2805 (2013).
13. Augello, M. A. *et al.* CHD1 Loss Alters AR Binding at Lineage-Specific Enhancers and Modulates Distinct Transcriptional Programs to Drive Prostate Tumorigenesis. *Cancer Cell* **35**, 817–819 (2019).
14. Zhang, Z. *et al.* Loss of CHD1 Promotes Heterogeneous Mechanisms of Resistance to AR-Targeted Therapy via Chromatin Dysregulation. *Cancer Cell* **37**, 584-598.e11 (2020).
15. Zhou, J. *et al.* Human CHD1 is required for early DNA-damage signaling and is uniquely regulated by its N terminus. *Nucleic Acids Res.* **46**, 3891–3905 (2018).
16. Kari, V. *et al.* Loss of CHD1 causes DNA repair defects and enhances prostate cancer therapeutic responsiveness. *EMBO Rep.* **17**, 1609–1623 (2016).
17. Hjorth-Jensen, K. *et al.* SPOP promotes transcriptional expression of DNA repair and replication factors to prevent replication stress and genomic instability. *Nucleic Acids Res.* **46**, 9484–9495 (2018).

18. Baca, S. C. *et al.* Punctuated evolution of prostate cancer genomes. *Cell* **153**, 666–677 (2013).

19. Ha, G. *et al.* TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res.* **24**, 1881–1893 (2014).

20. Han, M. *et al.* Biochemical (prostate specific antigen) recurrence probability following radical prostatectomy for clinically localized prostate cancer. *J. Urol.* **169**, 517–523 (2003).

21. Yuan, J. *et al.* Integrative comparison of the genomic and transcriptomic landscape between prostate cancer patients of predominantly African or European genetic ancestry. *PLoS Genet.* **16**, e1008641 (2020).

22. Oesper, L., Mahmoody, A. & Raphael, B. J. THetA: inferring intra-tumor heterogeneity from high-throughput DNA sequencing data. *Genome Biol.* **14**, R80 (2013).

23. 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* **13**, 1488–1501 (2018).

24. Shenoy, T. R. *et al.* CHD1 loss sensitizes prostate cancer to DNA damaging therapy by promoting error-prone double-strand break repair. *Ann. Oncol.* **28**, 1495–1507 (2017).

25. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).

780 26. Alexandrov, L. B. *et al.* The repertoire of mutational signatures in human
781 cancer. *Nature* **578**, 94–101 (2020).
782 27. Nik-Zainal, S. *et al.* Landscape of somatic mutations in 560 breast cancer
783 whole-genome sequences. *Nature* **534**, 47–54 (2016).
784 28. Póti, Á. *et al.* Correlation of homologous recombination deficiency induced
785 mutational signatures with sensitivity to PARP inhibitors and cytotoxic agents.
786 *Genome Biol.* **20**, 240 (2019).
787 29. Telli, M. L. *et al.* Homologous Recombination Deficiency (HRD) Score Predicts
788 Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with
789 Triple-Negative Breast Cancer. *Clin. Cancer Res.* **22**, 3764–3773 (2016).
790 30. Davies, H. *et al.* HRDetect is a predictor of BRCA1 and BRCA2 deficiency
791 based on mutational signatures. *Nat. Med.* **23**, 517–525 (2017).
792 31. Sztupinszki, Z. *et al.* Detection of Molecular Signatures of Homologous
793 Recombination Deficiency in Prostate Cancer with or without BRCA1/2
794 Mutations. *Clin. Cancer Res.* **26**, 2673–2680 (2020).
795 32. Zámborszky, J. *et al.* Loss of BRCA1 or BRCA2 markedly increases the rate of
796 base substitution mutagenesis and has distinct effects on genomic deletions.
797 *Oncogene* **36**, 746–755 (2017).
798 33. Feng, W. *et al.* Genetic determinants of cellular addiction to DNA polymerase
799 theta. *Nat Commun* **10**, 4286 (2019).
800 34. Murai, J. & Pommier, Y. PARP Trapping Beyond Homologous Recombination
801 and Platinum Sensitivity in Cancers. *Annu. Rev. Cancer Biol.* **3**, 131–150 (2019).

35. Barbieri, C. E. *et al.* The mutational landscape of prostate cancer. *Eur. Urol.* **64**, 567–576 (2013).

36. Yates, L. R. *et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat. Med.* **21**, 751–759 (2015).

37. Linch, M. *et al.* Intratumoural evolutionary landscape of high-risk prostate cancer: the PROGENY study of genomic and immune parameters. *Ann. Oncol.* **28**, 2472–2480 (2017).

38. Merseburger, A. S. *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* **10**, 223–228 (2003).

39. Furusato, B. *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* **13**, 228–237 (2010).

40. Karczewski, K. J. *et al.* The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res* **45**, D840–D845 (2017).

41. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

42. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

43. Favero, F. *et al.* Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol.* **26**, 64–70 (2015).

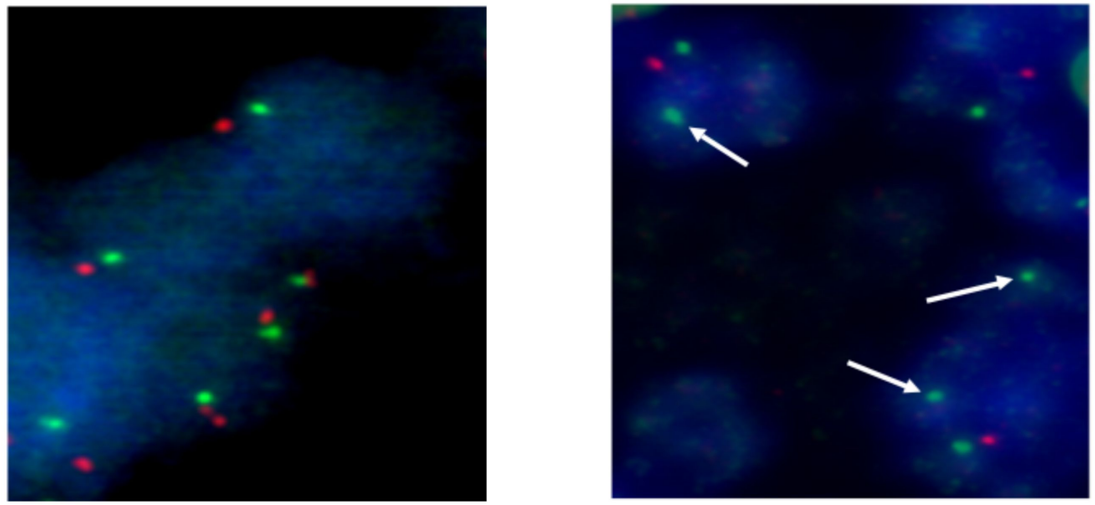
44. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297–1303 (2010).

45. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

46. Li, Q. & Wang, K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am. J. Hum. Genet.* **100**, 267–280 (2017).

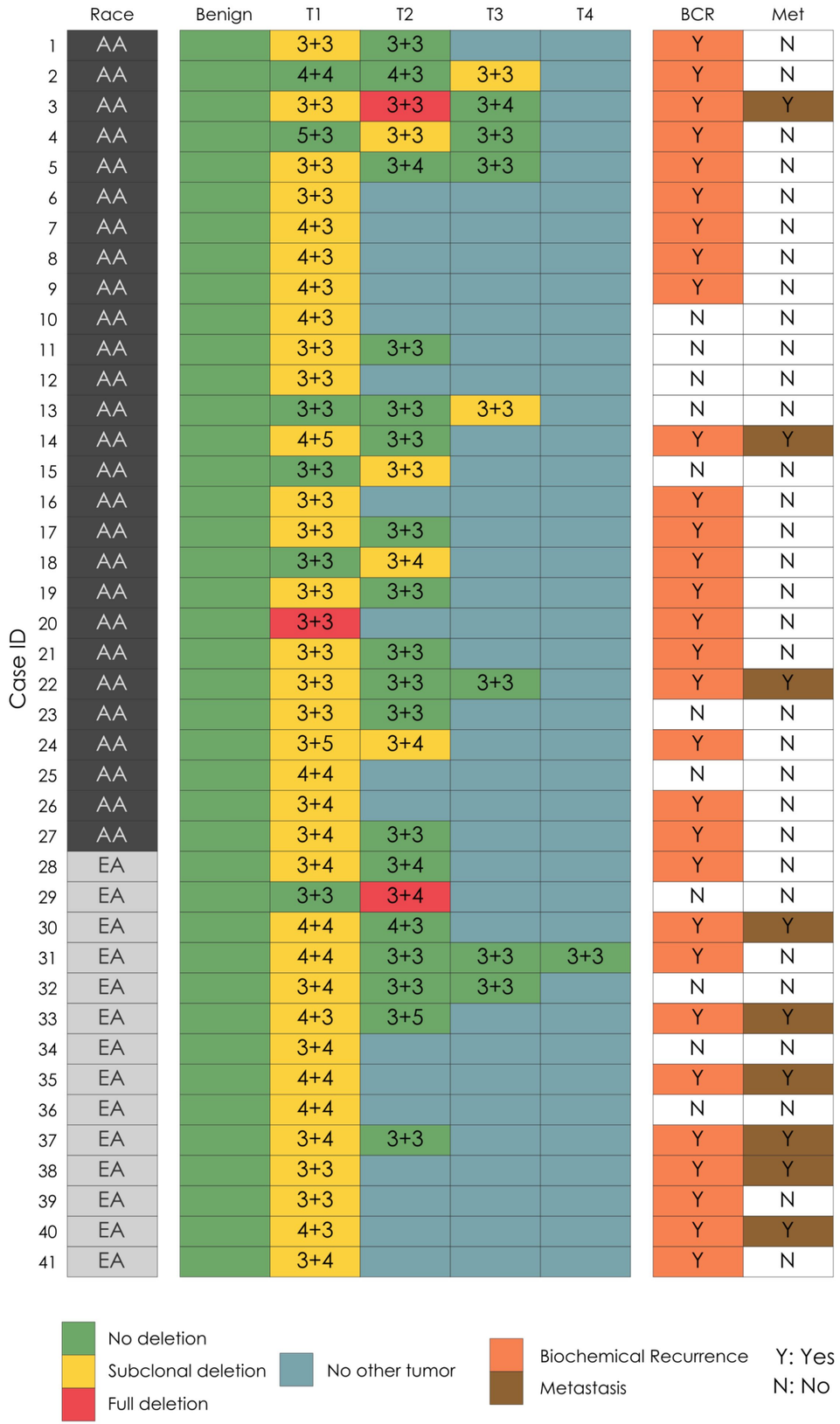
Figure 1

A



Race	No deletion (N = 159)	Deletion (N = 41)	p-value
AA (N = 91)	64 (70.3%)	27 (29.7%)	0.003
CA (N = 109)	95 (87.2%)	14 (12.8%)	

B



C

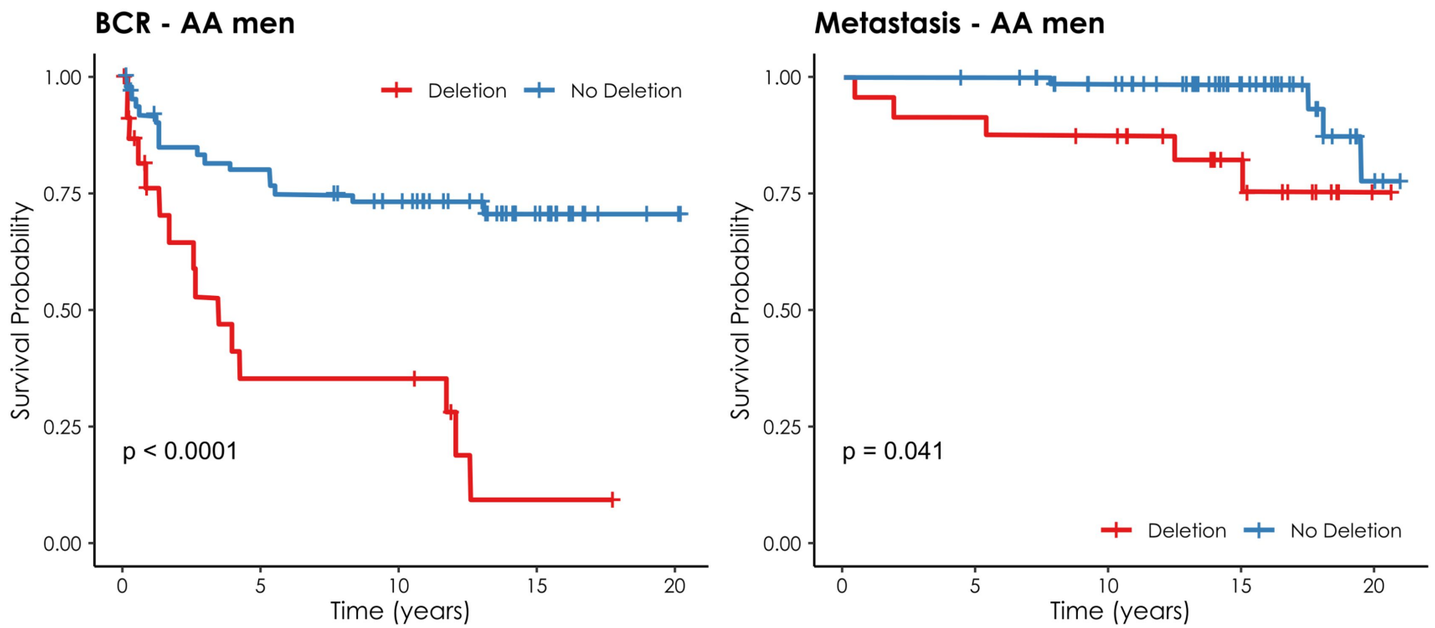
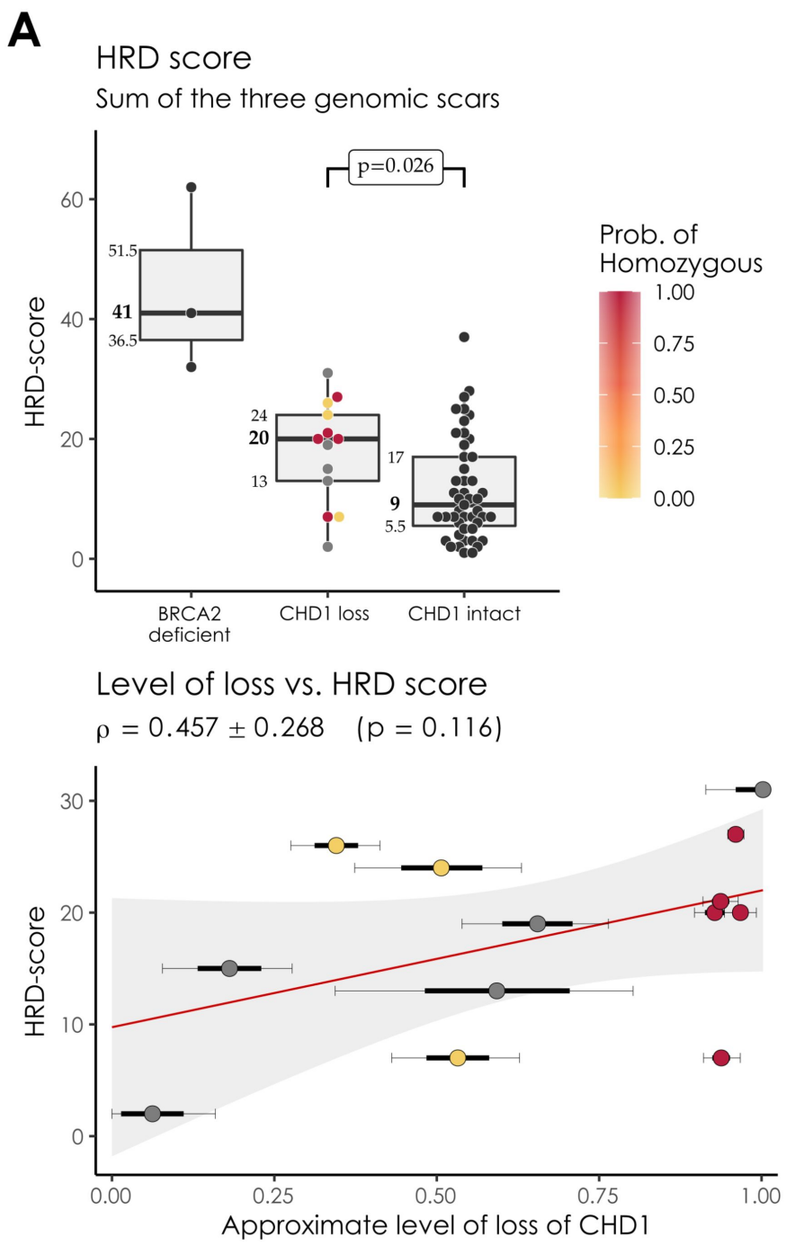
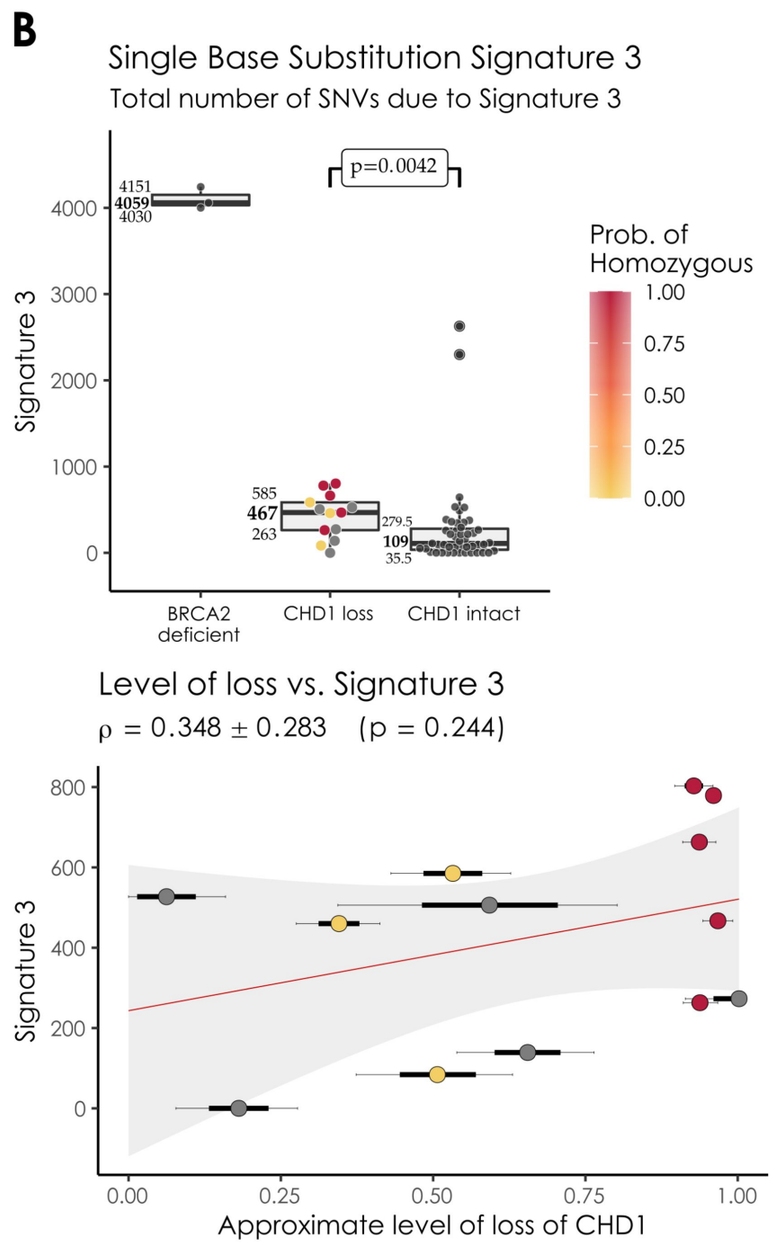


Figure 2


A
HRD score
Sum of the three genomic scars
p=0.026
HRD-score
51.5
41
36.5
24
20
13
17
9
5.5
Prob. of Homozygous
1.00
0.75
0.50
0.25
0.00
BRCA2 deficient
CHD1 loss
CHD1 intact
Level of loss vs. HRD score
ρ = 0.457 ± 0.268 (p = 0.116)
HRD-score
Approximate level of loss of CHD1


B
Single Base Substitution Signature 3
Total number of SNVs due to Signature 3
p=0.0042
Signature 3
4151
4059
4030
585
467
263
279.5
109
35.5
Prob. of Homozygous
1.00
0.75
0.50
0.25
0.00
BRCA2 deficient
CHD1 loss
CHD1 intact
Level of loss vs. Signature 3
ρ = 0.348 ± 0.283 (p = 0.244)
Signature 3
Approximate level of loss of CHD1

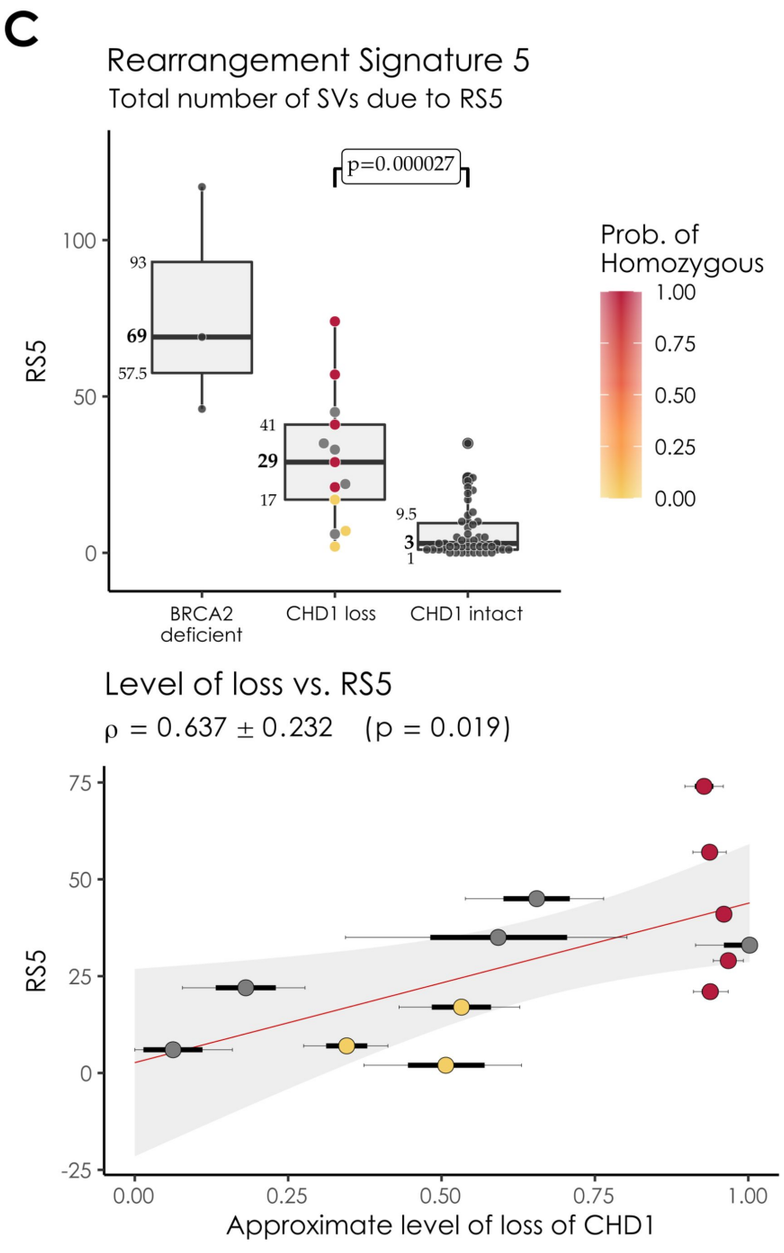
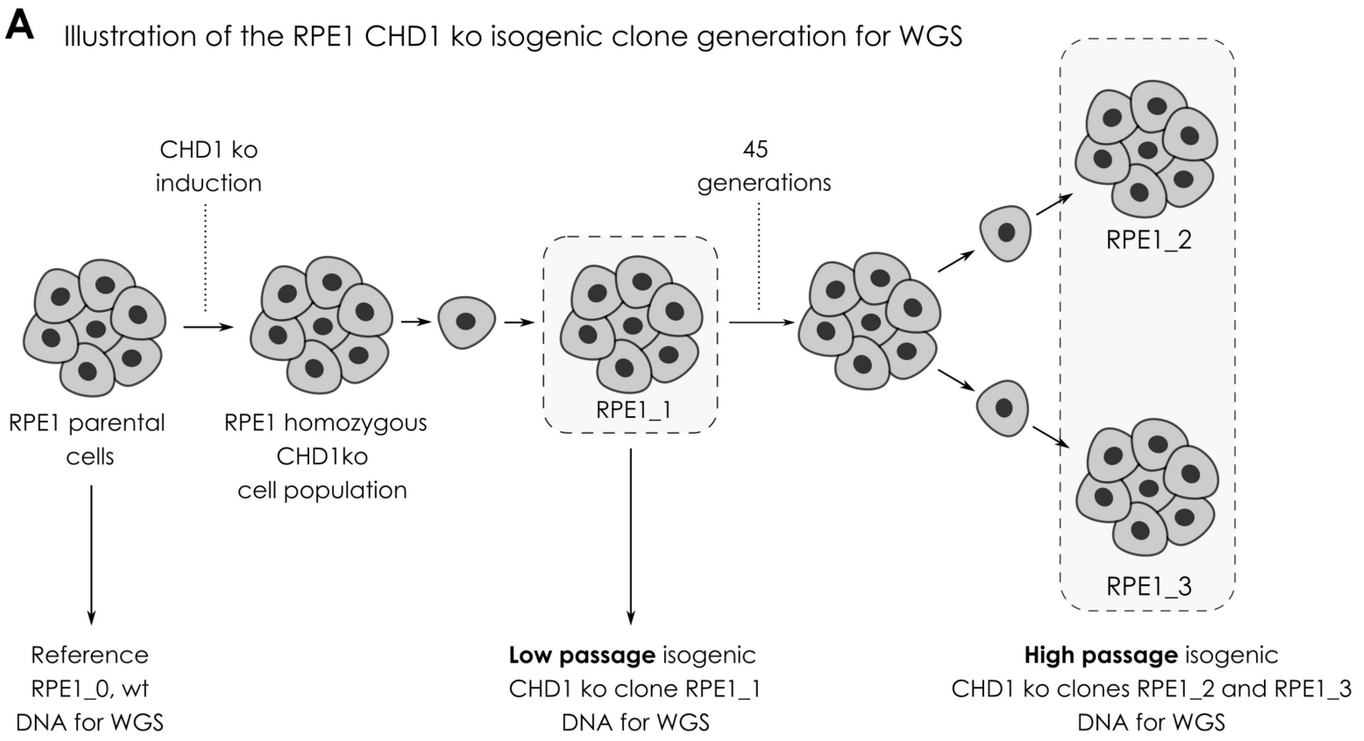

C
Rearrangement Signature 5
Total number of SVs due to RS5
p=0.000027
RS5
93
69
57.5
41
29
17
9.5
3
1
Prob. of Homozygous
1.00
0.75
0.50
0.25
0.00
BRCA2 deficient
CHD1 loss
CHD1 intact
Level of loss vs. RS5
ρ = 0.637 ± 0.232 (p = 0.019)
RS5
Approximate level of loss of CHD1

Figure 3


A
Illustration of the RPE1 CHD1 ko isogenic clone generation for WGS
CHD1 ko induction
45 generations
RPE1 parental cells
RPE1 homozygous CHD1ko cell population
RPE1_1
RPE1_2
RPE1_3
Reference RPE1_0, wt DNA for WGS
Low passage isogenic CHD1 ko clone RPE1_1 DNA for WGS
High passage isogenic CHD1 ko clones RPE1_2 and RPE1_3 DNA for WGS

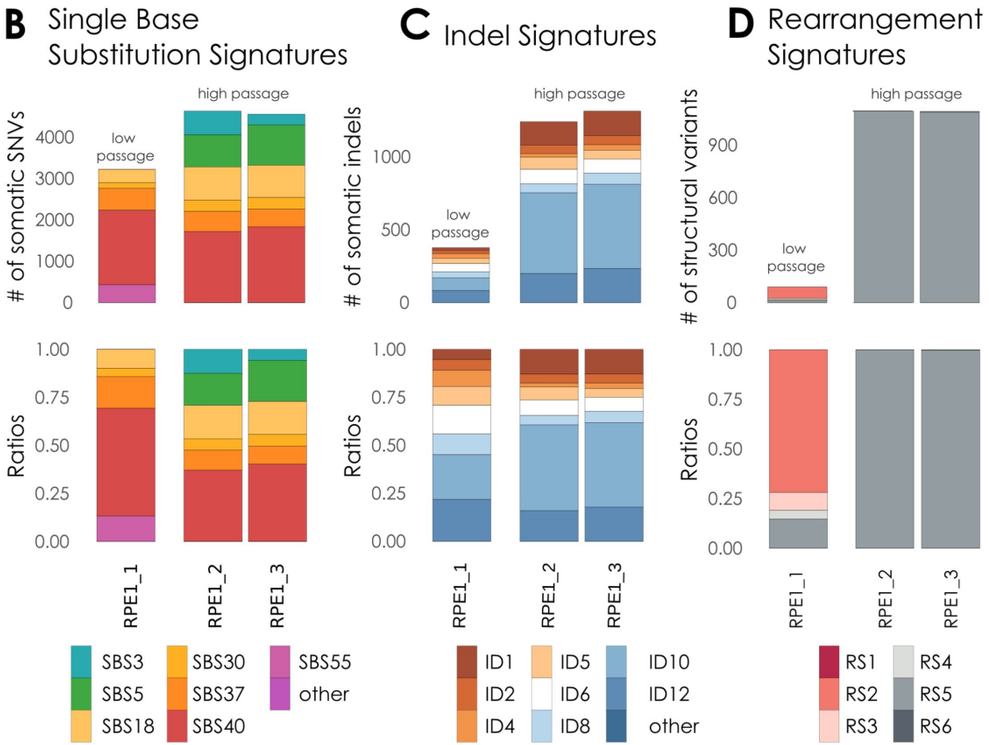

B
Single Base Substitution Signatures
C
Indel Signatures
D
Rearrangement Signatures
# of somatic SNVs
# of somatic indels
# of structural variants
Ratios
low passage
high passage
RPE1_1
RPE1_2
RPE1_3
SBS3
SBS5
SBS18
SBS30
SBS37
SBS40
SBS55
other
ID1
ID2
ID4
ID5
ID6
ID8
ID10
ID12
other
RS1
RS2
RS3
RS4
RS5
RS6

Figure 4

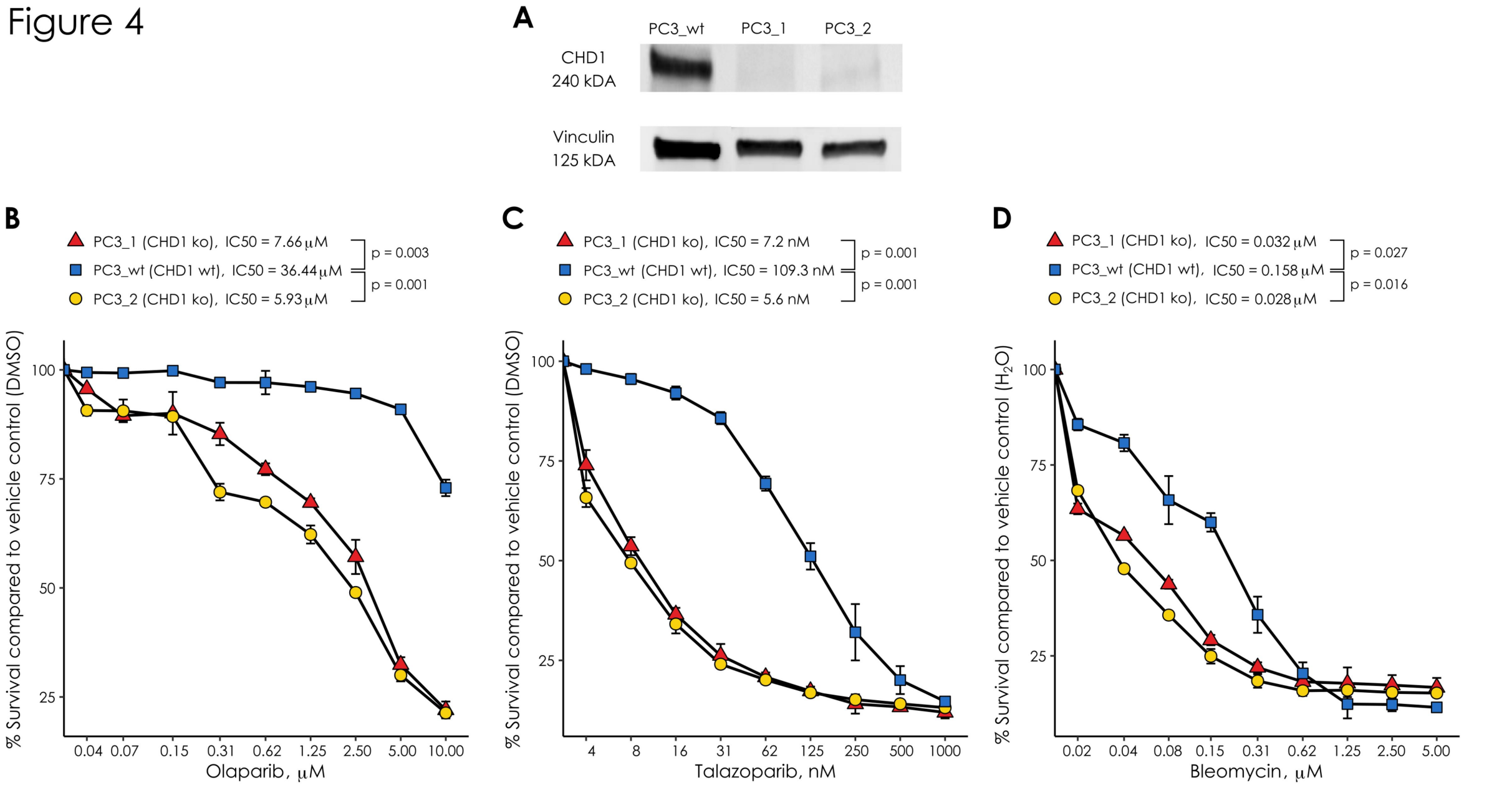
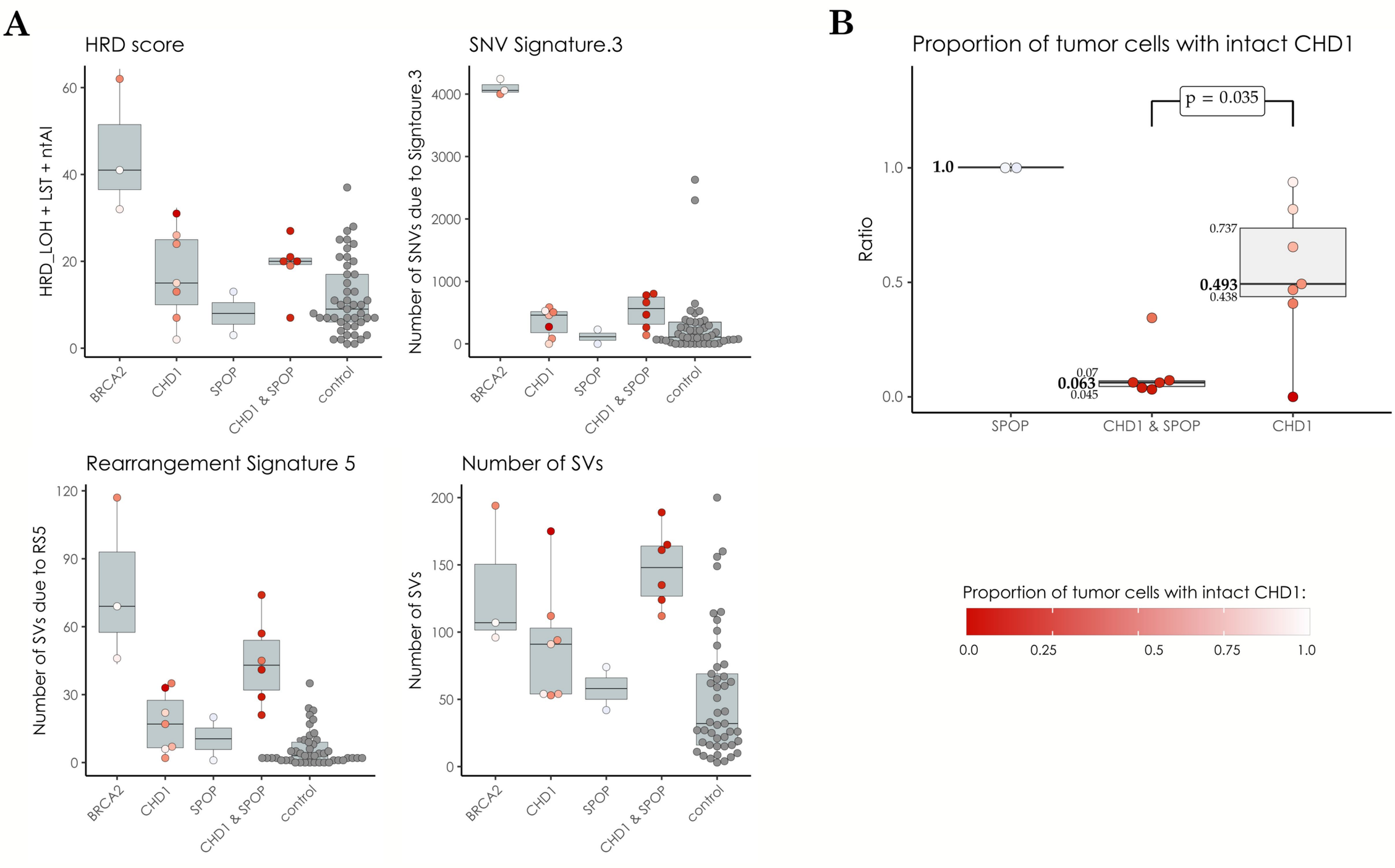

A
PC3_wt
PC3_1
PC3_2
CHD1
240 kDA
Vinculin
125 kDA
B
PC3_1 (CHD1 ko), IC50 = 7.66 μM
PC3_wt (CHD1 wt), IC50 = 36.44 μM
PC3_2 (CHD1 ko), IC50 = 5.93 μM
p = 0.003
p = 0.001
% Survival compared to vehicle control (DMSO)
100
75
50
25
0.04 0.07 0.15 0.31 0.62 1.25 2.50 5.00 10.00
Olaparib, μM
C
PC3_1 (CHD1 ko), IC50 = 7.2 nM
PC3_wt (CHD1 wt), IC50 = 109.3 nM
PC3_2 (CHD1 ko), IC50 = 5.6 nM
p = 0.001
p = 0.001
% Survival compared to vehicle control (DMSO)
100
75
50
25
4 8 16 31 62 125 250 500 1000
Talazoparib, nM
D
PC3_1 (CHD1 ko), IC50 = 0.032 μM
PC3_wt (CHD1 wt), IC50 = 0.158 μM
PC3_2 (CHD1 ko), IC50 = 0.028 μM
p = 0.027
p = 0.016
% Survival compared to vehicle control (H2O)
100
75
50
25
0.02 0.04 0.08 0.15 0.31 0.62 1.25 2.50 5.00
Bleomycin, μM

Figure 5


A
HRD score
HRD_LOH + LST + ntAI
60
40
20
0
BRCA2
CHD1
SPOP
CHD1 & SPOP
control
SNV Signature.3
Number of SNVs due to Signtaure.3
4000
3000
2000
1000
0
Rearrangement Signature 5
Number of SVs due to RS5
120
90
60
30
Number of SVs
200
150
100
50
B
Proportion of tumor cells with intact CHD1
p = 0.035
Ratio
1.0
0.5
0.0
0.737
0.493
0.438
0.07
0.063
0.045
SPOP
CHD1 & SPOP
CHD1
Proportion of tumor cells with intact CHD1:
0.0
0.25
0.5
0.75
1.0