The Journal of Molecular Diagnostics, Vol. 26, No. 6, June 2024

ELSEVIER

the Journal of Molecular Diagnostics
jmdjournal.org

Leveraging Off-Target Reads in Panel Sequencing for Homologous Recombination Repair Deficiency Screening in Tumor

Markus Ball,[*] Iordanis Ourailidis,[*] Klaus Kluck,[*] Michael Menzel,[*] Martina Kirchner,[*] Michael Allgäuer,[*] Timothy Kwang Yong Tay,[*†‡] Fabian Schnecko,[*] Anna-Lena Volckmar,[*] Hannah Goldschmid,[*] Olaf Neuman,[*] Stefan Fröhling,[§¶] Peter Schirmacher,[*‖] Jan Budczies,[*¶‖**] Albrecht Stenzinger,[*¶‖**] and Daniel Kazdal[*¶**]

From the Institute of Pathology,[*] Heidelberg University Hospital, Heidelberg, Germany; the Departments of Anatomical Pathology[†] and Molecular Pathology,[‡] Singapore General Hospital, Singapore; the Department of Translational Medical Oncology,[§] National Center for Tumor Diseases, Heidelberg, Germany; the German Cancer Consortium,[¶] Heidelberg, Germany; the Center for Personalized Medicine Heidelberg,[‖] Heidelberg, Germany; and the Translational Lung Research Center Heidelberg,[**] German Center for Lung Research, Heidelberg, Germany

Accepted for publication February 14, 2024.

Address correspondence to Daniel Kazdal, Ph.D., or Albrecht Stenzinger, M.D., Institute of Pathology, University Hospital Heidelberg, Im Neuenheimer Feld 224, 69120 Heidelberg, Germany.
E-mail: daniel.kazdal@med.uni-heidelberg.de or albrecht.stenzinger@med.uni-heidelberg.de.

Targeted tumor only sequencing has become a standard practice in cancer diagnostics. This study aims to develop an approach for robust copy number variant calling in tumor samples using only off-target region (OTR) reads. We also established a clinical use case for homologous recombination deficiency (HRD) score estimation (HRDest) using the sum of telomeric-allelic imbalance and large-scale state transition scores without the need for loss of heterozygosity information. A strong correlation was found between HRD score and the sum of telomeric-allelic imbalance + large-scale state transition in The Cancer Genome Atlas cohort ($\rho = 0.99$, $P < 2.2 \times 10^{-16}$) and in a clinical in-house cohort of 34 tumors ($\rho = 0.9$, $P = 5.1 \times 10^{-13}$) comparing whole-exome sequencing and targeted sequencing data. HRDest scores from 1086 clinical cases were compared with The Cancer Genome Atlas data set. There were no significant differences in HRD score distribution within the analyzed tumor types. As a control, commercially available HRD standards were also sequenced, and the HRDest scores obtained from the OTR reads were well within the HRD reference range provided by the manufacturer. In conclusion, OTR reads of tumor-only panel sequencing can be used to determine genome-wide copy number variant profiles and to approximate HRD scores. *(J Mol Diagn 2024, 26: 479–486; https://doi.org/10.1016/j.jmoldx.2024.02.008)*

Chromosomal microarray analysis based on oligo hybridization and/or single-nucleotide polymorphism information has been the main approach for genome-wide copy number variation (CNV) detection for years until next-generation sequencing–based assays, like whole-genome sequencing and whole-exome sequencing (WES), became more frequently used for genome-wide CNV detection. However, other approaches also exist. The R package CopywriteR[1] and SavvyCNV[2] describe the leveraging of off-target region (OTR) reads even for small gene panels to fill in the gaps between the enriched regions of such panels. Other publications showed the use of the discarded reads from cancer sample sequencing for the generation of germline research cohorts[3] or to demonstrate how OTR reads can be used in a broader context of OTR coverage analysis and mitochondrial DNA copy number and microbial load estimation.[4] Besides CNVs, the discovery and use of complex biomarkers[5–11] in clinical cancer diagnostics has increased in recent years. One such biomarker, the homologous recombination deficiency (HRD) score, is now regularly used to predict the response of ovarian carcinomas to poly (ADP-ribose) polymerase inhibitors.[12] The HRD score is a measure of homologous recombination repair (HRR) deficiency, which is a key driver of genomic instability in cancer cells.[13] With ShallowHRD,[14] an approach based on shallow whole-genome sequencing was described, demonstrating the feasibility of robust CNV detection with a genome-wide coverage of approximately 1×.

https://doi.org/10.1016/j.jmoldx.2024.02.008

The idea behind this method is to develop a screening and estimation method for the HRD score by leveraging on the existing OTR reads of a standard targeted sequencing panel without the need for additional sequencing. This approach mirrors a genome-wide CNV analysis and provides a global low-resolution cytogenetic representation that is suitable for genomic scar detection of large segments, similar to shallow whole-genome sequencing. We hypothesize that by only using the OTR reads of targeted gene panels, a robust HRD score estimation can be achieved. This study demonstrates the feasibility and accuracy, as well as the limitations, of this method in a cohort of clinical samples, and discuss potential implications for clinical practice.

Materials and Methods

Samples

The samples analyzed here are a retrospective cohort sequenced at the Universitätsklinik Heidelberg/University Hospital Heidelberg (UKHD; Heidelberg, Germany) and consist of sequencing data from 1843 formalin-fixed, paraffin-embedded clinical tissue samples of several different tumor types. All samples received comprehensive molecular workup using the TruSight Oncology 500 (TSO) panel (Illumina Inc., San Diego, CA) at the Institute of Pathology, UKHD, from 2019 to 2023. Subsequently, samples with a tumor cell content $<20\%$ were excluded. Sample and data processing protocols were approved by the ethics committee of Heidelberg University (S-315/2020).

To have WES-derived HRD scores as a ground truth for comparison, 34 cases previously analyzed by WES at the same institute[15] were also selected to form a reference cohort with a wide and even distribution of HRD scores ranging from 1 to 90. The HRD scores of these samples were determined by matched tumor-normal sequencing.

Additional reference material, such as the recently released and currently only commercially available HRD reference standard, Seraseq FFPE HRD Reference Material (SeraCare Life Sciences Inc., Milford, MA), was included, which consists of a negative (HRD, 31 ± 2), a low-positive (HRD, 54 ± 2), and a high-positive (HRD, 72 ± 3) sample, all three with tumor cell content of 65%. These were also sequenced using the targeted panel.

Library Preparation and Sequencing

The preparation of the TSO targeted panel DNA libraries was performed either manually, according to the manufacturer's protocol, or for samples that were processed later than June 2022, via an automated procedure using the TSO high-throughput library kit (both from Illumina Inc.) on a Biomek i7 liquid handler (Beckman Coulter Inc., Brea, CA) with an Illumina qualified workflow. For the generation of a panel of normal, 30 normal samples, including 12 fresh frozen and 18 formalin-fixed, paraffin-embedded samples, were additionally sequenced using the TSO assay.

TSO libraries were sequenced with a target of 40 to 50 million read pairs per sample independent from the sequencing platform, using either 300-cycle high-output kits on a NextSeq500 or SP or S1 200-cycle V1.5 kits on a NovaSeq 6000 (both from Illumina Inc.). TSO sequencing data were processed using the docker TruSight Oncology 500 version 2.2 Local App (Illumina Inc.).

For WES, the hybrid capture bait set Exome 2.0 plus comprehensive exome spike-in (Twist Bioscience, South San Francisco, CA) combined with Twist UMI Adapter System (Twist Bioscience) was used for library preparation. The libraries were sequenced on a NovaSeq 6000 (Illumina Inc.) using SP or S1 200-cycle V1.5 kits, with a target of 133 million read pairs for tumor and 66 million read pairs for the corresponding normal sample. Sequencing data were analyzed using the DRAGEN Bio-IT platform 3.10.9 (Illumina Inc.).

Off-Target Read Assessment

To be able to restrict the subsequent analysis to OTR reads, a bed file with the OTRs was generated using BEDTools version 2.26.0[16] as follows: i) the sex chromosomes and mitochondrial DNA were excluded, ii) the whole genome (hg19) was subdivided into 100-kb segments, iii) the TSO target regions padded bidirectionally by 250 bp were subtracted, and iv) contigs <100 kb were removed. The CNV workflow was performed according to Genome Analysis Toolkit (GATK) version 4.3.0.0 somatic CNV workflow[17–19] using the subtools CollectReadCounts, CreateReadCountPanelOfNormals, DenoiseReadCounts, PlotDenoisedCopyRatios, ModelSegments, PlotModeled Segments, and CallcopyRatioSegments.

HRD Score WES

For the HRD score calculation from WES, subcohort allele-specific copy numbers were estimated from matched tumor and normal BAM files, using Sequenza version 3.0.0.[20] Telomeric-allelic imbalance (TAI), large-scale state transitions (LSTs), loss of heterozygosity (LOH), and the HRD score ($=$ TAI $+$ LST $+$ LOH) were calculated from allele-specific copy numbers using a modified version of scarHRD.[15,21]

HRD Score TCGA

HRD scores for 9594 samples of The Cancer Genome Atlas (TCGA) cohort, based on single-nucleotide polymorphism array data using the ASCAT algorithm,[22] were obtained from a previous study of the authors' working group.[23] When comparing TCGA and UKHD data sets, the tumor subtypes included in each tumor group may differ. For example, TCGA-OV group consisted only of high-grade

serous ovarian cancer, whereas the UKHD-OV group was not restricted to a specific subtype of ovarian cancer. Furthermore, the UKHD cohort did not differentiate between colon adenocarcinoma and rectum adenocarcinoma. Therefore, the analysis of TCGA data combined these two groups to represent a colorectal cancer group. Appropriate subsets were generated when a detailed analysis required comparison of specific tumor types.

Statistical Analysis

For data analysis and visualization, the statistical language R version 4.1.3 (R Core Team, R Project for Statistical Computing, *https://www.r-project.org*) was used. To calculate the correlation of two continuous variables, the authors used Spearman rank-order correlation and reported the Spearman rank correlation coefficient.

HRD and estimated OTR-based HRD scores were compared between groups using the *U*-test (Wilcoxon rank-sum test). $P < 0.05$, after multiple testing correction using the Bonferroni method, was considered significant.

Results

This study explored the possibility of deriving genome-wide copy number profiles by comprehensive assessment and exploitation of the off-target reads generated by hybrid capture–based panel sequencing. The approach was verified by comparing the estimated HRD scores (HRDest) of the UKHD cohort, derived from such off-target read analysis, with WES-derived HRD scores and predetermined HRD scores of reference standards. In addition, the distributions of HRD/HRDest scores within different tumor types and by consideration of the HRR gene mutations status were compared between the UKHD and TCGA data sets.

First, the feasibility of genome-wide CNV calling was assessed by considering a combination of on-target regions and OTRs or OTRs only. When including on-target reads in the assay, the results led to batch effects because of the variable effectiveness of the enrichment reactions after library preparation (data not shown). Therefore, only the OTR reads were analyzed further, which has the additional benefit of being a panel independent approach. For the targeted panel used in this study, 24,755 segments were generated, corresponding to 85.8% of the total length of the genome, excluding sex and mitochondrial chromosomes (Figure 1).

A panel of normal of 30 samples was used for denoising of OTR read counts, according to GATK[17–19] best practices workflows for somatic copy number variant discovery. Using this approach, the OTR reads in the UKHD cohort, consisting of 1843 clinical targeted panel sequencing data sets, were analyzed.

The count data of deduplicated OTR reads in millions per sample ranged from 0.32 to 30.45, with a mean of 4.19 (SD = 2.17) and median of 3.69. After denoising, the samples showed a mean median absolute deviation of 0.118 (SD = 0.029) with a range from 0.064 to 0.330 and a median of 0.111 (Figure 2A). The Spearman correlation between off-target reads and median absolute deviation showed a negative ρ of -0.88 with a P value of 2.2×10^{-16}.

To remove low-quality samples that were unsuitable for this estimation approach, a cutoff value of 0.17 was determined for the median absolute deviation, based on subsampling experiments where the authors systematically reduced the number of OTR reads (Supplemental Figure S1). A total of 93% (1714 of 1843) of the samples processed in the UKHD cohort passed the quality threshold, which translates approximately to 2 million deduplicated OTR reads in the targeted panel.

The HRD score used clinically as a predictive biomarker is calculated from the sum of TAI, LST, and LOH. Because of the lack of coverage of single-nucleotide polymorphisms across the genome in targeted sequencing methods, only LST and TAI could be determined in this approach. However, in TCGA data set (9594 cases), a high correlation (Spearman $\rho = 0.99$, $P = 2 \times 10^{-25}$) was observed between the combination of LST and TAI and the complete HRD score (LOH + TAI + LST) (Figure 2B). The slope of 1.3 of the linear regression of this correlation can be used as a conversion factor to bridge the sum of TAI and LST to the HRD score. When applying a cutoff of 42, the positive and negative predictive values for the (sum of TAI and LST) $\times$ 1.3 in TCGA cohort were 0.91 and 0.96, respectively (Supplemental Table S1).

As proof of principle, 34 clinical WES samples, mainly ovarian cancer, that were previously sequenced with known HRD scores evenly distributed over the entire range of 1 to 90, were selected. These cases were then resequenced using the TSO targeted panel, and their OTR reads were analyzed to determine TAI and LST. For the WES analysis, the TAI, LST, and LOH were determined on the basis of matched tumor-normal sequencing, whereas the targeted panel approach consisted of tumor-only sequencing and a panel of normal.

First, considering only the WES data, the sum of TAI and LST was compared with the complete HRD score and a strong correlation (Spearman $\rho = 0.99$, $P < 2.21 \times 10^{-16}$) and a conversion factor of 1.3 were observed, confirming the results observed for TCGA data set (Figure 3A). Next, the sum of TAI and LST derived from the OTR reads generated by TSO targeted panel sequencing and the sum of TAI and LST derived from WES (Figure 3B) were compared. The correlation showed a high concordance (Spearman $\rho = 0.89$, $P = 1 \times 10^{-12}$), with a slope of 0.9 revealing a slight overestimation for the results obtained from the panel sequencing OTR reads. Finally, the OTR read–based TAI and LST counts were compared with the complete HRD scores from WES (Figure 3C) and a similar and significant correlation (Spearman $\rho = 0.89$, $P = 5.1 \times 10^{-13}$) was observed. On the basis of the slope of the regression line, a conversion function for an HRDest was derived as

HRDest = 1.2 × (TAI + LST). The difference in the bridging factor of 1.2 and 1.3 calculated on the basis of TCGA data (Figure 2B) or the WES-only data (Figure 3A) can be attributed to the overestimation using OTR reads (Figure 3B). Positive and negative predictive values for HRDest scores applying a cutoff of 42 were 0.89 and 0.75, respectively (Supplemental Table S1).

To evaluate the performance of HRDest as a predictor of HRD scores, the receiver operating characteristic was calculated for HRDest using the 34 clinical samples with known WES-based HRD scores. With 0.9 (0.79 − 1), the area under the curve indicated a highly effective classification (Figure 4).

To validate this approach, the first step was to compare the distribution of HRD/HRDest scores between TCGA and the UKHD data set, in a tumor type agnostic setting (Figure 5A) and separately for specific tumor types (Supplemental Figure S2). Cancer types with <30 samples or subsets/supersets of TCGA classification were removed, resulting in 4875 and 1023 samples from TCGA and the UKHD cohort, respectively. The distributions were similar and not significantly different ($P = 0.185$), considering the pan-cancer set as well as all analyzed tumor types [bladder urothelial carcinoma, breast invasive carcinoma (BRCA), cholangiocarcinoma, colorectal cancer (to represent the colorectal cancer group in the UKHD cohort, colon adenocarcinoma and rectum adenocarcinoma were combined for TCGA data), kidney renal clear cell carcinoma, lung adenocarcinoma, mesothelioma, ovarian serous cystadenocarcinoma (OV) (here only the high-grade serous ovarian cancer subset of the UKHD-OV samples was used to match TCGA data), pancreatic adenocarcinoma, skin cutaneous melanoma, stomach adenocarcinoma, and uterine corpus endometrial carcinoma] (Supplemental Table S2). To further elucidate the generated HRDest scores, all samples from both cohorts were categorized considering their mutational status in *BRCA1/2* and other HRR genes according to the HRR classification previously published[23] (H1a: *BRCA1/2* loss-of-function mutation; H1b: loss-of-function mutation in other HRR gene; H2a: *BRCA1/2* variants of unknown significance; H2b: variants of unknown significance in other HRR gene; and H3: no HRR gene mutations). The HRD/HRDest distribution was not significantly different for the pan-cancer approach or the OV samples considering all HRR categories, but was different for the H2b group of the OV samples (Figure 5, B and C). The highest median HRDest score was observed for both comparisons in the H1a group, followed by group H2a (excluding UKHD-OV with a single case), whereas the groups H1b and H2b showed similar median HRDest scores compared with the group with no HRR gene mutations (H3), which is in line with the described results for TCGA data.[23] On examining the H1 group in the pan-cancer UKHD cohort and distinguishing cases with potential monoallelic and biallelic loss-of-function alterations in *BRCA1/2*, a significant difference in the median HRDest was observed ($P = 0.002$) (Supplemental Figure S3). The median HRDest was 34.2 for monoallelic loss and 56.4 for biallelic loss, again consistent with previously published results for TCGA data set.[23]

To validate this approach of estimating HRD scores further, the authors calculated the HRDest scores by targeted panel sequencing of three commercially available Seraseq HRD reference standards: HRD-high-positive,

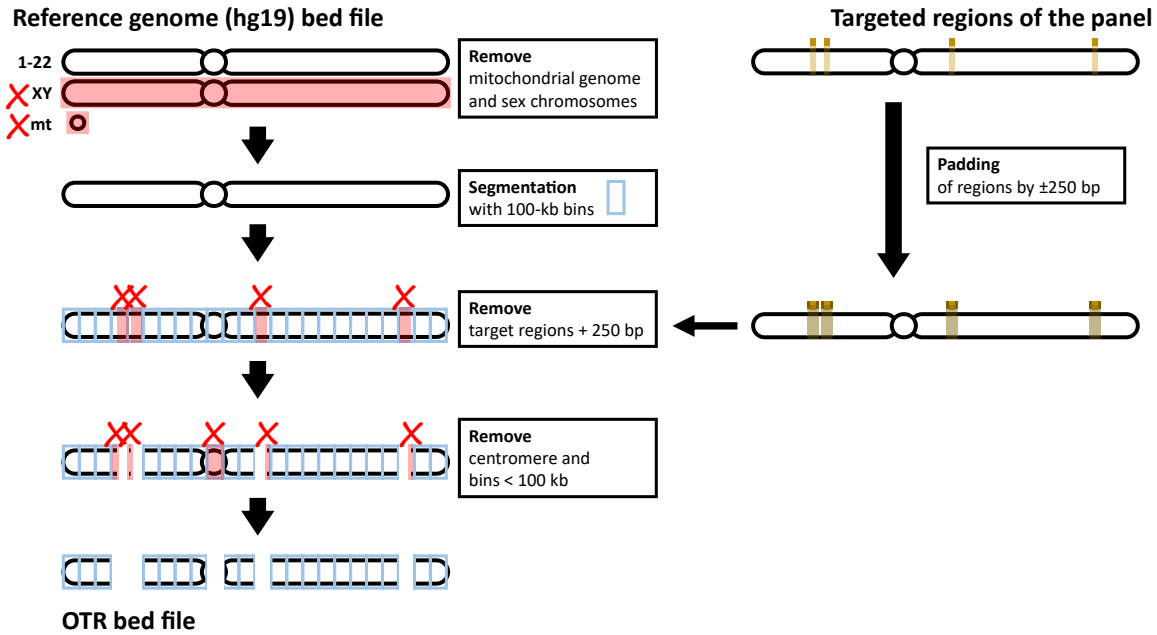

Figure 1 Schematic step-by-step procedure for generating an off-target region (OTR) bed file with the reference genome (**left side**) and the target regions in a targeted panel (**right side**). mt, mitochondrial.

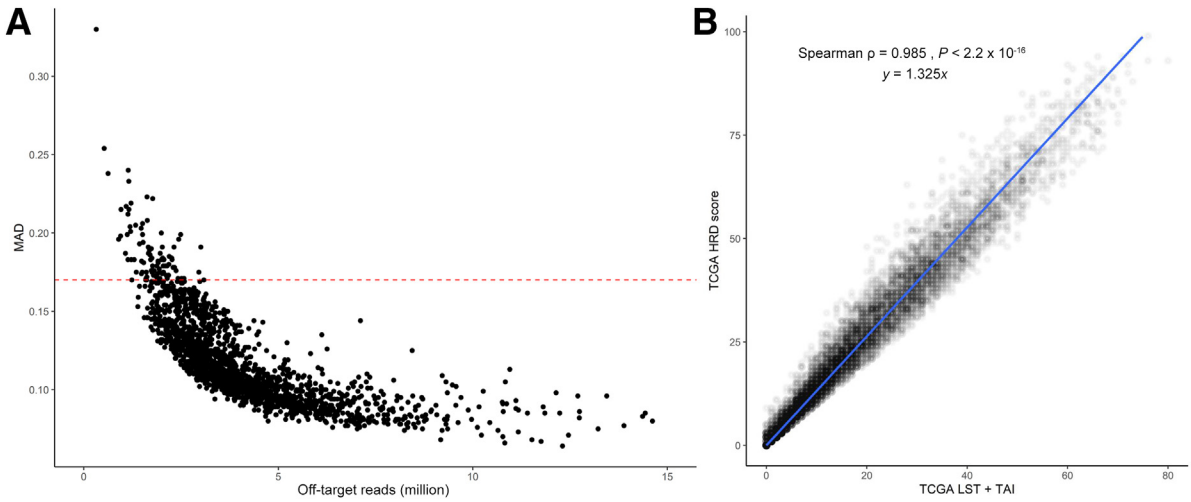

Figure 2 **A:** Representation of deduplicated reads of 1792 clinical targeted panel sequencing samples from University Hospital Heidelberg aligned in the off-target regions and the median absolute deviation (MAD) after denoising. The **red dotted line** marks the MAD threshold of 0.17 (Supplemental Figure S1). **B:** Scatterplot of 9594 The Cancer Genome Atlas (TCGA) samples showing the correlation of the homologous recombination deficiency (HRD) score [loss of heterozygosity + telomeric-allelic imbalance (TAI) + large-scale state transition (LST)] and the score (TAI + LST).

HRD-low-positive, and HRD-negative, with preset HRD scores of 72 ± 3, 54 ± 2, and 31 ± 2, respectively. With HRDest scores of 72, 56, and 29, all three samples were within the expected range. A visual genome-wide representation of the copy number profiles generated and published by SeraCare compared with the copy number profiles derived from the OTR reads shows a good agreement in the CNV patterns (Supplemental Figure S4).

Discussion

Targeted panel sequencing has become the standard in molecular diagnostics for patients with cancer. Large panels cover a genomic footprint of approximately 1 megabase or more, primarily aimed at classic oncogenes and tumor suppressors implicated in diagnosis and therapy management. Although on-target reads identify hot spot mutations (eg, *BRAF* V600E), sequence reads mapped outside of these targeted regions as a result of imperfect enrichment are usually ignored by classic analysis pipelines. However, such OTR reads are almost uniformly distributed across the complete genome, although their prevalence may be biased by factors such as unintended enrichment or GC content. They may also cover mitochondrial DNA, DNA from cells that form the tumor microenvironment, and microbial DNA, which, in addition, may provide useful information[1–4,24] that

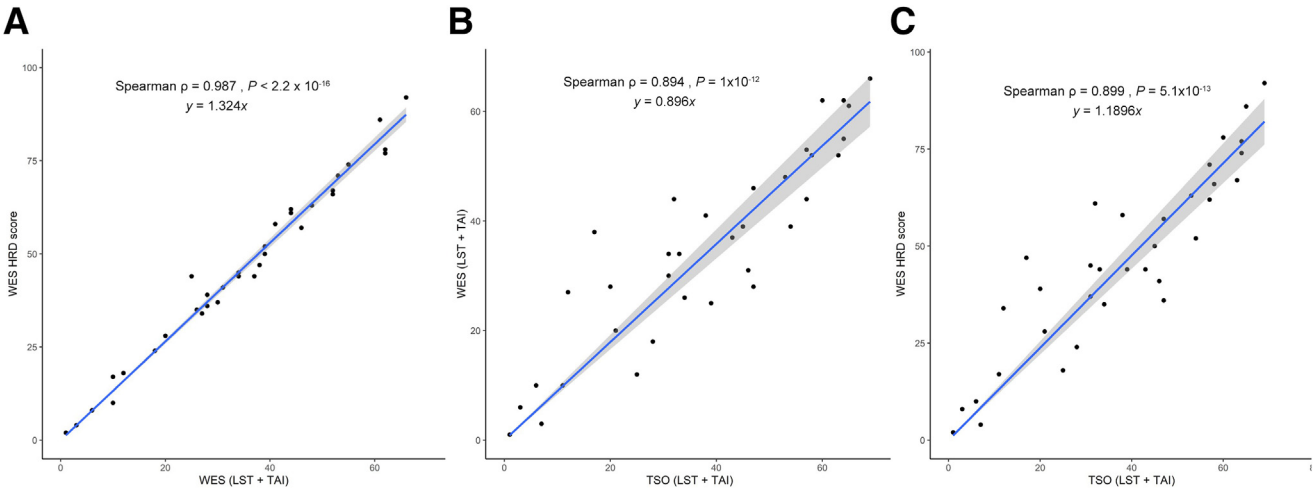

Figure 3 Correlation and regression of the telomeric-allelic imbalance (TAI) + large-scale state transition (LST) and homologous recombination deficiency (HRD) scores derived from whole-exome sequencing (WES) and targeted panel sequencing; the gray shaded area highlights the 95% CI. Depicted are the correlations of WES LST + TAI to WES HRD scores (**A**), of the targeted panel LST + TAI to WES LST + TAI scores (**B**), and of the targeted panel LST + TAI to WES HRD scores used for calibration (**C**), resulting in estimated HRD score function $y = 1.2x$. $n = 34$ (**A**–**C**). TSO, TruSight Oncology 500.

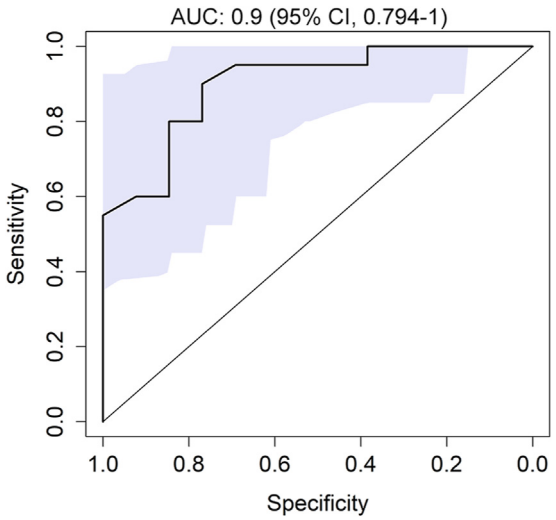

Figure 4 The receiver operating characteristic curve for 34 paired whole-exome sequencing (WES) and targeted panel samples illustrates the performance of the targeted panel-based model in distinguishing between homologous recombination deficiency (HRD)—positive and HRD-negative instances with WES results as ground truth; the gray shaded area represents the 95% CI. AUC, area under the curve.

complements the data obtained from targeted panel sequencing.

This study reports a method that leverages on OTR reads obtained from hybrid capture panel sequencing to comprehensively interrogate CNV profiles. Specifically, the study investigated off-target reads derived from a large 1.4-Mb targeted panel comprising 523 genes by using a bin size of 100-kb segments, providing sufficient resolution for the detection of nonfocal gains and deletions. To ensure the accuracy and precision of results, quality control parameters, such as a maximum mean absolute deviation of 0.17 (Supplemental Figure S1), were established, which, in our cohort, could be achieved usually with 2 million OTR reads after deduplication and a tumor cell content threshold of 20%. For application of this method with other targeted panels, quality parameters need to be selected individually.

Subsequently, in a proof-of-principle study, this approach was used to robustly estimate the HRD score, a biomarker that was approved recently by the US Food and Drug Administration,[12] although LOH could not be assessed using this approach. On the basis of previously published genomic scar signatures,[25] this study demonstrated the

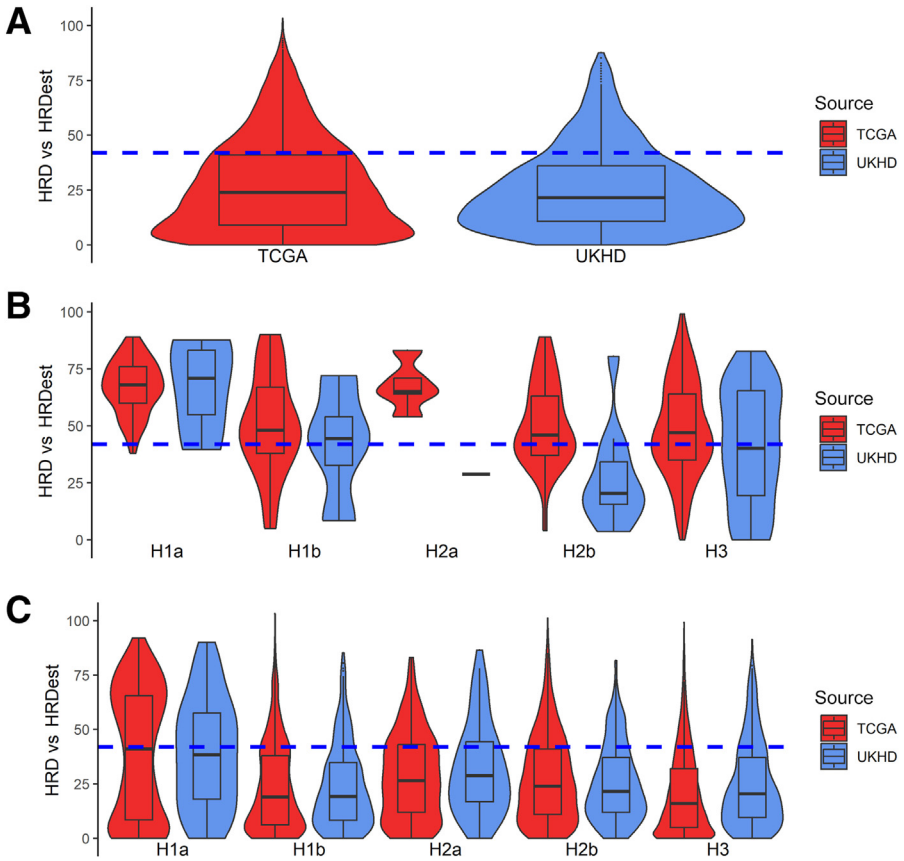

Figure 5 **A**: Pan-cancer distribution of homologous recombination deficiency (HRD) for The Cancer Genome Atlas (TCGA) and estimated HRD (HRDest) for the University Hospital Heidelberg (UKHD) cohort. In detail comparison of TCGA and UKHD cohort split by the homologous recombination repair (HRR)—status classification: H1a: *BRCA* deletion; H1b: non-*BRCA* HRR-gene deletion; H2a: *BRCA* variant of uncertain significance; H2b: non-*BRCA* HRR-gene variant of uncertain significance; H3: no mutations in HRR genes for ovarian cancer. **B:** HRDest scores for UKHD and TCGA HRD scores. **C:** Pan-cancer HRDest scores for UKHD and TCGA HRD scores. The **blue dashed lines** represent the threshold of 42. $n = 77$ (**B**, UKHD); $n = 411$ (**B**, TCGA); $n = 1086$ (**C**, UKHD); $n = 4875$ (**C**, TCGA).

combination of TAI and LST alone actually correlates well with the complete HRD score (LOH + TAI + LST), indicating that this method can be used as a robust proxy for the complete HRD score. Similar results were also reported by Eeckhoutte et al[14] when they used shallow whole-genome sequencing. In evaluating this approach to estimate HRD scores based on OTR reads, the results (HRDest scores) were compared with HRD scores derived from WES and demonstrated a significant correlation between these two methods, indicating that OTRs can provide a reliable resource for HRD score estimation in large cohorts of panel-sequenced clinical samples. In addition, no significant differences were observed in the distributions of single-nucleotide polymorphism−array−derived HRD scores in TCGA compared with the HRDest score distribution when considering all analyzed cancer types together or individually; similar results were obtained considering the HRR classification, further supporting the validity of this approach. There was high concordance between samples with *BRCA1/2* loss-of-function mutations and the significantly higher HRDest values observed for cases with a biallelic loss. Finally, when analyzing reference standards with known CNV profiles and HRD scores, the derived HRDest scores did not show any deviations from the known HRD values.

In summary, this proof-of-principle study illustrates a potential direct application of the analysis of OTR reads in clinical care. Although the application of this method to directly calculate HRD scores for therapy response prediction requires further study and validation, this study demonstrates that these readily available off-target reads obtained in panel sequencing approaches can be used to identify potentially HRD-positive cases for various clinical purposes, including enrichment for clinical trials. In these cases, further orthogonal testing by approved clinical tests could be used to corroborate the result derived from OTR analysis. Together with other studies in the field, this study contributes to a better understanding of off-target reads derived from targeted panel sequencing data and highlights their potential applications.

Disclosure Statement

O.N. reports personal fees from Novartis outside the submitted work. A.-L.V. reports personal fees from Astra Zeneca outside the submitted work. J.B. reports a grant from the German Cancer Aid outside the submitted work. S.F. reports consulting or advisory board membership for Bayer, Illumina, and Roche; honoraria from Amgen, Eli Lilly, PharmaMar, and Roche; research funding from AstraZeneca, Pfizer, PharmaMar, and Roche; and travel or accommodation expenses from Amgen, Eli Lilly, Illumina, PharmaMar, and Roche. P.S. reports personal fees from Bristol Myers Squibb, Merck Sharp & Dohme (MSD), Incyte, Janssen, Amgen, Novartis, Roche, and AstraZeneca outside the submitted work. A.S. reports grants and personal fees from Bayer and Bristol Myers Squibb; grants from Chugai; and personal fees from AstraZeneca, MSD, Takeda, Seattle Genetics, Novartis, Illumina, Thermo Fisher, Eli Lilly, and Takeda outside the submitted work. D.K. reports personal fees from AstraZeneca, Bristol Myers Squibb, Pfizer, Lilly, Agilent, and Takeda outside the submitted work. The remaining authors declare no conflict of interest.

Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.jmoldx.2024.02.008*.

References

1. Kuilman T, Velds A, Kemper K, Ranzani M, Bombardelli L, Hoogstraat M, Nevedomskaya E, Xu G, de Ruiter J, Lolkema MP, Ylstra B, Jonkers J, Rottenberg S, Wessels LF, Adams DJ, Peeper DS, Krijgsman O: CopywriteR: DNA copy number detection from off-target sequence data. Genome Biol 2015, 16:49
2. Laver TW, De Franco E, Johnson MB, Patel KA, Ellard S, Weedon MN, Flanagan SE, Wakeling MN: SavvyCNV: genome-wide CNV calling from off-target reads. PLoS Comput Biol 2022, 18: e1009940
3. Gusev A, Groha S, Taraszka K, Semenov YR, Zaitlen N: Constructing germline research cohorts from the discarded reads of clinical tumor sequences. Genome Med 2021, 13:179
4. Mangul S, Brito JJ, Groha S, Zaitlen N, Gusev A: Seeing beyond the target: leveraging off-target reads in targeted clinical tumor sequencing to identify prognostic biomarkers. bioRxiv 2021. [Preprint] doi: 10.1101/2021.05.28.446240
5. Marabelle A, Fakih M, Lopez J, Shah M, Shapira-Frommer R, Nakagawa K, Chung HC, Kindler HL, Lopez-Martin JA, Miller WH Jr, Italiano A, Kao S, Piha-Paul SA, Delord JP, McWilliams RR, Fabrizio DA, Aurora-Garg D, Xu L, Jin F, Norwood K, Bang YJ: Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. Lancet Oncol 2020, 21: 1353−1365
6. Marabelle A, Le DT, Ascierto PA, Di Giacomo AM, De Jesus-Acosta A, Delord JP, Geva R, Gottfried M, Penel N, Hansen AR, Piha-Paul SA, Doi T, Gao B, Chung HC, Lopez-Martin J, Bang YJ, Frommer RS, Shah M, Ghori R, Joe AK, Pruitt SK, Diaz LA Jr: Efficacy of pembrolizumab in patients with noncolorectal high microsatellite instability/mismatch repair-deficient cancer: results from the phase II KEYNOTE-158 study. J Clin Oncol 2020, 38:1−10
7. Stenzinger A, Endris V, Budczies J, Merkelbach-Bruse S, Kazdal D, Dietmaier W, Pfarr N, Siebolts U, Hummel M, Herold S, Andreas J, Zoche M, Togel L, Rempel E, Maas J, Merino D, Stewart M, Zaoui K, Schlesner M, Glimm H, Frohling S, Allen J, Horst D, Baretton G, Wickenhauser C, Tiemann M, Evert M, Moch H, Kirchner T, Buttner R, Schirmacher P, Jung A, Haller F, Weichert W, Dietel M: Harmonization and standardization of panel-based tumor mutational burden measurement: real-world results and recommendations of the quality in pathology study. J Thorac Oncol 2020, 15:1177−1189
8. Budczies J, Allgauer M, Litchfield K, Rempel E, Christopoulos P, Kazdal D, Endris V, Thomas M, Frohling S, Peters S, Swanton C, Schirmacher P, Stenzinger A: Optimizing panel-based tumor mutational burden (TMB) measurement. Ann Oncol 2019, 30:1496−1506
9. Geoerger B, Kang HJ, Yalon-Oren M, Marshall LV, Vezina C, Pappo A, Laetsch TW, Petrilli AS, Ebinger M, Toporski J, Glade-

Bender J, Nicholls W, Fox E, DuBois SG, Macy ME, Cohn SL, Pathiraja K, Diede SJ, Ebbinghaus S, Pinto N: Pembrolizumab in paediatric patients with advanced melanoma or a PD-L1-positive, advanced, relapsed, or refractory solid tumour or lymphoma (KEYNOTE-051): interim analysis of an open-label, single-arm, phase 1-2 trial. Lancet Oncol 2020, 21:121−133

10. Le DT, Kim TW, Van Cutsem E, Geva R, Jager D, Hara H, Burge M, O'Neil B, Kavan P, Yoshino T, Guimbaud R, Taniguchi H, Elez E, Al-Batran SE, Boland PM, Crocenzi T, Atreya CE, Cui Y, Dai T, Marinello P, Diaz LA Jr, Andre T: Phase II open-label study of pembrolizumab in treatment-refractory, microsatellite instability-high/mismatch repair-deficient metastatic colorectal cancer: keynote-164. J Clin Oncol 2020, 38:11−19
11. Kazdal D, Endris V, Allgauer M, Kriegsmann M, Leichsenring J, Volckmar AL, Harms A, Kirchner M, Kriegsmann K, Neumann O, Brandt R, Talla SB, Rempel E, Ploeger C, von Winterfeld M, Christopoulos P, Merino DM, Stewart M, Allen J, Bischoff H, Meister M, Muley T, Herth F, Penzel R, Warth A, Winter H, Frohling S, Peters S, Swanton C, Thomas M, Schirmacher P, Budczies J, Stenzinger A: Spatial and temporal heterogeneity of panel-based tumor mutational burden in pulmonary adenocarcinoma: separating biology from technical artifacts. J Thorac Oncol 2019, 14:1935−1947
12. Ray-Coquard I, Pautier P, Pignata S, Perol D, Gonzalez-Martin A, Berger R, Fujiwara K, Vergote I, Colombo N, Maenpaa J, Selle F, Sehouli J, Lorusso D, Guerra Alia EM, Reinthaller A, Nagao S, Lefeuvre-Plesse C, Canzler U, Scambia G, Lortholary A, Marme F, Combe P, de Gregorio N, Rodrigues M, Buderath P, Dubot C, Burges A, You B, Pujade-Lauraine E, Harter P; PAOLA-1 Investigators Olaparib plus bevacizumab as first-line maintenance in ovarian cancer. N Engl J Med 2019, 381:2416−2428
13. Powell SN, Bindra RS: Targeting the DNA damage response for cancer therapy. DNA Repair (Amst) 2009, 8:1153−1165
14. Eeckhoutte A, Houy A, Manie E, Reverdy M, Bieche I, Marangoni E, Goundiam O, Vincent-Salomon A, Stoppa-Lyonnet D, Bidard FC, Stern MH, Popova T: ShallowHRD: detection of homologous recombination deficiency from shallow whole genome sequencing. Bioinformatics 2020, 36:3888−3889
15. Menzel ME, Endris V, Schwab C, Kluck K, Neumann O, Beck S, Ball M, Schaaf C, Fröhling S, Lichtner P, Schirmacher P, Kazdal D, Stenzinger A, Budczies J: Accurate tumor purity determination is critical for the analysis of homologous recombination deficiency (HRD). Transl Oncol 2023, 35:101706
16. Quinlan AR, Hall IM: BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010, 26:841−842
17. O'Connor BD, Van der Auwera GA: Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. Sebastopol, CA: O'Reilly Media, 2020
18. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ: A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011, 43:491−498
19. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010, 20:1297−1303
20. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, Szallasi Z, Eklund AC: Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann Oncol 2015, 26:64−70
21. Sztupinszki Z, Diossy M, Krzystanek M, Reiniger L, Csabai I, Favero F, Birkbak NJ, Eklund AC, Syed A, Szallasi Z: Migrating the SNP array-based homologous recombination deficiency measures to next generation sequencing data of breast cancer. NPJ Breast Cancer 2018, 4:16
22. Van Loo P, Nordgard SH, Lingjaerde OC, Russnes HG, Rye IH, Sun W, Weigman VJ, Marynen P, Zetterberg A, Naume B, Perou CM, Borresen-Dale AL, Kristensen VN: Allele-specific copy number analysis of tumors. Proc Natl Acad Sci U S A 2010, 107:16910−16915
23. Rempel E, Kluck K, Beck S, Ourailidis I, Kazdal D, Neumann O, Volckmar AL, Kirchner M, Goldschmid H, Pfarr N, Weichert W, Hubschmann D, Frohling S, Sutter C, Schaaf CP, Schirmacher P, Endris V, Stenzinger A, Budczies J: Pan-cancer analysis of genomic scar patterns caused by homologous repair deficiency (HRD). NPJ Precis Oncol 2022, 6:36
24. Petrova-Drus K, Quesada AE, Bowman AS, Ptashkin R, Yao J, Arcila ME, Ho C, Moung C, Regalado J, Benayed R, Benhamida JK, Galera PK, Dogan A, Vanderbilt C: Quantitative off-target detection of Epstein-Barr virus-derived DNA in routine molecular profiling of hematopoietic neoplasms by panel-based hybrid-capture next-generation sequencing. J Mol Diagn 2022, 24:69−78
25. Marquard AM, Eklund AC, Joshi T, Krzystanek M, Favero F, Wang ZC, Richardson AL, Silver DP, Szallasi Z, Birkbak NJ: Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. Biomark Res 2015, 3:9