

HRDex: a tool for deriving homologous recombination deficiency (HRD) scores from whole exome sequencing data

John Pluta^{1*}, Ryan Hausler^{2*}, Brad Wubbenhorst¹, Heena Desai², Susan M. Domchek^{1,4}, Katherine L. Nathanson^{1,3,4}, Kara N. Maxwell^{2,3,4}

1. Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, 19104
2. Division of Hematology/Oncology, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, 19104
3. Department of Genetics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, 19104
4. Abramson Cancer Center, University of Pennsylvania, Philadelphia, PA, 19104

Correspondence to Kara N. Maxwell, MD PhD, Division of Hematology/Oncology, Department of Medicine, 421 Curie Blvd, Room 810, Philadelphia, PA 19104, USA, Tel: 215-898-6944, Fax: 215-573-7039, Email: kara.maxwell@pennmedicine.upenn.edu

Keywords HRD, homologous recombination deficiency, LST, genomic LOH, allelic imbalance, exome sequencing, *BRCA* deficiency

Word count Abstract: 286 words; Manuscript: 2001 words; 2 Main Figure/Tables; 47 references

Abstract

Background: Breast and ovarian tumors in patients with biallelic *BRCA1* and *BRCA2* mutations either by germline mutations accompanied by allele-specific loss of heterozygosity (LOH) or truncal somatic mutations respond to PARP inhibition. The repair of double stranded DNA breaks in tumors these tumors leads to homologous recombination deficiency (HRD), which can be measured using a variety of genomic and transcriptomic signatures. However, the optimal biomarker for *BRCA* deficiency is unknown.

Methods: We developed HRDex to determine HRD and its composite scores from allele specific copy number data analysis of whole exome sequencing (WES) data and examined the discriminatory ability of HRDex and other genomic and transcriptomic measures to identify *BRCA* deficiency in breast and ovarian tumors from The Cancer Genome Atlas (TCGA).

Results: HRDex scores have high correlation with SNP array based HRD scores in both breast and ovarian cancers. HRDex scores have high discriminatory accuracy to distinguish *BRCA* deficient breast tumors, similar to SNP array based scores (AUC 0.87 vs 0.90); however, discriminatory ability for ovarian tumors was lower (AUC 0.79 vs 0.90). HRD-LST had the best discriminatory ability of the three composite HRD scores. HRDex had higher discriminatory ability for identification of *BRCA* deficiency than RNA expression based scores (eCARD, tp53, RPS and PARPi7) in breast and ovarian tumors. Tumor mutational burden (TMB) was associated with *BRCA* deficiency in breast but not ovarian cancer. Combining HRDex score with mutational signature 3 modestly increased discriminatory ability for *BRCA* deficient breast and ovarian tumors (breast: AUC 0.90 vs 0.87; ovarian: AUC 0.83 vs 0.79).

Conclusions: WES based HRD scores perform similarly to SNP array HRD scores, and better than other genomic or transcriptomic signatures, for identification of tumors with *BRCA* deficiency due to biallelic *BRCA* loss.

Introduction

Tumors develop homologous recombination deficiency (HRD) due to a variety of mechanisms, including both germline and somatic inactivation of BRCA1 and BRCA2. Repair of double stranded DNA breaks in tumors with HRD occurs by error prone mechanisms leading to genomic instability which can be measured using a variety of genomic and transcriptomic signatures^{1,2}. HRD has clinical relevance as when tumor cells are stressed with further DNA damage, for example due to stalled replication forks resulting from trapping of PARP on DNA by inhibitors, cell death occurs by synthetic lethality³. PARP inhibitors have FDA approved indications in ovarian, breast, pancreatic and prostate cancer patients with germline *BRCA1/2* mutations⁴⁻¹¹. However, inactivation of a number of HR genes by genetic mutation or epigenetic silencing may lead to HRD in *BRCA1/2* wildtype tumors, but which specific HR genes should be interrogated is unclear¹. On the other hand, not all *BRCA1/2* alterations may lead to HRD, ie in the absence of locus specific loss of heterozygosity^{12,13}. Therefore, a more generalizable phenotypic readout of HRD is needed.

Biomarkers of *BRCA* deficiency have long been sought¹⁴. Array based comparative genomic hybridization (aCGH) signatures can distinguish tumors with copy number changes seen in germline *BRCA1* and/or *BRCA2* tumors^{15,16}. Transcriptomic based measures of DNA repair deficiency have also been developed (ie eCARD¹⁷, PARPi7¹⁸, RPS¹⁹, tp53 score²). The most widely used HRD measure from tumors was originally developed using SNP arrays^{20,21}. These HRD scores identify three discrete gross chromosomal abnormalities: genomic LOH (HRD-LOH)²², telomeric allelic imbalance (HRD-TAI)²³, and large state transitions (HRD-LST)²⁴. The combination HRD score has been extensively validated, demonstrating ability to predict *BRCA* deficiency²⁵ and response to both platinum chemotherapy²⁶ and PARP inhibitor^{27,28}. However, at present, exome or other targeted based next generation sequencing (NGS) approaches, without aCGH, SNP arrays, or RNA expression analyses, are commonly performed

clinically. Therefore, we developed an algorithm called HRDex to determine HRD scores from allele specific copy number data analysis of WES data.

Methods

TCGA breast and ovary level 1 whole exome sequencing data (including BAM) were obtained from The Cancer Genome Atlas (TCGA) with a dbGAP approved project²⁹. Downloaded TCGA BAM files were aligned to the hg19 assembly of the human genome. All exonic single nucleotide and insertion/deletion variants were identified using a combination of GATK³⁰, Mutect³¹ and VarScan2³². Variants underwent initial quality control filtering according to GATK best practices³⁰. Germline and somatic *BRCA1/2* alterations were identified as previously described²⁹. *BRCA* deficient samples were defined as tumor-normal pairs with a germline loss of function *BRCA1* or *BRCA2* mutation with loss of heterozygosity (LOH) or a somatic *BRCA1* or *BRCA2* mutation with allele frequency (AF)>50%. Tumors with germline or somatic mutations in other genes in the homologous recombination pathway were excluded. This resulted in splitting of 708 TCGA breast cancer samples into 699 *BRCA* proficient and 39 *BRCA* deficient and splitting of 152 TCGA ovarian cancer samples into 99 *BRCA* proficient and 53 *BRCA* deficient.

HRDex was created to use an input file of genomic segments; where each segment is defined by its start and end position in physical coordinates with the copy number of each allele at said positions. Sequenza³³ (v2.1.2, release date 10/10/2015) was used to create input files for the TCGA cohorts in this study. As the Sequenza segmentation algorithm can return false breaks (e.g. reporting two segments when in truth there is only one), if two adjacent segments had the same copy number and number of A and B alleles, and the gap between them was less than 3 Mb, this was considered a false break, and the two segments were joined. The length of the newly joined segment was defined as the distance between the start point of the first segment and the end point of the second.

To create the HRDex component scores and the sum, loss of heterozygosity (LOH) events were defined as segments that have a B allele copy number of 0, length greater than 15 Mb, and length less than 90% of the total chromosomal length. Large state transition (LST) events were defined as two adjacent segments with a gap of at least 3 Mb, where each segment is greater than 10 Mb and does not cross the centromere. Non-telomeric allelic imbalance (NTAI) events were defined as any segment that is not in either telomere, that does not cross the centromere, and has a segment size of greater than 11 Mb; segments with copy number equal to ploidy value were removed.

The ploidy of a tumor can be expressed as the number of genomic doublings that occur: $p=2^n$; where p is the ploidy, and n is the number of genomic doublings. HRDex component scores were normalized to ploidy by dividing the total number of events by the number of genomic doublings. Ploidy was estimated from the data using Sequenza and averaged over the genome, therefore, a correction factor was included to account for continuous values. We defined a ploidy correction factor, k :

$$k = \left(2^{\frac{\log(p)}{\log(2)} - 1} \right)^{-1}$$

HRDex.LST and HRDex.LOH were normalized by multiplying the number of events by k (HRDex.LSTr and HRDex.LOHr). HRD.NTAI was normalized by subtracting the main copy number segments first, and then multiplying by k (HRDex.NTAIm). HRDex is the sum of the three component scores.

For public data analyses and comparisons of HRDex scores to SNP array based scores, SNP array based HRD scores, tumor mutational burden, aneuploidy score, expression signatures and mutational signatures for TCGA BRCA and OV datasets were obtained from Ref².

ROC curves comparing the two data sets were computed, using leave-one-out cross validation to estimate out of sample performance. Area under the curves (AUC) of the two scores were statistically compared using deLong's test.

Results

HRD scores have been defined as a sum of three genomic events, LOH, LST and NTAI²⁵. Because these events may increase in number based on increasing ploidy status of a tumor and not because of deficiency in the HR pathway, we tested whether there was a relationship between SNP array based HRD scores and ploidy in TCGA pancancer data². The relationship between ploidy and SNP array based HRD scores differed between cancer types, with stronger correlations seen in certain cancer types such as uterine (UCEC), sarcoma (SARC), breast (BRCA) and stomach (STAD). Furthermore, certain cancer types had no correlation between HRD score and ploidy (**Supplementary Figure 1a**).

Because we noted a relationship between HRD and ploidy in certain cancer types, including breast, we created a tool to calculate ploidy-normalized HRD scores from allele specific copy number based analysis of whole exome sequencing data, called HRDex. HRDex is a freely available R package (<https://github.com/maxwell-lab/HRDex>), and outputs three component scores HRD.LOH, HRD.NTAI, HRD.LST and the sum of these three metrics, HRD.Score (**Supplementary Table 1**). HRDex scores normalized for ploidy remove the coorelation between HRD and ploidy, as shown for breast (**Supplementary Figure 1b**).

In order to test the HRDex algorithm, TCGA breast and ovary tumors were stratified into *BRCA* deficient and proficient tumors based on analysis of germline and somatic *BRCA1/2* mutations. Only tumors with germline or somatic *BRCA1/2* mutations with biallelic loss were considered “*BRCA* deficient”. Tumors with other germline or somatic mutations in the homologous recombination pathway, germline *BRCA1/2* tumors without biallelic loss, and tumors with low VAF somatic *BRCA1/2* mutations were excluded. Remaining tumors were

“*BRCA* proficient”. HRD scores based on SNP-array data were an excellent classifier of *BRCA* deficiency in TCGA-Breast, as expected (AUC=0.9, **Figure 1a**). HRDex had a high correlation with SNP array based HRD and its component scores in TCGA breast tumors ($r^2 = 0.68-0.87$, **Supplementary Figure 2a**). HRDex, before and after ploidy normalization performed almost as well as SNP array based HRD in TCGA-Breast, with AUCs of 0.87 and 0.88, respectively (**Figure 1a**). The performance of SNP array based HRD versus HRDex was not significantly different ($p = 0.68$). HRD.score and each component were significantly higher in *BRCA* deficient versus proficient breast tumors, with the most significant difference seen for HRD.LST (**Figure 1b**). Accordingly, the AUC for HRD.LST was similar to summed HRD.Score, although all component scores had AUC>0.80 (**Figure 1c**).

HRD scores based on SNP-array were also an excellent classifier of *BRCA* deficiency in TCGA-Ovary, as expected (AUC=0.9, **Figure 1c**). HRDex had a high correlation with SNP array based HRD and its component scores in TCGA ovarian tumors, although lower than for breast ($r^2 = 0.42-0.79$, **Supplementary Figure 2b**). HRDex, before and after ploidy normalization performed well in TCGA-Ovary, with AUCs of 0.81 and 0.82, respectively (**Figure 1c**). The performance of SNP array versus HRDex was not significantly different ($p = 0.99$). HRD.score and each component were significantly higher in *BRCA* deficient versus proficient ovary tumors, although the magnitude of differences were smaller in ovary compared to breast (**Figure 1e**). The AUC for the component HRD scores were lower in ovary, particularly for HRD-NTAI (**Figure 1f**).

Finally, we determined the discriminatory accuracy of the published DNA Damage Response (DDR) resource scores² for distinguishing *BRCA* deficient versus proficient breast and ovarian tumors and compared to HRDex (**Table 1**). In TCGA-Breast, tumor mutational burden (TMB) had the highest AUC (0.98) for identifying *BRCA* deficient tumors; however, in TCGA-Ovary, TMB was a poor classifier (AUC=0.67). Aneuploidy score (defined as the sum of all aneuploid chromosome arms) was a poor classifier of *BRCA* deficient tumors in both TCGA-

Breast and TCGA-ovary. Of DDR expression based measures, eCARD performed the best in TCGA-Breast (AUC:0.72); whereas PARPi7 performed the best in TCGA-Ovary (AUC:0.65). However, all DDR expression signatures had lower AUCs than genomic measures. In both TCGA-Breast and Ovarian cancer, mutational signature 3 was associated with *BRCA* deficiency (AUC 0.76 and 0.75). Combining HRDex or HRDex.LST score with mutational signature 3 modestly increased discriminatory ability for *BRCA* deficient tumors (breast: AUC 0.90 vs 0.87; ovarian: AUC 0.83 vs 0.79) versus the HRDex score alone.

Discussion

Homologous recombination deficiency (HRD) scores were originally developed using SNP array based data for the purpose of distinguishing germline *BRCA1/2* related tumors²²⁻²⁴. SNP arrays are rarely performed clinically or in genomics research due to the decreasing costs of next generation sequencing (NGS) based approaches. While whole genome sequencing (WGS) approaches are known to identify tumors with HRD³⁴⁻³⁶, WGS is still cost-prohibitive in most clinical and research settings. Given the large volume of whole exome sequencing (WES) data that is publicly available, we developed HRDex to derive HRD scores from allele specific copy number data derived from WES. We show that HRDex has comparable discriminatory accuracy to SNP array based HRD scores.

Currently, germline and/or somatic mutational analysis by NGS is the most commonly used tool to identify potential patients for platinum and/or PARP inhibitor eligibility³⁷; however pan-cancer studies suggest mutational analyses alone will miss 30-40% of HRD tumors^{34,38}. A number of trials have shown that HRD score is associated with response to neoadjuvant platinum based chemotherapy in breast cancer^{20,25,26,39}. HRD scores may also predict responsiveness to platinum and PARP inhibitor in ovarian cancer^{27,40} and in tumors beyond breast and ovarian cancer⁴¹. Therefore, using NGS based HRD measures, such as HRDex or

other NGS based approaches⁴², may expand identification of patients responsive to platinum or PARP inhibitors beyond just mutational analysis³⁷.

It is important to note that the HRD score has not predicted platinum or PARPi responsiveness in all trials^{43,44}, possibly due to reversion of HR function despite maintenance of genomic scars of prior HRD. Functional assays, such as RAD51 foci⁴⁵⁻⁴⁸ or ex vivo approaches⁴⁹, may be superior in determining platinum and PARPi sensitivity but are not yet scalable in a clinical environment. A recent analysis suggests that incorporating *BRCA1* and *RAD51C* methylation assays into genomic analyses of all HR-related genes will find the vast majority of HRD tumors⁵⁰; however, these assays are also not yet standardized.

HRD scores are important biomarkers of PARP inhibitor and platinum responsiveness, and useful in understanding the underlying biology of human tumors. HRD scores, such as HRDex, derived from readily available NGS based approaches, are useful in both clinical and research settings for the determination of HRD and study of PARP inhibitor responsiveness.

Table 1: Performance of DDR scores in discriminating *BRCA* proficient and deficient tumors

Source	Score	TCGA BRCA	TCGA Ovary
WES Scores	HRDex.Score	0.874	0.794
	HRDex.LSTr	0.870	0.799
	HRDex.LOHr	0.830	0.731
	HRDex.NTAIm	0.748	0.574
	Aneuploidy Score	0.547	0.590
	TMB (Mutect)	0.977	0.672
SNP Array Scores	DDR HRD Sum	0.888	0.896
	DDR HRD LST	0.896	0.884
	DDR HRD TAI	0.861	0.800
	DDR HRD LOH	0.857	0.830
Expression signatures	eCARD	0.721	0.564
	tp53_score	0.685	0.622
	RPS	0.641	0.517
	PARPi7	0.591	0.647
Mutational Signatures	Signature.3 (BRCA)	0.761	0.753
	Signature.1A (Aging)	0.680	0.786
	Signature.13 (APOBEC)	0.569	0.515
	Signature.7 (UV)	0.542	0.506
	Signature.1B (Aging)	0.536	0.541
	Signature.2 (APOBEC)	0.533	0.544
	Signature.16 (Unk)	0.527	0.505
	Signature.21 (Unk)	0.527	0.521
	Signature.10 (POLE)	0.526	0.512
	Signature.4 (Smoking)	0.517	0.525
	Signature.9-null	0.515	0.496
	Signature.15 (MMR)	0.510	0.531
	Signature.12 (Unk)	0.506	0.500
	Signature.18 (Unk)	0.504	0.525
	Signature.8 (Unk)	0.503	0.566
	Signature.11 (Alkylators)	0.503	0.503
Combinations	Signature.5 (Unk)	0.502	0.508
	Signature.20 (MMR)	0.501	0.516
	Signature.19 (Unk)	0.501	0.505
	Signature.17 (Unk)	0.500	0.500
	Signature.14 -(Unk)	0.500	0.500
	Signature.6 (MMR)	0.496	0.507
	HRDex Score + Signature.3	0.895	0.834
	HRDex LST + Signature.3	0.891	0.843

Figure 1: Performance of HRDex in TCGA breast and ovarian tumors. **(a)** Discriminatory performance of HRDex scores for *BRCA* deficient versus proficient TCGA breast tumors by HRDex, ploidy normalized HRDex, and SNP array based HRD. **(b)** Violin plots of component and total HRDex scores in *BRCA* proficient and deficient TCGA breast tumors. **(c)** Discriminatory performance of HRDex-total, HRDex-LOH, HRDex-LST, HRDex-NTAI component scores for *BRCA* deficient versus proficient breast tumors. **(d)** Discriminatory performance of HRDex scores for *BRCA* deficient versus proficient TCGA ovarian tumors by HRDex, ploidy normalized HRDex, and SNP array based HRD. **(e)** Violin plots of component and total HRDex scores in *BRCA* proficient and deficient TCGA ovarian tumors. **(f)** Discriminatory performance of HRDex-total, HRDex-LOH, HRDex-LST, HRDex-NTAI component scores for *BRCA* deficient versus proficient ovarian tumors.

Availability of data and material: HRDex scripts are available at: <https://github.com/maxwell-lab/HRDex>

Conflict of Interest Disclosure: The authors have declared that no conflict of interest exists.

Acknowledgements: This work is supported by the National Cancer Institute (K08CA215312, KNM), the Burroughs Wellcome Foundation (#1017184, KNM), Basser Center for *BRCA* (KNM, SMD, KLN), V Foundation for Cancer Research (KLN), Gray Foundation (SMD, KLN) the Konner Family Foundation (KNM), and the Breast Cancer Research Foundation (SMD, KLN).

Authors' contributions. Study concept and design were performed by KNM and KLN. Data acquisition was performed by BW, and computational program created by BW, JP, and KNM. Data analysis and interpretation was performed by JP, RH, HD, KNM, SMD and KLN. Writing was performed by all authors.

References

1. Hoppe MM, Sundar R, Tan DSP, Jeyasekharan AD. Biomarkers for Homologous Recombination Deficiency in Cancer. *J Natl Cancer Inst.* 2018;110(7):704-713.
2. Knijnenburg TA, Wang L, Zimmermann MT, et al. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell reports.* 2018;23(1):239-254 e236.
3. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. *Science.* 2017;355(6330):1152-1158.
4. Tutt A, Robson M, Garber JE, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. *Lancet.* 2010;376(9737):235-244.
5. Gelmon KA, Tischkowitz M, Mackay H, et al. Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: a phase 2, multicentre, open-label, non-randomised study. *Lancet Oncol.* 2011;12(9):852-861.
6. Kaufman B, Shapira-Frommer R, Schmutzler RK, et al. Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. *J Clin Oncol.* 2015;33(3):244-250.
7. Morgan RD, Clamp AR, Evans DGR, Edmondson RJ, Jayson GC. PARP inhibitors in platinum-sensitive high-grade serous ovarian cancer. *Cancer Chemother Pharmacol.* 2018;81(4):647-658.
8. Shroff RT, Hendifar A, McWilliams RR, et al. Rucaparib Monotherapy in Patients With Pancreatic Cancer and a Known deleterious BRCA Mutation. *JCO Precis Oncol.* 2018;2018.
9. Zimmer AS, Gillard M, Lipkowitz S, Lee JM. Update on PARP Inhibitors in Breast Cancer. *Curr Treat Options Oncol.* 2018;19(5):21.

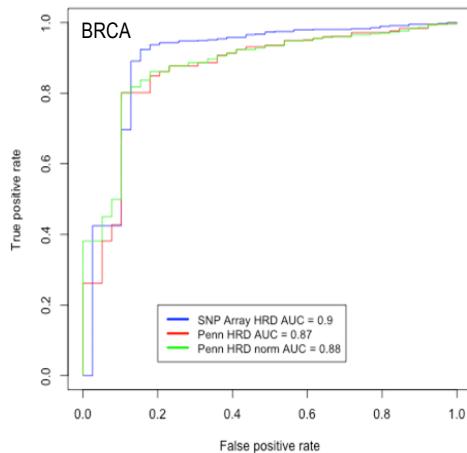
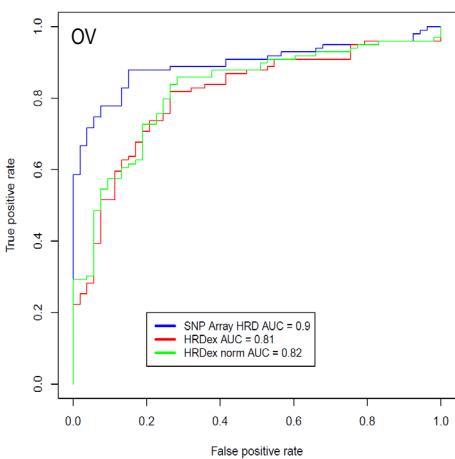
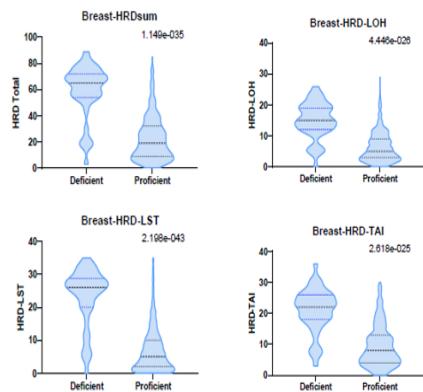
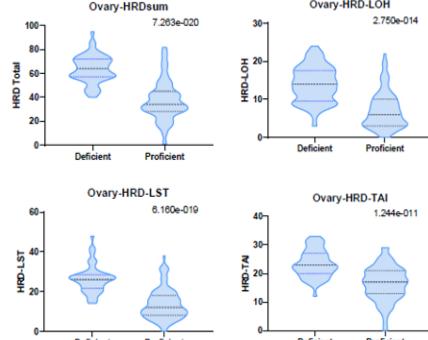
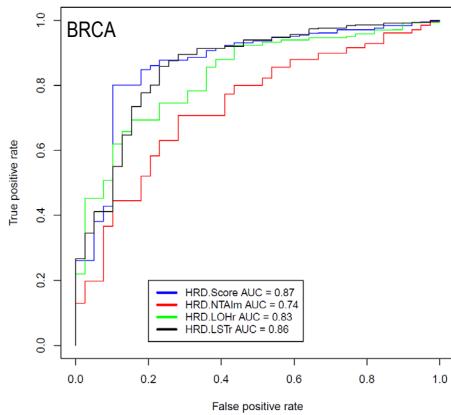
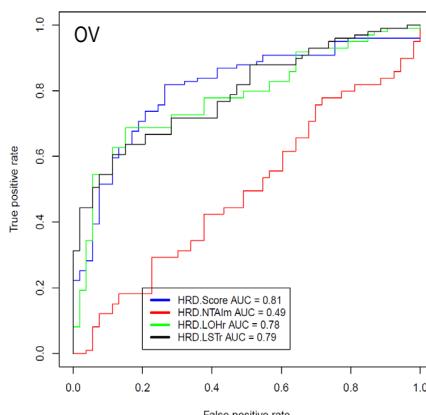
10. Abida W, Patnaik A, Campbell D, et al. Rucaparib in Men With Metastatic Castration-Resistant Prostate Cancer Harboring a BRCA1 or BRCA2 Gene Alteration. *J Clin Oncol.* 2020;38(32):3763-3772.
11. de Bono J, Mateo J, Fizazi K, et al. Olaparib for Metastatic Castration-Resistant Prostate Cancer. *N Engl J Med.* 2020;382(22):2091-2102.
12. Maxwell KN, Wubbenhorst B, De Sloover D, et al. BRCA locus specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. In: Communications AaN, ed2017.
13. Riaz N, Blecua P, Lim RS, et al. Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. *Nature communications.* 2017;8(1):857.
14. Stecklein SR, Sharma P. Tumor homologous recombination deficiency assays: another step closer to clinical application? *Breast Cancer Res.* 2014;16(4):409.
15. Vollebergh MA, Lips EH, Nederlof PM, et al. Genomic patterns resembling BRCA1- and BRCA2-mutated breast cancers predict benefit of intensified carboplatin-based chemotherapy. *Breast Cancer Res.* 2014;16(3):R47.
16. Vollebergh MA, Lips EH, Nederlof PM, et al. An aCGH classifier derived from BRCA1-mutated breast cancer and benefit of high-dose platinum-based chemotherapy in HER2-negative breast cancer patients. *Ann Oncol.* 2011;22(7):1561-1570.
17. Zimmermann MT, Jiang G, Wang C. Single-sample expression-based chemo-sensitivity score improves survival associations independently from genomic mutations for ovarian cancer Patients. *AMIA Jt Summits Transl Sci Proc.* 2016;2016:94-100.
18. Daemen A, Wolf DM, Korkola JE, et al. Cross-platform pathway-based analysis identifies markers of response to the PARP inhibitor olaparib. *Breast Cancer Res Treat.* 2012;135(2):505-517.

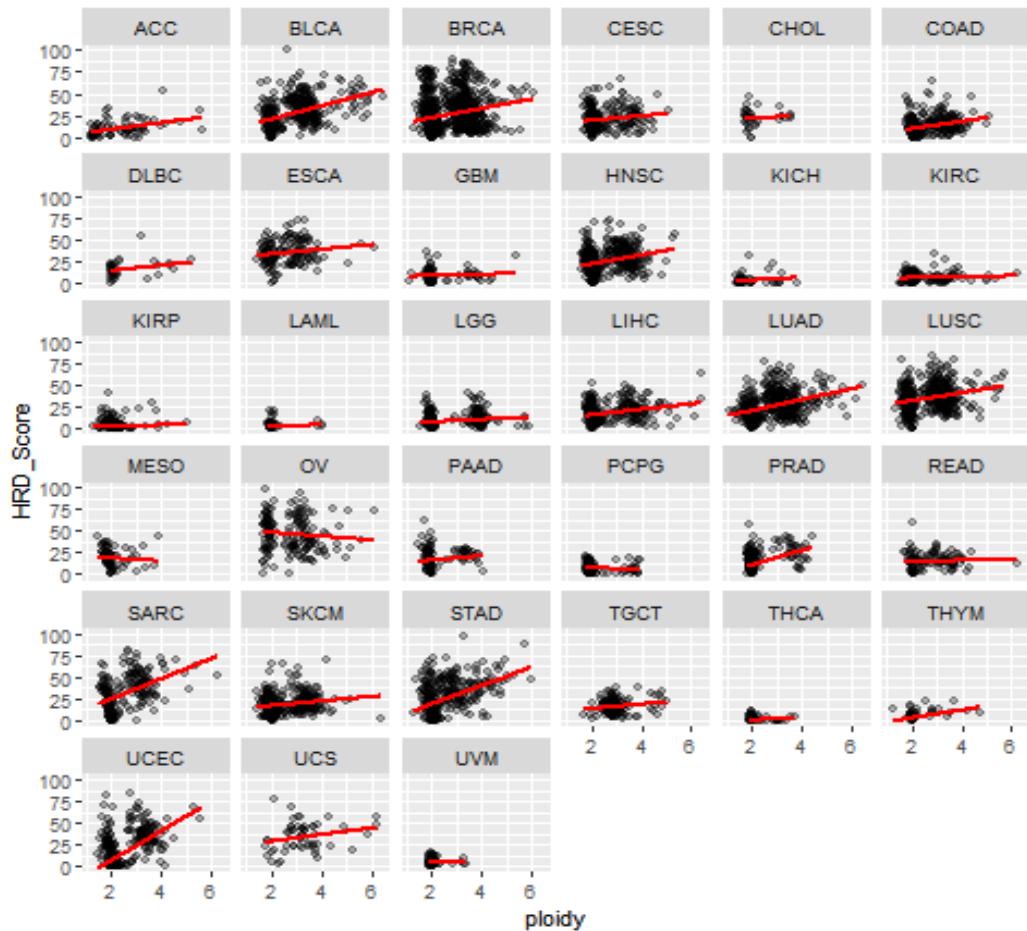
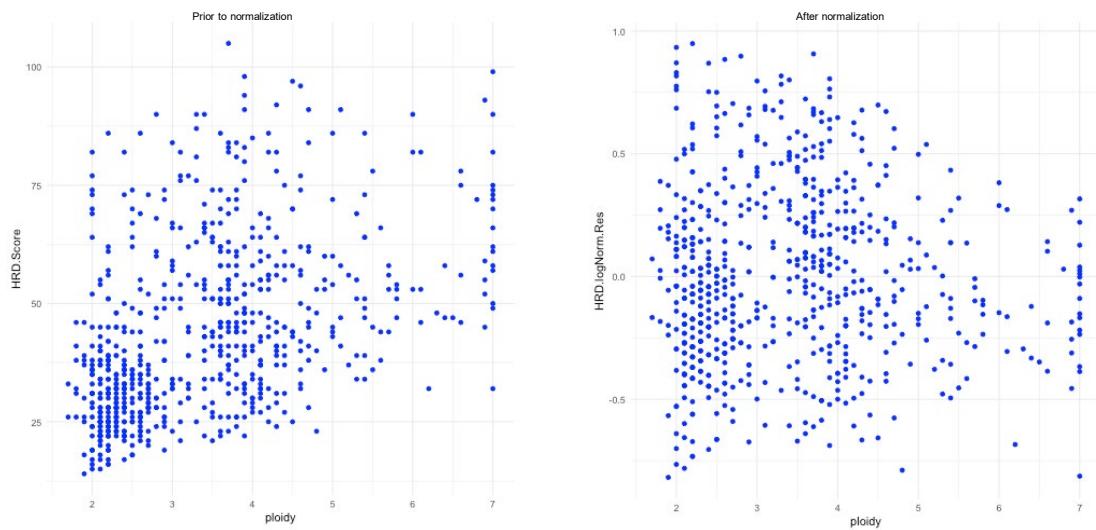
19. Pitroda SP, Pashtan IM, Logan HL, et al. DNA repair pathway gene expression score correlates with repair proficiency and tumor sensitivity to chemotherapy. *Sci Transl Med.* 2014;6(229):229ra242.
20. Telli ML, Timms KM, Reid J, et al. Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. *Clin Cancer Res.* 2016;22(15):3764-3773.
21. Timms KM, Abkevich V, Hughes E, et al. Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. *Breast Cancer Res.* 2014;16(6):475.
22. Abkevich V, Timms KM, Hennessy BT, et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br J Cancer.* 2012;107(10):1776-1782.
23. Birkbak NJ, Wang ZC, Kim JY, et al. Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. *Cancer discovery.* 2012;2(4):366-375.
24. Popova T, Manie E, Rieunier G, et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res.* 2012;72(21):5454-5462.
25. Telli ML, Hellyer J, Audeh W, et al. Homologous recombination deficiency (HRD) status predicts response to standard neoadjuvant chemotherapy in patients with triple-negative or BRCA1/2 mutation-associated breast cancer. *Breast Cancer Res Treat.* 2018;168(3):625-630.
26. Loibl S, Weber KE, Timms KM, et al. Survival analysis of carboplatin added to an anthracycline/taxane-based neoadjuvant chemotherapy and HRD score as predictor of response-final results from GeparSixto. *Ann Oncol.* 2018;29(12):2341-2347.

27. Lheureux S, Lai Z, Dougherty BA, et al. Long-Term Responders on Olaparib Maintenance in High-Grade Serous Ovarian Cancer: Clinical and Molecular Characterization. *Clin Cancer Res.* 2017;23(15):4086-4094.
28. Mirza MR, Monk BJ, Herrstedt J, et al. Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. *N Engl J Med.* 2016;375(22):2154-2164.
29. Maxwell KN, Wubbenhorst B, Wenz BM, et al. BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. *Nature communications.* 2017;8(1):319.
30. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-1303.
31. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol.* 2013;31(3):213-219.
32. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012;22(3):568-576.
33. Favero F, Joshi T, Marquard AM, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol.* 2015;26(1):64-70.
34. Davies H, Glodzik D, Morganella S, et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat Med.* 2017;23(4):517-525.
35. de Luca XM, Newell F, Kazakoff SH, et al. Using whole-genome sequencing data to derive the homologous recombination deficiency scores. *NPJ breast cancer.* 2020;6:33.
36. Degasperi A, Zou X, Amarante TD, et al. Substitution mutational signatures in whole-genome-sequenced cancers in the UK population. *Science.* 2022;376(6591).

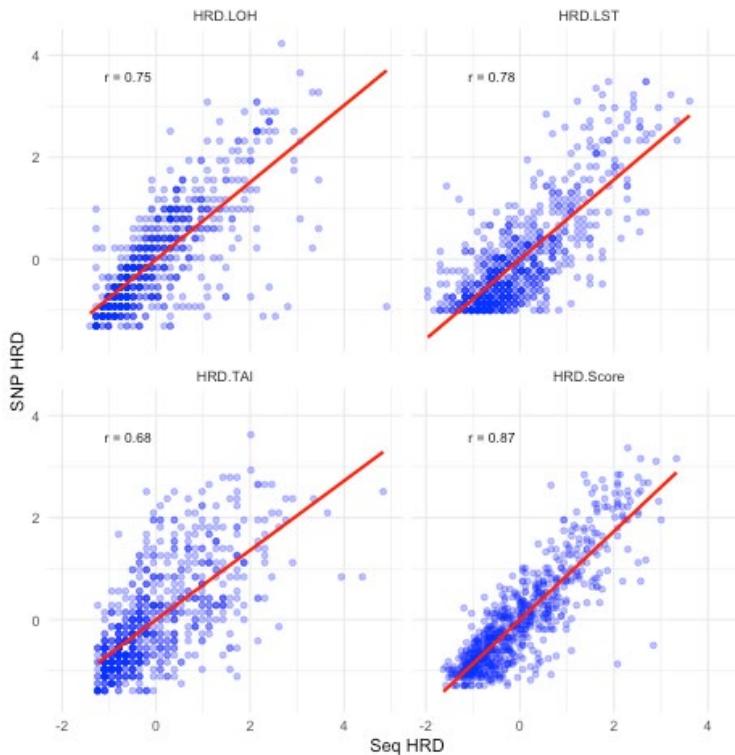
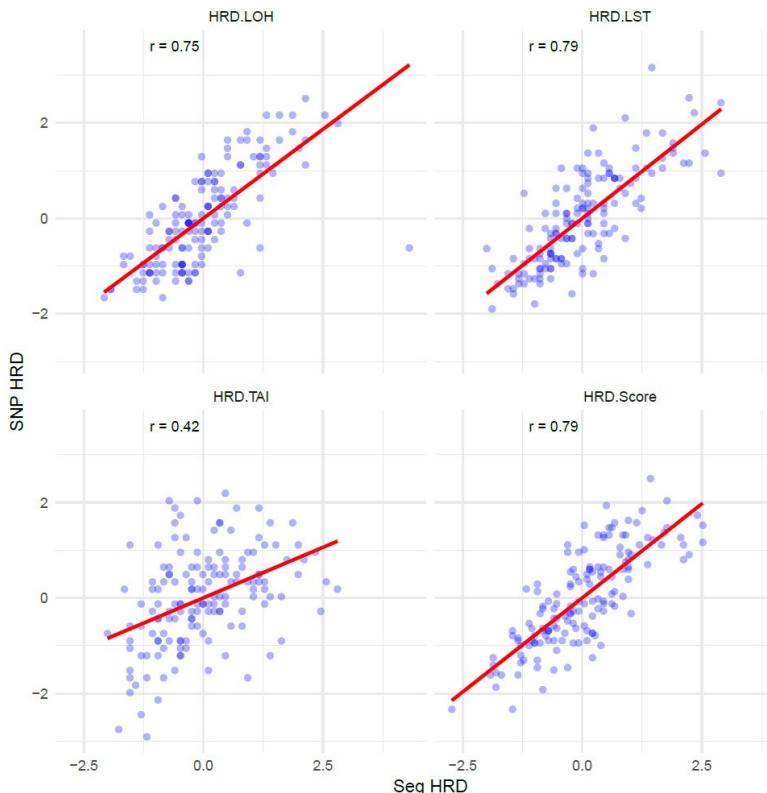
37. van Wijk LM, Nilas AB, Vrieling H, Vreeswijk MPG. RAD51 as a functional biomarker for homologous recombination deficiency in cancer: a promising addition to the HRD toolbox? *Expert review of molecular diagnostics*. 2022;22(2):185-199.
38. Nguyen L, J WMM, Van Hoeck A, Cuppen E. Pan-cancer landscape of homologous recombination deficiency. *Nature communications*. 2020;11(1):5584.
39. Kaklamani VG, Jeruss JS, Hughes E, et al. Phase II neoadjuvant clinical trial of carboplatin and eribulin in women with triple negative early-stage breast cancer (NCT01372579). *Breast Cancer Res Treat*. 2015;151(3):629-638.
40. Stronach EA, Paul J, Timms KM, et al. Biomarker Assessment of HR Deficiency, Tumor BRCA1/2 Mutations, and CCNE1 Copy Number in Ovarian Cancer: Associations with Clinical Outcome Following Platinum Monotherapy. *Mol Cancer Res*. 2018;16(7):1103-1111.
41. Kadouri L, Rottenberg Y, Zick A, et al. Homologous recombination in lung cancer, germline and somatic mutations, clinical and phenotype characterization. *Lung Cancer*. 2019;137:48-51.
42. Sztupinszki Z, Diossy M, Krzystanek M, et al. Migrating the SNP array-based homologous recombination deficiency measures to next generation sequencing data of breast cancer. *NPJ breast cancer*. 2018;4:16.
43. Tutt A, Tovey H, Cheang MCU, et al. Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAneSS subgroups: the TNT Trial. *Nat Med*. 2018;24(5):628-637.
44. Galland L, Ballot E, Mananet H, et al. Efficacy of platinum-based chemotherapy in metastatic breast cancer and HRD biomarkers: utility of exome sequencing. *NPJ breast cancer*. 2022;8(1):28.

45. Castroviejo-Bermejo M, Cruz C, Llop-Guevara A, et al. A RAD51 assay feasible in routine tumor samples calls PARP inhibitor response beyond BRCA mutation. *EMBO Mol Med.* 2018;10(12).
46. Meijer TG, Nguyen L, Van Hoeck A, et al. Functional RECAP (REpair CAPacity) assay identifies homologous recombination deficiency undetected by DNA-based BRCAness tests. *Oncogene.* 2022;41(26):3498-3506.
47. Guffanti F, Alvisi MF, Anastasia A, et al. Basal expression of RAD51 foci predicts olaparib response in patient-derived ovarian cancer xenografts. *Br J Cancer.* 2022;126(1):120-128.
48. Pellegrino B, Herencia-Ropero A, Llop-Guevara A, et al. Preclinical In Vivo Validation of the RAD51 Test for Identification of Homologous Recombination-Deficient Tumors and Patient Stratification. *Cancer Res.* 2022;82(8):1646-1657.
49. Tumiati M, Hietanen S, Hynnen J, et al. A Functional Homologous Recombination Assay Predicts Primary Chemotherapy Response and Long-Term Survival in Ovarian Cancer Patients. *Clin Cancer Res.* 2018;24(18):4482-4493.
50. Stancz P, Hamel N, Sigel KM, Foulkes WD, Karlic R, Polak P. The Great Majority of Homologous Recombination Repair-Deficient Tumors Are Accounted for by Established Causes. *Front Genet.* 2022;13:852159.

a.**d.****b.****e.****c.****f.****Figure 1**

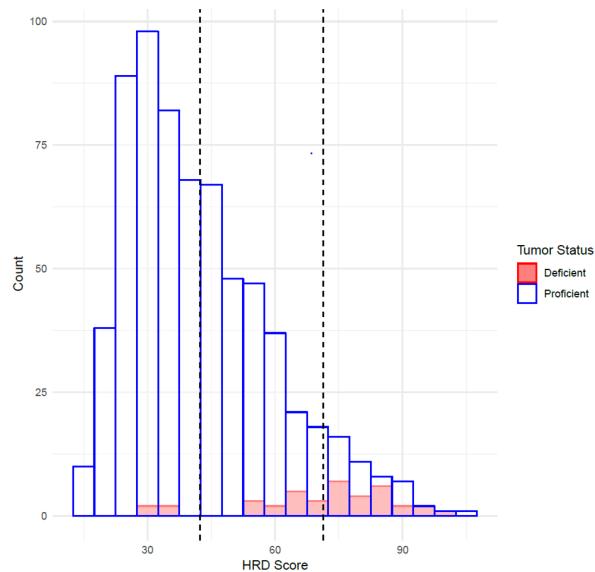
a.**b.**

Supplementary Figure 1: Development and performance of HRDex. (a): Analysis of pan-cancer TCGA data showing relationship of ploidy to raw HRD score as generated by SNP array data in different tumor types. (b) Correlation of HRDex score with ploidy prior to and after normalization in TCGA breast.

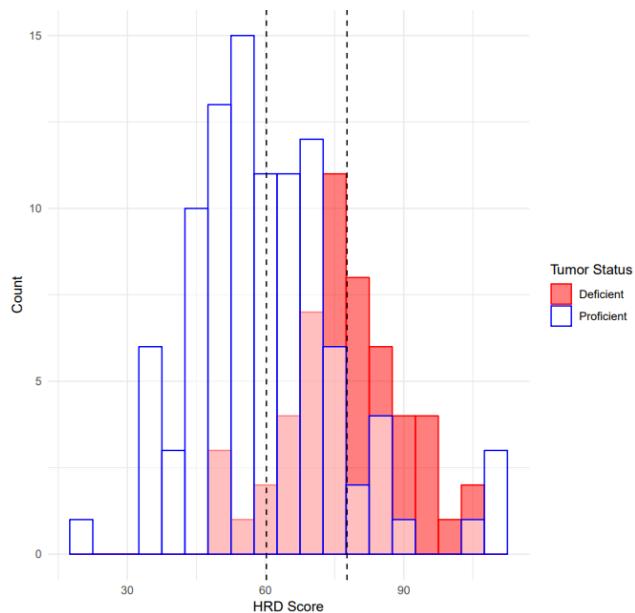
a.**b.**

Supplementary Figure 2: (a) Correlation plot of HRD score, LOH, LST and NTAI between pancancer SNP array data and HRDex derived for breast. (b) Correlation plot of HRD score, LOH, LST and NTAI between pancancer SNP array and HRDex derived for ovary.

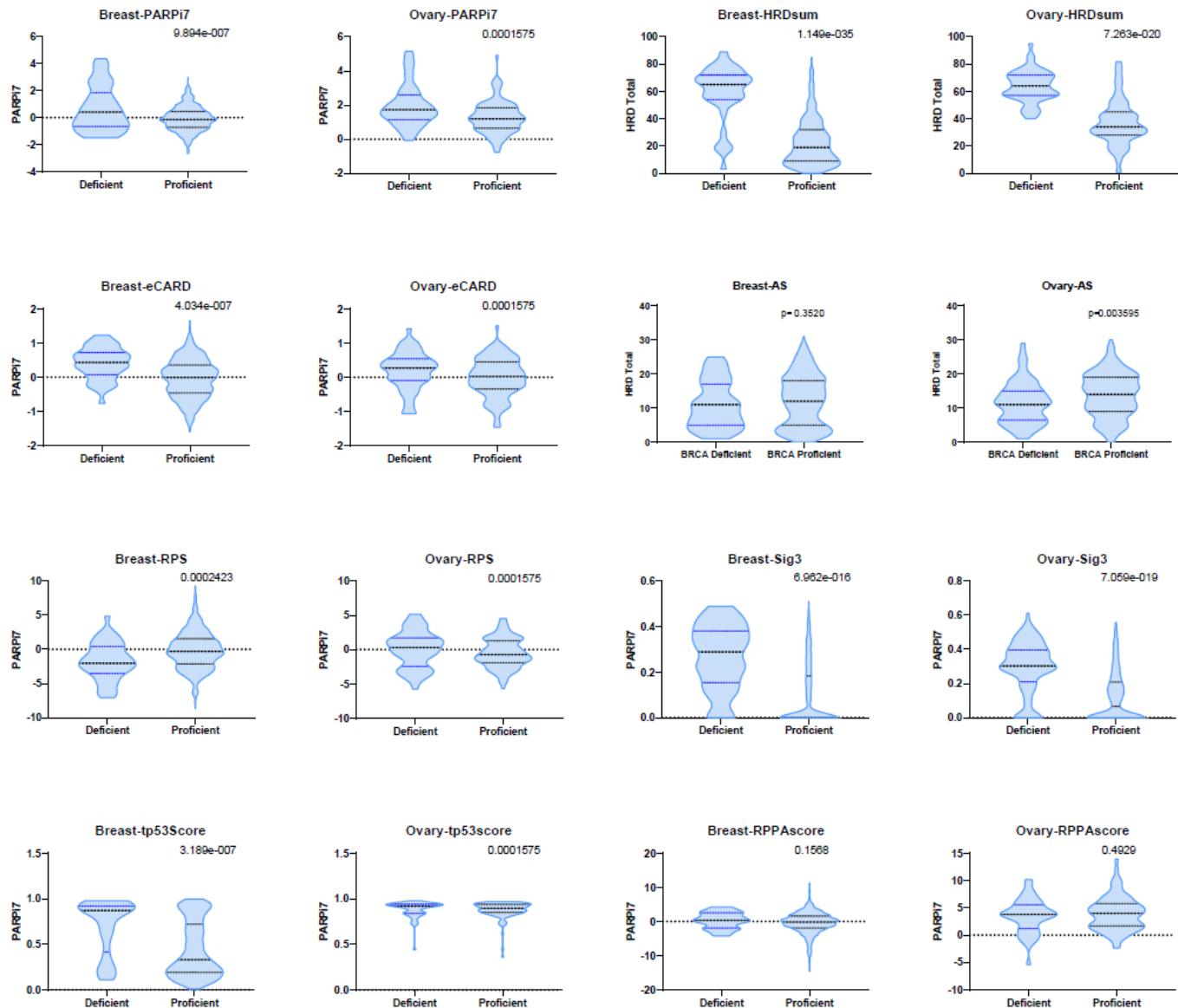
a.



b.



Supplementary Figure 3: (a) Histogram of HRD scores by HRDex in TCGA breast (b) Histogram of HRD scores by HRDex in TCGA ovary



Supplementary Figure 4: Violin plots of DNA Damage Response (DDR) scores from pan-cancer analysis in breast and ovarian tumors classified by *BRCA* status.