Clinical Cancer Research

Homologous Recombination Deficiency and Platinum-Based Therapy Outcomes in Advanced Breast Cancer

Eric Y. Zhao[1], Yaoqing Shen[1], Erin Pleasance[1], Katayoon Kasaian[1], Sreeja Leelakumari[1], Martin Jones[1], Pinaki Bose[1], Carolyn Ch'ng[1], Caralyn Reisle[1], Peter Eirew[2], Richard Corbett[1], Karen L. Mungall[1], Nina Thiessen[1], Yussanne Ma[1], Jacqueline E. Schein[1], Andrew J. Mungall[1], Yongjun Zhao[1], Richard A. Moore[1], Wendie Den Brok[3], Sheridan Wilson[3], Diego Villa[3], Tamara Shenkier[3], Caroline Lohrisch[3], Stephen Chia[3], Stephen Yip[4], Karen Gelmon[3], Howard Lim[3], Daniel Renouf[3], Sophie Sun[3], Kasmintan A. Schrader[2,5], Sean Young[1,4], Ian Bosdet[1,4], Aly Karsan[1,4], Janessa Laskin[3], Marco A. Marra[1,5], and Steven J.M. Jones[1,5]

Abstract

Purpose: Recent studies have identified mutation signatures of homologous recombination deficiency (HRD) in over 20% of breast cancers, as well as pancreatic, ovarian, and gastric cancers. There is an urgent need to understand the clinical implications of HRD signatures. Whereas *BRCA1/2* mutations confer sensitivity to platinum-based chemotherapies, it is not yet clear whether mutation signatures can independently predict platinum response.

Experimental Design: In this observational study, we sequenced tumor whole genomes (100× depth) and matched normals (60×) of 93 advanced-stage breast cancers (33 platinum-treated). We computed a published metric called HRDetect, independently trained to predict *BRCA1/2* status, and assessed its capacity to predict outcomes on platinum-based chemotherapies. Clinical endpoints were overall survival (OS), total duration on platinum-based therapy (TDT), and radiographic evidence of clinical improvement (CI).

Results: HRDetect predicted *BRCA1/2* status with an area under the curve (AUC) of 0.94 and optimal threshold of 0.7. Elevated HRDetect was also significantly associated with CI on platinum-based therapy (AUC = 0.89; $P = 0.006$) with the same optimal threshold, even after adjusting for *BRCA1/2* mutation status and treatment timing. HRDetect scores over 0.7 were associated with a 3-month extended median TDT ($P = 0.0003$) and 1.3-year extended median OS ($P = 0.04$).

Conclusions: Our findings not only independently validate HRDetect, but also provide the first evidence of its association with platinum response in advanced breast cancer. We demonstrate that HRD mutation signatures may offer clinically relevant information independently of *BRCA1/2* mutation status and hope this work will guide the development of clinical trials. *Clin Cancer Res; 23(24); 7521–30. ©2017 AACR.*

Introduction

Genomic instability and mutagenesis are hallmarks of human cancers that can arise from deficient DNA repair processes. One such process, homologous recombination (HR) involves strand invasion by homologous sequences to facilitate error-free repair of double strand breaks and interstrand crosslinks (1). Mutations in genes responsible for HR are prevalent among human cancers. The *BRCA1* and *BRCA2* genes are centrally involved in HR, DNA-damage repair, end resection, and checkpoint signaling (2). Inherited mutations in *BRCA1* and *BRCA2* account for 5% to 10% of all breast cancers, conferring an up to 85% lifetime risk (3, 4). There is emerging evidence suggesting that germline *BRCA1*- and *BRCA2*-mutated cancers are associated with sensitivity to platinum-based chemotherapy and PARP inhibitors (5–9). This is further supported by resistance to platinum-based agents arising from secondary mutations that cause somatic reversion of germline *BRCA1/2* variants (10).

HR deficiency (HRD) is complex, and its myriad causes are not fully understood. However, examining characteristic patterns of mutation, collectively known as mutation signatures or genomic scars, can provide an aggregate functional metric of pathway function. For example, *BRCA1* and *BRCA2* are associated with characteristic copy-number variant (CNV) patterns (11), which have been suggested to independently predict platinum sensitivity in primary breast cancer (12). However, a clinical trial in advanced stage triple-negative breast cancer did not verify this association (5). Meanwhile, new genomic correlates have refined

[1]Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada. [2]Department of Molecular Oncology, British Columbia Cancer Agency, Vancouver, British Columbia, Canada. [3]Department of Medical Oncology, British Columbia Cancer Agency, Vancouver, British Columbia, Canada. [4]Department of Pathology and Laboratory Medicine, The University of British Columbia, Vancouver, British Columbia, Canada. [5]Department of Medical Genetics, The University of British Columbia, Vancouver, British Columbia, Canada.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Corresponding Author: Steven J.M. Jones, British Columbia Cancer Agency, Suite 100, 570 West 7th Avenue, Vancouver, BC V5Z 4S6, Canada. Phone: 604-877-6083; Fax: 604-876-3561; E-mail: sjones@bcgsc.ca

doi: 10.1158/1078-0432.CCR-17-1941

Translational Relevance

Cancer cells that lack DNA repair abilities are more sensitive to a common class of drugs called platinum-based chemotherapy. Successfully identifying patients whose cancers lack DNA repair function could substantially improve the use of such therapies, but is an ongoing challenge. Recent research has revealed that sequencing the whole genome of a tumor can reveal the mutational "scars" left behind by deficient DNA repair. In this article, we analyze the whole genomes of 93 breast cancers to identify these mutational scars. We then demonstrate that this approach can, with good accuracy, identify patients who respond well to platinum-based chemotherapies. As whole genome sequencing becomes more affordable, this presents an exciting opportunity to improve breast cancer treatment.

the detection of HRD. Large-scale genome profiling across thousands of cancers has revealed characteristic patterns of mutation giving rise to millions of somatic single-nucleotide variants (SNV; ref. 13) and structural variants (SV; ref. 14). Recent efforts aggregated six HRD-associated signatures into a single score called HRDetect to accurately classify breast cancers by their *BRCA1* and *BRCA2* status (15).

With this improved capability to quantify "BRCA-ness," there is substantial interest in its therapeutic implications in breast cancer (13, 15–18). Importantly, these measures may be able to identify *BRCA1*- and *BRCA2*-intact but HR-deficient tumors to guide eligibility for HRD-targeted clinical trials and treatment decision-making. However, there is not yet direct evidence that aggregated genomic scar metrics predict platinum sensitivity. In this observational biomarker study, we perform whole-genome sequencing (WGS) to identify HRD mutation signatures in a cohort of 93 patients with advanced-stage breast cancers and associate them with molecular, pathologic, and clinical features. Using HRDetect, we aggregate HRD signatures and demonstrate their association with clinical benefit on platinum-based chemotherapy.

Materials and Methods

Patient samples, ethics, and data policy

A total of 93 study participants with advanced-stage breast cancer underwent tumor biopsies at the BC Cancer Agency (BCCA) and collaborating hospitals as part of the Personalized oncogenomics (POG) Project, the first 100 cases of which were described in an earlier publication (19). The study was approved by the University of British Columbia Research Ethics Board (REB# H12-00137 and H14-00681-A019). Written informed consent, including potential publication of findings, was obtained from patients prior to genomic profiling. Patient information was anonymized, and each was assigned an alphanumeric identification code. This study includes data from the first 93 verified breast cancer cases that underwent whole genome characterization and met quality assurance standards. Whole-genome sequencing and RNA-seq data (.bam files) have been submitted to the European Genome-Phenome Archive (EGA; www.ebi.ac.uk/ega/home) under the study accession number EGAS00001001159. Dataset accession IDs are provided in Supplementary Table S5.

Sample collection, preparation, and sequencing

Biopsy samples were embedded in optimal cutting temperature (OCT) compound and sectioned. Pathology review was completed for each specimen, including assessment of tumor content. Genome libraries from tumor and peripheral blood (normal control) as well as transcriptome libraries from tumor were constructed using Illumina protocols. Whole genome and transcriptome sequencing was performed on an Illumina HiSeq2000 or HiSeq2500 sequencer. The details of library construction and sequencing have been previously described (20, 21).

Bioinformatic analysis

Sequencing reads were aligned to the human reference genome (GSCh37, available from http://www.bcgsc.ca/downloads/genomes/9606/hg19/1000genomes/bwa_ind/genome) by the BWA aligner (v0.5.7; refs. 22, 23). Somatic SNVs and small insertions/deletions were processed using samtools (24) and Strelka (v0.4.6.2; ref. 25). CNVs were called using CNASeq (v0.0.6) as described in ref. 26 and LOH by APOLLOH (v0.1.1; ref. 27). The matched normal genome was used to subtract germline variants and to report cancer risk variants in 98 select actionable genes, preapproved by an ethics committee. Germline variant pathogenicity was estimated according to established ACMG guidelines (28) using a local curated variant database and custom-built risk calculator established by the BCCA Cancer Genetics Laboratory. Transcriptomes were repositioned using JAGuaR (version 2.0.3; ref. 29). Differential expression analysis was performed by comparing RPKM expression levels against a compendium of 16 normal tissues from the Illumina BodyMap 2.0 project (available from ArrayExpress, query ID: E-MTAB-513) as described in ref. 26. Intrinsic subtypes were determined by performing Spearman rank-order correlations on the expression of genes in the PAM50 gene set (30) for each breast cancer subtype between sequenced samples and 823 breast cancers derived from The Cancer Genome Atlas (31). For each sample, the subtype with the greatest correlation coefficient was taken as the intrinsic subtype (Supplementary Fig. S1). One tumor sample did not pass quality control for RNA-seq and was excluded from analyses involving intrinsic subtypes.

Determining HRDetect scores

HRDetect scores were computed by aggregating six mutation signatures associated with HRD: (i) SNV signature 3/V9, (ii) SNV signature 8/V6, (iii) SV signature 3/R1, (iv) SV signature 5/R5, (v) the HRD index, and (vi) the fraction of deletions with microhomology. All signatures were normalized and log transformed as previously described (15), and HRDetect scores were computed using a logistic model with the same intercept and coefficients as those reported in the previously trained model, without any retraining or adjustment (15). The intercept was −3.364 and the coefficients were 1.611, 0.091, 1.153, 0.847, 0.667, and 2.398, respectively, for the six HRD signatures. The sections that follow detail the computation of the six component signatures. All non-public computer code used to compute HRDetect or its component signatures will be made available upon request.

Single nucleotide variant mutation signatures

Somatic SNVs called by Strelka were used for mutation signature calculation. SNVs were categorized based on 6 variant types and 16 trinucleotide context subtypes to yield a total of 96 mutation classes. Mutation signatures were deciphered using a

published framework (32), which uses nonnegative matrix factorization (NMF) to infer both the operative signatures prevalent across the 93-genome cohort and the relative exposure of each signature to each genome. Exposures are modeled as the number of mutations contributed by a mutation signature. Fractional exposure was defined as the proportion of a genome's total mutation burden contributed by a particular signature. Signature stability estimates were obtained by bootstrap resampling with 1,008 iterations (84 iterations over 12 cores). Solutions with a 7- to 10-signature model were found to best maximize signature stability and minimize Frobenius reconstruction error (Supplementary Fig. S3A). Among these, a 9-signature model was selected as it yielded one signature with maximal cosine similarity to the previously described HRD-associated Signature 3. The similarity of signatures to 30 previously described mutational signatures (available from http://cancer.sanger.ac.uk/cosmic/signatures) was quantified using the cosine similarity metric (Supplementary Fig. S3C).

Structural variant mutation signatures

Large-scale somatic SVs were reconstructed by *de novo* assembly of tumor and normal reads using ABySS and Trans-ABySS (33). Candidate SVs were realigned to the reference genome to resolve breakpoints. Additionally, we used DELLY (v0.6.1) to obtain an independent SV set by reference-based analysis of split and paired end reads (34). Germline events were filtered out by subtracting SVs found in the matched normal genome. SVs detected by the two methods were merged to yield a high-quality consensus set, containing an intersection of variants called by both methods with matching pairs of breakpoint loci each separated by no more than 20 base pairs.

The 32-parameter SV mutation catalog vectors were computed by binning variants based on breakpoint clustering, SV type, and SV length (14), yielding a 32 by 93 catalogue matrix. This matrix was decomposed by NMF (like with SNV signatures) using a 6-signature model, which was chosen to maximize signature stability and minimize Frobenius reconstruction error (Supplementary Fig. S3B). Pairwise comparisons of newly deciphered mutation signatures to six previously described signatures were performed by cosine similarity metric (Supplementary Fig. S3D).

Calculation of the HRD index

For each cancer genome, the HRD index was computed as the arithmetic sum of LOH, TAI, and LST scores. CNV and LOH analysis pipelines yielded coordinates segmenting whole genomes by allele-specific copy-number ratios. We created an R package called HRDtools that computes LOH, TAI, and LST scores based on the genome-wide CNV profile. Because the HRD index relies upon large-scale events, HRDtools first filters out small events occurring within contiguous events at least 100 times larger. The three scores are then determined based on published guidelines (11)

Analysis of deletion microhomology

Somatic deletions were detected based on sequence alignment using Strelka. The 3′ and 5′ flanking deletions were obtained. The microhomology fraction was determined as the proportion of deletions which were larger than three base pairs and demonstrated overlapping microhomology at the breakpoints.

Review of clinical case data

Retrospective chart review was performed to obtain treatment history and clinical response to chemotherapy regimens. We queried a province-wide registry of oncology therapeutic records (35) to obtain dates of (1) birth, (2) death if applicable, (3) most recent cancer diagnosis, and (4) start and end dates of all platinum-based chemotherapy regimens administered to treat the most recent cancer diagnosis along with therapies used in combination. Treatment timelines and clinical response are presented in Supplementary Fig S2. All patients were treated as part of standard cancer care either prior to, during, or after the sequencing biopsy. Platinum-treated patients were given standard doses of cisplatin (30 mg/m^2 on days 1 and 8 of a 21-day cycle) or carboplatin (calculated in milligrams as glomerular filtration rate + 25, multiplied by 6 for monotherapy or 5 in combination regimens).

To assess therapeutic benefit, three outcomes were chosen: overall survival (OS), total duration on platinum-based therapy (TDT), and clinical response based on imaging. Overall survival was assessed in patients treated after sequencing ($n = 19$) and was computed as the duration from first post-biopsy dose of platinum-based chemotherapy to death. TDT was examined as a surrogate for therapy effectiveness. To improve relevance to the present diagnosis, TDT included only treatment regimens occurring within 2 years of sequencing biopsy ($n = 33$; Supplementary Fig. S2).

Clinical imaging reports were reviewed to evaluate platinum response including fludeoxyglucose positron emission tomography and computed tomography obtained during or within two months after the period of platinum-based therapy, compared with pretreatment comparison scans. Treatment response was classified as follows: (1) clinical improvement (CI), any tumor shrinkage of one or more lesions with no evidence of growth or new lesions; (2) stable disease (SD), either no change in lesions or decreased size of some lesions with growth of others; or (3) progressive disease (PD), disease progression with no associated tumor shrinkage. The best observed response per regimen was recorded.

Results

Somatic mutation signatures

Using a published framework (32), we deciphered the mutation signatures of 1,182,840 somatic SNVs and 11,393 SVs from the whole genomes of 93 advanced-stage breast cancers.

Of the nine resulting SNV signatures, numbered V1–V9 (Fig. 1A), six closely match previously described mutation signatures (Supplementary Fig. S3C). V9 (signature 3) and V6 (signature 8) are associated with HRD (13–15). V4 (signature 1) is associated with aging (36). V1 (signature 2) and V2 (signature 13) are associated with APOBEC deaminase activity. V3 (signature 17) has been observed across many cancers, but its etiology is unclear.

The three remaining signatures, V5, V7, and V8, represent novel breast cancer mutational signatures. V5 predominantly displays C>T mutations in CpCpT and CpCpC contexts and was present in only three cancers. V7 is characterized by high pyrimidine transition rate with enrichment in NpCpG and NpTpG contexts and was observed across many tumors spanning histological and molecular subtypes. V8 demonstrated moderate enrichment of all base substitution types when flanked by T and A bases, and was present at low levels across many tumors. These signatures may

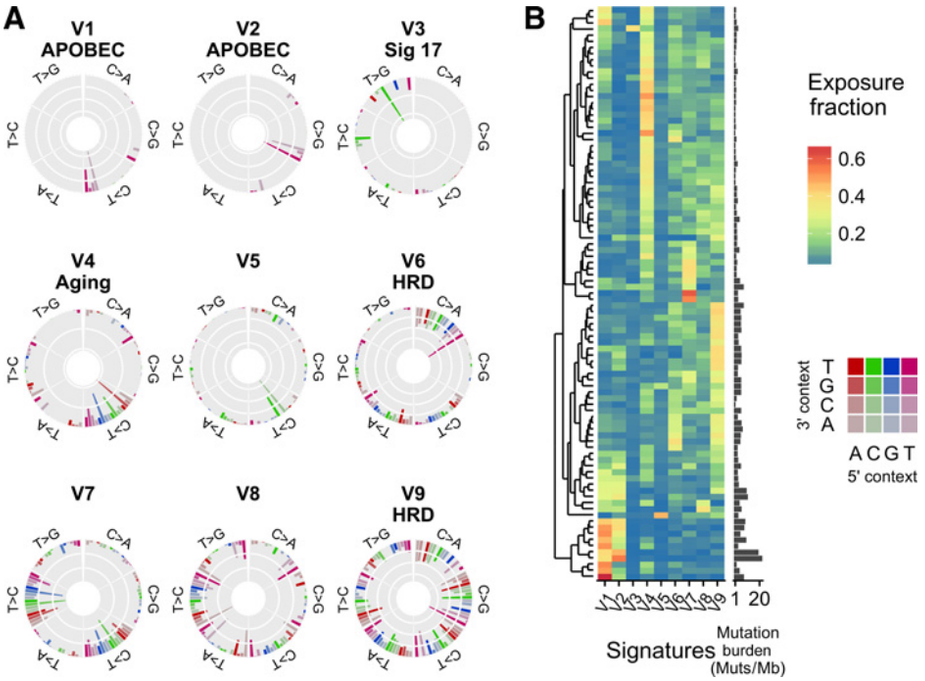

Figure 1.
Nine signatures of single nucleotide variation deciphered from 93 breast cancer whole genomes. **A,** Signatures are visualized according to relative frequencies of mutations grouped by base change and 3′/5′ context. Six of nine signatures match previously published mutation signatures (cosine similarity > 0.9), five of which are associated with hypothetical etiologies. **B,** Fractional exposures and mutation burdens across the patient cohort, ordered by hierarchical clustering, reveal groups defined by APOBEC, aging, and HRD-associated signatures.

reflect the advanced, recurrent, and drug-treated nature of our cohort, whereas previous mutation signatures have been derived from primary untreated cancers. Further study is necessary to verify etiology.

Hierarchical clustering revealed that most cases of high SNV burden were driven by APOBEC or HRD associated processes (Fig. 1B), which together were dominant in 46 (49%) of the 93 sequenced breast cancers. The aging mutation signature was ubiquitous across cancers, and was the dominant signature in 31 (33%) cases, all of which had low mutation burden (<5 SNVs per Mb; Supplementary Fig. S5).

Ninety-two samples with available RNA-seq data were classified into intrinsic subtypes based on expression profiles of PAM50 (30) genes. Nonparametric analysis demonstrated significant differences in signatures V2, V3, V8, and V9 across subtypes (Supplementary Table S1). *Post hoc* pairwise Dunn tests revealed elevated V3, V8, and V9 within basal-like cancers (Supplementary Fig. S1), suggesting that diverse mutagenic etiologies, including HRD, underlie this subtype. Elevated signature V9 was also most common among triple-negative tumors.

The six deciphered SV signatures, numbered R1–R6, closely resembled the six previously described breast cancer signatures (ref. 14; Supplementary Figs. S6 and S7). R1–R4 and R6 uniquely matched previously described signatures. By visual inspection, R5 matches previously described rearrangement signature 5 albeit with more nonclustered translocations.

Genomic findings associated with HRD

Alongside these four SNV and SV mutation signatures, we measured two additional HRD-associated patterns of somatic mutation. The HRD index measures the frequency of large-scale loss-of-heterozygosity (LOH), telomeric allelic imbalance (TAI), and large-scale transition (LST) events (11) and was computed using allelic copy-number ratios inferred from read alignment frequencies. The proportion of small deletions associated with microhomology was determined by comparing sequences flanking deletion breakpoints. As per a published method (15), all six scores were log transformed, normalized, and combined into a single HRDetect predictor. This was performed using a logistic predictor with the same coefficients as those reported by Davies and colleagues (15) to ensure consistency with the trained model.

Nineteen breast cancers had high HRDetect scores (>0.7), 37 had moderate scores (0.005–0.7), and 37 had low scores (< 0.005). All cancers underwent genome-wide characterization of germline and somatic point mutations, insertions and deletions, and copy loss in gene regions and splice sites. Across the 93 breast cancers, HRDetect predicted pathogenic germline and somatic variants in *BRCA1* and *BRCA2* with high accuracy and an optimal differentiating threshold of 0.74 (Fig. 3B). These findings closely agree with the previously established threshold of 0.70 (15). Because VUS have previously not been associated with increased HRDetect (15), we classified VUS as nonpathogenic mutations for the purposes of this analysis. Elevated HRDetect scores were observed in all tumors with observed *BRCA1*/*BRCA2* frame shifts, nonsense mutations, homozygous deletions, or splice variants identified as likely pathogenic in ClinVar (Fig. 2). There were 11 cases with germline missense VUS. The most common of these was *BRCA2* T1915M, which had a global minor allele frequency (GMAF) of 1.14% and has been reported both to reduce (37) and contribute to (38) breast cancer risk. In our study, seven breast cancers (BR004, BR027, BR032, BR036, BR064, BR074, and BR086) harbored germline *BRCA2* T1915M, of which three (BR004, BR036, and BR086) were homozygous in the tumor and displayed a wide range of HRDetect scores (0, 0.04, and 0.62, respectively). However, BR086 exhibited coincident homozygous deletion of *RAD51*, which may account for the elevated score.

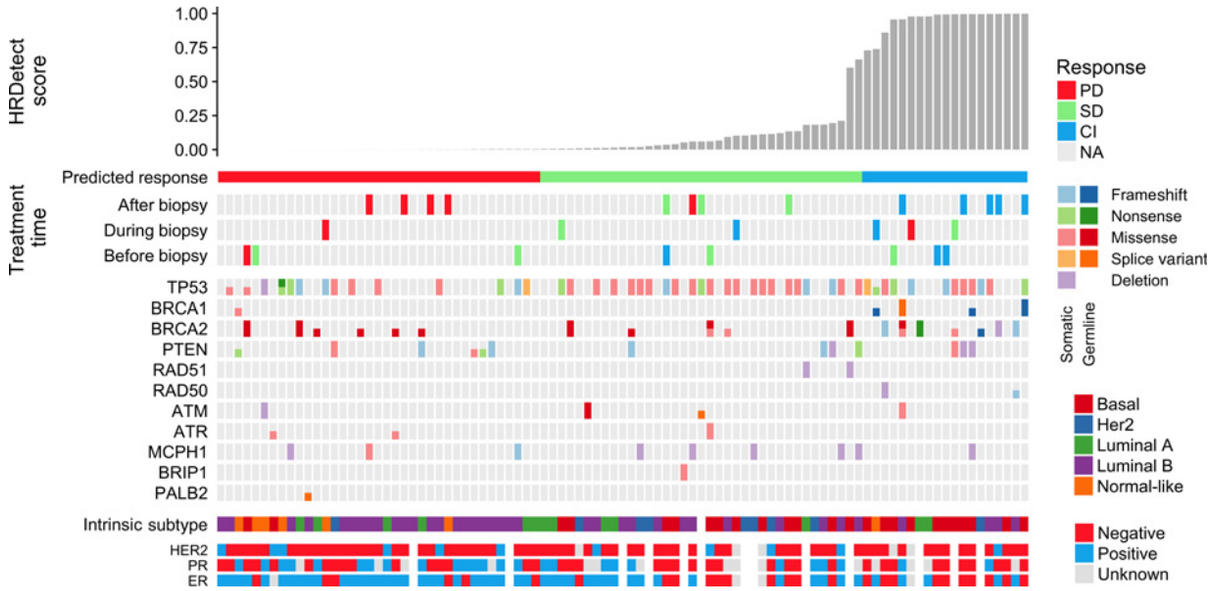

Figure 2.

HRDetect scores, mutations in key homologous recombination genes, and outcomes on platinum-based therapy. Six distinct mutation signatures associated with homologous recombination deficiency (HRD) were deciphered from 93 breast cancer whole genomes and aggregated into a single HRDetect score. Radiology reports during and after treatment regimens involving platinum-based chemotherapy were reviewed for evidence of clinical improvement (CI), stable disease (SD), or progressive disease (PD). Analysis of receiver-operator characteristic curves suggested HRDetect thresholds of 0.7 for CI and 0.005 for SD, indicated here by a color bar.

These data therefore do not suggest pathogenicity of *BRCA2* T1915M.

A number of other genes involved in HR demonstrated tentative associations with HRDetect scores. Elevated HRDetect was observed in three cases with homozygous deletion of *PTEN* as well as one case with two coincident *PTEN* missense mutations (F278L and P38S). However, one case with homozygous *PTEN* A126D somatic mutation was associated with a low HRDetect score. Homozygous deletions in *RAD50*, *RAD51*, and *MCPH1* were observed in some tumors with moderate or high HRDetect scores. *MCPH1* is a potential cancer susceptibility gene (39) whose deletion may be a poor prognostic marker (40). Although recurrently deleted in our cohort, its link to HRD signatures was inconsistent.

High HRDetect scores were also associated with triple negative and basal-like breast cancers (Table 1). Of 19 samples with high HRDetect, 11 (58%) were classified as basal-like. Among low HRDetect samples, only 2 (5%) were basal like. Luminal B and normal-like tumors were more likely to have low HRDetect scores, whereas most (7/9) HER2-like tumors displayed moderate HRDetect. Receptor status was assessed by immunohistochemistry and retrieved from pathology records, which were available for 79 tumors at primary and 76 at relapse (Supplementary Fig. S1). High HRDetect was inversely associated with positive receptor status in all three receptors. Fifty percent of high HRDetect tumors were triple negative, compared with only 6% of primary and 15% of metastatic low HRDetect tumors.

HRD mutation signatures are associated with platinum outcomes

High HRDetect scores were significantly associated with clinical improvement on platinum-based chemotherapy, even after adjusting for *BRCA1*/*BRCA2* status and treatment timing ($P = 0.006$, $n = 26$; Supplementary Table S4). HRDetect demonstrated areas under the ROC curve of 0.89 for CI and 0.86 for SD, which exceeded those of its component signatures (Fig. 3B and C; Supplementary Table S3). Optimal thresholds of 0.005 for predicting stable disease (SD) and 0.7 for predicting clinical improvement (CI) were chosen (Fig. 3B and C). Sensitivity, specificity, precision, and recall were computed for both thresholds and are reported in Supplementary Table S2.

Biallelic loss of *BRCA1* or *BRCA2* was also associated with clinical improvement on platinum-based chemotherapy (Fig. 3A) but was observed in only 3 of 26 treated patients with available imaging. By comparison, 11 patients demonstrated HRDetect scores above 0.7, of whom 8 experienced CI, 2 experienced SD, and 1 had disease progression. Therefore, HRDetect scores correctly identified 5 additional patients without biallelic loss of *BRCA1* or *BRCA2* who benefited from platinum-based therapy. In a joint logistic model, *BRCA1* and *BRCA2* status did not contribute significantly to the predictive value of HRDetect (Supplementary Table S4).

Effects of HRDetect on overall survival and treatment duration

Of patients treated post-biopsy with platinum-based chemotherapy, there was a statistically significant difference in OS depending upon HRDetect ($P = 0.04$, $n = 33$). Five patients with predicted CI (HRDetect > 0.7) demonstrated a median survival of 384 days, 8 with predicted SD (0.7 > HRDetect > 0.005) had a median survival of 160 days, and 6 patients with predicted PD (HRDetect < 0.005) had a median survival of 122 days. This difference should be interpreted with caution due to small sample size, but represents a promising trend which warrants further study.

Table 1. Summary of patient molecular and clinical characteristics by HRDetect status

	HRDetect status			
	Low < 0.005	**Moderate 0.005–0.7**	**High > 0.7**	**Total**
Sample counts				
Total count	37	37	19	93
Treated count	9	13	11	33
Treated and imaged	8	7	11	26
Pathogenic BRCA1/2 variant	0	0	7	7
Response to platinum-based therapy				
CI	0	2	8	10
SD	2	4	2	8
PD	6	1	1	8
Median TDT	56 days ($n = 9$)	71 days ($n = 13$)	143 days ($n = 11$)	
Median OS	122 days ($n = 6$)	160 days ($n = 8$)	384 days ($n = 5$)	
Intrinsic subtype				
Basal	2	12	11	24
HER2 amplified	1	7	1	9
Luminal A	6	5	2	13
Luminal B	22	12	4	38
Normal-like	6	0	1	7
Primary receptor status				
ER (positive/negative)	31/3	20/9	7/8	58/20
PR (positive/negative)	18/4	11/10	4/12	33/26
HER2 (positive/negative)	4/23	4/22	0/14	8/59
Triple negative	2 (6%)	8 (28%)	8 (50%)	18
Metastatic receptor status				
ER (positive/negative)	27/6	17/10	5/10	49/26
PR (positive/negative)	15/15	9/13	2/10	26/38
HER2 (positive/negative)	6/28	4/22	1/13	11/63
Triple negative	5 (15%)	6 (21%)	8 (50%)	19

In addition to OS, TDT was used as a surrogate for clinical response. In practice, platinum-based chemotherapy is typically continued in responding patients until disease progression or significant toxicity. Supplementary Fig. S2 verifies that, in 26 patients with available imaging, patients with reported radiographic response were more likely to undergo a longer duration of

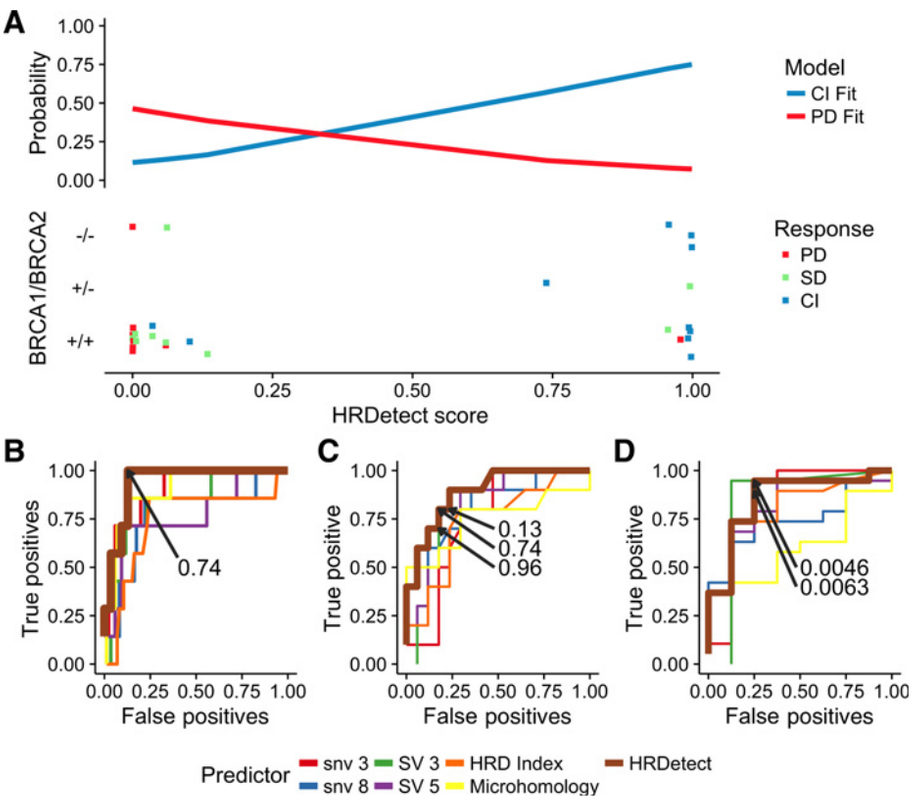

Figure 3.
Association of platinum-based treatment outcomes with HRDetect, an aggregate of six homologous recombination deficiency (HRD) mutation signatures. **A,** The HRDetect score is significantly associated with clinical improvement (CI) on platinum-based chemotherapy (logistic regression, adjusted for *BRCA1/2* status and treatment timing, $P = 0.006$). There was also a trend between low HRDetect and progressive disease (PD; $P = 0.112$). Moreover, of 8 *BRCA1/2*-intact cases with elevated HRDetect score, 5 responded favorably to platinum-based chemotherapy. Receiver-operator characteristic for **B,** BRCA status and (**C, D**) therapeutic outcomes on platinum-based chemotherapy (C: CI; D: stable disease, SD). These suggest optimal HRDetect thresholds of 0.7 and 0.005 for CI and SD, respectively. Specific near-threshold HRDetect values are labeled. In all three ROC curves, HRDetect had a superior area under the curve than its six constituent mutation signatures.

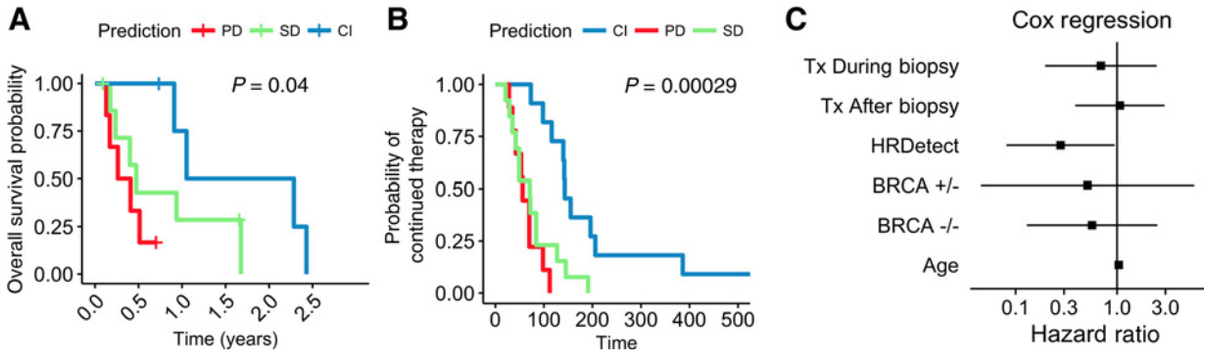

Figure 4.

Homologous recombination deficiency is associated with extended overall survival (OS) and total duration on platinum-based therapy (TDT). **A,** Among patients treated after the sequencing biopsy ($n = 19$), OS was computed as the duration between first post-biopsy treatment and death. There was a statistically significant ($P = 0.04$) difference between patients predicted to be CI (HRDetect > 0.7), SD (0.7 > HRDetect > 0.005), and PD (HRDetect < 0.005). **B,** Platinum-treated patients ($n = 33$) with different predicted treatment outcomes also experienced significantly different TDT as part of standard care for advanced breast cancer. **C,** Multivariate Cox survival model demonstrated a significant association between HRDetect and TDT independently of *BRCA1/2* mutation status.

treatment. HRDetect scores were significantly associated with extended TDT with a hazard ratio of 0.28 (0.081–0.95; $P = 0.04$, $n = 33$), after adjusting for *BRCA1* and *BRCA2* mutation status, timing of treatment, and patient age (Fig. 4B). Tumors were classified based on HRDetect scores into predicted treatment response categories. There was a significant difference in TDT ($P < 0.001$, $n = 33$; Fig. 4A) between patients with predicted CI (median 143 days), SD (median 71 days), and PD (median 56 days). This amounts to an estimated 3-month difference in median TDT between high HRD and low HRD cases.

Feasibility of HRD analysis in personalized medicine

The development of precision oncology initiatives (19, 41–43) has necessitated genome analysis pipelines compatible with "N of 1" cases. One challenge of mutation signature analysis by nonnegative matrix factorization (NMF) is the reliance upon large cohorts of sequenced tumors. This has led to techniques to determine the most likely composition of signatures for a single isolated sample (44). HRD analysis provides a promising target for personalized treatment decision-making. Thus, in addition to cohort-based *de novo* signature discovery, we also computed individual-tumor best-fit signature exposure profiles for SNV signatures 3 (V9) and 8 (V6) and SV signatures 3 (R1) and 5 (R5) using nonnegative least squares (NNLS; details in Materials and Methods). We then recomputed HRDetect scores using these individualized NNLS signature exposures to assess accuracy.

HRDetect scores and all four HRD-associated SNV and SV signatures demonstrated high concordance between NMF and NNLS approaches based on Pearson linear regression (r > 0.9; Fig. 5). Using the selected thresholds of 0.005 for SD and 0.7 for CI, 86 out of 93 cancers were concordantly classified by NNLS and NMF, including all cases predicted to experience CI. NNLS reclassified 4 cancers from PD to SD, and 3 from SD to PD.

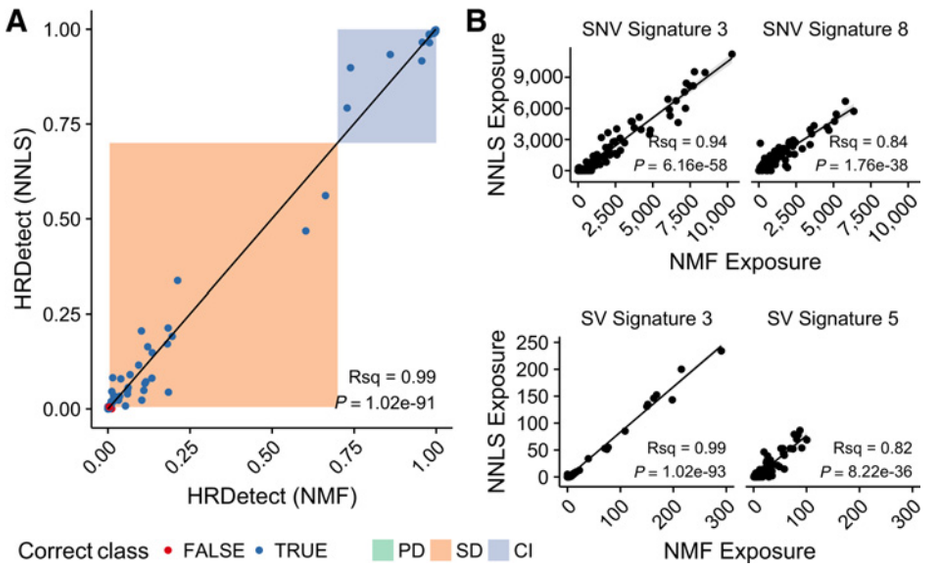

Figure 5.

Mutation signatures of single-nucleotide variation and structural variation were deciphered using two approaches. The first, nonnegative matrix factorization (NMF), deciphers signatures and associated exposures *de novo* across a cancer cohort. The second, nonnegative least squares (NNLS), estimates the best-fit exposure vector using a consensus set of known mutation signatures. **A,** HRDetect scores were computed using both NMF- and NNLS-derived mutation signatures. Scores derived from the two approaches were strongly correlated (Pearson R squared = 0.99) and demonstrated high classification concordance based on selected thresholds. **B,** Individual HRD-associated mutation signatures were concordant between the two approaches (Pearson R squared > 0.82 for all signatures).

These findings demonstrate that NNLS-based "N of 1" computation of mutation signature exposures provides robust HRD estimates concordant with a cohort-based NMF approach. This is promising for the application of HRD biomarkers in sequencing-driven treatment guidance.

Discussion

In this retrospective study, HRD mutation signatures were associated with clinical benefit on platinum-based chemotherapy in advanced-stage breast cancer. Specifically, we demonstrated that HRDetect, the same model independently trained to predict *BRCA1* and *BRCA2* status with high sensitivity and specificity (15), was also significantly associated with favorable response to platinum chemotherapy response and longer TDT. Moreover, we identified an optimal HRDetect threshold of 0.7, which agrees with the previously established cutoff for *BRCA1*/*BRCA2* status (15). Therefore, our findings both independently validate the HRDetect model and provide promising evidence for its clinical relevance.

A key limitation of this study is the ability to establish causation. As this was an observational cohort of advanced-stage breast cancers undergoing standard chemotherapy treatments, some patients were sequenced during or after courses of platinum-based chemotherapy. To mitigate the impacts of tumor evolution, we limited analyses to patients sequenced within two years of treatment. Another significant challenge when studying treated tumors is that platinum-associated mutagenesis may affect the mutation signature profile, especially in cancers biopsied after treatment. A few factors help to mitigate this challenge, but cannot entirely rule out platinum-induced mutagenesis. First, we adjusted for the treatment timing in statistical analyses of the association between HRDetect and clinical outcomes. Second, there has been reproducible evidence of HRD-associated signatures in cohorts of predominantly primary tumors which are a close match to the signatures we deciphered (11, 13–15, 45). Lastly, the aggregation of six distinct signatures into a more robust metric should help minimize any platinum-induced mutagenesis impacting any one signature in particular. Notably, the investigation of advanced-stage breast cancers is an important feature of this study. Whereas a previous trial did not find that the HRD index alone was predictive in advanced breast cancer (5), our findings renew promise for aggregated metrics such as HRDetect. However, studying advanced-stage tumors inevitably introduces potential confounders such as variable treatment histories. Therefore, well-designed prospective clinical trials are needed to further validate HRDetect as a predictive biomarker.

HRD is common among breast cancers. Based on our HRDetect predictive thresholds, 19 cases (20%) showed potentially targetable high HRD status (HRDetect > 0.70). An additional 37 cancers (40%) showed moderate HRD status consistent with stable disease on platinum-based chemotherapy (HRDetect > 0.005). By comparison, biallelic germline and somatic mutations were detected in only 11 cases, and known pathogenic variants in only 7. Similarly, an analysis of 560 breast cancer genomes, which additionally examined promoter hypermethylation, estimated the frequency of BRCA-null breast cancers at 14% (14). The analysis of HRD signatures may identify patients who could benefit from platinum-based therapy otherwise undetected on *BRCA1/2* screening. These signatures may also have implications for PARP inhibitor sensitivity, which exploit a synthetic lethal interaction between *PARP-1* and the HR pathway. Germline mutations in *BRCA1* and *BRCA2* are associated with improved response to PARP inhibitors (9). Additional translational research incorporating WGS is necessary to reveal whether HRD mutation signatures are similarly associated with PARP inhibitor response.

However, clinical translation of HRD mutation signatures requires sufficient capture of somatic SNVs and SVs to infer the processes underlying mutagenesis. While HRDetect improves upon the accuracy of the clinically employed LOH, TAI, and LST metrics, it requires WGS, which currently poses technical and financial challenges for clinical use. Further research to develop predictive models that exclude SV signatures may enable application on cancer exomes or other targeted sequencing methods, which can capture sufficient somatic mutations for SNV signature but not SV signature analysis. Additionally, orthogonal HRD assays, for example using gene set expression profiling (46), may also serve as lower cost parameters for treatment prediction. Nevertheless, as sequencing costs fall, WGS provides unique opportunities to integrate diverse markers of genomic instability and mutagenesis within a single protocol. Moreover, we demonstrated that NNLS mutation signature analysis enables accurate "*N of 1*" HRD signature investigation for genome-driven personalized medicine initiatives.

Quantifying HRD signatures supplements existing knowledge and paradigms of cancer detection and stratification. HRDetect scores were associated not only with *BRCA1* and *BRCA2*, but also potentially with other genes such as *PTEN*. This approach provides a functional indicator for mutations whose impact on gene function is uncertain, potentially expanding the repertoire of known causative variants which comprise hereditary cancer screening. Additionally, we observed that HRD signatures were more common in, but not exclusive to, triple-negative and basal-like breast cancers. This agrees with previous work (14) and helps to situate HRD in the context of other widely used breast cancer markers. A topic for future investigation is the value of screening basal-like and triple-negative breast cancers for signatures of HRD.

Breast cancer remains the most common cancer diagnosis in women worldwide. It is evident that a substantial proportion are driven in some part by HRD. Here, we have quantified the relationship between aggregated HRD signatures and measures of success on platinum-based chemotherapy, providing the basis for further investigation of this putative predictive biomarker in prospective trials. In doing so, this study demonstrates the potential for mutation signatures to guide clinical therapy in a precision oncology setting.

Disclosure of Potential Conflicts of Interest

K.A. Gelmon reports receiving speakers bureau honoraria from Astra Zeneca and Pfizer. S. Young reports receiving speakers bureau honoraria from and is a consultant/advisory board member for Astra-Zeneca. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: E.Y. Zhao, E. Pleasance, S. Chia, K. Gelmon, H. Lim, K.A. Schrader, J. Laskin, S.JM. Jones

Development of methodology: E.Y. Zhao, A.J. Mungall, S. Chia, S. Yip, J. Laskin

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): E. Pleasance, M. Jones, J.E. Schein, A.J. Mungall, Y. Zhao, R.A. Moore, S. Wilson, D. Villa, T. Shenkier, C. Lohrisch, S. Chia, S. Yip, K. Gelmon, H. Lim, D. Renouf, S. Sun, J. Laskin

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): E.Y. Zhao, Y. Shen, K. Kasaian, S. Leelakumari, P. Bose, C. Reisle, P. Eirew, R. Corbett, K.L. Mungall, Y. Ma, S. Chia, S. Sun, K.A. Schrader, S. Young, I. Bosdet, A. Karsan, J. Laskin

Writing, review, and/or revision of the manuscript: E.Y. Zhao, Y. Shen, E. Pleasance, M. Jones, P. Bose, A.J. Mungall, R.A. Moore, W. Den Brok, D. Villa, T. Shenkier, C. Lohrisch, S. Chia, K. Gelmon, H. Lim, D. Renouf, S. Sun, K.A. Schrader, I. Bosdet, A. Karsan, J. Laskin, M.A. Marra, S.JM. Jones
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C. Ch'ng, N. Thiessen, Y. Ma, H. Lim, J. Laskin
Study supervision: J. Laskin, M.A. Marra, S.JM. Jones
Other [bioinformatics support (analysis coordination)]: N. Thiessen

Acknowledgments

The authors thank Robyn Roscoe, Alexandra Fok, Katherine Mui, Jessica Nelson, and Payal Simpahimalani. This work would not be possible without the participation of our patients and families and the generous support of the BC Cancer Foundation. We also acknowledge contributions toward equipment and infrastructure from Canada Foundation for Innovation and the BC Knowledge Development Fund. The results published here are in whole or part based upon data generated by The Cancer Genome Atlas managed by the NCI and NHGRI. Information about TCGA can be found at http://cancergenome.nih.gov.

Grant Support

E.Y. Zhao is supported by a BCCA-CIHR-UBC MD/PhD Studentship, a UBC 4-Year Doctoral Fellowship, and a CIHR Vanier Canada Graduate Scholarship.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 6, 2017; revised August 14, 2017; accepted September 26, 2017; published online December 15, 2017.

References

1. Li X, Heyer WD. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res 2008;18:99–113.
2. Joosse SA. BRCA1 and BRCA2: a common pathway of genome protection but different breast cancer subtypes. 2012;12:372–2.
3. Canadian Cancer Society. BRCA gene mutations [Internet]. 2014. Available from: http://www.cancer.ca/en/cancer-information/cancer-101/what-is-a-risk-factor/genetic-risk/brca-gene-mutations/.
4. National Cancer Institute. BRCA1 and BRCA2: Cancer risk and genetic testing [Internet]. 2014. Available from: http://www.cancer.gov/cancertopics/factsheet/Risk/BRCA.
5. Tutt A, Ellis P, Kilburn L, Gilett C, Pinder S, Abraham J, et al. Abstract S3-01: The TNT trial: a randomized phase III trial of carboplatin (C) compared with docetaxel (D) for patients with metastatic or recurrent locally advanced triple negative or BRCA1/2 breast cancer (CRUK/07/012). Cancer Res 2015;75:S3–01.
6. Von Minckwitz G, Hahnen E, Fasching PA, Hauke J, Schneeweiss A, Salat C, et al. Pathological complete response (pCR) rates after carboplatin-containing neoadjuvant chemotherapy in patients with germline BRCA (gBRCA) mutation and triple-negative breast cancer (TNBC): results from GeparSixto. ASCO annual meeting proceedings. 2014;1005.
7. Arun B, Bayraktar S, Liu DD, Barrera AM, Atchley D, Pusztai L, et al. Response to neoadjuvant systemic therapy for breast cancer in BRCA mutation carriers and noncarriers: a single-institution experience. J Clin Oncol 2011;29:3739–46.
8. Byrski T, Gronwald J, Huzarski T, Grzybowska E, Budryk M, Stawicka M, et al. Pathologic complete response rates in young women with BRCA1-positive breast cancers after neoadjuvant chemotherapy. J Clin Oncol 2010;28:375–9.
9. Robson M, Im SA, Senkus E, Xu B, Domchek SM, Masuda N, et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. N Engl J Med 2017;377:523–33.
10. Norquist B, Wurz KA, Pennil CC, Garcia R, Gross J, Sakai W, et al. Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas. J Clin Oncol 2011;29:3008–15.
11. Timms KM, Abkevich V, Hughes E, Neff C, Reid J, Morris B, et al. Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. Breast Cancer Res 2014;16:1.
12. Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, et al. Homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. Clin Cancer Res 2016;22:3764–73.
13. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA Jr, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.
14. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47–54.
15. Davies H, Glodzik D, Morganella S, Yates LR, Staaf J, Zou X, et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nature Med 2017;23:517–25.
16. Jacot W, Theillet C, Guiu S, Lamy PJ. Targeting triple-negative breast cancer and high-grade ovarian carcinoma: refining BRCAness beyond BRCA1/2 mutations? Future Oncol 2015;11:557–9.
17. Stecklein SR, Sharma P. Tumor homologous recombination deficiency assays: another step closer to clinical application? Breast Cancer Res 2014;16:409.
18. Lips EH, Mulder L, Oonk A, Kolk LE van der, Hogervorst FB, Imholz AL, et al. Triple-negative breast cancer: BRCAness and concordance of clinical features with BRCA1-mutation carriers. Br J Cancer 2013;108:2172–7.
19. Laskin J, Jones S, Aparicio S, Chia S, Ch'ng C, Deyell R, et al. Lessons learned from the application of whole-genome analysis to the treatment of patients with advanced cancers. Cold Spring Harb Mol Case Stud 2015;1:a000570.
20. Sheffield BS, Tinker AV, Shen Y, Hwang H, Li-Chang HH, Pleasance E, et al. Personalized oncogenomics: clinical experience with malignant peritoneal mesothelioma using whole genome sequencing. PloS one 2015;10:e0119689.
21. Bose P, Pleasance ED, Jones M, Shen Y, Ch'ng C, Reisle C, et al. Integrative genomic analysis of ghost cell odontogenic carcinoma. Oral Oncol 2015;51:e71–5.
22. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25:1754–60.
23. Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 2010;26:589–95.
24. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25:2078–9.
25. Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor–normal sample pairs. Bioinformatics 2012;28:1811–7.
26. Jones SJ, Laskin J, Li YY, Griffith OL, An J, Bilenky M, et al. Evolution of an adenocarcinoma in response to selection by targeted kinase inhibitors. Genome Biol 2010;11:1–12.
27. Ha G, Roth A, Lai D, Bashashati A, Ding J, Goya R, et al. Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Res 2012;22:1995–2007.
28. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–23.
29. Butterfield YS, Kreitzman M, Thiessen N, Corbett RD, Li Y, Pang J, et al. JAGuaR: junction alignments to genome for RNA-seq reads. PloS one 2014;9:e102398.
30. Bastien RR, Rodríguez-Lescure Á, Ebbert MT, Prat A, Munárriz B, Rowe L, et al. PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers. BMC Med Genomics 2012;5:1.
31. The Cancer Genome Atlas. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61–70.

32. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. Cell Reports 2013;3:246–59.
33. Robertson G, Schein J, Chiu R, Corbett R, Field M, Jackman SD, et al. De novo assembly and analysis of RNA-seq data. Nature Methods 2010;7: 909–12.
34. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 2012;28:i333–9.
35. Wu J, Ho C, Laskin J, Gavin D, Mak P, Duncan K, et al. The development of a standardized software platform to support provincial population-based cancer outcomes units for multiple tumour sites: OaSIS—Outcomes and Surveillance Integration System. Stud Health Technol Inform 2013; 183:98–103.
36. Alexandrov L, Jones P, Wedge D, Sale J, Campbell P, Nik-Zainal S, et al. Clock-like mutational processes in human somatic cells. Nat Commun 2015;47:1402–7.
37. Serrano-Fernández P, Dębniak T, Górski B, Bogdanova N, Dörk T, Cybulski C, et al. Synergistic interaction of variants in CHEK2 and BRCA2 on breast cancer risk. Breast Cancer Res Treat 2009;117:161–5.
38. Johnson N, Fletcher O, Palles C, Rudd M, Webb E, Sellick G, et al. Counting potentially functional variants in BRCA1, BRCA2 and ATM predicts breast cancer susceptibility. Hum Mol Genet 2007;16:1051–7.
39. Mantere T, Winqvist R, Kauppila S, Grip M, Jukkola-Vuorinen A, Tervasmäki A, et al. Targeted next-generation sequencing identifies a recurrent mutation in MCPH1 associating with hereditary breast cancer susceptibility. Eng C, editor. PLOS Genetics 2016;12:e1005816.
40. Tsuneizumi M, Emi M, Hirano A, Utada Y, Tsumagari K, Takahashi K, et al. Association of allelic loss at 8p22 with poor prognosis among breast cancer cases treated with high-dose adjuvant chemotherapy [Internet]. 2002. 75–82. Available from: http://www.sciencedirect.com/science/article/pii/S0304383502000101.
41. Meric-Bernstam F, Farhangfar C, Mendelsohn J, Mills GB. Building a personalized medicine infrastructure at a major cancer center. J Clin Oncol 2013;31:1849–57.
42. Mestan KK, Ilkhanoff L, Mouli S, Lin S. Genomic sequencing in clinical trials. J Transl Med 2011;9:222.
43. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 2017;23:703–713.
44. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol 2016;17:31.
45. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, et al. Mutational processes molding the genomes of 21 breast cancers. Cell 2012;149:979–93.
46. Mulligan JM, Hill LA, Deharo S, Irwin G, Boyle D, Keating KE, et al. Identification and validation of an anthracycline/cyclophosphamide–based chemotherapy response assay in breast cancer. JNCI 2014;106:1.