AcornHRD: an HRD algorithm which was highly associated with anthracycline-based neoadjuvant chemotherapy in Chinese breast cancer

Jia-Ni Pan
Zhejiang Chinese Medicine University，Hangzhou 310053，
Pu-Chun Li
Wenzhou Medical University，Wenzhou 325035，
Meng Wang
AcornMed Biotechnology Co., Ltd
Ming-Wei Li
AcornMed Biotechnology Co., Ltd
Xiao-Wen Ding
Department of Breast Surgery, The Cancer Hospital of the University of Chinese Academy of Sciences(Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer，IBMC，
Tao Zhou
AcornMed Biotechnology Co., Ltd
Hui-Na Wang
AcornMed Biotechnology Co., Ltd
Yun-Kai Wang
AcornMed Biotechnology Co., Ltd
Li-Bin Chen
AcornMed Biotechnology Co., Ltd
Rong Wang
Department of Breast Medical Oncology, The Cancer Hospital of the University of Chinese Academy of Sciences(Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer，IBMC，
Wei-Wu Ye
Department of Breast Medical Oncology, The Cancer Hospital of the University of Chinese Academy of Sciences(Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer，IBMC，
Feng Lou
AcornMed Biotechnology Co., Ltd
Xiao-Jia Wang
Department of Breast Medical Oncology, The Cancer Hospital of the University of Chinese Academy of Sciences(Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer，IBMC，
Wen-Ming Cao（caowm@zjcc.org.cn）

Department of Breast Medical Oncology, The Cancer Hospital of the University of Chinese Academy of Sciences(Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer（IBMC）

Research Article

Keywords: HRD algorithm, breast cancer, anthracycline, neoadjuvant chemotherapy

Posted Date: June 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1770144/v1

Abstract

Background

Homologous recombination deficiency (HRD) can result from BRCA dysfunction and is associated with platinum sensitivity, PARP inhibitor, and other DNA-damaging drugs. There are many commercial HRD detection assays, but there is still no uniform standard in China. This study aimed to develop and validate an HRD scoring algorithm.

Methods

Ninety-six in-house BC samples and 6 HRD positive standard cells were analyzed by whole-genome sequencing (WGS). Besides, 122 BCs from the TCGA database were down-sampled to ~ 1X WGS. We constructed a algorithm named AcornHRD for HRD score based on WGS at low coverage as input data to estimate large-scale copy number alteration (LCNA) events on the genome. The sensitivity and specificity were compared between our algorithm and the ShallowHRD. A clinical cohort of 50 BCs (15 cases carrying *BRCA* mutation) was used to assess the association between HRD status and anthracyclines-based neoadjuvant treatment outcomes.

Results

A 100kb-window was defined as the optimal size using 41 in-house cases and the TCGA dataset. HRD positive threshold was determined as HRD score ≥ 10 using 55 in-house BCs with *BRCA* mutation to achieve 95% sensitivity. The sensitivity and specificity of AcornHRD were both 100%, while those of the ShallowHRD were 40% and 100%, respectively. Meanwhile, AcornHRD sensitivity was superior to ShallowHRD (87% vs 13%) in the clinical cohort. *BRCA* status was significantly associated with HRD status by AcornHRD and ShallowHRD (P = 0.00838 and P = 0.00284, respectively). However, AcornHRD had a higher positive concordance rate than the ShallowHRD algorithm (70% vs 60%). In the clinical cohort of neoadjuvant treatment, the HRD positive group was more likely to respond to anthracycline-based chemotherapy than the HRD negative group, with outcomes of pCR (OR = 9.5, 95% CI: 1.11–81.5, p = 0.04) and residual cancer burden score of 0 or 1 (RCB0/1) (OR = 10.29, 95% CI: 2.02–52.36, p = 0.005). Among 35 patients lacking *BRCA* mutations, the HRD positive group tended to have RCB0/1 responses compared to the HRD negative group (OR = 6.0, 95% CI: 1.00–35.91, p = 0.05).

Conclusion

Here, we developed a stable algorithm for the HRD score. A promising assay for clinical application to predict the sensitivity of DNA-damaging drugs.

Introduction

Breast cancer susceptibility genes *BRCA1* and *BRCA2* are involved in homologous recombination (HR) and play a pivotal role in the repair of DNA double-strand breaks [1]. Cancers with loss of HR function due to inactivation of *BRCA1/2* and other HRR genes are known to be sensitive to platinum and poly (adenosine diphosphate-ribose) polymerase (PARP) inhibitors [2–4]. Germline *BRCA* mutations just account for 5.3% of all breast cancers [5], and Elizabeth G et al. [6] showed that homologous recombination deficiency (HRD) is approximately 18% in breast cancer. HRD testing will allow more precise treatment recommendations and benefit populations for platinum and PARP inhibitors (PARPi). Moreover, the conclusion has been confirmed in multiple clinical trials of ovarian cancer. Patients with *BRCA* wild-type but positive HRD have an equal benefit from PARPi compared with *BRCA* mutations, based on the results of the PRIMA study and the PAOLA-1 study [7, 8]. There are commercial HRD detection methods abroad, but there is no uniform standard in China so far. Therefore, we developed an HRD scoring algorithm based on the Chinese population to precisely guide medication and screen benefit populations.

In current practice, anthracycline-based regimens and sequential administration of taxanes are the most commonly used chemotherapy regimens in neoadjuvant and adjuvant settings. Prior studies have shown that platinum chemotherapy agents are active in the treatment of breast cancer with a germline *BRCA* mutation and/or HRD [9–11]. In the neoadjuvant setting, a single-arm prospective study using cisplatin monotherapy reported a pathologic complete response (pCR) rate of 61% among *BRCA1*-mutated breast cancer patients [12]. Moreover, the GeparSixto trial demonstrated that the pCR rates were 33.9% and 63.5% in the paclitaxel plus liposomal doxorubicin (PM) group and PM plus carboplatin group, respectively, among the HRD breast cancer [11]. Conversely, the INFORM trial results showed that anthracycline-based regimens are also effective in HER2-negative *BRCA*-mutated breast cancer. The pCR rate was 18% and 26% in the single-agent cisplatin group and doxorubicin-cyclophosphamide group, respectively, which yielded a risk ratio (RR) of 0.70 (90%CI, 0.39–1.2) [13]. Moreover, it was recently reported that breast cancers with high HRD scores are more sensitive to anthracycline in the neoadjuvant setting [14, 15].

This study aimed to develop an HRD scoring algorithm. We construct an operational model and compare HRD status between different algorithms. To validate the accuracy of this HRD scoring algorithm, we evaluated the correlation of HRD scores with *BRCA* mutations was assessed first, and then the correlation of HRD scores with pCR to anthracycline-based neoadjuvant chemotherapy (NAC) was assessed.

Materials And Method

DNA extractions, library preparation and sequencing

Five HRD positive and 1 HRD negative standards (Cat No. CBP90023) stored at -20°C from Nanjing Cobioer biosciences CO., LTD, prepared for genome-wide DNA extraction. All genomic DNA (gDNA)

samples were extracted using the Genomic DNA Extraction Kit (Item No DP304). Forty-one in-house collected whole genomic DNA samples (cohort I) and 55 in-house collected whole genomic DNA samples with *BRCA* mutations (cohort II) from 85 patients with Breast cancer were also prepared for library construction. According to the quantitative results of the QUIBT tool, 200ng gDNA was used for library construction. Then 200ng gDNA for each sample was transferred to a 50μL Covaris tube and segmented to the main peak of 300-350bp using the Covaris M220 instrument. Subsequently, segmented DNA was end-repaired, A-tailed, and ligated with custom adapters in reaction pooling. The ligation product was amplified (6 cycles) and purified using AmpureXP beads (Agencourt/Beckman Coulter). After purification, the library was quantified using Qubit 4 fluorimeter and Qubit dsDNA HS Assay Kit (Thermofisher). Finally, library fragment quality control was performed using Agilent 2100 Bioanalyzer and Agilent 2100 DNA 1000 Kit. Each library is programmed to generate ~3.5Gb bases.

Filter and variant calling

FASTP tool [16] was applied for FASTQ file quality control to remove those reads with the adaptor, low-quality bases. High-quality reads were aligned into the human genome (hg19) with Burrows-Wheeler Aligner [17]. Duplicate reads generated by PCR were marked using Picard (broadinstitute.github.io/picard/). Moreover, local realignment around known InDels and base quality were recalibrated and then removed duplicate reads using the Sentieon tool [18]. Finally, base alternatives and InDels detected by Sentieon were annotated using Annovar [19]. A series of 122 aligned bam files (Supplementary Table 1) downloaded from the TCGA breast cancer database (www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga) were down-sampled to ~1X whole-genome sequencing (WGS) with SamBamba software [20]. Then all bam files also were processed with the above pipelines. *BRCA* positive status implies that any mutation was detected in either *BRCA1* or *BRCA2* for each sample. Identified criteria of mutation in *BRCA* were as follows: 1) For somatic mutation, mutation information was collected from the TCGA database. 2) For germline mutation identified by in-house analysis pipeline, its status was annotated as likely pathogenic or pathogenic by either InterVar [21] or ClinVar [22], and the depth of supporting allele reads was greater than 3.

Workflow of the algorithm for HRD evaluation

An in-house algorithm (defined as AcornHRD) of low-depth WGS copy-number variation (CNV) detection was constructed to evaluate HRD status for patients with cancer. HRD score was predicted based on the large-scale copy number alteration (LCNA) events, and the methodology is similar to the LST in SNP arrays. We set out to evaluated the HRD score, defined as the number of LCNAs. Tumors with HRD scores $\geq$ 10 were defined as HRD positive (See Results section for more details). AcornHRD adopted the pattern of the sliding window to detect LCNA events on the genome. The detailed algorithm description was divided into two parts. Part one was to calculate the coverage depth in the window unit along the genome as follows: 1) The coverage depth in each window was calculated, and then GC correction and self-standardization were carried out for the coverage depth value of each window. 2) Next, the coverage depth value of each window was normalized using an in-house constructed baseline. 3) Finally, the

circular binary segmentation (CBS) method was utilized to calculate the depth ratio value and median depth ratio value for each window with the R procedure (bioconductor.org/packages/release/bioc/html/DNAcopy.html). The ratio result file was used as input data to estimate the HRD status for each sample. Part two was to detect HRD status as follows: 1) Firstly, to minimize the impacts from highly complex genomic regions (such as centromere regions, telomere regions and highly repetitive regions) and sex chromosomes, those overlap windows are removed. 2) Then, those ratios of sequencing depth in every window were processed with log2 fold change. 3) Subsequently, those windows were merged into large segments with chromosome arm information and processed depth ratio. 4) Segments larger than 100kb were used for smoothing simulation in building larger fragments. 5) Finally, the above segment larger than 10Mb is defined as LCNA event.

Validation by clinical breast cancer samples

We retrospectively reviewed the medical records of 1449 patients with primary breast cancer who visited the Cancer Hospital Affiliated to the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital) from February 2008 to October 2020 and completed a 98-gene panel genetic screening. Fifty patients who received anthracycline-based NAC and underwent subsequent surgery (mastectomy or breast-conserving surgery) were included in the statistical analysis (see Figure 1 and Supplementary Table 2 for details). All of them received NAC with epirubicin (75 mg/m2) and cyclophosphamide (600 mg/m2), followed by docetaxel (80-100 mg/m2) every 3 weeks for 8 cycles. All biopsied tumor samples for histological and HRD examination were obtained from patients before NAC and kept by fixed-formalin paraffin-embedded (FFPE). The study was reviewed and approved by the Ethical Committee of Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital) and was performed in accordance with the Declara tion of Helsinki.

Statistical analyses

The mode values used to identify the befitting window size were counted using an in-house script. Chi-Squared test was performed using the Scipy package (scipy.org) of Python (version 3).

Results

AcornHRD algorithm construction

AcornHRD was based on the results of sequencing depth ratio as input data to estimate LCNA events on the genome. For those samples of low sequencing depth, a fitness window size seems particularly important. To address the question of optimal window size, we adopted seven different window sizes (40kb, 100kb, 300kb, 500kb, 800kb, 1Mb and 1.4Mb) to estimate the number of LCNA based on the samples from cohort I. For each sample, the mode of the count of LCNA in different window sizes was calculated. There were up to all samples in every window size. The statistical results showed that the 100kb window covered the most 31 (75.6%) samples (Figure 2A). Besides, the 500kb-window size had a close value (73.2%) regarding the number of samples. To validate the stability of the 100kb-window size,

breast cancer samples from the TCGA cohort were utilized to identify LCNA events. As expected, the number of samples covered by the 100kb-window size still remained the maximum number (Figure 2B). Therefore, the 100kb-window size had good stability and was defined as the optimal window size for the following analysis.

Mutations in the *BRCA* are strongly associated with HRD positive [23, 24]. To verify AcornHRD sensitivity, the patients with *BRCA* mutation from cohort II were constructed into a test panel with a 100kb-window size and 50kb-step size. For a 95% confidence detection rate, the score of 10 was defined as the cut-off threshold value. Of 55 tumors, 53 (96.4%) were determined to be HRD positive with a score greater than or equal to 10 (Supplementary Table 3). The result of 95% sensitivity demonstrated a good degree of credibility.

Comparing HRD status between different algorithms

For a more comprehensive evaluation of the algorithm, ShallowHRD software [25] was joined into the later comparative analysis. Six standard samples were sequenced with whole genomic DNA as described in the method section. The sensitivity and specificity of AcornHRD are both 100%, while those of the ShallowHRD are 40% and 100%, respectively (Table 1). Subsequently, the clinical cohort including 15 *BRCA*-positive and 35 *BRCA*-negative patients from the clinical cohort, was used to detect HRD status. The HRD results of the clinical cohort showed that AcornHRD sensitivity is far superior to ShallowHRD, but its specificity is not as good (Supplementary Table 4 and Supplementary Table 5).

Correlation between *BRCA* mutations and HRD status

Variation located in the *BRCA* as a prominent hallmark had been studied in many cancers [26-30]. As previous studies have reported, *BRCA* mutations are the most critical factors in patients with breast cancer to assess risk [31, 32]. HRD, another tumor biomarker, is being used in guiding therapy in more and more studies [33-36]. *BRCA* mutations are known to be strongly associated with HRD; therefore, we applied *BRCA* mutation status to compare the accuracy of the two HRD assessment methods, TCGA cohort (2 samples without somatic mutation information were filtered out) was analyzed using AcornHRD and ShallowHRD software [25], respectively. Firstly, the mutations of *BRCA* genes were confirmed in tumor sequencing reads by in-house calling variation pipeline (more details presented in Methods Part). Of the 120 patients, 20 harbored *BRCA* mutations. Subsequently, the HRD status of 120 tumor tissue samples was identified by two algorithms. The results of AcornHRD (Table 2) and ShallowHRD (Table 3) both show that *BRCA* status is significantly correlated with HRD status (P=0.00838 and P=0.00284, respectively). However, the positive agreement rate of AcornHRD is higher than the ShallowHRD algorithm, which is 70% (14/20) and 60% (12/20), respectively (Table 2 and Table 3). In summary, AcornHRD is more stable in the application performance of three different cohort of WGS data, which is superior to the published algorithm.

The HRD in clinical cohort

High HRD score was also significantly correlated with *BRCA* mutation (Table 4) in the 50 clinical cohort. Among the 50 patients who received anthracycline-based neoadjuvant therapy, 28 had high HRD scores, and 22 had low HRD scores. The breast cancer samples selected for clinical study were all HER-2 negative, including 24 TNBC samples and 26 ER and/or PR positive samples. High HRD score was significantly correlated with TNBC and high Ki-67 expression (Table 4). HRD score high trend to ER negative and PR negative (Table 4).

Correlation between HRD status and NAC efficacy

In this study, HRD positive includes either a high HRD score or a BRCA mutation, whereas HRD negative includes a low HRD score and no BRCA mutation. Of the 50 patients, 30 were identified as HRD positive. Moreover, pCR and residual tumor burden (RCB 0/1) are both important indicators for tumor efficacy evaluation, of which pCR (RCB 0) is the main evaluation indicator.

Regarding the HRD status of all patients (n = 50), patients with HRD positive were more likely to respond to standard NAC containing anthracyclines than HRD negative patients, with a pCR (RCB 0) outcome (OR = 9.5, 95% CI 1.11 - 81.5, p = 0.04) (Table 5). Similar results were observed for the endpoint RCB of 0/1. In the entire cohort of 50 patients, patients with HRD positive were more likely to achieve RCB 0/1 compared to non-deficient patients (OR = 10.29, 95% CI 2.02 - 52.36, p = 0.005) (Table 6). This applied to a cohort of 35 patients lacking germline BRCA mutations; patients with HRD trended toward an RCB 0/1 response compared with HRD negative patients (OR = 6.0, 95% CI 1.00 - 35.91, p = 0.05) (Table 6).

Discussion

Genomic scar analysis is a very important HRD detection method. When non-homologous end joining (NHEJ) repair is initiated, it will leave traces of damage repair in the genome, that is, "genomic scar [37]". Cells with HRD cannot repair DSBs in such a reliable manner as HR and therefore cause genomic scarring, and these quantifiable genomic alterations help to reverse the HRD status of the cell. There are three main types of genomic scars caused by HRD: LOH, TAI and LST [24]. To date, FDA has approved two products for clinical testing of HRD, Myriad myChoice ® CDx (myriad-oncology.com/mychoice-cdx) and FoundationFocus ™ CDx BRCA LOH (www.accessdata.fda.gov/cdrh_docs/pdf16/p160018c.pdf). Both products use the detection of *BRCA* gene mutations combined with the genomic scar to assess HRD status. The former contained the *BRCA* genes coding region and 54, 091 population Single Nucleotide Polymorphisms (SNPs). The Genomic Instability Score (GIS) was obtained by comprehensively calculating three indicators: LOH, TAI and LST, while GIS ≥ 42 was considered positive for genomic instability status [35, 38]. The latter calculated the proportion of fragments with LOH in this genome by covering 3500 SNPs in 324 genes on 22 chromosomes, and LOH accounted for ≥ 16%, that is, “high LOH [39]”. The above two commercial kits lack large-sample prospective clinical study data applied to the Chinese population, and it is urgent to promote the development and clinical validation of corresponding kits in China.

It has been confirmed that LST is feasible and has unique advantages for assessing HRD [24, 40–42]. LST is referred to the number of chromosomal breaks between flanking regions of at least 10 Mb [40, 43]. It has been reported that LST genomic signature accurately identified tumors with HRD and displayed excellent performance in a TNBC cohort reaching almost 100% in sensitivity and specificity for HRD detection, where HRD was defined as *BRCA* inactivation [40, 41]. LST also had better HRD evaluation performance in low-depth sequencing compared to LOH [44].

shallowHRD is a software tool based on mining copy number alterations profile from TCGA breast cancer that displayed a high performance for HRD detection in breast cancers in low coverage genomic data [25]. Fundamental to evaluating the HRD status is the robust determination of copy number data, which can be obtained using either SNP arrays, whole exome sequencing (WES) or WGS. Compared SNP array-, WES- to WGS-derived CNV have shown that WGS provides much more homogenous distribution of quality parameters (genotype quality, coverage depth) [45–47]. Studies indicated an excellent agreement (93.75%) between the original and downsampled WGS-derived HR classification status [44]. WGS at low coverage robustly detects CNV, even in FFPE samples and liquid biopsies [48], at low cost and with easy-storable data outputs.

LCNA identified with Shallow whole-genome sequencing is more and more popular in many diagnosis institutions. However, low sequencing depth also brings some challenges. Since the shallowHRD data are based on Western cases, it is unclear whether it is applicable to Chinese patients [25]. In the initial use of shallowHRD, it was found that its performance of sensitivity in detecting HRD was poor. In addition, when considering lower coverage genomes, the sensitivity to fully characterize somatic variations (single-nucleotide variants (SNVs), breakpoints and CNVs) becomes compromised, especially in tumors of low cellularity or when sequencing data present strong GC bias. Moreover, the uniformity of sequencing should be high; otherwise, it will be accompanied by serious noise pollution. The low coverage sequencing has to be balanced with the sensitivity and uniformity for robustly calling somatic mutations.

To address the questions, we developed a algorithm named AcornHRD, which detects LCNA events based on a low-depth detection algorithm of ~ 1x WGS reads. Compared with similar software shallowHRD, AcornHRD achieved a good capacity for HRD detection and improved the obvious disadvantage of shallowHRD of low sensitivity in the standards and Chinese breast cancer cohort. Moreover, its sensitivity is much better. In summary, AcornHRD's performance in evaluating HRD is better than that of shallowHRD. However, it is worth mentioning that different HRD assessment methods and their algorithms are non-equivalent, and AcornHRD needs to be further compared with the two FDA-approved products.

We investigated the relationship between the high HRD score and clinicopathological features of breast cancer, high HRD score was significantly correlated with *BRCA* mutation and high Ki-67 expression, and trend to ER negative and PR negative. Thus, the phenotype of high HRD score tumors is considered to be biologically aggressive. High HRD score was significantly more prevalent in the triple-negative breast

cancer (TNBC) subtype than in the other three subtypes, which is consistent with previously reported results [24, 41, 49–51].

It has been shown that anthracycline-based regimens are effective in HER2-negative *BRCA*-mutated breast cancer [11, 13, 52–54]. Telli et al. [14] reported that HRD high TNBC identified by next-generation sequencing was more sensitive to anthracyclines in the neoadjuvant setting. Recently, it has been reported that HRD tumors are more likely to benefit from anthracyclines, and HRD scores may be a clinically useful marker of chemosensitivity based on subtypes [51, 55]. In contrast, Imanishi et al. [49] reported the opposite result. We retrospectively analyzed the association of HRD score with the response to neoadjuvant anthracycline therapy in patients with HER2-negative breast cancer. The results showed that HRD score was significantly associated with RCB 0/I and pCR in the total population cohort (n = 50), and similar results were obtained in the cohort of patients lacking germline *BRCA* mutations (n = 35).

Previous studies have shown that response to neoadjuvant platinum-based therapy (pCR and RCB0/1) is significantly associated with HRD status in TNBC [11, 35, 56], suggesting the clinical utility of HRD scoring in selecting breast tumors that are more likely to respond to platinum-based regimens. Conversely, the GeparOLA study [57] found that neoadjuvant Olaparib-based therapy could bring a higher pCR rate than carboplatin-based regimen (55.1% vs. 48.6%) HER2-negative or TNBC with HRD. Consistent results were obtained in younger (< 40 years) and HR-positive patients (76.2% vs 45.5%, 52.6% vs 20.0%, respectively), suggesting that HRD may have a good application prospect in predicting the efficacy of PARPi for neoadjuvant treatment.

Conclusion

In conclusion, we developed an HRD score algorithm, named AcornHRD, which has high sensitivity and specificity and can accurately predict the efficacy of anthracycline-based NAC. Furthermore, AcornHRD can be employed in clinical applications and translational research, such as screening patients for clinical research and the practice of DNA-damaging drugs.

Abbreviations

Homologous recombination deficiency (HRD), homologous recombination (HR), pathological complete response (pCR), *BRCA*-mutated breast cancers (BC), whole-genome sequencing (WGS), large-scale copy number alteration (LCNA), platinum and poly (adenosine diphosphate-ribose) polymerase (PARP), loss of heterozygosity (LOH), telomeric allelic imbalance (TAI), large-scale state transition (LST), pathologic complete response (pCR), risk ratio (RR), fixed-formalin paraffin-embedded (FFPE), non-homologous end joining (NHEJ), single-nucleotide variants (SNVs), the triple-negative breast cancer (TNBC)

Declarations

Ethics approval and consent to participate

The research has been approved by the Ethical Committee of Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital).

Consent for publication

Informed consent was obtained from all individual participants included in the study.

Availability of data and materials

All the data supporting the results of this manuscript are available in the article and Supplemental Tables 1, 2 and 3. Further inquiries can be directed to the corresponding authors on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The study was funded by Key Research-Development Program of Zhejiang Province (grant number: 2019C04001), Natural Science Foundation of Zhejiang Province (grant number: LY21H160005), Medical and Health Research Project of Zhejiang Province (grant number: 2021KY566).

Authors' contributions

Jia-Ni Pan: Conceptualization, Software, Formal Analysis, Writing - Original Draft;

Pu-Chun Li: Data Curation, Methodology, Writing - Original Draft;

Meng Wang: Software, Methodology, Validation;

Ming-Wei Li: Visualization, Conceptualization, Writing - Original Draft;

Xiao-Wen Ding: Investigation, Funding Acquisition;

Tao Zhou: Formal Analysis, Software;

Hui-Na Wang: Project Administration;

Yun-Kai Wang: Validation, Investigation;

Li-Bin Chen: Software, Visualization;

Rong Wang: Resources, Supervision;

Wei-Wu Ye: Validation, Resources;

Feng Lou: Formal analysis, Validation, Software;

Xiao-Jia Wang: Visualization, Resources, Supervision;

Wen-Ming Cao: Conceptualization, Funding Acquisition, Supervision, Writing - Review & Editing.

Acknowledgments

Not applicable.

Authors' information

panjiani756@163.com; 958998326@qq.com；mengwang@acornmed.com；mingweili@acornmed.com; dingxw@zjcc.org.cn ; taozhou@acornmed.com；huinawang@acornmed.com； yunkaiwang@acornmed.com；libinchen@acornmed.com; jiangsuwangrong2006@126.com； zjhzy76@163.com；fenglou@acornmed.com; wxiaojia0803@163.com； caowm@zjcc.org.cn

References

1. Severson TM, Peeters J, Majewski I, Michaut M, Bosma A, Schouten PC, Chin SF, Pereira B, Goldgraben MA, Bismeijer T *et al*: **BRCA1-like signature in triple negative breast cancer: Molecular and clinical characterization reveals subgroups with therapeutic potential**. Mol Oncol 2015, **9**(8):1528–1538.
2. Yip CH, Newman LA: **American Society of Clinical Oncology, American Society for Radiation Oncology, and Society of Surgical Oncology Guideline for Management of Hereditary Breast Cancer**. JAMA Surg 2021, **156**(3):284–285.
3. Tung NM, Boughey JC, Pierce LJ, Robson ME, Bedrosian I, Dietz JR, Dragun A, Gelpi JB, Hofstatter EW, Isaacs CJ *et al*: **Management of Hereditary Breast Cancer: American Society of Clinical Oncology, American Society for Radiation Oncology, and Society of Surgical Oncology Guideline**. J Clin Oncol 2020, **38**(18):2080–2106.
4. Robson M, Im SA, Senkus E, Xu B, Domchek SM, Masuda N, Delaloge S, Li W, Tung N, Armstrong A *et al*: **Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation**. N Engl J Med 2017, **377**(6):523–533.
5. Cancer Genome Atlas N: **Comprehensive molecular portraits of human breast tumours**. Nature 2012, **490**(7418):61–70.
6. Turner NC: **Signatures of DNA-Repair Deficiencies in Breast Cancer**. N Engl J Med 2017, **377**(25):2490–2492.
7. Gonzalez-Martin A, Pothuri B, Vergote I, DePont Christensen R, Graybill W, Mirza MR, McCormick C, Lorusso D, Hoskins P, Freyer G *et al*: **Niraparib in Patients with Newly Diagnosed Advanced Ovarian Cancer**. N Engl J Med 2019, **381**(25):2391–2402.
8. Ray-Coquard I, Pautier P, Pignata S, Perol D, Gonzalez-Martin A, Berger R, Fujiwara K, Vergote I, Colombo N, Maenpaa J *et al*: **Olaparib plus Bevacizumab as First-Line Maintenance in Ovarian Cancer**. N Engl J Med 2019, **381**(25):2416–2428.

9. Silver DP, Richardson AL, Eklund AC, Wang ZC, Szallasi Z, Li Q, Juul N, Leong CO, Calogrias D, Buraimoh A *et al*: **Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer**. J Clin Oncol 2010, **28**(7):1145–1153.
10. Tutt A, Tovey H, Cheang MCU, Kernaghan S, Kilburn L, Gazinska P, Owen J, Abraham J, Barrett S, Barrett-Lee P *et al*: **Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT Trial**. Nat Med 2018, **24**(5):628–637.
11. Loibl S, Weber KE, Timms KM, Elkin EP, Hahnen E, Fasching PA, Lederer B, Denkert C, Schneeweiss A, Braun S *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.
12. Byrski T, Huzarski T, Dent R, Marczyk E, Jasiowka M, Gronwald J, Jakubowicz J, Cybulski C, Wisniowski R, Godlewski D *et al*: **Pathologic complete response to neoadjuvant cisplatin in BRCA1-positive breast cancer patients**. Breast Cancer Res Treat 2014, **147**(2):401–405.
13. Tung N, Arun B, Hacker MR, Hofstatter E, Toppmeyer DL, Isakoff SJ, Borges V, Legare RD, Isaacs C, Wolff AC *et al*: **TBCRC 031: Randomized Phase II Study of Neoadjuvant Cisplatin Versus Doxorubicin-Cyclophosphamide in Germline BRCA Carriers With HER2-Negative Breast Cancer (the INFORM trial)**. J Clin Oncol 2020, **38**(14):1539–1548.
14. Telli ML, Hellyer J, Audeh W, Jensen KC, Bose S, Timms KM, Gutin A, Abkevich V, Peterson RN, Neff C *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.
15. Davies H, Glodzik D, Morganella S, Yates LR, Staaf J, Zou X, Ramakrishna M, Martin S, Boyault S, Sieuwerts AM *et al*: **HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures**. Nat Med 2017, **23**(4):517–525.
16. Chen S, Zhou Y, Chen Y, Gu J: **fastp: an ultra-fast all-in-one FASTQ preprocessor**. Bioinformatics 2018, **34**(17):i884-i890.
17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: **The Sequence Alignment/Map format and SAMtools**. Bioinformatics 2009, **25**(16):2078–2079.
18. View ORCID ProfileDonald Freed, Rafael Aldana, Jessica A. Weber, Edwards JS: **The Sentieon Genomics Tools - A fast and accurate solution to variant calling from next-generation sequence data**. *bioRxiv* 2017(9).
19. Wang K, Li M, Hakonarson H: **ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data**. Nucleic acids research 2010, **38**(16):e164.
20. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P: **Sambamba: fast processing of NGS alignment formats**. Bioinformatics 2015, **31**(12):2032–2034.
21. Li Q, Wang K: **InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines**. American journal of human genetics 2017, **100**(2):267–280.

22. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J *et al*: **ClinVar: public archive of interpretations of clinically relevant variants**. Nucleic acids research 2016, **44**(D1):D862-868.
23. Moschetta M, George A, Kaye SB, Banerjee S: **BRCA somatic mutations and epigenetic BRCA modifications in serous ovarian cancer**. Ann Oncol 2016, **27**(8):1449–1455.
24. Timms KM, Abkevich V, Hughes E, Neff C, Reid J, Morris B, Kalva S, Potter J, Tran TV, Chen J *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.
25. Eeckhoutte A, Houy A, Manie E, Reverdy M, Bieche I, Marangoni E, Goundiam O, Vincent-Salomon A, Stoppa-Lyonnet D, Bidard FC *et al*: **ShallowHRD: detection of homologous recombination deficiency from shallow whole genome sequencing**. Bioinformatics 2020, **36**(12):3888–3889.
26. Gumaste PV, Penn LA, Cymerman RM, Kirchhoff T, Polsky D, McLellan B: **Skin cancer risk in BRCA1/2 mutation carriers**. The British journal of dermatology 2015, **172**(6):1498–1506.
27. Hawsawi YM, Al-Numair NS, Sobahy TM, Al-Ajmi AM, Al-Harbi RM, Baghdadi MA, Oyouni AA, Alamer OM: **The role of BRCA1/2 in hereditary and familial breast and ovarian cancers**. Molecular genetics & genomic medicine 2019, **7**(9):e879.
28. Hu Y, Guo M: **Synthetic lethality strategies: Beyond BRCA1/2 mutations in pancreatic cancer**. Cancer science 2020, **111**(9):3111–3121.
29. Lowery MA: **Genotype-phenotype correlation in BRCA1/2 mutation-associated pancreatic cancer**. British journal of cancer 2020, **122**(3):293–294.
30. Sopik V, Phelan C, Cybulski C, Narod SA: **BRCA1 and BRCA2 mutations and the risk for colorectal cancer**. Clinical genetics 2015, **87**(5):411–418.
31. Concolino P, Capoluongo E: **Detection of BRCA1/2 large genomic rearrangements in breast and ovarian cancer patients: an overview of the current methods**. Expert review of molecular diagnostics 2019, **19**(9):795–802.
32. Engel C, Fischer C, Zachariae S, Bucksch K, Rhiem K, Giesecke J, Herold N, Wappenschmidt B, Hübbel V, Maringa M *et al*: **Breast cancer risk in BRCA1/2 mutation carriers and noncarriers under prospective intensified surveillance**. Int J Cancer 2020, **146**(4):999–1009.
33. Frey MK, Pothuri B: **Homologous recombination deficiency (HRD) testing in ovarian cancer clinical practice: a review of the literature**. Gynecologic oncology research and practice 2017, **4**:4.
34. Lotan TL, Kaur HB, Salles DC, Murali S, Schaeffer EM, Lanchbury JS, Isaacs WB, Brown R, Richardson AL, Cussenot O *et al*: **Homologous recombination deficiency (HRD) score in germline BRCA2- versus ATM-altered prostate cancer**. Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc 2021, **34**(6):1185–1193.
35. Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, Szallasi Z, Barry WT, Winer EP, Tung NM *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.

36. Zhao EY, Shen Y, Pleasance E, Kasaian K, Leelakumari S, Jones M, Bose P, Ch'ng C, Reisle C, Eirew P *et al*: **Homologous Recombination Deficiency and Platinum-Based Therapy Outcomes in Advanced Breast Cancer**. Clin Cancer Res 2017, **23**(24):7521–7530.
37. McLornan DP, List A, Mufti GJ: **Applying synthetic lethality for the selective targeting of cancer**. N Engl J Med 2014, **371**(18):1725–1735.
38. Jenner ZB, Sood AK, Coleman RL: **Evaluation of rucaparib and companion diagnostics in the PARP inhibitor landscape for recurrent ovarian cancer therapy**. Future Oncol 2016, **12**(12):1439–1456.
39. Coleman RL, Oza AM, Lorusso D, Aghajanian C, Oaknin A, Dean A, Colombo N, Weberpals JI, Clamp A, Scambia G *et al*: **Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial**. Lancet 2017, **390**(10106):1949–1961.
40. Popova T, Manie E, Rieunier G, Caux-Moncoutier V, Tirapo C, Dubois T, Delattre O, Sigal-Zafrani B, Bollet M, Longy M *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.
41. Manie E, Popova T, Battistella A, Tarabeux J, Caux-Moncoutier V, Golmard L, Smith NK, Mueller CR, Mariani O, Sigal-Zafrani B *et al*: **Genomic hallmarks of homologous recombination deficiency in invasive breast carcinomas**. Int J Cancer 2016, **138**(4):891–900.
42. Isakoff SJ, Mayer EL, He L, Traina TA, Carey LA, Krag KJ, Rugo HS, Liu MC, Stearns V, Come SE *et al*: **TBCRC009: A Multicenter Phase II Clinical Trial of Platinum Monotherapy With Biomarker Assessment in Metastatic Triple-Negative Breast Cancer**. J Clin Oncol 2015, **33**(17):1902–1909.
43. Natrajan R, Mackay A, Lambros MB, Weigelt B, Wilkerson PM, Manie E, Grigoriadis A, A'Hern R, van der Groep P, Kozarewa I *et al*: **A whole-genome massively parallel sequencing analysis of BRCA1 mutant oestrogen receptor-negative and -positive breast cancers**. J Pathol 2012, **227**(1):29–41.
44. de Luca XM, Newell F, Kazakoff SH, Hartel G, McCart Reed AE, Holmes O, Xu Q, Wood S, Leonard C, Pearson JV *et al*: **Using whole-genome sequencing data to derive the homologous recombination deficiency scores**. NPJ Breast Cancer 2020, **6**:33.
45. Smolander J, Khan S, Singaravelu K, Kauko L, Lund RJ, Laiho A, Elo LL: **Evaluation of tools for identifying large copy number variations from ultra-low-coverage whole-genome sequencing data**. BMC Genomics 2021, **22**(1):357.
46. Guan Y, Wang X, Guan K, Wang D, Bi X, Xiao Z, Xiao Z, Shan X, Hu L, Ma J *et al*: **Copy number variation of urine exfoliated cells by low-coverage whole genome sequencing for diagnosis of prostate adenocarcinoma: a prospective cohort study**. BMC Med Genomics 2022, **15**(Suppl 2):104.
47. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, Shang L, Boisson B, Casanova JL, Abel L: **Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants**. Proc Natl Acad Sci U S A 2015, **112**(17):5473–5478.
48. Van Roy N, Van Der Linden M, Menten B, Dheedene A, Vandeputte C, Van Dorpe J, Laureys G, Renard M, Sante T, Lammens T *et al*: **Shallow Whole Genome Sequencing on Circulating Cell-Free DNA**

Allows Reliable Noninvasive Copy-Number Profiling in Neuroblastoma Patients. Clin Cancer Res 2017, **23**(20):6305–6314.

49. Imanishi S, Naoi Y, Shimazu K, Shimoda M, Kagara N, Tanei T, Miyake T, Kim SJ, Noguchi S: **Clinicopathological analysis of homologous recombination-deficient breast cancers with special reference to response to neoadjuvant paclitaxel followed by FEC**. Breast Cancer Res Treat 2019, **174**(3):627–637.
50. 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.
51. Kim SJ, Sota Y, Naoi Y, Honma K, Kagara N, Miyake T, Shimoda M, Tanei T, Seno S, Matsuda H *et al*: **Determining homologous recombination deficiency scores with whole exome sequencing and their association with responses to neoadjuvant chemotherapy in breast cancer**. Transl Oncol 2021, **14**(2):100986.
52. Wang C, Zhang J, Wang Y, Ouyang T, Li J, Wang T, Fan Z, Fan T, Lin B, Xie Y: **Prevalence of BRCA1 mutations and responses to neoadjuvant chemotherapy among BRCA1 carriers and non-carriers with triple-negative breast cancer**. Ann Oncol 2015, **26**(3):523–528.
53. Hahnen E, Lederer B, Hauke J, Loibl S, Krober S, Schneeweiss A, Denkert C, Fasching PA, Blohmer JU, Jackisch C *et al*: **Germline Mutation Status, Pathological Complete Response, and Disease-Free Survival in Triple-Negative Breast Cancer: Secondary Analysis of the GeparSixto Randomized Clinical Trial**. JAMA Oncol 2017, **3**(10):1378–1385.
54. Pohl-Rescigno E, Hauke J, Loibl S, Mobus V, Denkert C, Fasching PA, Kayali M, Ernst C, Weber-Lassalle N, Hanusch C *et al*: **Association of Germline Variant Status With Therapy Response in High-risk Early-Stage Breast Cancer: A Secondary Analysis of the GeparOcto Randomized Clinical Trial**. JAMA Oncol 2020, **6**(5):744–748.
55. Sharma P, Barlow WE, Godwin AK, Pathak H, Isakova K, Williams D, Timms KM, Hartman AR, Wenstrup RJ, Linden HM *et al*: **Impact of homologous recombination deficiency biomarkers on outcomes in patients with triple-negative breast cancer treated with adjuvant doxorubicin and cyclophosphamide (SWOG S9313)**. Ann Oncol 2018, **29**(3):654–660.
56. Kaklamani VG, Jeruss JS, Hughes E, Siziopikou K, Timms KM, Gutin A, Abkevich V, Sangale Z, Solimeno C, Brown KL *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.
57. Fasching PA, Link T, Hauke J, Seither F, Jackisch C, Klare P, Schmatloch S, Hanusch C, Huober J, Stefek A *et al*: **Neoadjuvant paclitaxel/olaparib in comparison to paclitaxel/carboplatinum in patients with HER2-negative breast cancer and homologous recombination deficiency (GeparOLA study)**. Ann Oncol 2021, **32**(1):49–57.

Tables

Table1 HRD status of six standards by AcornHRD and ShallowHRD

Standards ID		AcornHRD		shallowHRD	
	Proven status	HRD score	HRD status	HRD score	HRD status
102109017T3	Positive	30	Positive	34	Positive
102109018T3	Positive	14	Positive	13	Negative
102109019T3	Positive	25	Positive	22	Positive
102109020T3	Positive	12	Positive	12	Negative
102109021T3	Positive	11	Positive	8	Negative
102109022T3	Negative	3	Negative	3	Negative

Table 2 The HRD status according to *BRCA* mutations by AcornHRD in TCGA cohort

HRD Status	*BRCA* Status		Total
	Mutated	Non-mutated	
Positive	14	38	52
Negative	6	62	78
Total	20	100	120

HRD positive = HRD score ≥10

Table 3 The HRD status according to *BRCA* mutations by ShallowHRD in TCGA cohort

HRD Status	*BRCA* Status		Total
	Mutated	Non-mutated	
Positive	12	26	38
Negative	8	74	82
Total	20	100	120

HRD positive = HRD score ≥10

Table4 Patient Characteristics and HRD score from clinical cohort

	HRD score high (n=28)	HRD score low (n=22)	OR	95%CI	P value
Age				0.58-5.62	0.31
>40 years	10	11	1		
≤40 years	18	11	1.8		
BMI(Kg/m^2)				0.54-6.67	0.32
<25	18	17	1		
≥25	10	5	1.89		
BRCA status				1.69-44.34	0.011
Non-mutated	15	20	1		
Mutated	13	2	8.67		
Menopause				0.11-2.72	0.73
Pre-	25	18	1		
Post-	3	4	0.54		
ER				0.10-1.02	0.05
Negative	18	8	1		
Positive	10	14	0.32		
PR				0.10-1.05	0.057
Negative	19	9	1		
Positive	9	13	0.33		
Ki-67				1.33-25.05	0.032
<20%	3	9	1		
≥20%	25	13	5.77		
Molecular subtype				1.02-10.72	0.042
Non-TNBC	11	15	1		
TNBC	17	7	3.31		

HRD score high: HRD score ≥10, HRD score low: HRD score 10

OR: odds ratio, CI: confidence interval, Pre-: Premenopause, Post-: Postmenopause

Table 5 Association of *BRCA* mutation and HRD status with pCR (RCB 0) from clinical cohort

All patients (n = 50)	pCR n (%)	Non- pCR n (%)	OR	95%CI	Logistic P value
BRCA status					
non-mutated	5 (14)	30 (86)	Reference		
mutated	6 (40)	9 (60)	4.0	0.99-16.24	0.052
HRD status					
Negative	1 (5)	19 (95)	Reference		
Positive	10 (33)	20 (67)	9.5	1.11-81.5	0.04
BRCA wild-type (n = 35)					
HRD negative	1 (5)	19 (95)	Reference		
HRD positive	4 (27)	11 (73)	6.91	0.68-69.86	0.102

HRD positive= HRD score ≥10 or *BRCA* mutation

Table 6 Association of *BRCA* mutation and HRD status with RCB from clinical cohort

All patients (n = 50)	RCB0/1 n (%)	RCB2/3 n (%)	OR	95%CI	Logistic P value
BRCA status					
non-mutated	8 (23)	27 (77)	Reference		
mutated	10 (67)	5 (33)	6.75	1.78-25.58	0.005
HRD status					
Negative	2 (10)	18 (90)	Reference		
Positive	16 (53)	14 (47)	10.29	2.02-52.36	0.005
BRCA wild-type (n = 35)					
HRD negative	2 (10)	18 (90)	Reference		
HRD positive	6 (40)	9 (60)	6.0	1.00-35.91	0.05

HRD positive= HRD score ≥10 or *BRCA* mutation

Figures

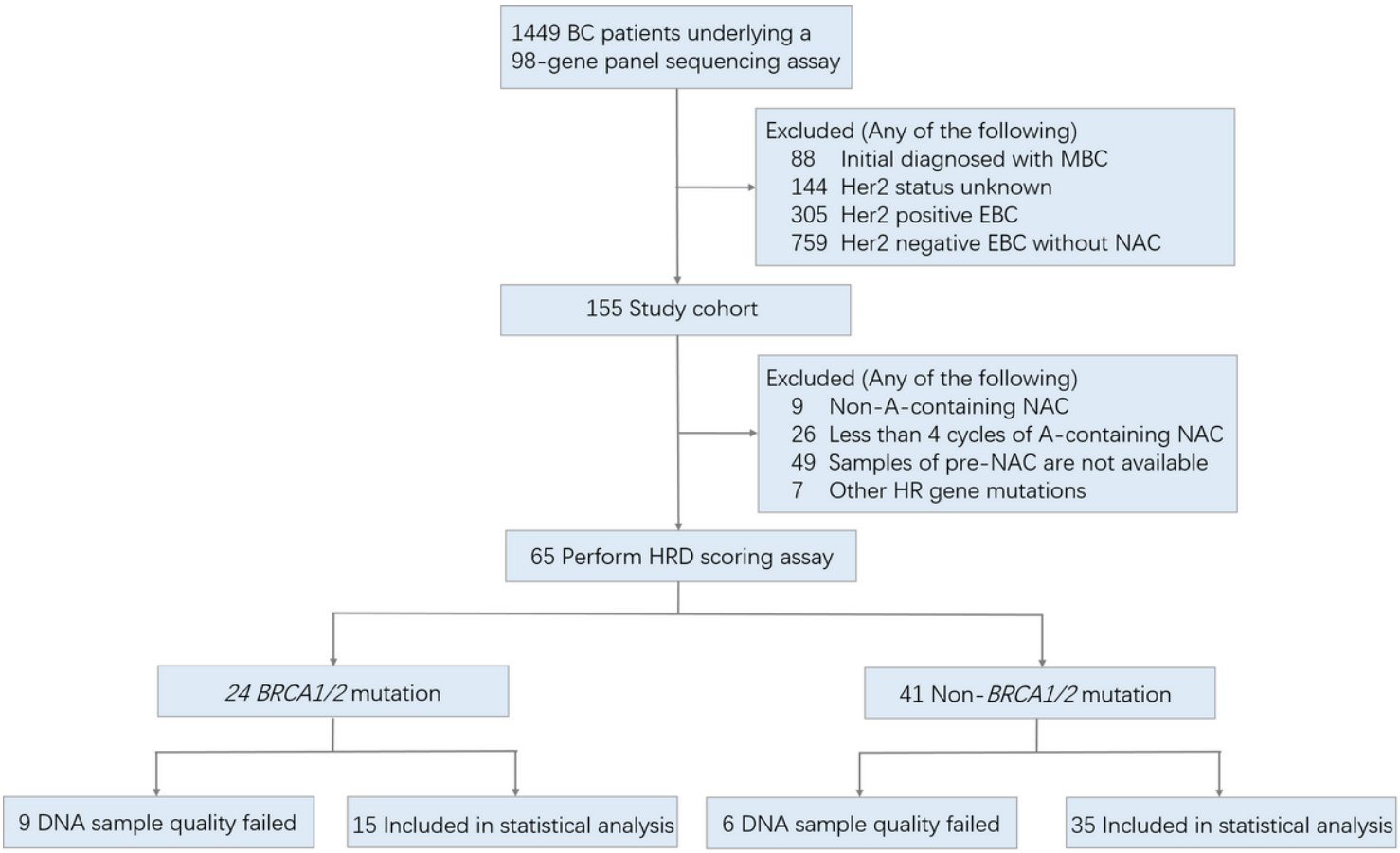

Figure 1

Patient selection criteria. Some patients met more than 1 excluded criteria. BC: breast cancer, MBC: metastatic breast cancer, EBC: early breast cancer, A: anthracycline.

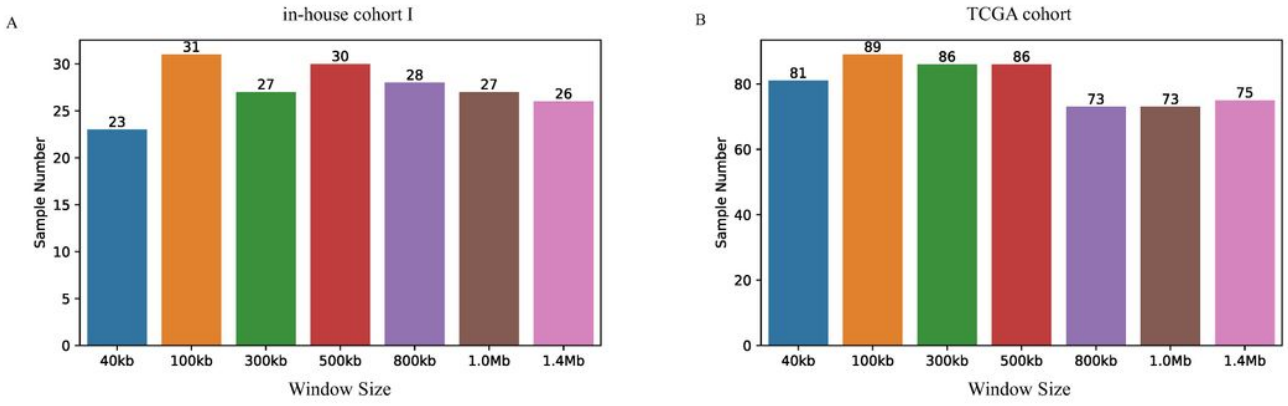

Figure 2

Frequency of the mode in seven different window sizes. In seven window sizes of each sample, the mode of the number of LCNVs is defined as 1 and the rest as 0. Then the frequency of mode samples is

calculated in each window. Horizontal axis represents seven kinds of window size and vertical Coordinates represents the number of samples. (A) Frequency of the mode in seven different window sizes in 41 breast cancer samples from an in-house breast cancer cohort. (B) Frequency of the mode in seven different window sizes in 122 breast cancer samples from TCGA database.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.xls