Application Note

Genomics

Agilent

Trusted Answers

Enabling Homologous Recombination Deficiency (HRD) Assessment Using the Agilent SureSelect Cancer CGP Assay

Authors

Akanksha Khare, MS,
Agilent Technologies, Inc

Abhilasha Cheruku, MS,
Insook Do, Hyunkyung Kong, PhD,
and Elly Lee, Psomagen

Michael Blum, PhD,
Samantha Brunel, PhD,
and Nicolas Philippe, PhD,
SeqOne Genomics

Gregoire Andre, PhD,
Pierluigi Iapicca, PhD,
and Christian Pozzorini, PhD,
SOPHiA GENETICS

Abstract

Evaluating biomarkers is essential for the molecular characterization of tumor samples. Homologous recombination deficiency (HRD) status can serve as a key biomarker for potentially predicting how responsive a tumor might be to poly adenosine diphosphate-ribose polymerase (PARP) inhibitors, which inhibit the repair mechanisms of cancer cells, leading to their death. This application note introduces a method to assess HRD status from solid tumor samples using a comprehensive genomic profiling (CGP) assay, the Agilent SureSelect Cancer CGP assay. The method includes modifications to its laboratory protocol by the addition of a low-pass whole genome sequencing (WGS) step and data analysis is demonstrated by two commercial informatics solutions.

This application note was developed in collaboration with:

psomagen

SeqOne
FROM GENOME TO PERSONALIZED MEDICINE

SOPHiA GENETICS™

Introduction

Major factors that underlie complex tumor biology include genomic instability and DNA repair defects. The failure to efficiently repair double-stranded DNA breaks through the homologous recombination repair (HRR) pathway results in the HRD phenotype. Alterations in HRR pathway genes can render tumors susceptible to PARP inhibitors and platinum-based chemotherapy, which specifically aim to eliminate cancer cells. However, a main challenge is the lack of a standardized method to define, measure, and report the HRD status of solid tumor samples, so that samples from different studies can be compared.

Key classes of somatic alterations can be detected from solid tumor samples by targeted, next-generation sequencing (NGS) solutions, the Agilent SureSelect Cancer assays. For comprehensive genomic profiling, the SureSelect Cancer CGP assay detects single nucleotide variants (SNVs), copy number variants (CNVs), insertions/deletions (indels), and translocations (TLs). It also assesses the immuno-oncology biomarkers, tumor mutational burden (TMB) and microsatellite instability (MSI), as well as RNA gene fusions. The SureSelect Cancer CGP assay uses NGS libraries that are enriched for expertly curated cancer genes, featuring a DNA panel (679 genes) and an RNA panel (80 genes), with flexible and modular workflows (Figure 1).

To augment the biomarker detection capabilities of the SureSelect Cancer CGP assay, we present a modified protocol that incorporates low-pass WGS of unenriched libraries with the DNA assay to enable HRD assessment. Low-pass WGS interrogates the entire human genome with shallow read coverage to measure genome-wide "genomic integrity"* or "genomic instability"*. We demonstrate that HRD assessment results are consistent when analyzed using two commercially available software platforms, provided by SOPHiA GENETICS and SeqOne Genomics.

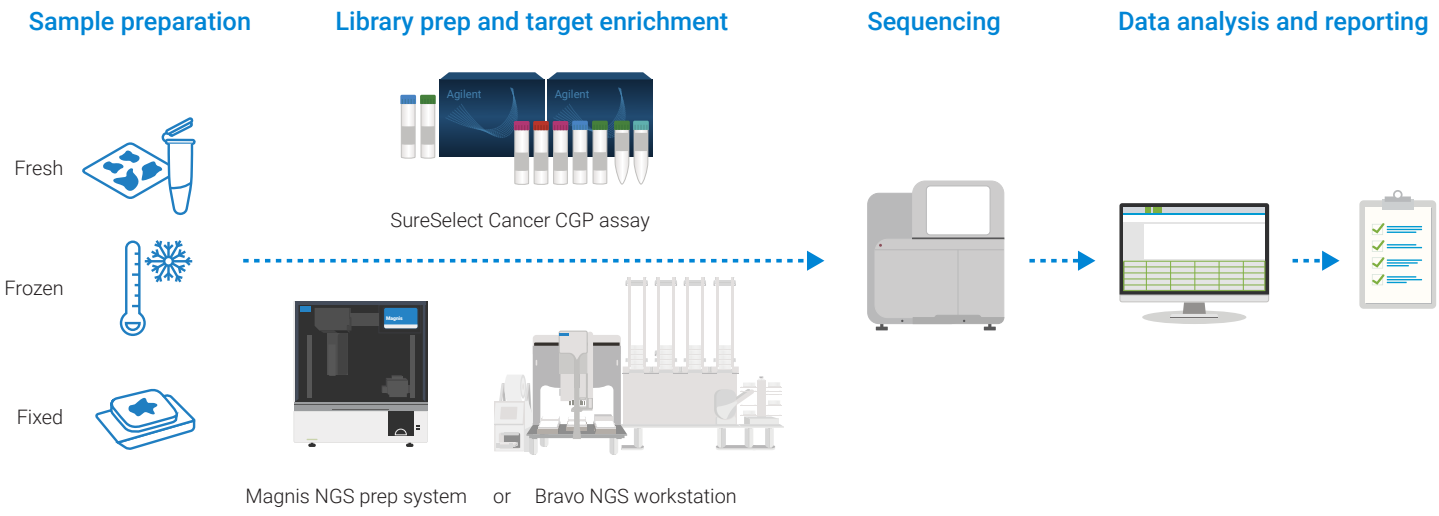

Figure 1. The Agilent SureSelect Cancer CGP assay offers flexible and modular workflows for biomarker detection. Library preparation and target enrichment can be carried out manually, or using automation platforms such as the Agilent Magnis NGS Prep system (a fully automated, walkaway system that requires only 15 minutes hands-on time) or the Agilent Bravo NGS workstation (an open and scalable liquid handling platform). Sequencing can be performed on compatible sequencers from Illumina, Element Biosciences, Pacific Biosciences, and MGI. Data analysis options include customer in-house and third-party software.

** The two distinct terms, "genomic integrity" and "genomic instability", are both abbreviated as "GI", by different software vendors. For SOPHiA GENETICS, "GI" refers to a "genomic integrity" index. For SeqOne Genomics, "GI" refers to a "genomic instability" status. To avoid confusion, this application note spells out the full terms instead of using the abbreviation.*

Materials and methods

The protocol for the SureSelect Cancer CGP assay[1] was modified to enable HRD assessment, as shown in Figure 2 and described as follows.

Figure 2. Modified workflow for the Agilent SureSelect Cancer CGP assay to enable HRD assessment. The standard laboratory protocol for the SureSelect Cancer CGP assay (fragmentation, library generation, target enrichment, and sequencing) follows the blue arrows. The workflow modifications to enable HRD assessment are shown by the green arrow. For each sample, an aliquot of the indexed, unenriched library is retained for HRD assessment through low-pass WGS. This aliquot of the unenriched library is then sequenced separately or pooled with the target enriched library to be sequenced together in a single run. The decision to sequence the libraries separately or pooled together is dependent on the capabilities of the selected data analysis platform and user preference. Sequencing is performed on an Illumina NovaSeq sequencer, and the data is normalized to 40 million reads per sample for CGP enriched libraries and unenriched libraries.

1. **Samples**: For the normal reference sample, 50 ng of Agilent OneSeq Human Reference DNA was used. Other well-characterized normal reference samples (HapMap NA12878 or NA24385) can also serve as normal control samples. To evaluate the HRD assessment method, Seraseq FFPE DNA reference samples with a range of known HRD status were used, each with a 50 ng DNA input (Table 1).

Table 1. Materials used in this study.

Material	Vendor	Part Number	Notes
OneSeq Human Reference DNA, Female	Agilent	5190-8850	Supplied at 200 ng/µL
Seraseq FFPE HRD High-Pos RM	SeraCare	0710-2643	Supplied in ready-to-use format (50 ng)
Seraseq FFPE HRD Low-Pos RM	SeraCare	0710-2645	Supplied in ready-to-use format (50 ng)
Seraseq FFPE HRD Negative RM	SeraCare	0710-2644	Supplied in ready-to-use format (50 ng)
SureSelect Cancer CGP Assay DNA Kit, 16 Reactions SureSelect Cancer CGP Assay DNA Kit, 96 Reactions	Agilent	G9967A G9967B	Includes SureSelect XT HS2 DNA library reagents and DNA enrichment probe for 679 genes. Enzymatic fragmentation reagents not included.
SureSelect XT HS Enzymatic Fragmentation Kit, 16 Reactions SureSelect Enzymatic Fragmentation Kit, 96 Reactions	Agilent	5191-4079 5191-4080	Enzymatic fragmentation reagents

2. **Library preparation**: For each sample, libraries were prepared according to the standard protocol in the SureSelect Cancer CGP assay user guide.[1] In this study, a total of four libraries (duplicate libraries for each of two data analysis methods) were created per sample (Table 1), for a total of 16 libraries. After performing quality control (QC) for each library, the required aliquot (500 to 1000 ng) of the unenriched (pre-capture) library was taken to perform target enrichment (see page 38 of the user guide), and the remaining portion of the unenriched library was retained for low-pass WGS (see step 4).
3. **Target enrichment and pooling of CGP enriched libraries for targeted sequencing**: For each library, target enrichment was performed according to the standard protocol.[1] In this study, 16 CGP-enriched libraries were generated. After target enrichment, all CGP-enriched libraries were pooled to create a pool with a final concentration of 10 nM in 20 µL (see page 47 of the user guide). This pool of CGP-enriched libraries was used for targeted sequencing.
4. **Pooling of unenriched libraries for HRD assessment with low pass WGS**: The eight unenriched libraries retained from step 2 above for each data analysis method were pooled to create a final pool with a concentration of 10 nM in 20 µL[1] (see page 49 of the user guide). This pool of unenriched libraries was used for low-pass WGS.
5. **Sequencing**: The two library pools created in steps 3 and 4 may be sequenced separately or combined and sequenced in a single run, depending on the capabilities of the chosen data analysis platform and user preference. The final pool was sequenced to 144 Gb to yield 40 million reads for the CGP-enriched libraries and 80 million reads for the combined CGP-enriched libraries and unenriched libraries.

Results

Formalin-fixed paraffin-embedded (FFPE) reference samples with known HRD status (high, low, or negative, see Table 1), along with the Agilent OneSeq Human Reference (non-FFPE) DNA, were used to test the addition of HRD assessment to the set of biomarkers detected by the SureSelect Cancer CGP assay. Sequencing results for low-pass WGS (HRD assessment) and targeted sequencing for somatic alteration assessment (data not shown) were analyzed using two leading software platforms, provided by SOPHiA GENETICS and SeqOne. Although each platform employs distinct algorithms and classification methods to determine genomic integrity or genomic instability status, both platforms have undergone rigorous validation against the reference PAOLA-1 cohort.[2] Importantly, each platform has the capability to process data from low-pass WGS and from panel-based assays.

HRD assessment by SOPHiA DDM

The SOPHiA DDM cloud-based infrastructure is flexible and scalable, enabling it to handle larger data volumes, and support new applications and functionalities based on customer needs. Powered by the deep-learning-based "GIInger" algorithm, which was validated on the PAOLA-1 cohort,[3] the SOPHiA platform[4] computes a Genomic Integrity index. This index serves as a primary indicator of genomic instability status, with a score above zero indicating a positive result for Genomic Instability.

The SOPHiA DDM analysis showed that reference samples with known HRD status yielded QA status of High, with four samples (High-Pos, Low-Pos) correctly called as positive Genomic Integrity status, and two samples (Negative) correctly called as negative Genomic Integrity status (Figure 3A). Overall, SOPHiA DDM results demonstrated strong concordance and reproducibility with SeraCare HRD reference samples.

The normalized WGS coverage, as produced by the GIInger algorithm and visualized by SOPHiA DDM, reflects the copy number variation for each sample (Figure 3B). The Agilent OneSeq Human Reference DNA is a germline sample that does not carry large copy number changes, as shown by its WGS coverage profile (Figure 3B). In the context of tumor samples analysis, a flat coverage profile can reflect either the total absence of large copy number changes (HRD-negative) or insufficient sample tumor content. Given the ambiguity of this scenario, for the Agilent OneSeq Human Reference DNA, the Genomic Integrity status was reported as "inconclusive" and QA status as "medium" (Figure 3A).

A

Sample	Sample Quality Assessment						SOPHiA GENETICS Genomic Integrity	
	Number of WGS Fragments [M]	Proportion of Coverage Outliers	Purity Ploidy Ratio	Residual Noise	Signal-to-Noise Ratio (SNR)	Quality Assessment (QA) Status	Genomic Integrity Index	Genomic Integrity Status
Agilent OneSeq Human Reference DNA - 1	14.3	0.0%	--	0.05	0.3	Medium	--	Inconclusive
Agilent OneSeq Human Reference DNA - 2	10.0	0.0%	--	0.06	0.29	Medium	--	Inconclusive
Seraseq FFPE HRD High-Pos - 1	14.2	0.0%	0.23	0.09	3.34	High	9.3	Positive
Seraseq FFPE HRD High-Pos - 2	22.0	0.1%	0.23	0.08	3.57	High	9.5	Positive
Seraseq FFPE HRD Low-Pos - 1	25.4	0.1%	0.18	0.09	3.67	High	5.6	Positive
Seraseq FFPE HRD Low-Pos - 2	21.7	0.1%	0.2	0.09	3.65	High	6.0	Positive
Seraseq FFPE HRD Negative - 1	31.4	0.8%	0.4	0.07	4.48	High	-7.9	Negative
Seraseq FFPE HRD Negative - 2	20.5	0.4%	0.4	0.07	4.15	High	-8.1	Negative

B

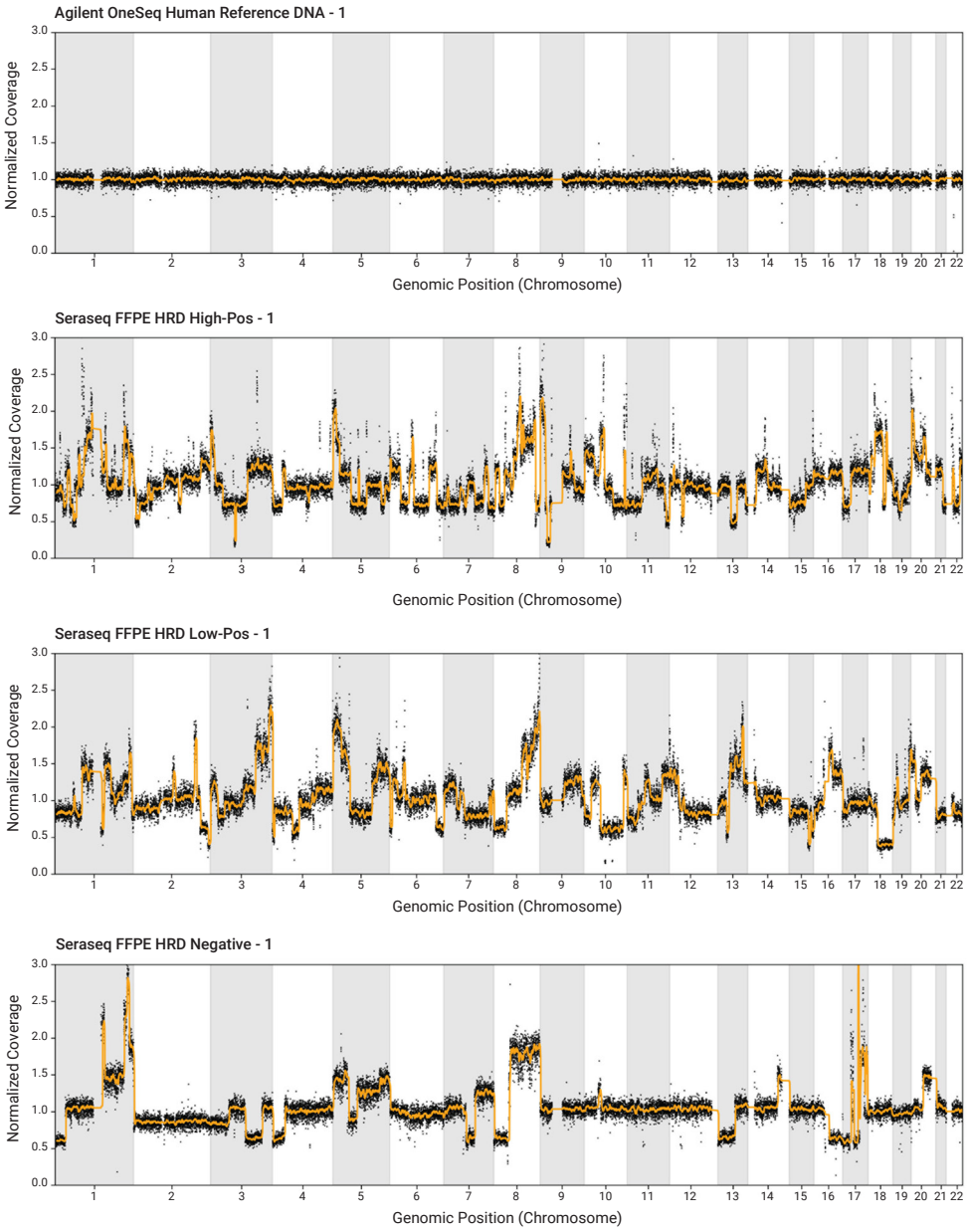

Figure 3. HRD assessment using the SOPHiA DDM software solution shows high concordance and reproducible detection of HRD status for Seraseq HRD reference samples. A. Seraseq FFPE HRD reference materials comprising of High-Positive, Low-Positive, and Negative samples, along with the Agilent OneSeq Reference (non-FFPE) DNA, were tested in replicates using the modified Agilent SureSelect Cancer CGP assay protocol, as described in the materials and methods. The normalized WGS coverage profile undergoes **Quality Assessment** (QA), in which the metrics are computed and one of the following QA statuses is assigned to the sample - High quality: the quality of the data is sufficient to confidently compute a **Genomic Integrity Index** and a **Genomic Integrity status**; Medium quality: the quality of the data is lower (compared to high quality) and, as a consequence, the deep learning algorithm may not succeed in computing a **Genomic Integrity Index**. Quality Assessment is based on the combination of purity ploidy ratio, residual noise, and signal-to-noise ratio. **Number of WGS fragments**: the total number of properly mapped DNA fragments available for the raw coverage WGS profile calculation. If the number of WGS fragments is smaller than 4 million, the QA status is deemed low. **Proportion of Coverage Outliers**: percentage of WGS regions considered for **Genomic Integrity** analysis which feature an artifactual and excessive localized coverage which is compensated by the coverage normalization algorithm. If the proportion of coverage outliers is larger than 20%, the QA status is deemed low. **Purity/ploidy ratio**: the ratio between sample tumor content and sample ploidy, estimated by measuring the strength of the signal induced in the normalized WGS coverage profile by a copy number change. If the purity/ploidy ratio estimation fails or if the estimated value is lower than 0.1 (suggesting insufficient tumor content), the QA status is deemed medium. **Residual noise**: is computed by measuring the standard deviation of the normalized WGS coverage profile with respect to the smoothed WGS coverage profile. If residual noise is larger than 0.17, the QA status is deemed low. **SNR**: strength of the signal induced in the normalized WGS coverage profile by all copy number aberrations present in the sample divided by the residual noise. Samples with SNR smaller than 0.55 will be classified as medium quality. SNR is also considered by the algorithm to assign the **Genomic Integrity** Negative* status. **B**. Normalized WGS coverage profile for the Agilent OneSeq Human Reference DNA sample and each Seraseq HRD reference control across genomic position (chromosome) computed by GIInger and used to compute the Genomic Integrity Index and status. White and gray shading denotes chromosome boundaries.

This design strategy mitigates the risk of False Negatives in samples with insufficient tumor content.

HRD assessment by SeqOne

SeqOne provides a cloud-based platform for comprehensive, scalable bioinformatic solutions for genomic analysis, supporting somatic and germline testing. The SeqOne SomaHRD pipeline combines machine learning algorithms with advanced genomic instability metrics for precise HRD assessment. Validated on the PAOLA-1 cohort[5], the SeqOne platform efficiently processes low-pass WGS data, enabling highly cost-effective and accurate HRD detection.[6]

The SeqOne platform offers transparency[7] into the calculation of genomic instability scores with Large Genomic Alteration (LGA) and Loss of Parental Copy (LPC) (Figure 4), as well as copy number alterations (CNA) for *CCNE1* and *RAD51B* (Figure 4A), whose amplifications are linked to poor responses to PARP inhibitors. The SomaHRD pipeline generates HRD probability scores from 0 to 100%, with scores of >50% indicating HRD-positive samples, while scores of <50% indicating HRD-negative status. HRD probabilities are calculated using the genomic instability statistics (LGA, LPC). Gene amplification values are included as supplementary information only. The platform also includes comprehensive quality metrics and coverage profiles with detected LGA events for each sample (Figure 4B), allowing confident interpretation of results.

Using the reference materials with known HRD status, the SeqOne platform correctly identified all HRD-positive samples with scores between 78 and 82%, and all HRD-negative samples with scores below 11% (Figure 4A). For the Agilent OneSeq Human Reference DNA sample, a "warning" was noted for the low tumor fraction, accurately reflecting that these were germline samples without tumoral content. The final decision between a negative or inconclusive result is left to the analyst. This clear distinction and transparent data presentation improves the reliability of HRD status determination. The SomaHRD pipeline showed excellent reproducibility with less than 3% variation between replicate samples. In summary, the SeqOne HRD platform demonstrated highly concordant results with HRD reference samples, robust reproducibility, and transparency in its analysis process.

A

Sample	Genomic Instability Features		Amplifications		Quality			SeqOne HRD Status	
	LGA	LPC	*CCNE1* copies	*RAD51B* copies	sWGS coverage	%correct mapping	Low tumor fraction	HRD status	HRD+ probability
Agilent OneSeq Human Reference DNA – 1	0	0	4.27	2.28	2.38	99.9	Warning	Negative	1%
Agilent OneSeq Human Reference DNA – 2	1	0	5.12	2.42	2.05	99.99	Warning	Negative	2%
Seraseq FFPE HRD High-Pos – 1	18	24	1.67	1.91	2.05	99.8	No	Positive	82%
Seraseq FFPE HRD High-Pos – 2	18	25	1.71	1.94	2.33	99.7	No	Positive	82%
Seraseq FFPE HRD Low-Pos - 1	23	10	1.96	2.4	1.97	99.7	No	Positive	79%
Seraseq FFPE HRD Low-Pos - 2	23	10	2.02	2.43	2.09	99.7	No	Positive	78%
Seraseq FFPE HRD Negative - 1	6	9	1.94	2.14	2.97	99.8	No	Negative	8%
Seraseq FFPE HRD Negative - 2	7	9	1.8	2.1	2.06	99.7	No	Negative	11%

B

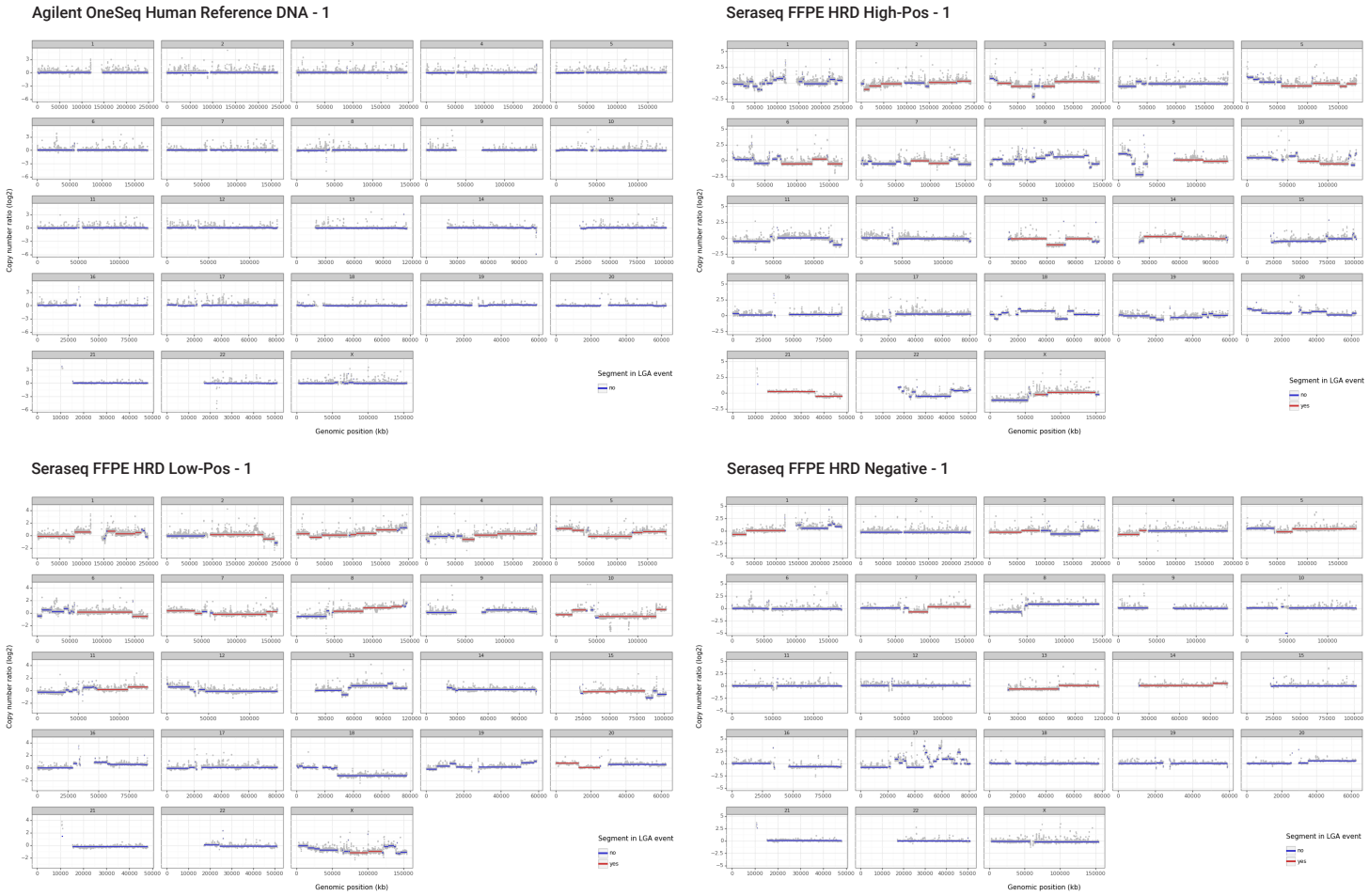

Figure 4. HRD score validation using the SeqOne platform demonstrates high concordance and reproducible detection of HRD status for Seraseq HRD reference materials. A. Seraseq FFPE HRD reference materials comprising of High-Positive, Low-Positive, and Negative samples, along with the Agilent OneSeq Reference (non-FFPE) DNA, were tested in replicates using the modified Agilent SureSelect Cancer CGP assay protocol. Genomic Instability is captured with two scores: **Large Genomic Alteration** (LGA, defined as the number of copy number breakpoints, where a breakpoint represents a change in copy number between two genomic segments of at least 10 Mb long and at most 3 Mb apart) and **Loss of Parental Copy** (LPC, the number of haploid segments of at least 10 Mb long). Gene amplification values for *CCNE1* and *RAD51B* are included as supplementary information only and are not used to compute HRD probabilities. Exact concordance was observed between the HRD status determined by SeqOne and the SeraCare reference classifications across all sample types. The average genomic coverage is indicated on the table as **shallow WGS (sWGS) coverage**. The results underscore the robust performance of the assay in accurately identifying HRD status, with minimal variation between replicates. The high level of agreement across technical replicates further highlights the platform's reliability and reproducibility, confirming its suitability for determining HRD status. **B**. Chromosomal coverage profiles are displayed across all chromosomes for the Agilent OneSeq Reference DNA sample and each Seraseq HRD reference control sample. The red segments in each plot indicate the genomic regions flanking the breakpoints that contribute to the calculation of LGA. In contrast, the blue segments represent regions that do not contribute to the LGA calculation.

Conclusion

This application note describes a robust method for assessing HRD status based on a modified protocol of the Agilent SureSelect Cancer CGP assay and low-pass WGS. High concordance was observed between the data analyzed by two software platforms, from SOPHiA GENETICS and SeqOne, demonstrating the robustness and reliability of this method. The low-pass WGS for HRD assessment method could also be extended for use with tissue-specific assays, the (Agilent SureSelect Cancer Tumor-Specific assays, Agilent SureSelect CD HRR17 panel[5]) and custom assays (Agilent SureSelect Cancer Custom panels), for greater flexibility.

The HRD assessment method described here can be used as part of an expanded approach for biomarker detection to help decipher the genomic complexity of tumors. Ultimately, this method can enable an enhanced understanding of tumor biology and the mechanisms underlying genomic instability.

References

1. SureSelect Cancer CGP Assay User Guide. Agilent Technologies User Manual, 2024, publication G9966-90000.
2. Ray-Coquard, I.; Pautier, P.; Pignata, S.; Pérol, D.; González-Martín, A.; Berger, R.; Fujiwara, K.; Vergote, I.; Colombo, N.; Mäenpää, J.; et al. Olaparib plus Bevacizumab as First-Line Maintenance in Ovarian Cancer. *New England Journal of Medicine* [Online] **2019**, *381* (25), 2416–2428. https://doi.org/10.1056/NEJMoa1911361.
3. Pozzorini, C.; Andre, G.; Coletta, T.; Buisson, A.; Bieler, J.; Ferrer, L.; Kempfer, R.; Saintigny, P.; Harlé, A.; Vacirca, D.; et al. GIInger Predicts Homologous Recombination Deficiency and Patient Response to PARPi Treatment from Shallow Genomic Profiles. *Cell Reports Medicine* [Online] **2023**, 4 (12), 101344. https://doi.org/10.1016/j.xcrm.2023.101344.
4. SOPHiA DDM™ Homologous Recombination Deficiency (HRD) Solution. SOPHiA GENETICS. https://www.sophiagenetics.com/clinical/oncology/solid-tumors/homologous-recombination-deficiency-solutions/.
5. Boidot, R.; Blum, M.G.B.; Wissler, M.-P.; Céline Gottin; Ruzicka, J.; Chevrier, S.; Delhomme, T. M.; Audoux, J.; Jeanniard, A.; Just, P.-A.; et al. Clinical Evaluation of a Low-Coverage Whole-Genome Test for Detecting Homologous Recombination Deficiency in Ovarian Cancer. *European Journal of Cancer* [Online] **2024**, *202*, 113978–113978. https://doi.org/10.1016/j.ejca.2024.113978.
6. How to perform fast and accurate HRD testing while mitigating costs. SeqOne. https://seqone.com/hrd/.
7. Bouvier, C. Case Study: North West GLH Validation and Implementation of In-House HRD testing with SeqOne's SomaHRD. SeqOne. https://seqone.com/news/case-study-north-west-glh-validation-and-implementation-of-in-house-hrd-testing-with-seqones-somahrd/.

www.agilent.com

Published in the USA, January 31, 2025
5994-7867EN

Agilent
Trusted Answers