INSTRUCTIONS FOR USE

16, 32 and 48 Samples

SOPHiA DDM™

Homologous Recombination Deficiency Solution

Using the SOPHiA GENETICS™ Universal Library Prep

SOPHiA DDM™

SUMMARY INFORMATION

Product Name	SOPHiA DDM™ Homologous Recombination Deficiency Solution Using the SOPHiA GENETICS™ Universal Library Prep
Product Type	Bundle Solution
Product Family	Molecular kit + analytics
Product Version	v1.0
Sample Type	Somatic DNA isolated from formalin-fixed paraffin-embedded (FFPE) tumor tissue specimens
Document ID	SG-06843
Document Version	v2.1
Revision Date	Jun 2026

This Instructions For Use is applicable to all SOPHiA DDM™ versions.

Please read the Instructions For Use thoroughly before using this product.

SOPHiA GENETICS SA
ZA La Pièce 12
1180 Rolle
Switzerland

REF

BS0127ILLRSMY10-16;
BS0127ILLRSMY10-32;
BS0127ILLRSMY10-48

RUO

PRODUCT CODES

	FULL PRODUCT CODE	BOX 1	BOX 2	LIBRARY PREPARATION KIT
REF	BS0127ILLRSMY10-16	B1.U1.0027.R-16	B2.0000.R-16	400232
	BS0127ILLRSMY10-32	B1.U1.0027.R-32	B2.0000.R-32	400232
	BS0127ILLRSMY10-48	B1.U1.0027.R-48	B2.0000.R-32	400234

DISCLAIMER

This document and its contents are the property of SOPHiA GENETICS SA and its affiliates ("SOPHiA GENETICS") and are intended solely for the contractual use by its customer in connection with the use of the product(s) described herein and for no other purpose. This document and its contents shall not be used or distributed for any other purpose and/or otherwise communicated, disclosed, or reproduced, or referenced to in any way whatsoever without the prior written consent of SOPHiA GENETICS.

SOPHiA GENETICS does not convey any license under its patent, trademark, copyright, or common-law rights nor similar rights of any third parties by this document. The instructions in this document must be strictly and explicitly followed by qualified and adequately trained personnel to ensure the proper and safe use of the product(s) described herein. All of the contents of this document must be fully read and understood before using such product(s).

FAILURE TO COMPLETELY READ AND EXPLICITLY FOLLOW ALL OF THE INSTRUCTIONS CONTAINED HEREIN MAY RESULT IN DAMAGE TO THE PRODUCT(S), INJURY TO PERSONS, INCLUDING TO USERS OR OTHERS, AND DAMAGE TO OTHER PROPERTY.

SOPHiA GENETICS DOES NOT ASSUME ANY LIABILITY ARISING OUT OF THE IMPROPER USE OF THE PRODUCT(S) DESCRIBED HEREIN (INCLUDING PARTS THEREOF OR SOFTWARE).

TRADEMARKS

Illumina® and NextSeq® are registered trademarks of Illumina, Inc.

NovaSeq™ is a trademark of Illumina, Inc.

SOPHiA GENETICS™ and SOPHiA DDM™ are trademarks of SOPHiA GENETICS SA and/or its affiliate(s) in the U.S. and/or other countries. All other names, logos, and other trademarks are the property of their respective owners. UNLESS SPECIFICALLY IDENTIFIED AS SUCH, SOPHiA GENETICS USE OF THIRD-PARTY TRADEMARKS DOES NOT INDICATE ANY RELATIONSHIP, SPONSORSHIP, OR ENDORSEMENT BETWEEN SOPHiA GENETICS AND THE OWNERS OF THESE TRADEMARKS. Any references by SOPHiA GENETICS to third party trademarks are to identify the corresponding third-party goods and/or services and shall be considered nominative fair use under the trademark law.

REVISION HISTORY

DOCUMENT ID / VERSION	DATE	DESCRIPTION OF CHANGE
SG-06843 – v2.1	Jun 2026	• Minor text modifications across the document • Replaced section 6.3.4 - **Combined Gene-level and Exon-Level CNV Analysis** with **DNA analysis: Gene-level And Exon-level CNV Detection Via MUSKAT™**. For full technical details of this module, refer to **SG-08061**. • *10 Support:* : Contact information has been updated.
SG-06843 – v2.0	Dec 2024	• *7.1.4 Review And Edit BRCA Status*: Added warning regarding the inclusion of CNVs in the BRCA variant report • *7.1.6 Define HRR Status*: Added warning regarding the inclusion of non-BRCA CNVs in the HRR status report • Replaced section *7.3.4 Included Analysis Modules* - ***Gene Amplification/CNVs*** with new section *7.3.4 Included Analysis Modules* - ***Combined Gene-level and Exon-Level CNV Analysis*** • *7.3.4 Included Analysis Modules - Proposed BRCA status*: Added warning regarding BRCA status computation • *8.1 SOPHiA DDM™ Results Location* and *8.2 SOPHiA DDM™ Downloadable Files*: Updated screenshots and extended list of downloadable files • *8.3 SOPHiA DDM™ Results Visualization*: Added warning regarding the download of CNV results • *9.2.2 Module-Specific Limitations*: Updated limitations for CNV detection module • Added new section *12.1 Appendix I: Genomic Coordinates of the CNV "Target Regions" of the HRD Solution* • Minor formatting changes and typo corrections across the document
SG-06843 – v1.1	Aug 2024	• *7.2.1 Data Generation*: Added Illumina® NovaSeq™ X as a compatible sequencer
SG-06843 – v1.0	Jun 2024	• Initial release

TABLE OF CONTENTS

1 Product Introduction **1**

2 General Statement of the Test Principles and Procedure **2**

3 Product Components **3**

4 Library Preparation and Sequencing **4**

5 Data Upload **5**

5.1 System Requirements and Data Management Procedures *5*

5.1.1 Dependencies 5

5.1.2 Usage of USB Keys 5

5.1.3 Network Drives and Shared Folders 5

5.2 Naming Conventions *6*

5.3 Illumina® File Specifications *7*

5.4 Pipeline Selection in KIT Column *8*

5.5 Illumina® Folder Configuration *9*

5.6 Sample Upload Requirements *10*

6 Analysis **11**

6.1 Analysis Workflow *11*

6.1.1 Workflow 11

6.1.2 Start an HRD Interpretation 12

6.1.3 Review and Edit HRD Status 12

6.1.4 Review and Edit BRCA Status 12

6.1.5 Review and Edit Genomic Integrity Status (GIS) 13

6.1.6 Define HRR Status 13

6.1.7 Define CCNE1 Gene Amplification Status 14

6.1.8 Report additional variants not in the scope of BRCA or HRR status analysis 14

6.1.9 Generate HRD Report 14

6.1.10 Considerations Regarding Inconclusive Results 14

6.1.11 Considerations Regarding GI Negative* Statuses 17

6.2 Analysis Prerequisites *21*

6.2.1 Data Generation 21

6.2.2 Run Composition 21

6.2.3 NGS Data Demultiplexing Instructions 21

6.3 Analysis Description and Parameters *22*

6.3.1 Resource Files 22

6.3.2 Target Regions ... 22
6.3.3 Raw Data Pre-Processing ... 23
6.3.4 Included Analysis Modules ... 24
7 Pipeline Deliverables ... 39
7.1 SOPHiA DDM™ Downloadable Files ... 39
7.1.1 Run Level ... 39
7.1.2 Sample Level ... 39
7.2 Variant Annotation ... 40
7.2.1 Description of FVT content ... 40
8 Warnings, Limitations, and Precautions ... 43
8.1 Warnings ... 43
8.1.1 General Warnings ... 43
8.2 Limitations ... 44
8.2.1 General Limitations ... 44
8.2.2 Module-Specific Limitations ... 44
9 Symbols ... 47
10 Support ... 48
11 Appendices ... 49
11.1 Appendix I: Genomic Coordinates of CNV "Target Regions" of the HRD Solution ... 49

1 PRODUCT INTRODUCTION

The intended use of the SOPHiA DDM™ Homologous Recombination Deficiency Solution is to facilitate research in solid tumor samples from Ovarian Cancers. Users will be able to detect Homologous Recombination Deficiency (HRD) through both: i) the direct observation of genomic aberrations, which are known or suspected causes of HRD, ii) the indirect detection of HRD, via the observation of genomic integrity loss.

The product includes a Next Generation Sequencing (NGS) kit and protocol to allow the user to perform, using a single workflow, targeted sequencing, and low-pass Whole Genome Sequencing (lpWGS) on DNA extracted from FFPE tissue.

The product includes a bioinformatics pipeline for: i) the detection of somatic and germline mutations in *BRCA1*, *BRCA2*, as well as in other genes involved in Homologous Recombination Repair (HRR), ii) the calculation of a genomic integrity (GI) index and determination of a GI status, known to be impaired in HRD samples, and iii) the detection of gene amplification events. After performing tertiary annotation, the bioinformatics pipeline computes and proposes to the end-user a sample BRCA status, GI status and HRD status.

The product includes SOPHiA DDM™ Platform to allow the user to visualize the bioinformatics pipeline results, review and, if needed, overwrite, the proposed BRCA, GI and HRD statuses. SOPHiA DDM™ also allows the user to assess and report the status of non-*BRCA* genes involved in HRR as well as the *CCNE1* gene amplification status, which have been statistically associated with HRD. These two additional biomarkers have been shown in the literature to be associated with HRD. For these two biomarkers, the bioinformatics pipeline will not compute and propose a status.

The SOPHiA DDM™ HRD product shall be used with FFPE samples from Ovarian cancers and in combination with the Illumina® NextSeq® 500/550 sequencer. The product shall be used through the SOPHiA DDM™ Platform to enable data upload, result visualization, as well as reporting of the sample HRD status.

2 GENERAL STATEMENT OF THE TEST PRINCIPLES AND PROCEDURE

The SOPHiA DDM™ Homologous Recombination Deficiency Solution product is intended for the preparation of DNA libraries to assess the HRD status of an ovarian tumor based on DNA extracted from FFPE tissue. The NGS kit should allow users to perform the following to determine HRD status in their own laboratory with a single workflow:

- Targeted sequencing (TS) of *BRCA1*, *BRCA2*, as well as other genes involved in HRR or linked to HRD.
- Low-pass Whole Genome Sequencing (lpWGS).

The analytical algorithms will:

- Process TS data to identify genomic aberrations which are known or suspected to cause HRD.
- Process lpWGS data via a deep-learning algorithm capable of quantifying genomic integrity.

3 PRODUCT COMPONENTS

The product consists of three major components: DNA library preparation and capture kit, the bioinformatics pipeline, and the SOPHiA DDM™ Platform.

The purpose of the DNA library preparation and capture kit is to allow the preparation of libraries from FFPE samples suitable for sequencing on an Illumina® sequencing platform.

The purpose of the bioinformatics pipeline is to analyze the sequencing data obtained from an Illumina® NextSeq® 500/550 sequencer and return variant calls, gene amplification calls, genomic integrity analysis results (i.e., secondary analysis), as well as proposed GI, BRCA, and HRD statuses, while the purpose of the SOPHiA DDM™ Platform is to host the pipeline and serve as the interface for the upload of sequencing data and tertiary analysis.

Sequencing data for analysis with the workflow should be generated starting with tumor FFPE samples by following the instructions for library preparation using the SOPHiA GENETICS™ Universal Library Prep and HRD capture kit and sequencing these samples on an Illumina® NextSeq® 500/550 sequencer.

4 LIBRARY PREPARATION and SEQUENCING

The SOPHiA DDM™ Homologous Recombination Deficiency Solution Product relies on the SOPHiA GENETICS™ Universal Library Prep protocol for library preparation and sequencing. Please see *3 Product Components* for more details.

For instructions on library preparation and sequencing, refer to the appropriate Instructions of Use (IFU), which accompanies the SOPHiA DDM™ HRD product. Depending on the library preparation setup (i.e., manual or automated on a liquid handling robot), this is one of the following documents:

- SG-03901 – Instructions for Use for SOPHiA GENETICS™ Universal Library Prep (*manual*)
- SG-03932 – Instructions for Use for SOPHiA GENETICS™ Universal Library Prep (*automated on Hamilton® NGS STAR*)
- SG-03931 – Instructions for Use for SOPHiA GENETICS™ Universal Library Prep (*automated on Hamilton® Clinical STARlet*)

The recommended number of post-capture amplification PCR cycles, sequencing depths per sample, and loading dilutions for various sequencers can be found in Appendix IV of the SOPHiA GENETICS™ Universal Library Prep IFU.

SOPHiA DDM™ Homologous Recombination Deficiency Solution requires combined sequencing of WGS libraries and captured libraries. Instructions can be found in Appendix V of the SOPHiA GENETICS™ Universal Library Prep IFU.

5 DATA UPLOAD

5.1 System Requirements And Data Management Procedures

5.1.1 Dependencies

1. A run is associated with the following:

    a) Computer

    b) User Account

    c) The SOPHiA DDM™ user that starts it.

2. The upload will be terminated if any of the following occurs during the upload:

    - The computer or laptop is shut down or the laptop lid is closed.
    - The user logs out of their account.
    - Files are removed from their original location before the upload is completed.

3. When terminated, SOPHiA DDM™ must be restarted to force the upload to start again. This must be carried out on the same computer, using the same User Account and the same SOPHiA DDM™ user that started the run.

5.1.2 Usage Of USB Keys

- When using a USB key to transfer data, it is recommended to copy the files onto the computer first and then start the run using the local copy.
- The USB key should not be removed before the copying is complete.
- To remove the key, use the "Safely Remove Hardware" option on Windows or "Eject" on Mac. This is also valid when making the copy from the sequencer.

5.1.3 Network Drives And Shared Folders

- It is recommended to copy the data from the network drive to the local computer first and subsequently upload using the local copy to prevent issues due to limited network bandwidth.
- When using a network drive, it is important not to disconnect from the local network before the run upload is completed.

5.2 Naming Conventions

The SOPHiA DDM™ Platform uses the Illumina® naming convention and folder configuration to find the FASTQ files. Therefore, it is strongly advised **NOT** to reorganize or rename the FASTQ files after they are copied from the sequencer, otherwise the run might fail.

5.3 Illumina® File Specifications

- A minimum of two files (R1 and R2) per analysis are required to be present in the same folder.
- More than two files per sample are acceptable, provided the files are always in pairs.
- Analysis files should use the standard Illumina® naming convention (SampleIdentifier1_S1_L001_R1_001.fastq.gz).
- All analysis folders in one run should be in the same folder – the main folder.
- Sample identifier name “SampleIdentifier1” should not contain an underscore sign “_”, a point “.” or any special characters.

5.4 Pipeline Selection in KIT Column

When prompted to specify/confirm the kit that was used for data generation, select SOPHiA DDM™ Homologous Recombination Deficiency Solution

All samples in the new analysis request must have been processed within the same capture and sequencing run and prepared with the same assay reagents.

Before finalizing the upload request, please ensure that the correct kit – "SOPHiA DDM Homologous Recombination Deficiency Solution [HRD_v2]" – is selected for all samples, in case the automatic kit selection may not work optimally. See the SOPHiA DDM™ operational manual (section "*Create a New Request*") for more information about the kit selection.

After completion of the analysis, users will receive a notification by email. If a notification is not received within 24 h from the initiation of the data upload process, please contact support.

5.5 Illumina® Folder Configuration

Before data upload, FASTQ files should be organized in a folder structure, as shown in the following example.

MainFolderRun1	
FolderSample1	FolderSample2
SampleIdentifier1_S11_L001_R1_001.fastq.gz	SampleIdentifier2_S15_L001_R1_001.fastq.gz
SampleIdentifier1_S11_L001_R2_001.fastq.gz	SampleIdentifier2_S15_L001_R2_001.fastq.gz
SampleIdentifier1_S11_L002_R1_001.fastq.gz	SampleIdentifier2_S15_L002_R1_001.fastq.gz
SampleIdentifier1_S11_L002_R2_001.fastq.gz	SampleIdentifier2_S15_L002_R2_001.fastq.gz

5.6 Sample Upload Requirements

The product was designed to upload and process the volume of data produced by an Illumina® NextSeq® 500/550 sequencer (Mid- and High-Output flow cells).

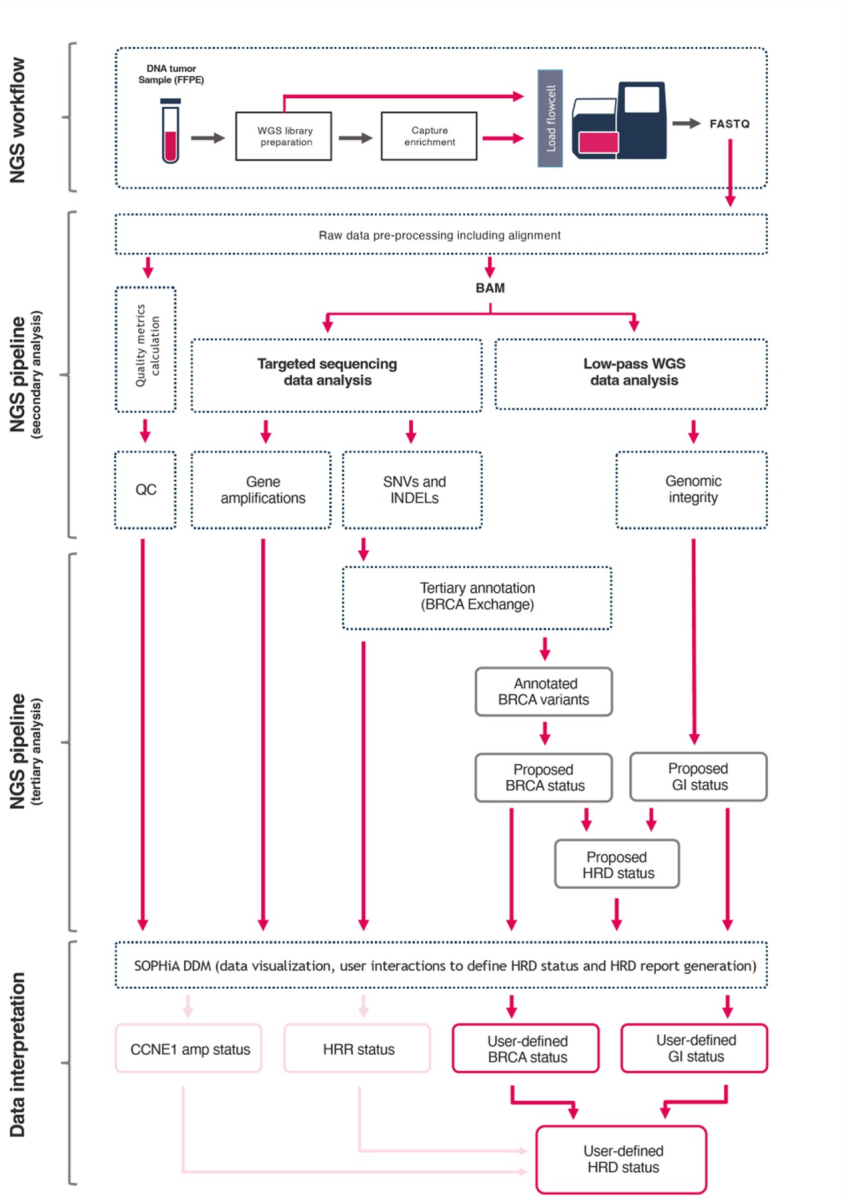
The maximal run size that can be uploaded is 200 GB. Uploading higher volumes in a single batch is not allowed. The pipeline has been shown to successfully process individual samples with up to 50 GB of data. Higher volumes per sample might cause the pipeline to fail.

6 ANALYSIS

6.1 Analysis Workflow

6.1.1 Workflow

The following schematic illustrates the workflow from the DNA samples isolated from FFPE material through the sequencing steps and finally into the proposed HRD status in SOPHiA DDM™.



6.1.2 Start An HRD Interpretation

To start an HRD interpretation:

- For SOPHiA DDM™ Desktop app: click on the HRD status button (lower-left corner of the Overview tab).
- For SOPHiA DDM™ Gen2 platform: go to the biomarker tab.

6.1.3 Review And Edit HRD Status

The following information is shown:

- The final HRD status. By default, the final status is based on the HRD status computed by the pipeline.
- The HRD status computed by the pipeline following the rules described in *Table 10* on page 31.
- The HRD-supporting Genomic Integrity (GII), BRCA, HRR, and CCNE1 status. By default, only the GII and BRCA statuses are selected to be reported together with HRD.

While the final HRD status is computed automatically by the pipeline, it can be manually edited by the user if needed.

6.1.4 Review And Edit BRCA Status

The following information is shown:

- The final BRCA status. By default, the final status is based on the BRCA status computed by the pipeline (see *Table 1* below.).
- A BRCA status computed by the pipeline.
- The list of BRCA variants supporting the BRCA status computed by the pipeline (only high confidence variants with positive, predicted positive, complex case or undetermined BRCA variant score will be prepopulated; these variants are reported if the BRCA status is selected for report).

The SOPHiA DDM™Gen2 platform allows users to visualize BRCA variant annotations with respect to multiple transcripts. However, the sample BRCA status computed by the pipeline is only computed based on the following pre-defined transcripts: BRCA1 (NM_000059.3) and BRCA2 (NM_007294.3).

Table 1: Translation of the BRCA status computed by the pipeline into the default final BRCA status.

BRCA STATUS COMPUTED BY THE PIPELINE	FINAL BRCA STATUS
Positive	Positive
Predicted Positive	Positive
Complex Case	Undetermined
Undetermined	Undetermined
Inconclusive	Undetermined
Predicted Negative	Negative
Negative	Negative

The following warnings can be displayed:

1. **False negative risk:** Less than 99% of the BRCA1, BRCA2 regions targeted by the panel are covered by >100x unique molecules. In this case, users should check low coverage regions to judge if the data quality is sufficient to confidently report a final BRCA status. Users can quickly check the % of BRCA1, BRCA2 regions covered by >100x unique molecules through the dedicated quality indicator (see *6.1.4* on the previous page). For more info about the quality indicator "BRCA COV PERCENTILE 100"
2. **Review clinical interpretation:** High confidence variants with BRCA score = complex case or undetermined were detected. In this case, users should look for them and, if they want, flag them as in scope for BRCA analysis.
3. **Review low confidence variants:** Some relevant BRCA variants have been detected by the pipeline with a low confidence warning. In this case, users may want to check those and bring them up for BRCA interpretation.

The final BRCA status can be edited by the user by adding or removing relevant variants through the SOPHiA DDM™ variant reporting functionality."

The variant reporting functionality for BRCA variants only extends to adding or removing SNPs and INDELs. To report variants of other types (e.g., CNVs), the "Comment" field must be used.

6.1.5 Review And Edit Genomic Integrity Status (GIS)

The following information is shown:

- The final GI status. By default, the final status is based on the proposed GI status following the rules described in *Table 2* below.
- The GI status computed by the pipeline for Ovarian Cancer (see *6.3.4 Included Analysis Modules*, *Proposed HRD Status*).
- The Genomic integrity index (see *6.3.4 Included Analysis Modules*, *Proposed HRD Status*).
- The Genomic integrity QA status (see *6.3.4 Included Analysis Modules*, *Proposed HRD Status*).
- The low pass WGS (lpWGS) coverage profile, showing the data used to compute the GI index.

Table 2: Translation of the GI status computed by the pipeline into the default final GI status

GI STATUS COMPUTED BY THE PIPELINE	FINAL GI STATUS
Positive	Positive
Negative	Negative
Negative*	Negative
Inconclusive	Undetermined
Rejected	Undetermined

6.1.6 Define HRR Status

This functionality allows users to report non-BRCA variants that they consider relevant for HRD assessment (e.g., deleterious variant in PALB2).

The following information is shown:

- The final HRR status. By default, the value is set to “Undetermined”; the pipeline does not compute a HRR gene status.
- The list of HRR status supporting variants. This list contains all variants reported by the user within the scope of HRR gene status determination. By default, this list is empty as the system does not pre-populate HRR variants.

The final HRR status can be edited by the user. Users can add or remove variants in scope of HRR status using theSOPHiA DDM™ variant reporting functionality.

The variant reporting functionality for reporting HRR variants only extends to adding or removing SNPs and INDELs. To report variants of other types (e.g., CNVs), the “Comment” field must be used.

6.1.7 Define CCNE1 Gene Amplification Status

This functionality allows users to report the presence of CCNE1 amplifications as supporting evidence for HRD negativity. The information shown is inherited from the CNV analysis results.

The information shown includes:

- The final CCNE1 status defined by the user. By default, the value is set to “Undetermined”.
- The CCNE1 copy number detected by the CNV module (see section *1 Combined Gene-Level and Exon-Level CNV Analysis*; also see section 7 of the SOPHiA DDM™ Platform Operation Manual).
- For samples with a CN > 3.25 (i.e., if a gene amplification is reported), the user can further interpret the magnitude of the CCNE1 copy number gain by visually inspecting the lpWGS CCNE1 copy number profile plot (see section *lpWGS CCNE1 gene amplification plot*). The plot is not available for samples with a GI rejected status.

6.1.8 Report Additional Variants Not In The Scope Of BRCA Or HRR Status Analysis

Users can add or remove additional variants (i.e., variants which are not in scope of BRCA status or HRR status) using theSOPHiA DDM™ variant reporting functionality.

6.1.9 Generate HRD Report

To generate the HRD report, please refer to the SOPHiA DDM™ Platform Operation Manual.

- The first part of the report is dedicated to the display of HRD, GII, and BRCA statuses (if selected).
- The second part of the report is dedicated to HRR and CCNE1 statuses (if selected).
- The third part of the report is dedicated to any additional variants that might be reported out of scope of HRD (see section *6.1.8 Report additional variants not in the scope of BRCA or HRR status analysis*.

6.1.10 Considerations Regarding Inconclusive Results

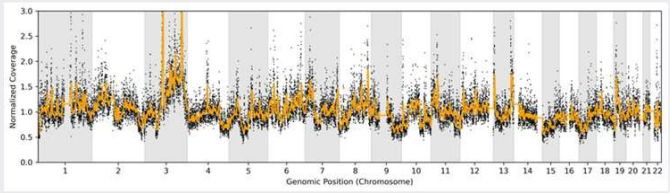
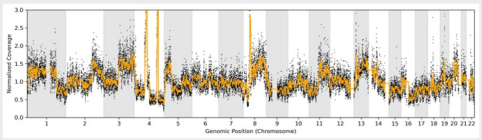
In Case Of GI Rejection

In case of a GI rejected status, we suggest the user to consult the sample QA metrics (available in the pipeline GI report) as well as the *Table 3* on the next page to identify the underlying cause.

Table 3: Recommendations in case of GI Rejection

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
Insufficient coverage due to issues in balancing WGS with capture.	• Check the Sample QA section of the pipeline GI report. A sample is rejected when the number of WGS fragments is below 4 M. • The Sample QA section of the pipeline GI report shows the percentage of WGS fragments. Compare this metric for all samples that were in the same capture pool. Identify if all the samples in the pool have an unusually low value (<50% of WGS fragments).	The WGS pool was not sequenced to the sufficient depth. We recommend re-sequencing the sample pool. If needed adjust the ratio between capture and WGS pool to obtain enough WGS reads.
Insufficient coverage due to issues balancing the samples multiplexed in the same run.	• Check the Sample QA section of the pipeline GI report. A sample is rejected when the number of WGS fragments is below 4 M. • Check the Sample QA section of the pipeline GI report. Compare the number of WGS fragments (M) between samples in the same run. Check if the reads distribution in the run is balanced, or if there are outliers with very few or too many reads (3-fold).	If you observe a strong imbalance in the sequencing run: We recommend re-sequencing your samples, ensuring to quantify your libraries carefully and to adjust all samples to the same molarity before pooling.
Insufficient coverage due to poor yield of the sequencing run.	• Check the Sample QA section of the pipeline GI report. A sample is rejected when the number of WGS fragments is below 4 M. • Go to the pipeline QA report and check the total number of reads produced by the sequencing run. Compare this with the expected number of reads (~260 M reads for a NextSeq® 500/550 Mid-Output run, ~800 M reads for a NextSeq® 500/550 High-Output run).	Repeat the sequencing run adjusting the final sequencing library loading concentration to avoid under- or over-clustering of the run.

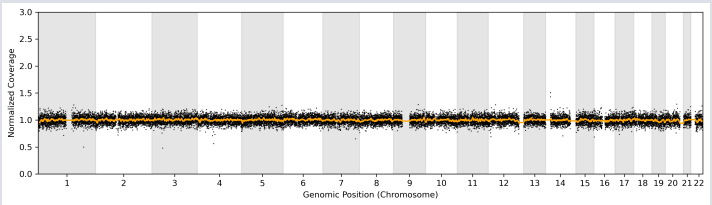
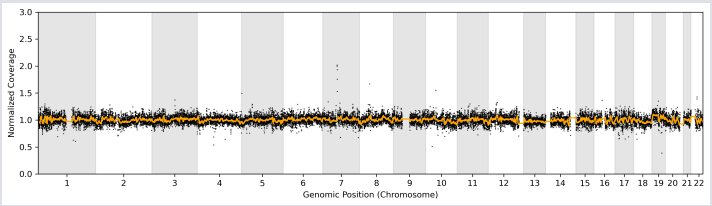
Recommendations in case of GI Rejection (continued)

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
Excess of residual noise	• Check the Sample QA section of the pipeline GI report. Residual noise ≥0.17 leads to sample rejection if purity-ploidy ratio is not detected. *Examples of lpWGS coverage profiles with excessive residual noise, in which PPR was not detected.*  	High levels of noise can be a sign of low-quality input DNA, possibly resulting from poor DNA extraction To improve the library preparation efficiency, repeat the library preparation, doubling the amount of input DNA and, if possible, starting with a new DNA extraction
High proportion of coverage outliers	• Check the Sample QA section of the pipeline GI report. A sample is rejected if the proportion of coverage outliers is larger or equal to 20%. • A high proportion of coverage outliers affecting the lpWGS coverage profile may result from sub-optimal execution of the capture step or from low DNA quality. • To distinguish the two scenarios, compare the proportion of coverage outliers for all samples that were in the same capture pool. Identify if all samples in the pool have an unusually high value ≥20%).	• In case all samples in the pool are rejected due to coverage outliers we recommend repeating the entire workflow for the sample pool paying close attention to the capture instructions. • In case only a minority of the samples in the pool are rejected due to coverage outliers, we recommend repeating the entire workflow only for the rejected samples, doubling the amount of input DNA.

In Case Of GI Inconclusive Status

In case of a GI inconclusive status, we suggest the user to consult the sample QA metrics (available in the pipeline GI report) as well as *Table 4* on the next page to identify the underlying cause.

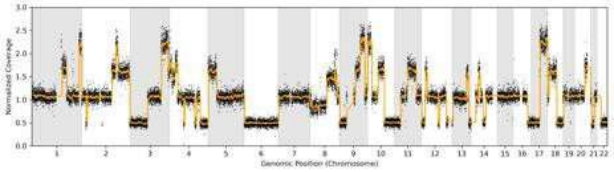
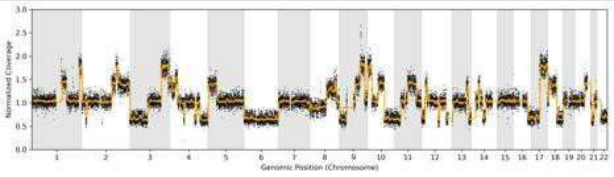
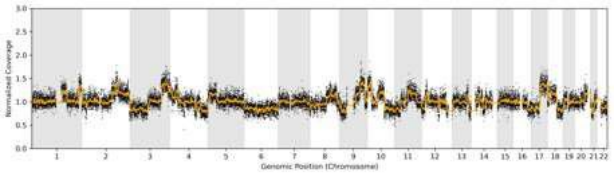
Table 4: Recommendations in case of GI Inconclusive status

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
The sample does not feature large copy number aberrations	• Check the Sample QA section of the pipeline GI report. A sample is classified GI inconclusive if signal to noise ratio is insufficient (i.e. SNR smaller than 0.55). • Perform a visual inspection of the lpWGS profile in the GI report. The smoothed coverage profile (orange line) should be flat, with no sign of large CN aberrations. The insufficient SNR is due to an absence of signal which could result either from: i) the absence of CN changes in the sample or, ii) from extremely low tumor content. *Examples of lpWGS coverage profiles without large CN aberrations*   	• If the sample tumor content is larger than 30%, it is likely that the sample does not feature large copy number aberrations. Repeating the experiment will not affect the results of the GI analysis. • If there are doubts about the sample tumor content, we recommend repeating the entire wetlab workflow, starting with a sample with higher tumor content.

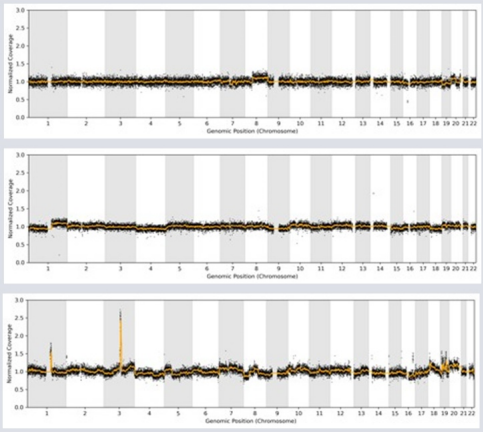
6.1.11 Considerations Regarding GI Negative* Statuses

GI Negative* statuses are medium confidence negative calls with an increased false negative risk. GI Negative* calls result from low signal-to-noise ratio in the NGS data, possibly reflecting insufficient sample tumor content. *Table 5* on the next page illustrates how insufficient tumor content impacts SNR.

Table 5: Considerations regarding case of GI Negative* statuses

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
Insufficient tumor content	• Verify that sample tumor content estimated before processing the DNA sample is larger than 30%. • Check the GI report and confirm that the PPR was either lower than 0.1 or not detected. • Note that GI analysis does not perform tumor content estimation. *Illustration of lpWGS profile obtained by decreasing sample tumor content. The magnitude of the coverage changes between segments decreases, with decreasing tumor content.* Tumor Content = 100%  Tumor Content = 66%  Tumor Content = 33%  *Examples of lpWGS coverage profiles for insufficient tumor content samples in which PPR was not detected.*	If there are doubts about the sample tumor content, we recommend repeating the entire wetlab workflow, starting with a sample with higher tumor content.

Considerations regarding case of GI Negative* statuses (continued)

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
		

In Case Of BRCA Inconclusive And/or Presence Of BRCA FN Risk

Assess the severity of the FN risk by checking the % of the BRCA1/2 regions covered with >100x unique molecules. This can be done by looking at the quality indicator “BRCA Cov Percentile 100” . A good sample is expected to have a score of at least 99%.

Note: When reading the troubleshooting guidelines in *Table 6* below, multiple root causes can apply. Make sure to read the entire guide and identify the most relevant cause.

Table 6: Recommendations in case of BRCA undetermined associated with a warning of FN risk

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
Poor coverage uniformity	• Check the coverage heterogeneity (<0.01) in the pipeline QA report • Identify if the coverage uniformity is low in all samples of the capture pool or just in individual samples.	• If you observe a pool effect: The capture efficiency is low, we recommend repeating the entire batch of samples, and pay close attention to the IFU section describing the hybridization and capture procedure. • If you observed the effect in an individual sample: The library preparation efficiency and/or sample quality is low. To improve the library quality, we recommend doubling the amount of input DNA and repeating the entire workflow, starting with a new DNA extraction, if possible.
Insufficient number of fragments mapping to the target regions of the panel	• Check the Sample QA section of the pipeline GI report. Calculate the number of fragments (M) on the target regions of the panel by subtracting the number of WGS fragments from the total number of fragments. Each sample should have at least 2 M fragments mapping to the target regions of the panel. • Identify if the number of fragments mapping to	• If you observe a pool effect: The capture pool was not sequenced to the sufficient depth, we recommend resequencing the sample pool. If needed, adjust the ratio between capture and WGS pool to obtain enough reads mapping to the target regions. • If you observed the effect in an individual

Recommendations in case of BRCA undetermined associated with a warning of FN risk (continued)

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
	the target regions of the panel is low in all samples of the capture pool or just in individual samples. • Check if the reads distribution in the capture is balanced, or if there are outliers with very few or too many reads (3-fold difference).	sample: Repeat the entire workflow for the sample and ensure to quantify your libraries carefully before pooling and capture. • If you observe a strong imbalance in the capture pool: Ensure to quantify your libraries carefully before pooling and capture. If you know the quality of your input DNA, you can improve the reads balance by pooling similar quality samples in the same pool. ***Warning:*** *The first two scenarios only apply if the coverage uniformity of the samples was of sufficient quality.*
Poor library diversity	• Check the Mapping Statistics in the pipeline QA report to determine the duplicate fraction in the sample. PCR duplicates should be similar for all your samples.	• High percentages of PCR duplicates suggest low library preparation efficiencies. We recommend repeating the entire workflow and, if possible, by doubling the amount of input DNA.
Small library fragment size	• Check if the library size is unusually low in comparison with the other samples processed. You can obtain the average library fragment size via high-resolution capillary electrophoresis, which is used to check the library quality at the end of the library preparation workflow.	• Small library sizes of individual samples can be a sign of low- quality input DNA (assuming multichannel pipettes were used to process samples according to good laboratory practice). To improve the library preparation efficiency, we recommend doubling the amount of input DNA.

6.2 Analysis Prerequisites

6.2.1 Data Generation

Sequencing data for analysis with the workflow must be generated starting with FFPE tissue-extracted DNA samples by following the SOPHiA GENETICS™ Universal Library Prep protocol and sequencing these samples on an Illumina® NextSeq® 500/550 sequencer.

The recommended sequencing depths per sample and loading dilutions for various sequencers can be found in the Appendix IV section of the Instructions for Use document for the SOPHiA GENETICS™ Universal Library Prep (SG-03901). The recommended reads per sample are reproduced in *Table 7* below:

Table 7: Recommended sequencing depths for SOPHiA DDM™ Homologous Recombination Deficiency Solution

READ LENGTH (BP)	RECOMMENDED TOTAL READS PER SAMPLE	RECOMMENDED READ-PAIRS (FRAGMENTS) PER SAMPLE
2 x 150	32 million	16 million

Increasing the number of reads per sample is expected to provide more confident and sensitive variant calls, especially in regions with otherwise low read depth. Refer to section *8 Warnings, Limitations, and Precautions* for more details.

Note: SOPHiA DDM™ Homologous Recombination Deficiency Solution is compatible with sequencing data generated by Illumina® NextSeq® 500/550 and NovaSeq™ X instruments, i.e., at the end of the analysis, the user will have access to the same type of results. However, please note that the analytical performance of this product has only been verified with Illumina® NextSeq® 500/550 instruments.

6.2.2 Run Composition

For high-quality results based on SOPHiA DDM™ analytics, 16 million fragments or read pairs (i.e., 32 million reads) per sample are recommended. This would be reflected in a total of 24 samples for a NextSeq® 500/550 High-Output run or 8 samples for a NextSeq® 500/550 Mid-Output run. If these specifications are not followed, please refer to section *8 Warnings, Limitations, and Precautions* for details on the maximum file size.

6.2.3 NGS Data Demultiplexing Instructions

The user should perform NGS data demultiplexing following the instructions provided by the user guide of the Illumina®:

- NextSeq® 500/550 sequencer (e.g., NextSeq® 550 Systems Guide, Document #15069765 v07)

As indicated by Illumina®, standalone demultiplexing for third-party analyses can be performed with bcl2fastq v2.0 or higher (bcl2fastq2 Conversion Software v2.20 Software Guide, Document #15051736 v03).

6.3 Analysis Description and Parameters

6.3.1 Resource Files

Alignment and variant calling are performed against the GRCh37 reference genome (also referred to as hg19). Variant coordinates are also provided against GRCh38 (also referred to as hg38), via a conversion operation termed liftOver.

The liftOver is performed using Picard and relies on the mapping of regions between the two assemblies (i.e., chain file) provided by UCSC.

Publicly available sources for all these requirements can be found at:

GRCh38 – hg38 reference genome:
https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/001/405/GCA_000001405.15_GRCh38/seqs_for_alignment_pipelines.ucsc_ids/GCA_000001405.15_GRCh38_no_alt_plus_hs38d1_analysis_set.fna.gz.

GRCh37 – hg19 reference genome:
https://storage.googleapis.com/genomics-public-data/references/b37/Homo_sapiens_assembly19.fasta.gz

Chain files:
http://hgdownload.soe.ucsc.edu/goldenPath/hg38/liftOver/hg38ToHg19.over.chain.gz
ftp://ftp.ensembl.org/pub/assembly_mapping/homo_sapiens/GRCh37_to_GRCh38.chain.gz

Picard:
https://github.com/broadinstitute/picard, version 2.23.8

6.3.2 Target Regions

See the table below for the target regions of the SOPHiA DDM™ Homologous Recombination Deficiency Solution gene panel.

GENE	BASES COVERED BY HRD_V2	TOTAL BASES ON CODING REGION	% OF GENE COVERED BY HRD_V2
AKT1	387	4329	8.94
ATM	9171	9171	100
BARD1	7311	7311	100
BRCA1	21078	21078	100
BRCA2	10257	10257	100
BRIP1	3750	3750	100
CCNE1	4605	4605	100
CDK12	8919	8919	100
CHEK1	8250	8250	100
CHEK2	5849	5907	99.02
ESR1	856	11664	7.34
FANCA	9537	9537	100
FANCD2	13128	13128	100

GENE	BASES COVERED BY HRD_V2	TOTAL BASES ON CODING REGION	% OF GENE COVERED BY HRD_V2
FANCL	2271	2271	100
FGFR1	5523	21459	25.7
FGFR2	4909	26337	18.6
FGFR3	1531	6933	22.1
MRE11	6294	6294	100
NBN	4284	4284	100
PALB2	3561	3561	100
PIK3CA	1147	3207	35.8
PPP2R2 A	2718	2718	100
PTEN	3564	3564	100
RAD51B	12525	12525	100
RAD51C	1539	1539	100
RAD51D	2685	2685	100
RAD54L	4488	4488	100
TP53	13338	13338	100

Table 8: Target regions of SOPHiA DDM™ HRD_v2

6.3.3 Raw Data Pre-Processing

The following is a brief description of the steps taken in the pipeline, up to and including alignment.

Pre-processing

- Collect quality metrics based on the raw FASTQ files.
- Truncate the FASTQ files to a maximum size of 2.5 GB.

Alignment

1. Cut adapters and trim low-quality ends from reads (base quality below Q20).
2. Align reads to the hg19 human reference genome in paired-end mode.
3. Compute alignment statistics and coverage metrics on the raw alignment files.
4. Trim overhanging adapter sequences.
5. Remove reads that have low mapping quality or low average base phred scores. A threshold of <30 is used in both cases.
6. Local coverage at any position within the target regions is limited to 30,000x and to 10,000x outside target regions. Excessive coverage is removed following a random down sampling.
7. Remove chimeric reads with hairpin loops.

8. Realign soft clips for better detections of long INDELs.
9. Assign reads to read groups based on start-end coordinates.
10. Annotate low coverage regions based on a threshold of 100 molecules (read groups).
11. Calculate statistics and coverage metrics on the processed alignment file.

6.3.4 Included Analysis Modules

Quality Indicators

For each analyzed sample, a panel of traffic-light-like quality indicators are displayed in SOPHiA DDM™ to inform on specific quality metrics of the sample. Each indicator is colored in green when the corresponding quality metric is within the expected range, or in red vice versa. For information on how to interpret the color codes of the quality indicator dots, please refer to Section 3.10 of the SOPHiA DDM™ Operation Manual. The following indicators are displayed for each sample:

- **Panel_cov_percentile_100:** The percentage of the gene panel covered at >100x molecules. The expected range for this metric is between 95% and 100%
- **BRCA_cov_percentile_100:** The percentage of the on-target BRCA1 and BRCA2 regions covered at >100x molecules. The expected range for this metric is between 99% and 100%. Whenever this indicator fails (i.e., below 99%) a FN risk warning will be triggered in the context of BRCA status analysis.
- **BRCA_cov_unif:** The coverage uniformity in the on-target BRCA1 and BRCA2 regions. Coverage uniformity is defined as the fraction of the genomic regions that have coverage within 5 and 1/5 of the median coverage. The expected range for this metric is between 0.9 and 1.
- **Group_size:** The group size is defined as the median number of reads in each read group. A high group size reflects low library conversion rate and high rate of duplications. The expected range for this metric is between 1 and 10.
- **DeamScore:** The deamination score is a score devised to reflect the rate of artificial deamination in a given FFPE sample. The expected range of this score is between 0 and 0.8.
- **FragLength:** The median DNA fragment length in a sample. The expected range for this metric is between 75 and 250.

SNVs/INDELs

The SNV/INDEL detection modules include local realignment algorithms and variant calling software that apply statistical tests to identified mismatches versus the reference genome, variant regularization functions, variant quantification functions, and variant filtering functions.

Variant Calling

- Identify variants (SNVs and INDELs up to 10,000 bp) by selecting positions in which the signal supporting the alternative allele is significantly different from background noise:
- Perform dedicated variant calling for long duplications and long insertions that cannot be identified by pileup of reads.
- Merge variants together if they are on the same allele (phasing).

- Re-quantify the variant fraction considering all the haplotypes in the neighboring region.
- Unify homopolymer annotation to long anchor standards (e.g., GAAA → GAA instead of GA → G).

Variant Filtering

1. Filter variants below the background noise level of the panel and sequencer (filter = high_background_noise).
2. Filter variants with variant fraction below 4% (filter = low_variant_fraction).
3. Filter variants with read coverage below 30 reads (filter = low_coverage).
4. Filter variants outside of the target regions (filter = off_target).
5. Filter INDELs in homopolymers if the homopolymer length is ≥10 bp (filter = homopolymer_region).
6. Calculate a score based on the fraction of C:G>T:A variants with low molecular support and apply a differentiated threshold depending on the sample specific low/high score (filter = low_molecular_support).
7. Remove variants beyond the target regions with a padding of 500 bp.
8. Remove duplications longer than 500 bp.
9. Filter variants with confidence score below 0 (filter = low_quality).

Variants without any filter associated to them are considered as “high confidence” calls and are the ones used for the assessment of analytical performance. Variants labeled with any of the filters mentioned above are considered “low confidence” and are shown with the sole purpose of helping the interpretation of all the weak signals present in the bam file.

Variant Annotation

The annotation system computes transcript-specific annotations following HGVS coordinate normalizations and notation guidelines (c.DNA and protein notation). This module also provides functional information for the variant’s coding consequence, along with positional and contextual information such as rank and distance to the closest exon, ref and alt codon sequence, and ref and alt amino acid sequence. Transcript- (RefSeq identifiers) and gene-level (HGNC symbol, OMIM gene number) information is also provided at that stage.

The annotation system then queries external databases via genomic coordinates matches to retrieve variant-level information including dbSNP identifiers, allele frequencies from GnomAD, 1000 genomes project, ExAC, ESP5400, CG69, prediction scores from dbNSFP (SIFT, PolyPhen2, MutationTaster) and clinical significance assertions from ClinVar. The system finally annotates variants with licensed catalogs such as:

- OMIM
- CKB (actionable evidence displayed in OncoPortal)
- The BRCA Exchange databases

Note: the BRCA1 and BRCA2 variants are annotated as per “BRCA_pathogenicity”; this score is obtained by aggregating the BRCA exchange’s “Pathogenicity_expert” and “Pathogenicity_all” fields, following the set of rules detailed in *Table 9* on the next page.

Table 9: Rules followed for aggregating BRCA_Exchange information into the “BRCA_Pathogenicity” displayed in the variant table

RULES FOLLOWED FOR AGGREGATING BRCA_EXHCANGE INFORMATION INTO THE “BRCA_PATHOGENICITY” DISPLAYED IN THE VARIANT TABLE
The BRCA_pathogenicity is assigned per variant, via classification rules consuming the "Pathogenicity_expert" and "Pathogenicity_all" attributes retrieved from BRCA exchange and applying rules according to the following precedence order:
1: Any variant annotated with 'pathogenic', 'likely pathogenic', 'risk factor', 'probable pathogenic' in **Pathogenicity_expert** gets a “BRCA_pathogenicity” as “pathogenic”.
2: Any variant annotated with 'likely benign', 'benign', 'probably not pathogenic', 'benign / little clinical significance' in **Pathogenicity_expert** gets a “BRCA_pathogenicity” as “benign”.
3: Any variant annotated with 'pathogenic', 'likely pathogenic', 'risk factor', 'probable pathogenic' in **Pathogenicity_all** gets a “BRCA_pathogenicity” as “pathogenic”.
4: Any variant annotated with 'likely benign', 'benign', 'probably not pathogenic', 'benign / little clinical significance' in **Pathogenicity_all** gets a “BRCA_pathogenicity” as “benign”.
5: Any variant with conflict between rules 3 and 4 gets a “BRCA_pathogenicity” as “unclear”.
6: Any variant without any “BRCA_pathogenicity” assigned as per rules 1 to 5; and annotated with 'uncertain significance', 'no known pathogenicity', 'variant of unknown significance' in **Pathogenicity_all** gets a “BRCA_pathogenicity” as “unclear”.
Rule traceability: The “pathogenic” and “benign” values, are further prefixed either with ‘ex:’ (expert) when they originate from priority rules 1 and 2 or with 'ag:' (aggregated) when they are generated by rules 3 and 4. As a result, the final available values for the “BRCA_pathogenicity” field are: • ex:pathogenic • ag:pathogenic • ex:benign • ag:benign • unclear • “” (in case there is no entry in BRCA exchange).

DNA Analysis: Gene-level And Exon-level CNV Detection Via MUSKAT™

Analysis purpose

Partial-gene copy number alterations, such as exon-level deletions or duplications, can have significant functional consequences in cancer. Tumor suppressor genes are particularly sensitive to loss-of-function events, and even deletions affecting a subset of exons can disrupt protein function and drive tumorigenesis. Detecting these subgenic events is therefore important for accurately assessing gene inactivation, which is critical for making clinically relevant interpretations from cancer genomic profiles.

The gene-level and exon-level (GLEL) CNV analysis performs CNV detection at both gene-level and exon-level resolution and is applied to a subset of 26 genes of the gene panel:

RAD54L,FANCL,BARD1,FANCD2,PPP2R2A,NBN,PTEN,MRE11,ATM,CHEK1,BRCA2,RAD51B, PALB2, FANCA, TP53, RAD51D, CDK12, BRCA1, RAD51C, BRIP1, CHEK2, PIK3CA*, FGFR3*, FGFR1*, FGFR2*, CCNE1*

Genes denoted with an asterix are tested for gene-level gains but not for gene-level losses and are not in scope of exon-level CNV analysis.

Technical overview of the analysis

The technical details of the GLEL module are documented in the GLEL Manual (SG-08061), which describes the module independently of the application. The table below provides the parameter values used for this specific application.

PARAMETER ID	NAME	DESCRIPTION	VALUE
QA1	GLEL coverage threshold	Samples with an “Average nb. fragments in genomic regions” lower than the value will be rejected.	30
QA2	GL residual noise threshold	Gene-level CNV analysis residual noise threshold. Samples with a value exceeding this threshold will be rejected.	0.25
QA3	GLEL gene residual noise threshold	Gene-level and exon-level CNV analysis gene residual noise threshold. Genes with a value exceeding this threshold will be rejected.	0.15
QA4	GLEL sample residual noise threshold	Gene-level and exon-level CNV analysis sample residual noise threshold. Samples with a value exceeding this threshold will be rejected.	Not applied
Th1	GL gain threshold	Gene-level coverage signal threshold for gain. Genes with a coverage signal greater than or equal to this value will be reported as “Gain”	3.25
Th2	GL suspected loss threshold	Gene-level coverage signal threshold for suspected loss. Genes with a coverage signal lower than or equal to this value will be reported as “Suspected loss”	1.2
Th3	GL loss threshold	Gene-level coverage signal threshold for loss. Genes with a coverage signal lower than or equal to this value will be reported as “Loss”	0.9

Description of the results

The GLEL module generates a CNV variant table containing the following fields for each gene analyzed:

- **Gene-level status**: Indicates whether the gene is affected by a whole-gene event or an exon-level event. The possible statuses are:
    - **Gain**: Tumor cells carry a gene-level gain.
    - **Loss**: Tumor cells carry a homozygous gene-level deletion (i.e., zero copies of the gene are present in tumor DNA).
    - **Suspected loss**: Tumor cells are suspected to carry a homozygous gene-level deletion. In samples with high tumor content and/or high tumor ploidy, a heterozygous gene-level deletion may be sufficient to trigger a Suspected loss call.

- **Exon-level**: The gene is not affected by a whole-gene CNV but carries an intragenic CNV.
- **Normal**: No gene-level CNVs nor intragenic CNVs are detected.
- **Rejected**: The gene-level status cannot be computed due to insufficient data quality.

- **Exon-level status**: Indicates any exon-level events affecting the gene and specifies their type. The possible statuses are:
  - **Normal**: No exon-level CNVs are detected.
  - **Exon-gain**: The gene is affected by an exon-level gain. The regions of the gene impacted by the event are identified and reported.
  - **Exon-loss**: The gene is affected by an exon-level loss. The regions of the gene impacted by the event are identified and reported.
  - **Suspected exon-loss**: The gene is affected by an exon-level CNV. The regions of the gene suspected to be impacted by a loss are identified. The biological interpretation of the exon-level CNV has low confidence. The gene could be affected by an exon-gain, and not by an exon-loss. For details, see SG-08061.
  - **Uncharacterized**: The gene is affected by an exon-level CNV. The regions of the gene impacted by the event could not be identified. The gene is affected either by an exon-gain or by an exon-loss.
  - **Complex rearrangement**: The exon-level coverage signal reflects the presence of multiple exon-level CNV events.
  - **Rejected**: The exon-level status cannot be computed due to insufficient data quality.
- **Segment regions, coverage levels and interpretations**: Reports and interprets the coverage levels values measured within a gene.
- **Gene QA status**: Reports if the exon-level coverage signal of the gene of interest is of insufficient quality. If Gene QA status is Rejected, the gene is rejected.

The current version of the GLEL-CNV module does not support tertiary annotation. In all outputs generated by the module, exon nomenclature follows an internal convention. The GLEL-CNV_regions.txt file, which can be downloaded from the module, provides a table that enables interpretation of exon nomenclature used.

Additionally, the GLEL module computes the following quality metrics:

- **Nb. fragments**: Total number of DNA fragments (in millions) available for CNV analysis. This QA metric is provided for information purposes only.
- **Average nb. fragments in genomic region**s: Average number of DNA fragments mapped to the target regions analyzed by the GLEL CNV module. The acceptance criterion is >= 30.
- **Coverage normalization status**: This QA metric reports whether the GLEL CNV module could not reliably compute the gene-level coverage signal, due to an abnormal coverage profile. If this QA metric is Fail, the QA status of the sample is Rejected.
- **Gene-level analysis residual noise**: Measure of noise in the data used to compute gene-level coverage signal. The acceptance criterion is <= 0.25.
- **Exon-level analysis residual noise**: Measure of noise in the data used to compute exon-level coverage signal.

- **QA status**: Indicates the outcome of the GLEL CNV analysis, as either successful completion or rejection due to insufficient data quality. The possible sample QA statuses are:
    - **Pass**: All QA metrics meet the acceptance criteria.
    - **Rejected**: NGS data are deemed of insufficient quality.

The results are made available in theSOPHiA DDM™ platform via downloadable files. The following table provides an overview of the output files produced.

FILE NAME	RUN-VS-SAMPLE SPECIFIC OUTPUT	DESCRIPTION
GLEL -CNV_results.tsv	Run	Text file containing GLEL CNV results for all samples in the run.
GLEL -CNV_QC.tsv	Run	Text file containing GLEL QC metrics for all samples in the run.
GLEL -CNV_Report.pdf	Run	PDF report describing GLEL CNV results for all samples in the run.
GLEL -CNV_Report_ sample_<SAMPLE ID>.pdf	Sample	PDF report describing GLEL CNV results for individual samples.
GLEL-CNV_regions.txt	Resource File	List of regions in scope for GLEL CNV analysis, including region name, genomic coordinates and associated transcript annotation.

Precautions and Limitations

- The copy number levels reported by the module quantify the copy number in the DNA sample and thus depend on the sample tumor content. At low tumor content, the signal from somatic CNVs may be diluted by the germline DNA possibly resulting in False Negatives.
- The module reports copy number levels, which are quantified relative to the average copy number across all genes in the panel.
- The ability to detect an intragenic CNV within a gene of interest depends on:
    - The size of the CNV (i.e., the number of genomic regions affected by the CNV).
    - The size of the gene (i.e., the total number of genomic regions considered by the GLEL CNV module for the gene of interest). Small intragenic CNVs occurring within small genes may be missed or may cause the gene of interest to be excluded from the analysis.
- The ability to detect CNVs depends on the magnitude of the event. For example, the ability to detect a copy number change from 2 to 3 (gain of an extra copy) is lower compared to detecting a change from 2 to 1 (heterozygous deletion).
- The GLEL CNV module uses cross-sample coverage normalization. If many samples share the same CNV in a given gene, the analysis for that gene can be biased.
- The GLEL CNV module is only applied to batches including at least 8 DNA samples. If the total number of non-rejected DNA samples in the batch is below 4, the whole batch is rejected. The performance of the module is expected to improve with an increased number of samples.
- For optimal performance of the module, all the samples included in the batch must be processed under the same laboratory condition.

- Processing of highly degraded FFPE samples may produce poor quality coverage signals, which can increase the risk of false positives, false negatives, or sample rejection.
- Genomic regions with extreme GC content may be associated with poor quality coverage signals, making these regions more prone to false positives and false negatives.
- The GLEL CNV module only considers a subset of genomic regions targeted by the DNA panel (refer to “GLEL-CNV _ regions.txt” downloadable file).

lpWGS CCNE1 Gene Amplification Plot

The copy number levels reported by the CNV module quantify the copy number in the DNA sample (effective CN), not the absolute copy number of the tumoral DNA present in the sample. The lpWGS CCNE1 amplification plots are designed to help the users estimate the absolute copy number of the tumoral DNA present in the sample upon visual inspection (see *Figure 1* below and *Figure 2* below).

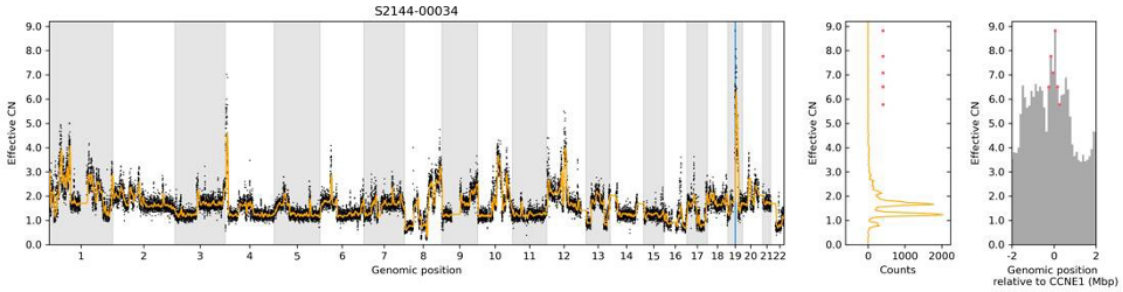


Figure 1: Example of CCNE1 amplification plot with a strong amplification

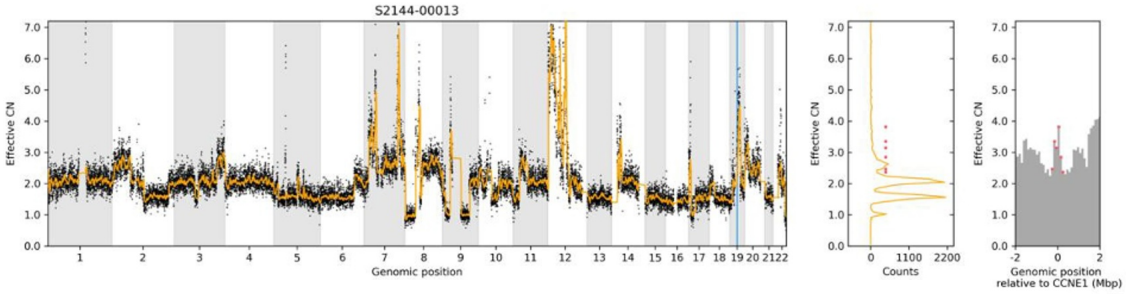


Figure 2: Example of CCNE1 amplification plot with a weak copy number gain

The left panel shows the lpWGS coverage profile after normalization (black: raw data at 100 kb resolution, orange: smoothed version of raw data), with the vertical blue line showing the genomic position of CCNE1.

The middle panel shows: i) Orange: the histogram of the smoothed lpWGS coverage profile shown in the left panel; ii) Red dots: the value of the normalized lpWGS coverage profile measured around the CCNE1 gene (±300 kb of CCNE1 genomic position). Distinct peaks in the histogram correspond to distinct copy number levels present in the tumor DNA. The position of the red dots with respect to the histogram peaks allows the user to interpret the magnitude of the CCNE1 CN gain present in the tumor DNA. In *Figure 1* above, the red dots are far away from all histogram peaks, suggesting the presence of a strong copy number gain in CCNE1. In *Figure 2* above, the red dots overlap with some histogram peaks, suggesting the presence of a weak copy number gain in CCNE1.

The right panel provides a zoomed-in view of the lpWGS coverage profile after normalization (same data as in left panel) around the CCNE1 genomic location (±2 Mb). The regions within ±300 kb of CCNE1 are highlighted by red dots (same red dots as shown in the middle panel).

Proposed HRD Status

The proposed HRD status is computed from the aggregation of the “Proposed GI status” and the “Proposed BRCA status”. The set of rules that determine the “Proposed HRD status”, represented in the table cells, can be inferred from *Table 10* below, where the “Proposed GI status” is represented along the input row and the “Proposed BRCA status” is represented along the input column.

HRD negativity is constrained by the type of genomic aberrations that are covered by the product. See limitations of proposed BRCA status determination for more information.

Table 10: Proposed HRD status computation

PROPOSED GI STATUS	PROPOSED BRCA STATUS						
	POSITIVE	PREDICTED POSITIVE	COMPLEX CASE	UNDETERMINED	PREDICTED NEGATIVE	NEGATIVE	INCONCLUSIVE
Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Inconclusive	Positive	Positive	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
Rejected	Positive	Positive	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
Negative*	Positive	Positive	Undetermined	Negative	Negative	Negative	Undetermined
Negative	Positive	Positive	Undetermined	Negative	Negative	Negative	Undetermined

Proposed BRCA Status

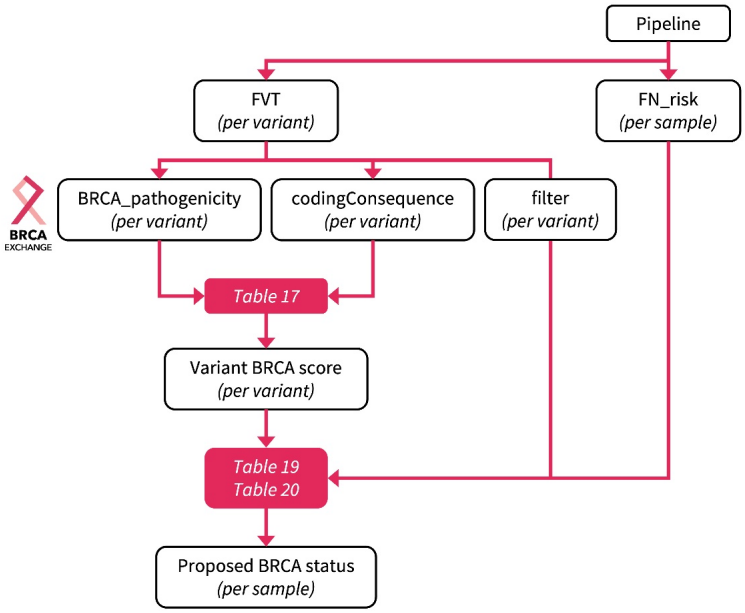
The “Proposed BRCA status” exposes the user to state-of-the-art clinical knowledge for BRCA1/2 variants while transparently conveying uncertainty in the variant detection and calling. The Proposed BRCA status thus combines clinical knowledge and variant detection metrics gathered at the variant and sample levels.

The system proceeds sequentially by:

1. Establishing the functional impact and clinical relevance of each BRCA1/2 variant detected in the sample.
2. Establishing the functional status of BRCA1/2, at the sample level, via a “Proposed BRCA status”.

The present section exposes the data flow, rules and intermediate steps involved in establishing the “Proposed BRCA status”. The nomenclatures and definitions of labels emitted by the pipeline are detailed in *Table 11* on page 33, *Table 12* on page 34, *Table 13* on page 36, and *Table 14* on page 36.

The BRCA status computation only considers SNPs and INDELs.



The analysis starts with the Full Variant Table (columns “BRCA_pathogenicity” and “codingConsequence”). The system establishes the functional impact (i.e., coding consequence) and clinical relevance (i.e., BRCA pathogenicity) of each variant detected in the sample. As a result, each BRCA variant gets assigned a “variant BRCA score” as per rules detailed in *Table 11* on the next page (with nomenclature and conventions listed in *Table 12* on page 34). Those rules classify each BRCA1/2 variant into:

- **Positive:** variant known as pathogenic as per clinical knowledge
- **Predicted positive:** variant with coding consequence predictive of loss-of-function
- **Complex case:** VUS with coding consequence requiring further inspection by end user
- **Predicted Negative:** variant with coding consequence predictive of no functional alteration
- **Negative:** variant known as benign as per clinical knowledge
- **Undetermined:** otherwise.

The system reports an “NA” category in case of absence of detected variant in BRCA.

The SOPHiA DDM™ Gen2 platform performs variant annotation with respect to multiple transcripts. For each BRCA1 and BRCA2 transcript considered, a corresponding “variant BRCA score” is computed. Given that variant coding consequence is transcript-dependent, “variant BRCA scores” for the same variant can differ across transcripts. Note that the overall sample BRCA status is computed based on “variant[CP4.1][FV4.2] BRCA scores” obtained considering the following transcripts: NM_000059.3 (BRCA1) NM_007294.3 (BRCA2).

Table 11: Rules used to establish the variant BRCA score

CODING CONSEQUENCE	BRCA PATHOGENICITY FLAGS			
	EX:PATHOGENIC OR AG:PATHOGENIC	EX:BENIGN OR AG:BENIGN	UNCLEAR	NA
3'UTR	Positive	Negative	Undetermined	Predicted negative
5'UTR	Positive	Negative	Undetermined	Predicted negative
intronic	Positive	Negative	Undetermined	Predicted negative
synonymous	Positive	Negative	Undetermined	Predicted negative
splice_*	Positive	Negative	Complex case	Complex case
inframe*	Positive	Negative	Complex case	Complex case
missense	Positive	Negative	Undetermined	Undetermined
no_stop	Positive	Negative	Undetermined	Undetermined
frameshift_ 5'	Positive	Negative	Predicted positive	Predicted positive
frameshift_ 3'	Positive	Negative	Undetermined	Undetermined
nonsense_5'	Positive	Negative	Predicted positive	Predicted positive
nonsense_3'	Positive	Negative	Undetermined	Undetermined
no variant detected for BRCA 1/2	NA			

Table 12: Nomenclature and conventions

TERM	DESCRIPTION
AA_position	Amino acid position of the variant along the BRCA1_refTranscript or BRCA2_refTranscript reference.
BRCA1_refTranscript	NM_007294.3
BRCA2_refTranscript	NM_000059.3
AA_threshold_BRCA1	AA_threshold_BRCA1 = most 3' AA_position of pathogenic variants as per the BRCAexchange catalog (version in production). Given that the most 3' pathogenic variant for BRCA1 according to BRCA_Exchange is p.(Ala1823_*1864del), this places AA_threshold_BRCA1 value that is aminoacid 1863, which corresponds to the full length of the protein.
AA_threshold_BRCA2	AA_cutoff = 3326 (AA coordinates), as per Mazoyer S et al., Nature Genetics 1996, 14:253-254. This is equivalent to genomic position hg19:13:32972627
codingConsequence	As per codingConsequence column of the Full Variant Table file produced by the annotation system.
BRCA_pathogenicity flags	As per BRCA_pathogenicity column of the Full Variant Table file produced by the annotation system. Pathogenicity assertions provided per variant, as retrieved from BRCA exchange after label recasting and error checking exposed in the annotation sections.
splice_*	any codingConsequence matching with the label "splice"
inframe*	any codingConsequence matching with the label "inframe"
frameshift_5'	BRCA1 : codingConsequence is "frameshift" and AA_position ≤ AA_threshold_BRCA1
	BRCA2 : codingConsequence IS "frameshift" AND AA_position ≤ AA_threshold_BRCA2
frameshift_3'	BRCA1 : codingConsequence IS "frameshift" AND AA_position > AA_threshold_BRCA1
	BRCA2 : codingConsequence IS "frameshift" AND AA_position > AA_threshold_BRCA2
nonsense_5'	BRCA1 : codingConsequence IS "nonsense" AND AA_position ≤ AA_threshold_BRCA1
	BRCA2 : codingConsequence IS "nonsense" AND AA_position ≤ AA_threshold_BRCA2
nonsense_3'	BRCA1 : codingConsequence IS "nonsense" AND AA_position > AA_threshold_BRCA1
	BRCA2 : codingConsequence IS "nonsense" AND AA_position > AA_threshold_BRCA2

The next step calls the “Proposed BRCA status”, reports the set of supporting variants and emits user warnings. The rule set is detailed in *Table 13* on page 36 and *Table 14* on page 36. For that, the bioinformatics pipeline takes into consideration:

1. The variant BRCA score (established in step 1), and ranks for all BRCA variants found in a sample according to their clinical relevance: “Positive” > “Predicted positive” > “Complex case” > “Undetermined” > ”Predicted negative” > “Negative” > “NA”.
2. The uncertainty in variant detection, considering separately high- and low-confidence variant calls.
3. The presence or absence of False Negative risk (see section *Quality Indicators*: BRCA_cov_percentile_1000).

Variants gathered from high-confidence calls are used to establish the proposed BRCA status. The sample is called as either BRCA:

1. Positive: score of the most relevant high-confidence variant is “Positive”.
2. Predicted positive: score of the most relevant high-confidence variant is “Predicted positive”.
3. Complex case: score of the most relevant high-confidence variant is “Complex case”.
4. Predicted negative: score of the most relevant high-confidence variant is “Predicted negative” and there is no FN risk.
5. Negative: score of all variants is “Negative” (or no variants are detected) and there is no FN risk.
6. Inconclusive: score of the most relevant high-confidence variant is “Predicted negative” or “Negative” and there is a FN risk.
7. Undetermined: none of the previous cases. Is a result of insufficient clinical knowledge regarding the reported BRCA1/2 variants (BRCA_Pathogenicity is “unclear”).

The warnings signal three possible situations to the user:

1. “Review low-confidence variants with relevant BRCA score”: One or more low-confidence variants with a relevant BRCA score have been detected.
2. “Review clinical interpretation”: the most relevant evidence found is a “Complex case” or “Undetermined” high-confidence variant.
3. “False negative risk”: no relevant evidence is found, but a False-Negative risk cannot be excluded for the sample.

The obtained “proposed BRCA status”, the list of supporting evidence, and the warnings are then used to initiate the user interfaces are detailed in section *7 Pipeline Deliverables*.

Table 13 on the next page and *Table 14* on the next page: Proposed BRCA status. The following set of rules are applied to aggregate *variant BRCA scores* into a sample-level *proposed BRCA status*. The *variant BRCA scores* of the obtained variants are combined into a sample-level *proposed BRCA status* according to rules that apply either to:

a) samples for risk of False-Negatives has not been identified (i.e., FN_risk is PASS) or

b) samples with False-negatives risk (i.e., FN_risk is FAIL).

The tables are formatted to expose rule inputs as rows (*BRCA score* of most clinically relevant variant in the high confidence calls) and columns (*BRCA score* of most clinically relevant variant in low confidence calls). Values in the table represent sample-level proposed BRCA status. Finally, the rule system emits a series of user warnings, listed at the bottom of *Table 14* on the next page.

Table 13: Sample FN risk: FAIL

BRCA SCORE OF THE MOST RELEVANT VARIANT CALLED WITH HIGH CONFIDENCE	BRCA SCORE OF THE MOST RELEVANT BRCA VARIANT CALLED WITH LOW CONFIDENCE						
	POSITIVE	PREDICTED POSITIVE	COMPLEX CASE	UNDETERMINED	PREDICTED NEGATIVE	NEGATIVE	NA
Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive
Complex case	Complex case[1,2,3]	Complex case[1,2,3]	Complex case[2,3]	Complex case[2,3]	Complex case[2,3]	Complex case[2,3]	Complex case[2,3]
Undetermined	Undetermined 1,2,3	Undetermined 1,2,3	Undetermined 1,2,3	Undetermined 2,3	Undetermined 2,3	Undetermined 2,3	Undetermined 2,3
Predicted negative	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3
Negative	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3
NA	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3	Inconclusive 1,3

Table 14: Sample FN risk: PASS

BRCA SCORE OF THE MOST RELEVANT VARIANT CALLED WITH HIGH CONFIDENCE	BRCA SCORE OF THE MOST RELEVANT BRCA VARIANT CALLED WITH LOW CONFIDENCE						
	POSITIVE	PREDICTED POSITIVE	COMPLEX CASE	UNDETEMINED	PREDICTED NEGATIVE	NEGATIVE	NA
Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive	Predicted positive
Complex case	Complex case 1,2	Complex case 1,2	Complex case 2	Complex case 2	Complex case 2	Complex case 2	Complex case 2
Undetermined	Undetermined 1,2	Undetermined 1,2	Undetermined 1,2	Undetermined[2]	Undetermined 2	Undetermined 2	Undetermined 2
Predicted negative	Predicted negative[1]	Predicted negative[1]	Predicted negative[1]	Predicted negative[1]	Predicted negative	Predicted negative	Predicted negative
Negative	Negative[1]	Negative[1]	Negative[1]	Negative[1]	Negative	Negative	Negative
NA	Negative[1]	Negative[1]	Negative[1]	Negative[1]	Negative	Negative	Negative

User warnings: [1]Review low-confidence variants with relevant *BRCA score*, [2]Review clinical interpretation, [3]False negative risk

Genomic Integrity Analysis

Genomic Integrity analysis aims at detecting HRD by assessing the degree of genomic instability. Functional homologous recombination repair is necessary for the error-free repair of double strand breaks and for maintaining genome integrity. A consequence of HRD is the loss of genomic integrity via the accumulation of genomic aberrations due to the cell's inability to repair double strand breaks. Our Genomic Integrity analysis aims at assessing these genomic aberrations through the use of Whole Genome Sequencing (WGS). More specifically, our method uses a deep learning algorithm that has been

specifically trained to recognize patterns of genomic instability in the WGS coverage profile. The algorithm analyzes low-pass WGS (lpWGS) data and produces a Genomic Integrity (GI) index that reflects the level of genomic integrity. A GI index above 0 indicates low genomic integrity. A GI index below 0 indicates high genomic integrity.

During GI analysis, the following algorithmic steps are performed sequentially for each sample of interest: first low pass WGS (lpWGS) data are preprocessed and undergo quality assessment, next the GI index is computed, and lastly a proposed GI status is assigned to the sample. These steps are described in the sections below.

Step 1: Data preprocessing and sample quality assessment:

1. WGS paired-end reads are mapped to the human reference genome and processed to trim adapters and low-quality base calls. The WGS coverage profile is computed and normalized.
2. The normalized WGS coverage profile undergoes QA based on the metrics defined at the end of this section. One of the following GI QA statuses is assigned to the sample:
    - *High quality:* the quality of the data is sufficient to confidently compute a GI index and a proposed GI status.
    - *Medium quality:* the quality of the data is lower (compared to high quality) and, consequently, the deep learning algorithm may not succeed in computing a GI index.
    - *Low quality:* the quality of the data does not meet the criteria required to compute a GI index.

Step 2: GI Index calculation:

The GI index is obtained by processing the normalized WGS coverage profile using a proprietary deep learning algorithm that has been trained to recognize patterns of genomic instability.

Step 3: Proposed GI Status determination:

A proposed GI status that applies to Ovarian Cancer samples is determined by combining the sample QA status and the GI index. Five outcomes are possible:

- *GI positive:* samples with a GI index larger or equal than 0.
- *GI negative:* samples with a GI index smaller than 0.
- *GI negative*:* samples with a GI index smaller than 0 but featuring an increased risk of false negative calls due to low signal to noise ratio.
- *GI inconclusive:* medium quality samples for which the deep learning algorithm did not succeed in computing a GI index due to insufficient signal to noise ratio.
- *GI rejected:* low quality samples discarded from GI analysis. Samples are rejected if the number of DNA fragments available for WGS coverage profile calculation is insufficient, if the noise of the WGS coverage profile is excessive, or if the proportion of coverage outliers is excessive.

Definition of QA and GI metrics:

- *Total nb. of fragments:* Total number of DNA fragments (paired-end reads) that are properly mapped.
- *Nb. WGS fragments:* Total number of DNA fragments available for the raw coverage WGS profile calculation. DNA fragments mapping to the genomic regions enriched for variant calling are excluded. This QA metric is compared against a threshold of 4 million to determine the GI QA status.
- *Percentage WGS fragments:* Fraction of the total number of WGS fragments over the total number of fragments.

- *Proportion of Coverage Outliers:* Percentage of WGS regions considered for GI analysis which feature an artefactual and excessive localized coverage which is compensated by the coverage normalization algorithm. This QA metric is compared against a threshold of 20% to determine the GI QA status.
- *Purity/ploidy ratio:* 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. This GI QA metric is compared against a threshold of 0.1 (suggesting low tumor content) to determine the GI QA status. The inability to measure purity-ploidy ratio from the data is also used to determine the GI QA status.
- *Residual noise:* Residual noise is computed by measuring the standard deviation of the normalized WGS coverage profile with respect to the smoothed WGS coverage profile. This QA metric is compared against a threshold of 0.17 to determine the GI QA status.
- *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. This GI QA metric is compared against a threshold of 0.55 to determine the GI QA status. SNR is also considered by the algorithm to assign the GI Negative* and GI Inconclusive status.
- *QA Status:* GI quality assessment status of a sample.
- *GI Index:* The GI index is a scalar value, ranging between −20 and 20, that reflects the level of genomic integrity. High GI indices reflect low levels of genomic integrity. Low GI indices reflect high levels of genomic integrity.
- *GI Status:* The proposed genomic integrity status of a sample.

7 PIPELINE DELIVERABLES

7.1 SOPHiA DDM™ Downloadable Files

7.1.1 Run Level

- **fastq.gz files:** The zipped input FASTQ files for all samples in the run.
- **HRD_v2-GLEL-CNV_report.pdf:** The report of gene-level and exon-evel CNV calls in PDF format.
- **HRD_v2-GLEL-CNV_results.tsv:** The report of gene-level and exon-evel CNV calls in .tsv format.
- **HRD_v2-GLEL-CNV_qc.tsv:** The CNV module QC metrics in .tsv format.
- **HRD_v2-GLEL-CNV_regions.txt:** The list of regions in scope for CNV analysis in text format.
- **QA-report.pdf:** The pipeline QA report in PDF format at the run level.
- **GI-report.pdf:** The report for the genome integrity analysis in PDF format.
- **GI_status_table.txt:** Text file summarizing the GI results.

7.1.2 Sample Level

- **fastq.gz files:** The zipped input FASTQ files for the considered sample.
- **bam and bai files:** The alignment file per sample and its index.
- **quality_indicators.txt:** Text file containing the quality indicator values (e.g., panel coverage metrics).
- **full_variant_table.txt/vcf:** The detected variants in text and vcf format.
- **cosmic_info_table.txt:** Text file containing annotation information from the COSMIC database.
- **flagged_regions.txt:** Text file that lists target regions flagged as low-coverage.
- **Exon_coverage_stats_v3.txt:** Text file containing the coverage statistics (computed in molecules) at the exon level.
- **QA-patient.pdf:** Sample-level QA report in PDF format.
- **HRD_v2-GLEL-CNV_report_sample_<sample_name>.pdf:** The report of gene-level and exon-level CNV calls in PDF format.

7.2 Variant Annotation

7.2.1 Description Of FVT Content

See below an example of the Full Variant Table (FVT) with a description of each field and an example where possible.

FIELD	DESCRIPTION	EXAMPLE
id	variant run ID, internal to FVT	1
annotation_id	variant ID in annotation database	60245
gene	HGNC gene symbol	BRCA2
overlapKnown	rsid of pathogenic clinvar entries	rs1555760738
type	variant type	SNP / INDEL
codingConsequence	consequence on protein	5'UTR
refGenome	reference genome	GRCh37/hg19
chromosome	chromosome	13
genome_position	genomic position (from variant caller)	32890572
depth	sequencing depth	12224
var_percent	variant fraction (relative depth of alternative allele)	49.32%
exon_rank	exon identifier	2
c.DNA	HGVS cDNA	c.-26G>A
protein	HGVS protein	
ref	ref in reference genome	G
alt	alternative allele	A
refNum	depth reference allele	6178
altNum	depth alternative allele	6029
refSeq	ref codon	
altSeq	alt codon	
refAA	ref amino-acid	
altAA	alt amino-acid	
tx_id	transcript ID in annotation database	35315
tx_name	transcript symbol in annotation database	NM_000059
refSeqId	transcript symbol in RefSeq	NM_000059
tx_version	transcript version in annotation database	3
refSeqIdVersion	transcript version in RefSeq	3
gene_boundaries	Exome / Intergenic qualifier	within

FIELD	DESCRIPTION	EXAMPLE
exon_id	exon legacy identifier	2
pos_in_exon	position in exon (strand specific)	14
dist2exon	position in intron (distance to closest exon)	0
filter	whether any quality filter should apply	.
dbSNP	dbSNP's rsid	rs1799943
g1000	Allele frequency	0.2093
esp5400	Allele frequency	0.2078
ExAC	Allele frequency	0.243
GnomAD	Allele frequency	0.2427
LJB_PhyloP	dbNSFP's precomputed PhyloP score	
LJB_SIFT	dbNSFP's precomputed SIFT score	
LJB_PolyPhen2	dbNSFP's precomputed PolyPhen2 score	
LJB_PolyPhen2_HumDiv	dbNSFP's precomputed PolyPhen2_HumDiv score	
LJB_LRT	dbNSFP's precomputed LRT score	
LJB_MutationTaster	dbNSFP's precomputed MutationTaster score	
LJB_GERP	dbNSFP's precomputed GERP score	
id_cosmic_coding	COSMIC ID coding variants	COSN20442808
id_cosmic_non_coding	COSMIC ID non-coding variants	
id_clinvar	Clinvar's rsid	rs1799943
CLNSIG	Clinvar's pathogenicity assertion	Benign
CLNREVSTAT	Clinvar's metadata	reviewed_by_expert_panel
gene_strand	strand	+
ref1	normalized ref (3' alnmt, in direction of strand)	G
alt1	normalized alternative allele (3' alnmt, in direction of strand)	A
first1	normalized genome_position (3' alnmt, in direction of strand)	32890572
last1	normalized genome_position(3' alnmt, in direction of strand)	32890572
multiTranscriptId		1
flagged_region_id		
depth_uniq	Number of unique molecular fragments at the position of a variant taking base Phred scores into account	400
refNum_uniq	Number of unique molecular fragments supporting the	203

FIELD	DESCRIPTION	EXAMPLE
	reference allele taking base quality (Phred Score) into account	
altNum_uniq	Number of unique molecular fragments supporting the alternative allele taking base quality (Phred Score) into account	197
matchStatus		exact
OMIM	mim2gene identifier	600185
hg38_chrom	Chromosome in hg38	1
hg38_pos	Genome position in hg38	46259775
hg38_ref	Reference allele in hg38	C
hg38_alt	Alternative allele in hg38	T
lift_diagnostic		PICARD
hg38_refGenome	Reference for hg38	GRCh38/hg38
sgid		00010001c698114b5b53cf6 0b6cdb5d02a62beba
hg38_sgid		00010101651f8aaff0d56960 c3121277b4d53606
BRCA_pathogenicity	“BRCA_pathogenicity”; as obtained by aggregating BRCA exchange’s “Pathogenicity_expert” and “Pathogenicity_all” fields, following aggregation rules detailed in section *5.2.4 Included Analysis Module* under *Variant Annotation*	ex:pathogenic

Table 15: Example of a full variant table (FVT)

8 WARNINGS, LIMITATIONS, and PRECAUTIONS

8.1 Warnings

8.1.1 General Warnings

1. This product is for research use only and not for use in diagnostic procedures.
2. For detailed instructions on the software, refer to the SOPHiA DDM™ User Manual.
3. If any part of the handling, protocol, sequencer, multiplexing, etc., is changed, the analyses are not covered by the described Instructions For Use parameters.
4. The SOPHiA DDM™ HRD solution has been fully validated for DNA FFPE samples from Ovarian Cancer with a DQN>3 (measured using Fragment Analyzer) and sequenced on Illumina® NextSeq® 500/550.

8.2 Limitations

8.2.1 General Limitations

1. The proposed GI status and the proposed HRD status computed by the pipeline only apply to Ovarian Cancer samples
2. The GI QA metrics and GI index reported to the end-users are obtained by rounding up the full precision values used for QA status and GI status calculations to the first or second decimal (depending on the metric).
3. The maximal run size that can be uploaded is 200 GB. Uploading higher volumes in a single batch is not allowed. The pipeline has been shown to successfully process individual samples with up to 50 GB of data. Higher volumes per sample might cause the pipeline to fail.
4. Even when sample multiplexing recommendations are followed, and the recommended average number of reads per sample (~32M) is achieved for a given sequencing run, read depth may be insufficient to provide sensitive and accurate results for various reasons, including: poor sample quality, poor NGS data quality, significant uneven read allocation between WGS and captured libraries, significant uneven read allocation between samples multiplexed in the same run, skewed read depth allocation across the gene panel (e.g., consistently lower read depth in AT rich regions).
5. Other available devices for high-resolution gel electrophoresis, such as Tapestation, are not suitable for the estimation of DQN on FFPE samples, due to interference of the lower marker with highly degraded DNA fragments.
6. Variant crosstalk between samples may occur due to sample-to-sample carryover in the steps before the library amplification. Such variant crosstalk is expected to manifest as potential False Positive variants: in most cases at insignificantly low variant fraction, i.e., below the variant calling threshold; nevertheless, variant cross-talk between samples at significant variant fraction leading to erroneous variant calls is expected to be rare but cannot be completely ruled out.

8.2.2 Module-Specific Limitations

Data Upload

1. The product was designed to upload and process the volume of data produced by an Illumina® NextSeq® 500/550 sequencer (Mid- and High-Output flow cells).

SNVs/INDELs

1. Variant detection in this product has been optimized for SNVs and short INDELs (up to 50 bp).
   Please note that any other type of alteration that can have an impact in the BRCA status can be missed by our algorithm.
2. Variant detection in this product has been optimized on the regions defined as “target regions”. Please note that regions outside this definition might present false negative and/or false positives.
3. DNA fragments carrying SNVs or INDELs may not be efficiently captured by the panel probes, possibly resulting in False Negatives. This may affect:
    1. Multiple nucleotide variants (MNVs)
    2. Delins

3. INDELs that are larger than 50 bp
4. DNA fragments carrying multiple variants.

4. Complex variants such as i) delins or ii) INDELs near SNVs or MNVs can create non-trivial soft-clip patterns that cannot be properly processed, resulting in false negatives.
5. Due to the complexity of the genomic context, some variants may be represented in different valid ways at the same time (e.g., different position within a repeated sequence). This would lead to a split assignment of the reads among the different representations of the variant, which would lead to an underestimation of the VAF and possibly to filtering the variant due to low variant fraction.
6. SNVs/INDELs in long repeat/low complexity regions may be missed due to the high levels of noise.
7. For a stable performance of the algorithm, we recommend coverage of unique molecules of at least 100x. Lower coverage will increase the risk of False Negative variant calls significantly.
8. Variants below 4% VAF are filtered and considered of low confidence to avoid an excess of false positives.
9. SNVs/INDELs in homopolymers of length 10 or higher cannot be called confidently, since their detection is confounded by high background noise for homopolymers of this length.
10. NGS data quality can be affected by the quality and quantity of the input DNA. High levels of DNA degradation reduce the library preparation efficiency and, thus, the library complexity. Further, errors and inaccuracies in the library preparation and capture workflow, as well as sequencing-related errors, can reduce the NGS data quality. Poor NGS data quality can confound the bioinformatics pipeline, possibly resulting in False Positives or False Negatives.
11. Genomic regions of interest with low-complexity nucleotide sequences, nucleotide bias, repeats of any length (e.g., mono-, di- or trinucleotide repeats, transposable elements, Alu repeats, etc.) or with significant sequence similarity to other genomic regions (e.g., pseudogenes and gene families) are at higher risk of variant calling artifacts including False Positive and False Negative variant calls.

Copy Number Variations (CNVs)

1. The CNV module has been tested exclusively on data from Ovarian Cancer FFPE DNA samples.
2. The copy number level reported by the CNV module for a genomic region of interest reflects the number of DNA copies in the sample relative to the average number of copies in other regions targeted by the panel. Thus, the reported copy number level reflects the copy number of tumor DNA in the sample, but it is affected by both sample tumor content and tumor ploidy.
3. The copy number levels reported by the CNV module are normalized so that the average copy number across the panel is equal to 2. Therefore, a copy number of 2 corresponds to the ploidy of the sample. However, this normalization approach may not accurately reflect the true copy number in highly aneuploid samples.
4. The ability to detect CNVs depends on the sample tumor content and tumor ploidy. CNVs may be missed in samples with low tumor content and/or low tumor ploidy.
5. The ability to detect CNVs depends on the magnitude of the event. For example, the ability to detect a copy number change from 2 to 3 (gain of an extra copy) is lower compared to detecting a change from 2 to 1 (heterozygous deletion).
6. The ability to detect an intragenic CNV within a gene of interest depends on:

i. The size of the CNV (i.e., the number of genomic regions affected by the CNV)

ii. The size of the gene (i.e., the total number of genomic regions considered by the CNV module for the gene of interest). Small intragenic CNVs occurring within small genes may be missed or may cause the gene of interest to be excluded from the analysis.

7. The CNV module performs coverage signal normalization across the samples being analyzed. When a significant portion of samples share the same CNV event in a genomic region of interest, the signal measured in that genomic region may be biased during normalization, potentially causing false positives and/or false negatives.
8. The CNV module is only applied to requests including at least 8 samples. If the total number of non-rejected samples in the batch is below 4, the whole request is rejected from the CNV analysis.
9. The CNV module only considers genomic regions targeted by the HRD_v2 panel, with some genomic regions excluded (see *11.1 Appendix I: Genomic Coordinates of CNV "Target Regions" of the HRD Solution*).

Proposed HRD Status

Proposed BRCA Status

1. Structural variants, including CN gains and losses affecting one or multiple exons of BRCA1 and BRCA2, are NOT within the scope of the Proposed BRCA status provided by this product.
2. Mobile element insertions are NOT within the scope of the Proposed BRCA status provided by this product.
3. The proposed BRCA status is computed based on the SNVs and short INDELs detected by the pipeline.
4. Other types of pathogenic genomic aberrations are not considered for the proposed BRCA status calculation.
5. The present product relies on data provided by "BRCA exchange", which is a third-party provider. The product is thereby exposed to licensing and data sharing constraints that are independent of SOPHiA GENETICS' control. Such constraints can introduce deviations between the observations reported in the product and those exposed via the "BRCA exchange" website.
6. The proposed BRCA status is computed based on the two following transcripts: NM_007294.3 (BRCA1) and NM_000059.3 (BRCA2).

Proposed Genomic Integrity Status

1. Processing samples with tumor content <30% is likely to result in GI inconclusive or GI negative* results but can also result in false negative results.
2. Low-pass WGS data do not allow to reliably distinguish samples lacking large copy number aberrations from samples having insufficient tumor content. Processing samples that do not feature large copy number alterations will likely produce GI inconclusive calls.

9 SYMBOLS

SYMBOL	TITLE
	Consult instructions for use
	Catalog number
	Manufacturer
	Research Use Only

10 SUPPORT

In case of difficulty using the SOPHiA DDM™ Platform, please consult the troubleshooting section of the "General information about usage of SOPHiA DDM™" document, visit our support portal, e-mail support@sophiagenetics.com, or contact our support line by telephone at:

- **EU:** +41 21 561 34 75
- **US:** +1 617 313 7957
- **FR:** +33 5 47 51 01 29
- **AU:** +61 2 51 10 7564

Any serious incident occurring in relation to the device should be promptly reported to SOPHiA GENETICS and the competent authorities of the member state where the user is established.

Do not use components that are damaged. Contact the SOPHiA GENETICS support team if there are any concerns with the kits using the support channels mentioned above.

11 APPENDICES

11.1 Appendix I: Genomic Coordinates of CNV “Target Regions” of the HRD Solution

CNV Target Regions

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
RAD54L	ex1-2	1	46713968	46714381	NM_003579	3	ex1-2	TRUE	TRUE
RAD54L	ex3	1	46715575	46715994	NM_003579	3	ex3	TRUE	TRUE
RAD54L	ex4	1	46724248	46724529	NM_003579	3	ex4	TRUE	TRUE
RAD54L	ex5	1	46725526	46725881	NM_003579	3	ex5	TRUE	TRUE
RAD54L	ex6-8	1	46726104	46727167	NM_003579	3	ex6-8	TRUE	TRUE
RAD54L	ex9	1	46733021	46733392	NM_003579	3	ex9	TRUE	TRUE
RAD54L	ex10	1	46736227	46736568	NM_003579	3	ex10	TRUE	TRUE
RAD54L	ex11-12	1	46738027	46738586	NM_003579	3	ex11-12	TRUE	TRUE
RAD54L	ex13-14	1	46738916	46739530	NM_003579	3	ex13-14	TRUE	TRUE
RAD54L	ex15	1	46739700	46739999	NM_003579	3	ex15	TRUE	TRUE
RAD54L	ex16	1	46740099	46740500	NM_003579	3	ex16	TRUE	TRUE
RAD54L	ex17-18	1	46743377	46744066	NM_003579	3	ex17-18	TRUE	TRUE
FANCL	ex14	2	58386790	58387045	NM_018062	3	ex14	TRUE	TRUE
FANCL	ex13	2	58387133	58387415	NM_018062	3	ex13	TRUE	TRUE
FANCL	ex12	2	58388547	58388869	NM_018062	3	ex12	TRUE	TRUE
FANCL	ex10-11	2	58389891	58390319	NM_018062	3	ex10-11	TRUE	TRUE
FANCL	ex9	2	58390459	58390762	NM_018062	3	ex9	TRUE	TRUE
FANCL	ex8	2	58392749	58393120	NM_018062	3	ex8	TRUE	TRUE
FANCL	ex7	2	58425615	58425907	NM_018062	3	ex7	TRUE	TRUE
FANCL	ex6	2	58431155	58431472	NM_018062	3	ex6	TRUE	TRUE
FANCL	ex5	2	58448967	58449288	NM_018062	3	ex5	TRUE	TRUE
FANCL	ex4	2	58453753	58454030	NM_018062	3	ex4	TRUE	TRUE
FANCL	ex3	2	58456839	58457120	NM_018062	3	ex3	TRUE	TRUE
FANCL	ex2	2	58459079	58459358	NM_018062	3	ex2	TRUE	TRUE
FANCL	ex1	2	58468243	58468607	NM_018062	3	ex1	FALSE	FALSE
BARD1	ex11	2	215593107	215593845	NM_000465	3	ex11	TRUE	TRUE
BARD1	ex10	2	215595004	215595363	NM_000465	3	ex10	TRUE	TRUE
BARD1	ex9	2	215609635	215609994	NM_000465	3	ex9	TRUE	TRUE
BARD1	ex8	2	215610332	215610691	NM_000465	3	ex8	TRUE	TRUE
BARD1	ex7	2	215617055	215617414	NM_000465	3	ex7	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
BARD1	ex6	2	215632022	215632501	NM_000465	3	ex6	TRUE	TRUE
BARD1	ex5	2	215633787	215634146	NM_000465	3	ex5	TRUE	TRUE
BARD1	ex4	2	215645154	215646353	NM_000465	3	ex4	TRUE	TRUE
BARD1	ex3	2	215656871	215657350	NM_000465	3	ex3	TRUE	TRUE
BARD1	ex2	2	215661683	215662052	NM_000465	3	ex2	TRUE	TRUE
BARD1	ex1	2	215673956	215674562	NM_000465	3	ex1	FALSE	FALSE
FANCD2	ex2	3	10070284	10070488	NM_001018115	2	ex2	TRUE	TRUE
FANCD2	ex3	3	10074466	10074705	NM_001018115	2	ex3	TRUE	TRUE
FANCD2	ex4	3	10076067	10076306	NM_001018115	2	ex4	TRUE	TRUE
FANCD2	ex5	3	10076354	10076507	NM_001018115	2	ex5	TRUE	TRUE
FANCD2	ex6	3	10076721	10076972	NM_001018115	2	ex6	TRUE	TRUE
FANCD2	ex7	3	10077878	10078117	NM_001018115	2	ex7	TRUE	TRUE
FANCD2	ex8	3	10080883	10081122	NM_001018115	2	ex8	TRUE	TRUE
FANCD2	ex9	3	10081347	10081586	NM_001018115	2	ex9	TRUE	TRUE
FANCD2	ex10	3	10083281	10083420	NM_001018115	2	ex10	TRUE	TRUE
FANCD2	ex11	3	10084218	10084372	NM_001018115	2	ex11	TRUE	TRUE
FANCD2	ex12	3	10084709	10084859	NM_001018115	2	ex12	TRUE	TRUE
FANCD2	ex13	3	10085143	10085301	NM_001018115	2	ex13	TRUE	TRUE
FANCD2	ex14	3	10085411	10085650	NM_001018115	2	ex14	FALSE	FALSE
FANCD2	ex15	3	10088152	10088432	NM_001018115	2	ex15	FALSE	FALSE
FANCD2	ex16	3	10089549	10089788	NM_001018115	2	ex16	FALSE	FALSE
FANCD2	ex17	3	10091004	10091243	NM_001018115	2	ex17	FALSE	FALSE
FANCD2	ex18	3	10094046	10094206	NM_001018115	2	ex18	FALSE	FALSE
FANCD2	ex19	3	10101953	10102112	NM_001018115	2	ex19	FALSE	FALSE
FANCD2	ex20	3	10103805	10103953	NM_001018115	2	ex20	FALSE	FALSE
FANCD2	ex21	3	10105416	10105655	NM_001018115	2	ex21	FALSE	FALSE
FANCD2	ex22	3	10105957	10106196	NM_001018115	2	ex22	FALSE	FALSE
FANCD2	ex23	3	10106367	10106606	NM_001018115	2	ex23	FALSE	FALSE
FANCD2	ex24	3	10107053	10107203	NM_001018115	2	ex24	TRUE	TRUE
FANCD2	ex25	3	10107523	10107688	NM_001018115	2	ex25	TRUE	TRUE
FANCD2	ex26	3	10108868	10109026	NM_001018115	2	ex26	TRUE	TRUE
FANCD2	ex27	3	10114530	10114690	NM_001018115	2	ex27	TRUE	TRUE
FANCD2	ex28	3	10114912	10115071	NM_001018115	2	ex28	TRUE	TRUE
FANCD2	ex29	3	10116166	10116405	NM_001018115	2	ex29	TRUE	TRUE
FANCD2	ex30	3	10119740	10119906	NM_001018115	2	ex30	TRUE	TRUE
FANCD2	ex31	3	10122728	10122967	NM_001018115	2	ex31	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
FANCD2	ex32	3	10123005	10123173	NM_001018115	2	ex32	TRUE	TRUE
FANCD2	ex33	3	10127467	10127636	NM_001018115	2	ex33	TRUE	TRUE
FANCD2	ex34	3	10128764	10129003	NM_001018115	2	ex34	TRUE	TRUE
FANCD2	ex35	3	10130108	10130251	NM_001018115	2	ex35	TRUE	TRUE
FANCD2	ex36	3	10130454	10130693	NM_001018115	2	ex36	TRUE	TRUE
FANCD2	ex37	3	10131938	10132107	NM_001018115	2	ex37	TRUE	TRUE
FANCD2	ex38	3	10133781	10134020	NM_001018115	2	ex38	TRUE	TRUE
FANCD2	ex39	3	10134869	10135108	NM_001018115	2	ex39	TRUE	TRUE
FANCD2	ex40	3	10135891	10136130	NM_001018115	2	ex40	TRUE	TRUE
FANCD2	ex41	3	10136802	10137041	NM_001018115	2	ex41	TRUE	TRUE
FANCD2	ex42	3	10137964	10138203	NM_001018115	2	ex42	TRUE	TRUE
FANCD2	ex43	3	10140333	10140932	NM_001018115	2	ex43	TRUE	TRUE
FANCD2	ex44	3	10142812	10143061	NM_001018115	2	ex44	TRUE	TRUE
PIK3CA	ex2	3	178916490	178917086	NM_006218	3	ex2	TRUE	FALSE
PIK3CA	ex5	3	178921215	178921694	NM_006218	3	ex5	TRUE	FALSE
PIK3CA	ex8-9	3	178927774	178928253	NM_006218	3	ex8-9	TRUE	FALSE
PIK3CA	ex10	3	178935910	178936209	NM_006218	3	ex10	TRUE	FALSE
PIK3CA	ex21	3	178951812	178952231	NM_006218	3	ex21	TRUE	FALSE
FGFR3	ex7	4	1803517	1803816	NM_000142	4	ex7	TRUE	FALSE
FGFR3	ex9	4	1805973	1806332	NM_000142	4	ex9	TRUE	FALSE
FGFR3	ex14	4	1807789	1807968	NM_000142	4	ex14	TRUE	FALSE
FGFR3	ex16	4	1808192	1808491	NM_000142	4	ex16	TRUE	FALSE
ESR1	ex5	6	152332761	152332942	NM_000125	3	ex5	FALSE	FALSE
ESR1	ex7	6	152415436	152415645	NM_000125	3	ex7	FALSE	FALSE
ESR1	ex8	6	152419826	152420020	NM_000125	3	ex8	FALSE	FALSE
PPP2R2A	ex1	8	26149226	26149452	NM_002717	3	ex1	FALSE	FALSE
PPP2R2A	in1	8	26150653	26150910	NM_002717	3	in1	TRUE	TRUE
PPP2R2A	ex2	8	26151072	26151367	NM_002717	3	ex2	TRUE	TRUE
PPP2R2A	ex3	8	26196296	26196613	NM_002717	3	ex3	TRUE	TRUE
PPP2R2A	ex4	8	26211874	26212260	NM_002717	3	ex4	TRUE	TRUE
PPP2R2A	ex5	8	26217575	26217908	NM_002717	3	ex5	TRUE	TRUE
PPP2R2A	ex6	8	26218380	26218778	NM_002717	3	ex6	TRUE	TRUE
PPP2R2A	ex7	8	26220089	26220475	NM_002717	3	ex7	TRUE	TRUE
PPP2R2A	ex8	8	26221080	26221517	NM_002717	3	ex8	TRUE	TRUE
PPP2R2A	ex9	8	26223721	26224069	NM_002717	3	ex9	TRUE	TRUE
PPP2R2A	ex10	8	26227540	26228039	NM_002717	3	ex10	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
FGFR1	ex14	8	38272205	38272504	NM_023110	2	ex14	TRUE	FALSE
FGFR1	ex12	8	38274760	38274999	NM_023110	2	ex12	TRUE	FALSE
FGFR1	ex9	8	38276983	38277342	NM_023110	2	ex9	TRUE	FALSE
FGFR1	ex3	8	38287132	38287491	NM_023110	2	ex3	TRUE	FALSE
NBN	ex16	8	90947616	90947975	NM_002485	4	ex16	TRUE	TRUE
NBN	ex15	8	90949069	90949428	NM_002485	4	ex15	TRUE	TRUE
NBN	ex14	8	90955358	90955717	NM_002485	4	ex14	TRUE	TRUE
NBN	ex13	8	90958183	90958662	NM_002485	4	ex13	TRUE	TRUE
NBN	ex12	8	90959884	90960243	NM_002485	4	ex12	TRUE	TRUE
NBN	ex11	8	90965336	90966089	NM_002485	4	ex11	TRUE	TRUE
NBN	ex10	8	90967317	90967916	NM_002485	4	ex10	TRUE	TRUE
NBN	ex9	8	90970789	90971268	NM_002485	4	ex9	TRUE	TRUE
NBN	ex8	8	90976484	90976903	NM_002485	4	ex8	TRUE	TRUE
NBN	ex7	8	90982470	90982949	NM_002485	4	ex7	TRUE	TRUE
NBN	ex6	8	90983280	90983639	NM_002485	4	ex6	TRUE	TRUE
NBN	ex5	8	90990320	90990679	NM_002485	4	ex5	TRUE	TRUE
NBN	ex4	8	90992772	90993251	NM_002485	4	ex4	TRUE	TRUE
NBN	ex3	8	90993407	90993886	NM_002485	4	ex3	TRUE	TRUE
NBN	ex2	8	90994837	90995196	NM_002485	4	ex2	TRUE	TRUE
NBN	ex1	8	90996561	90996920	NM_002485	4	ex1	FALSE	FALSE
PTEN	ex1	10	89623702	89624325	NM_000314	6	ex1	FALSE	FALSE
PTEN	ex2	10	89653644	89654003	NM_000314	6	ex2	TRUE	TRUE
PTEN	ex3	10	89685197	89685386	NM_000314	6	ex3	TRUE	TRUE
PTEN	ex4	10	89690730	89690919	NM_000314	6	ex4	TRUE	TRUE
PTEN	ex5	10	89692730	89693069	NM_000314	6	ex5	TRUE	TRUE
PTEN	ex6	10	89711786	89712065	NM_000314	6	ex6	TRUE	TRUE
PTEN	ex7	10	89717534	89717813	NM_000314	6	ex7	TRUE	TRUE
PTEN	ex8	10	89720637	89720923	NM_000314	6	ex8	TRUE	TRUE
PTEN	ex9	10	89724977	89725296	NM_000314	6	ex9	TRUE	TRUE
FGFR2	ex11	10	123247436	123247735	NM_000141	4	ex14	TRUE	FALSE
FGFR2	ex9	10	123257947	123258185	NM_000141	4	ex12	TRUE	FALSE
FGFR2	ex4	10	123279416	123279775	NM_000141	4	ex7	TRUE	FALSE
MRE11	ex20	11	94153109	94153468	NM_005591	3	ex20	TRUE	TRUE
MRE11	ex19	11	94162967	94163273	NM_005591	3	ex19	TRUE	TRUE
MRE11	ex18	11	94168822	94169181	NM_005591	3	ex18	TRUE	TRUE
MRE11	ex17	11	94170233	94170550	NM_005591	3	ex17	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
MRE11	ex16	11	94178838	94179197	NM_005591	3	ex16	TRUE	TRUE
MRE11	ex15	11	94180261	94180740	NM_005591	3	ex15	TRUE	TRUE
MRE11	ex14	11	94189264	94189623	NM_005591	3	ex14	TRUE	TRUE
MRE11	ex13	11	94192491	94192860	NM_005591	3	ex13	TRUE	TRUE
MRE11	ex12	11	94193972	94194331	NM_005591	3	ex12	TRUE	TRUE
MRE11	ex11	11	94197162	94197521	NM_005591	3	ex11	TRUE	TRUE
MRE11	ex10	11	94200839	94201198	NM_005591	3	ex10	TRUE	TRUE
MRE11	ex9	11	94203453	94203932	NM_005591	3	ex9	TRUE	TRUE
MRE11	ex8	11	94204563	94205042	NM_005591	3	ex8	TRUE	TRUE
MRE11	ex7	11	94209344	94209703	NM_005591	3	ex7	TRUE	TRUE
MRE11	ex6	11	94211702	94212181	NM_005591	3	ex6	TRUE	TRUE
MRE11	ex5	11	94212704	94213063	NM_005591	3	ex5	TRUE	TRUE
MRE11	ex4	11	94218900	94219379	NM_005591	3	ex4	TRUE	TRUE
MRE11	ex3	11	94223885	94224244	NM_005591	3	ex3	TRUE	TRUE
MRE11	ex2	11	94225748	94226107	NM_005591	3	ex2	TRUE	TRUE
ATM	ex2-3	11	108098298	108098618	NM_000051	3	ex2-3	TRUE	TRUE
ATM	ex4	11	108099900	108100100	NM_000051	3	ex4	TRUE	TRUE
ATM	ex5	11	108106229	108106613	NM_000051	3	ex5	TRUE	TRUE
ATM	ex6	11	108114675	108114850	NM_000051	3	ex6	TRUE	TRUE
ATM	ex7	11	108115510	108115759	NM_000051	3	ex7	TRUE	TRUE
ATM	ex8	11	108117686	108117929	NM_000051	3	ex8	TRUE	TRUE
ATM	ex9	11	108119637	108119856	NM_000051	3	ex9	TRUE	TRUE
ATM	ex10	11	108121285	108121805	NM_000051	3	ex10	TRUE	TRUE
ATM	ex11	11	108122559	108122871	NM_000051	3	ex11	TRUE	TRUE
ATM	ex12	11	108123472	108123711	NM_000051	3	ex12	TRUE	TRUE
ATM	ex13	11	108124536	108124771	NM_000051	3	ex13	TRUE	TRUE
ATM	ex14	11	108126935	108127074	NM_000051	3	ex14	TRUE	TRUE
ATM	ex15	11	108128201	108128401	NM_000051	3	ex15	TRUE	TRUE
ATM	ex16	11	108129638	108129877	NM_000051	3	ex16	TRUE	TRUE
ATM	ex17	11	108137893	108138074	NM_000051	3	ex17	TRUE	TRUE
ATM	ex18	11	108139117	108139356	NM_000051	3	ex18	TRUE	TRUE
ATM	ex19-20	11	108141712	108142138	NM_000051	3	ex19-20	TRUE	TRUE
ATM	ex21-22	11	108143177	108143584	NM_000051	3	ex21-22	TRUE	TRUE
ATM	ex23	11	108150054	108150395	NM_000051	3	ex23	TRUE	TRUE
ATM	ex24	11	108151576	108151900	NM_000051	3	ex24	TRUE	TRUE
ATM	ex25	11	108153432	108153611	NM_000051	3	ex25	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
ATM	ex26	11	108154949	108155206	NM_000051	3	ex26	TRUE	TRUE
ATM	ex27	11	108158265	108158504	NM_000051	3	ex27	TRUE	TRUE
ATM	ex28	11	108159683	108159837	NM_000051	3	ex28	TRUE	TRUE
ATM	ex29	11	108160324	108160533	NM_000051	3	ex29	TRUE	TRUE
ATM	ex30	11	108163296	108163535	NM_000051	3	ex30	TRUE	TRUE
ATM	ex31	11	108163994	108164233	NM_000051	3	ex31	TRUE	TRUE
ATM	ex32	11	108165581	108165855	NM_000051	3	ex32	TRUE	TRUE
ATM	ex33	11	108167941	108168189	NM_000051	3	ex33	TRUE	TRUE
ATM	ex34	11	108170436	108170677	NM_000051	3	ex34	TRUE	TRUE
ATM	ex35	11	108172370	108172527	NM_000051	3	ex35	TRUE	TRUE
ATM	ex36	11	108173575	108173761	NM_000051	3	ex36	TRUE	TRUE
ATM	ex37	11	108175397	108175584	NM_000051	3	ex37	TRUE	TRUE
ATM	ex38	11	108178577	108178780	NM_000051	3	ex38	TRUE	TRUE
ATM	ex39	11	108180882	108181047	NM_000051	3	ex39	TRUE	TRUE
ATM	ex40	11	108183062	108183301	NM_000051	3	ex40	TRUE	TRUE
ATM	ex41-42	11	108186474	108186908	NM_000051	3	ex41-42	TRUE	TRUE
ATM	ex43	11	108188045	108188264	NM_000051	3	ex43	TRUE	TRUE
ATM	ex44	11	108190615	108190854	NM_000051	3	ex44	TRUE	TRUE
ATM	ex45	11	108191968	108192207	NM_000051	3	ex45	TRUE	TRUE
ATM	ex46	11	108195788	108196277	NM_000051	3	ex46	TRUE	TRUE
ATM	ex47	11	108196730	108196957	NM_000051	3	ex47	TRUE	TRUE
ATM	ex48	11	108198338	108198511	NM_000051	3	ex48	TRUE	TRUE
ATM	ex49	11	108199743	108199970	NM_000051	3	ex49	TRUE	TRUE
ATM	ex50	11	108200936	108201153	NM_000051	3	ex50	TRUE	TRUE
ATM	ex51	11	108202108	108202347	NM_000051	3	ex51	TRUE	TRUE
ATM	ex52	11	108202601	108202769	NM_000051	3	ex52	TRUE	TRUE
ATM	ex53	11	108203484	108203632	NM_000051	3	ex53	TRUE	TRUE
ATM	ex54	11	108204564	108204713	NM_000051	3	ex54	TRUE	TRUE
ATM	ex55	11	108205691	108205841	NM_000051	3	ex55	TRUE	TRUE
ATM	ex56	11	108206510	108206749	NM_000051	3	ex56	TRUE	TRUE
ATM	ex57	11	108213874	108214135	NM_000051	3	ex57	TRUE	TRUE
ATM	ex58	11	108216394	108216640	NM_000051	3	ex58	TRUE	TRUE
ATM	ex59	11	108217929	108218168	NM_000051	3	ex59	TRUE	TRUE
ATM	ex60	11	108224430	108224669	NM_000051	3	ex60	TRUE	TRUE
ATM	ex61	11	108225450	108225689	NM_000051	3	ex61	TRUE	TRUE
ATM	ex62	11	108235753	108235950	NM_000051	3	ex62	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
ATM	ex63	11	108236047	108236240	NM_000051	3	ex63	TRUE	TRUE
CHEK1	ex1	11	125495564	125496019	NM_001114121	2	5utr	FALSE	FALSE
CHEK1	ex2	11	125496554	125496839	NM_001114121	2	ex2	TRUE	TRUE
CHEK1	ex3	11	125497392	125497835	NM_001114121	2	ex3	TRUE	TRUE
CHEK1	ex4-5	11	125499017	125499466	NM_001114121	2	ex4-5	TRUE	TRUE
CHEK1	ex6	11	125502948	125503356	NM_001114121	2	ex6	TRUE	TRUE
CHEK1	ex7	11	125505214	125505628	NM_001114121	2	ex7	TRUE	TRUE
CHEK1	ex8	11	125507234	125507549	NM_001114121	2	ex8	TRUE	TRUE
CHEK1	ex9-11	11	125513576	125514648	NM_001114121	2	ex9-11	TRUE	TRUE
CHEK1	ex12	11	125523543	125523853	NM_001114121	2	ex12	TRUE	TRUE
CHEK1	ex13	11	125525010	125525325	NM_001114121	2	ex13	TRUE	TRUE
BRCA2	ex1	13	32889507	32889914	NM_000059	3	5utr	FALSE	FALSE
BRCA2	ex2	13	32890488	32890775	NM_000059	3	ex2	TRUE	TRUE
BRCA2	ex3	13	32893104	32893551	NM_000059	3	ex3	TRUE	TRUE
BRCA2	ex4	13	32899103	32899432	NM_000059	3	ex4	TRUE	TRUE
BRCA2	ex5-7	13	32900126	32900863	NM_000059	3	ex5-7	TRUE	TRUE
BRCA2	ex8	13	32903470	32903739	NM_000059	3	ex8	TRUE	TRUE
BRCA2	ex9	13	32904970	32905259	NM_000059	3	ex9	TRUE	TRUE
BRCA2	ex10	13	32906297	32907740	NM_000059	3	ex10	TRUE	TRUE
BRCA2	ex11	13	32910292	32915534	NM_000059	3	ex11	TRUE	TRUE
BRCA2	ex12	13	32918585	32918876	NM_000059	3	ex12	TRUE	TRUE
BRCA2	ex13	13	32920854	32921143	NM_000059	3	ex13	TRUE	TRUE
BRCA2	ex14	13	32928887	32929536	NM_000059	3	ex14	TRUE	TRUE
BRCA2	ex15	13	32930482	32930856	NM_000059	3	ex15	TRUE	TRUE
BRCA2	ex16	13	32931769	32932176	NM_000059	3	ex16	TRUE	TRUE
BRCA2	ex17	13	32936549	32936941	NM_000059	3	ex17	TRUE	TRUE
BRCA2	ex18	13	32937206	32937781	NM_000059	3	ex18	TRUE	TRUE
BRCA2	ex19	13	32944429	32944777	NM_000059	3	ex19	TRUE	TRUE
BRCA2	ex20	13	32945000	32945347	NM_000059	3	ex20	TRUE	TRUE
BRCA2	ex21	13	32950697	32951038	NM_000059	3	ex21	TRUE	TRUE
BRCA2	ex22-24	13	32953344	32954409	NM_000059	3	ex22-24	TRUE	TRUE
BRCA2	ex25	13	32968716	32969180	NM_000059	3	ex25	TRUE	TRUE
BRCA2	ex26	13	32970924	32971283	NM_000059	3	ex26	TRUE	TRUE
BRCA2	ex27	13	32972187	32973020	NM_000059	3	ex27	TRUE	TRUE
RAD51B	ex2	14	68290151	68290454	NM_001321809	1	ex2	TRUE	TRUE
RAD51B	ex3	14	68292071	68292404	NM_001321809	1	ex3	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
RAD51B	ex4	14	68301687	68302024	NM_001321809	1	ex4	TRUE	TRUE
RAD51B	ex5	14	68331552	68331966	NM_001321809	1	ex5	TRUE	TRUE
RAD51B	ex6	14	68352476	68352897	NM_001321809	1	ex6	TRUE	TRUE
RAD51B	ex7	14	68353554	68354098	NM_001321809	1	ex7	TRUE	TRUE
RAD51B	ex8	14	68758422	68758807	NM_001321809	1	ex8	TRUE	TRUE
RAD51B	ex9	14	68878031	68878354	NM_001321809	1	ex9	TRUE	TRUE
RAD51B	ex10	14	68934779	68935078	NM_001321809	1	ex10	TRUE	TRUE
RAD51B	in10_1	14	68937129	68937371	NM_001321809	1	in10_1	TRUE	TRUE
RAD51B	in10_2	14	68944194	68944498	NM_001321809	1	in10_2	TRUE	TRUE
RAD51B	in10_3	14	68963731	68963967	NM_001321809	1	in10_3	TRUE	TRUE
RAD51B	in10_4	14	69006750	69007042	NM_001321809	1	in10_4	TRUE	TRUE
RAD51B	in10_5	14	69061061	69061514	NM_001321809	1	in10_5	TRUE	TRUE
RAD51B	in10_6	14	69069266	69069519	NM_001321809	1	ex11	TRUE	TRUE
RAD51B	in10_7	14	69077610	69078093	NM_001321809	1	3utr	TRUE	TRUE
RAD51B	ex11	14	69149504	69149783	NM_001321809	1	dn	TRUE	TRUE
AKT1	ex3	14	105246339	105246638	NM_001014431	1	ex3	FALSE	FALSE
PALB2	ex13	16	23614646	23615190	NM_024675	3	ex13	TRUE	TRUE
PALB2	ex12	16	23618989	23619468	NM_024675	3	ex12	TRUE	TRUE
PALB2	ex11	16	23625189	23625548	NM_024675	3	ex11	TRUE	TRUE
PALB2	ex10	16	23632573	23633026	NM_024675	3	ex10	TRUE	TRUE
PALB2	ex9	16	23634065	23634549	NM_024675	3	ex9	TRUE	TRUE
PALB2	ex8	16	23635193	23635552	NM_024675	3	ex8	TRUE	TRUE
PALB2	ex7	16	23637423	23637902	NM_024675	3	ex7	TRUE	TRUE
PALB2	ex6	16	23640351	23640710	NM_024675	3	ex6	TRUE	TRUE
PALB2	ex5	16	23640836	23641915	NM_024675	3	ex5	TRUE	TRUE
PALB2	ex4	16	23646063	23647862	NM_024675	3	ex4	TRUE	TRUE
PALB2	ex2-3	16	23649043	23649642	NM_024675	3	ex2-3	TRUE	TRUE
PALB2	ex1	16	23652245	23652604	NM_024675	3	ex1	FALSE	FALSE
FANCA	ex42-43	16	89804904	89805436	NM_000135	2	ex42-43	TRUE	TRUE
FANCA	ex41	16	89805498	89805737	NM_000135	2	ex41	TRUE	TRUE
FANCA	ex40	16	89805804	89806043	NM_000135	2	ex40	TRUE	TRUE
FANCA	ex39	16	89806148	89806532	NM_000135	2	ex39	TRUE	TRUE
FANCA	ex38	16	89807123	89807362	NM_000135	2	ex38	TRUE	TRUE
FANCA	ex37	16	89809157	89809396	NM_000135	2	ex37	TRUE	TRUE
FANCA	ex36	16	89811342	89811504	NM_000135	2	ex36	TRUE	TRUE
FANCA	ex34-35	16	89812967	89813388	NM_000135	2	ex34-35	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
FANCA	ex33	16	89815042	89815200	NM_000135	2	ex33	TRUE	TRUE
FANCA	ex32	16	89816105	89816344	NM_000135	2	ex32	TRUE	TRUE
FANCA	ex31	16	89818519	89818658	NM_000135	2	ex31	TRUE	TRUE
FANCA	ex30	16	89824945	89825183	NM_000135	2	ex30	TRUE	TRUE
FANCA	ex29	16	89828314	89828513	NM_000135	2	ex29	TRUE	TRUE
FANCA	ex28	16	89831267	89831617	NM_000135	2	ex28	TRUE	TRUE
FANCA	ex27	16	89833524	89833670	NM_000135	2	ex27	TRUE	TRUE
FANCA	ex26	16	89836219	89836458	NM_000135	2	ex26	TRUE	TRUE
FANCA	ex25	16	89836549	89836692	NM_000135	2	ex25	TRUE	TRUE
FANCA	ex24	16	89836887	89837126	NM_000135	2	ex24	TRUE	TRUE
FANCA	ex23	16	89838035	89838274	NM_000135	2	ex23	TRUE	TRUE
FANCA	ex22	16	89839654	89839817	NM_000135	2	ex22	TRUE	TRUE
FANCA	ex21	16	89842067	89842306	NM_000135	2	ex21	TRUE	TRUE
FANCA	ex19-20	16	89845174	89845441	NM_000135	2	ex19-20	TRUE	TRUE
FANCA	ex18	16	89846252	89846391	NM_000135	2	ex18	TRUE	TRUE
FANCA	ex16-17	16	89849237	89849535	NM_000135	2	ex16-17	TRUE	TRUE
FANCA	ex15	16	89851237	89851397	NM_000135	2	ex15	TRUE	TRUE
FANCA	ex14	16	89857758	89857997	NM_000135	2	ex14	TRUE	TRUE
FANCA	ex13	16	89858286	89858525	NM_000135	2	ex13	TRUE	TRUE
FANCA	ex12	16	89858797	89859036	NM_000135	2	ex12	TRUE	TRUE
FANCA	ex11	16	89862289	89862451	NM_000135	2	ex11	TRUE	TRUE
FANCA	ex10	16	89865330	89865666	NM_000135	2	ex10	TRUE	TRUE
FANCA	ex9	16	89865910	89866149	NM_000135	2	ex9	TRUE	TRUE
FANCA	ex8	16	89869639	89869778	NM_000135	2	ex8	TRUE	TRUE
FANCA	ex7	16	89871660	89871829	NM_000135	2	ex7	TRUE	TRUE
FANCA	ex6	16	89874653	89874798	NM_000135	2	ex6	TRUE	TRUE
FANCA	ex5	16	89877090	89877235	NM_000135	2	ex5	TRUE	TRUE
FANCA	ex4	16	89877288	89877527	NM_000135	2	ex4	TRUE	TRUE
FANCA	ex3	16	89880903	89881046	NM_000135	2	ex3	TRUE	TRUE
FANCA	ex2	16	89882260	89882419	NM_000135	2	ex2	TRUE	TRUE
FANCA	ex1	16	89882920	89883090	NM_000135	2	ex1	FALSE	FALSE
TP53	ex10	17	7572807	7573118	NM_000546	5	ex11	TRUE	TRUE
TP53	ex9	17	7573817	7574153	NM_000546	5	ex10	TRUE	TRUE
TP53	ex7-8	17	7576427	7577274	NM_000546	5	ex8-9	TRUE	TRUE
TP53	ex6	17	7577369	7577718	NM_000546	5	ex7	TRUE	TRUE
TP53	ex4-5	17	7577995	7578684	NM_000546	5	ex5-6	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
TP53	ex2-3	17	7579192	7580032	NM_000546	5	ex2-4	TRUE	TRUE
RAD51D	ex9-10	17	33427847	33428482	NM_001142571	1	ex9-10	TRUE	TRUE
RAD51D	ex7-8	17	33430088	33430687	NM_001142571	1	ex7-8	TRUE	TRUE
RAD51D	ex6	17	33433272	33433627	NM_001142571	1	ex6	TRUE	TRUE
RAD51D	ex4-5	17	33433895	33434605	NM_001142571	1	ex4-5	TRUE	TRUE
RAD51D	in3	17	33443801	33444166	NM_001142571	1	ex3	TRUE	TRUE
RAD51D	ex3	17	33445331	33445780	NM_001142571	1	in2	TRUE	TRUE
RAD51D	ex2	17	33446038	33446278	NM_001142571	1	ex2	TRUE	TRUE
RAD51D	ex1	17	33446412	33446771	NM_001142571	1	ex1	FALSE	FALSE
CDK12	ex1	17	37618106	37619504	NM_016507	3	ex1	FALSE	FALSE
CDK12	ex2	17	37627020	37628129	NM_016507	3	ex2	TRUE	TRUE
CDK12	ex3	17	37646699	37647097	NM_016507	3	ex3	TRUE	TRUE
CDK12	ex4	17	37648932	37649253	NM_016507	3	ex4	TRUE	TRUE
CDK12	ex5	17	37650667	37651003	NM_016507	3	ex5	TRUE	TRUE
CDK12	ex6	17	37657340	37657803	NM_016507	3	ex6	TRUE	TRUE
CDK12	ex7	17	37665848	37666101	NM_016507	3	ex7	TRUE	TRUE
CDK12	ex8	17	37667672	37667993	NM_016507	3	ex8	TRUE	TRUE
CDK12	ex9	17	37671874	37672171	NM_016507	3	ex9	TRUE	TRUE
CDK12	ex10	17	37673583	37673920	NM_016507	3	ex10	TRUE	TRUE
CDK12	ex11	17	37676102	37676451	NM_016507	3	ex11	TRUE	TRUE
CDK12	ex12	17	37680817	37681248	NM_016507	3	ex12	TRUE	TRUE
CDK12	ex13	17	37682006	37682680	NM_016507	3	ex13	TRUE	TRUE
CDK12	ex14	17	37686744	37687682	NM_016507	3	ex14	TRUE	TRUE
BRCA1	ex24	17	41197555	41197930	NM_007294	3	ex23	TRUE	TRUE
BRCA1	ex23	17	41199560	41199805	NM_007294	3	ex22	TRUE	TRUE
BRCA1	ex22	17	41201028	41201321	NM_007294	3	ex21	TRUE	TRUE
BRCA1	ex21	17	41202970	41203245	NM_007294	3	ex20	TRUE	TRUE
BRCA1	ex20	17	41208959	41209262	NM_007294	3	ex19	TRUE	TRUE
BRCA1	ex19	17	41215240	41215501	NM_007294	3	ex18	TRUE	TRUE
BRCA1	ex18	17	41215781	41216078	NM_007294	3	ex17	TRUE	TRUE
BRCA1	ex17	17	41219541	41219830	NM_007294	3	ex16	TRUE	TRUE
BRCA1	ex16	17	41222835	41223366	NM_007294	3	ex15	TRUE	TRUE
BRCA1	ex15	17	41226238	41226648	NM_007294	3	ex14	TRUE	TRUE
BRCA1	ex14	17	41228395	41228742	NM_007294	3	ex13	TRUE	TRUE
BRCA1	ex13	17	41231205	41231564	NM_007294	3	in12	TRUE	TRUE
BRCA1	ex12	17	41234311	41234703	NM_007294	3	ex12	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
BRCA1	ex11	17	41242789	41243159	NM_007294	3	ex11	TRUE	TRUE
BRCA1	ex10	17	41243335	41246994	NM_007294	3	ex10	TRUE	TRUE
BRCA1	ex9	17	41247753	41248050	NM_007294	3	ex9	TRUE	TRUE
BRCA1	ex8	17	41249151	41249484	NM_007294	3	ex8	TRUE	TRUE
BRCA1	ex7	17	41251682	41252007	NM_007294	3	ex7	TRUE	TRUE
BRCA1	ex6	17	41256104	41256429	NM_007294	3	ex6	TRUE	TRUE
BRCA1	ex5	17	41256775	41257084	NM_007294	3	ex5	TRUE	TRUE
BRCA1	ex4	17	41258351	41258660	NM_007294	3	ex4	TRUE	TRUE
BRCA1	ex3	17	41267633	41267906	NM_007294	3	ex3	TRUE	TRUE
BRCA1	ex2	17	41275924	41276223	NM_007294	3	ex2	TRUE	TRUE
BRCA1	ex1	17	41277088	41277611	NM_007294	3	5utr	FALSE	FALSE
RAD51C	ex1	17	56769856	56770260	NM_058216	2	ex1	FALSE	FALSE
RAD51C	ex2	17	56772182	56772732	NM_058216	2	ex2	TRUE	TRUE
RAD51C	ex3	17	56773944	56774330	NM_058216	2	ex3	TRUE	TRUE
RAD51C	ex4	17	56780447	56780825	NM_058216	2	ex4	TRUE	TRUE
RAD51C	ex5	17	56787106	56787461	NM_058216	2	ex5	TRUE	TRUE
RAD51C	ex6	17	56798023	56798282	NM_058216	2	ex6	TRUE	TRUE
RAD51C	ex7	17	56801291	56801572	NM_058216	2	ex7	TRUE	TRUE
RAD51C	ex8	17	56809701	56810016	NM_058216	2	ex8	TRUE	TRUE
RAD51C	ex9	17	56811369	56811694	NM_058216	2	ex9	TRUE	TRUE
BRIP1	ex20	17	59760543	59761622	NM_032043	2	ex20	TRUE	TRUE
BRIP1	ex19	17	59763044	59763643	NM_032043	2	ex19	TRUE	TRUE
BRIP1	ex18	17	59770713	59771082	NM_032043	2	ex18	TRUE	TRUE
BRIP1	ex17	17	59793188	59793547	NM_032043	2	ex17	TRUE	TRUE
BRIP1	ex16	17	59820280	59820649	NM_032043	2	ex16	TRUE	TRUE
BRIP1	ex15	17	59821570	59822054	NM_032043	2	ex15	TRUE	TRUE
BRIP1	ex14	17	59853573	59854052	NM_032043	2	ex14	TRUE	TRUE
BRIP1	ex13	17	59857470	59857834	NM_032043	2	ex13	TRUE	TRUE
BRIP1	ex12	17	59858098	59858571	NM_032043	2	ex12	TRUE	TRUE
BRIP1	ex11	17	59861446	59861925	NM_032043	2	ex11	TRUE	TRUE
BRIP1	ex10	17	59870844	59871203	NM_032043	2	ex10	TRUE	TRUE
BRIP1	ex9	17	59876321	59876800	NM_032043	2	ex9	TRUE	TRUE
BRIP1	ex8	17	59878485	59878964	NM_032043	2	ex8	TRUE	TRUE
BRIP1	ex7	17	59885644	59886243	NM_032043	2	ex7	TRUE	TRUE
BRIP1	ex6	17	59924330	59924809	NM_032043	2	ex6	TRUE	TRUE
BRIP1	ex5	17	59926374	59926733	NM_032043	2	ex5	TRUE	TRUE

CNV Target Regions (continued)

GENE	REGION	CHR.	START	END	TRANSCRIPT ID	TRANSCRIPT VERSION	TRANSCRIPT REGION	COVERED BY GENE-LEVEL ANALYSIS	COVERED BY EXON-LEVEL ANALYSIS
BRIP1	ex4	17	59934236	59934715	NM_032043	2	ex4	TRUE	TRUE
BRIP1	ex3	17	59937007	59937486	NM_032043	2	ex3	TRUE	TRUE
BRIP1	ex2	17	59938674	59939033	NM_032043	2	ex2	TRUE	TRUE
CCNE1	ex2	19	30303384	30303712	NM_001238	3	ex2-3	TRUE	FALSE
CCNE1	ex3	19	30303790	30304029	NM_001238	3	ex4	TRUE	FALSE
CCNE1	ex4-5	19	30307997	30308500	NM_001238	3	ex5-6	TRUE	FALSE
CCNE1	ex6	19	30311546	30311842	NM_001238	3	ex7	TRUE	FALSE
CCNE1	ex7	19	30312582	30312771	NM_001238	3	ex8	TRUE	FALSE
CCNE1	ex8	19	30312811	30313090	NM_001238	3	ex9	TRUE	FALSE
CCNE1	ex9-10	19	30313123	30313521	NM_001238	3	ex10-11	TRUE	FALSE
CCNE1	ex11	19	30314464	30314743	NM_001238	3	ex12	TRUE	FALSE
CHEK2	ex16	22	29083668	29084109	NM_007194	3	ex15	TRUE	TRUE
CHEK2	ex15	22	29084983	29085342	NM_007194	3	ex14	TRUE	TRUE
CHEK2	ex14	22	29089903	29090262	NM_007194	3	ex13	TRUE	TRUE
CHEK2	ex13	22	29090951	29091382	NM_007194	3	ex12	TRUE	TRUE
CHEK2	ex12	22	29091606	29091975	NM_007194	3	ex11	TRUE	TRUE
CHEK2	ex11	22	29092753	29093112	NM_007194	3	ex10	TRUE	TRUE
CHEK2	ex10	22	29095696	29096055	NM_007194	3	ex9	TRUE	TRUE
CHEK2	ex9	22	29099314	29099673	NM_007194	3	ex8	TRUE	TRUE
CHEK2	ex8	22	29105884	29106211	NM_007194	3	ex7	TRUE	TRUE
CHEK2	ex7	22	29107771	29108130	NM_007194	3	ex6	TRUE	TRUE
CHEK2	ex6	22	29115244	29115603	NM_007194	3	ex5	TRUE	TRUE
CHEK2	ex4-5	22	29120805	29121552	NM_007194	3	ex3-4	TRUE	TRUE
CHEK2	ex3	22	29126384	29126579	NM_007194	3	in2	TRUE	TRUE
CHEK2	ex2	22	29130302	29130911	NM_007194	3	ex2	TRUE	TRUE
CHEK2	ex1	22	29137610	29137969	NM_007194	3	5utr	FALSE	FALSE

Document Approvals

Approved Date: 30 Jun 2026

Approval Verdict: Approve	Technical Approval 26-Jun-2026 09:36:55 GMT+0000
QA Approval Verdict: Approve	Quality Specialist Quality Assurance Approval 30-Jun-2026 08:21:49 GMT+0000