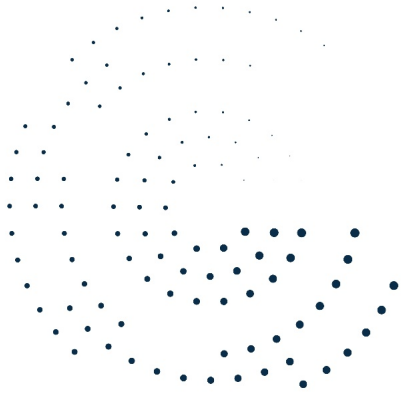
USER MANUAL

SOPHiA DDM™

GIInger Genomic Integrity Solution



SOPHiA DDM™

SUMMARY INFORMATION

Product Name	SOPHiA DDM™ GIInger Genomic Integrity Solution
Product Type	Drylab Solution
Product Family	Analytics
Product Version	v2.0
Sample Type	Somatic DNA isolated from formalin-fixed paraffin-embedded (FFPE) tumor tissue specimens
Document ID	SG-03803
Document Version	v1.4
Revision Date	Jun 2026

This User Manual is applicable to all SOPHiA DDM™ versions.

Please read the User Manual thoroughly before using this product.

SOPHiA GENETICS SA
ZA La Pièce 12
1180 Rolle
Switzerland

REF DL0121ILLRSM

RUO

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

NovaSeq™ is a trademark of Illumina, Inc.

Agilent Fragment Analyzer™ is a trademark of Agilent Technologies, 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-03803 – v1.4	Jun 2026	• *Document updates*: Minor changes to the text were performed and the following modifications were made:
		  • *7 Pipeline Deliverables*: Section fully rewritten with a list of **run-level** and **sample-level** output files.
		  • *10 Support*: Contact information has been updated.

SG-03803 – v1.3	May 2024	• Added instructions regarding Library Pooling For Low-Pass Whole Genome Sequencing (lpWGS) to section *6.2.1 Data Generation*. • Minor document restructuring and formatting / typo corrections.
SG-03803 – v1.2	Dec 2023	• Section *7.4.1 Genomic Integrity Status Equivalence With Comparator Assay*: Added information regarding additional equivalence study. • Section *8.1.2 Module-Specific Warnings*: Modified warning (2.) and inserted new warning (3.).
SG-03803 – v1.1	Nov 2023	• Section *7.1 Results Location* and *7.2 File Composition*: Updated text and screenshots to reflect that GI results will now also be made available via sample-specific PDF reports in which the Sample ID is provided un-anonymized.
SG-03803 – v1.0	Apr 2023	• Initial release

TABLE OF CONTENTS

1 PRODUCT INTRODUCTION

The SOPHiA DDM™ GIInger Genomic Integrity Solution is a bioinformatics workflow intended to facilitate the analysis of solid tumor samples from Ovarian Cancers. Users can detect Genomic Instability via the analysis of paired-end Whole Genome Sequencing (WGS) data. The presence of Genomic Instability in the DNA of a tumor sample is associated with Homologous Recombination Deficiency (HRD).

The product includes a bioinformatic pipeline for the calculation of a Genomic Integrity Index (GII) and the determination of a Genomic Integrity Status (GIS). The bioinformatic pipeline is integrated into the SOPHiA DDM™ Platform, which supports NGS data upload and results visualization.

The SOPHiA DDM™ GIInger Solution is for use on DNA samples extracted from formalin-fixed paraffin-embedded (FFPE) Ovarian Cancer samples and sequenced in WGS paired-end mode on Illumina® NovaSeq 6000. The solution is also technically compatible with paired-end WGS data obtained from Illumina® NextSeq 500/550 instruments.

The SOPHiA DDM™ GIInger Solution is a bioinformatics workflow intended for Research Use Only (RUO).

2 GENERAL STATEMENT OF THE TEST PRINCIPLES AND PROCEDURE

The SOPHiA DDM™ GIInger Genomic Integrity Solution is intended for the analysis of low-pass WGS data (paired-end, coverage ~1x) obtained via Illumina® NextSeq 500/550 and Illumina® NovaSeq 6000 instruments from DNA samples extracted from formalin-fixed paraffin-embedded (FFPE) Ovarian Cancer tissue.

The product performs data QA and determines the Genomic Integrity Status of Ovarian Cancer samples.

3 PRODUCT COMPONENTS

3.1 Components Included

The product consists of two major components: the analytical pipeline and the SOPHiA DDM™ Platform. The purpose of the pipeline is to analyze the sequencing data and return Genomic Integrity analysis results, while the Platform's purpose is to host the pipeline and serve as the interface for uploading sequencing data and visualizing results.

3.2 Equipment And Material Required (Not Provided)

The SOPHiA DDM™ GIInger Genomic Integrity Solution is a dry lab solution. All consumables, reagents, and equipment required for the generation of sequencing data are not included in the product.

4 INSTALLATION

4.1 Minimum System Requirements

4.1.1 Operating System

- Microsoft Windows 7 (64 bit) or later
- Microsoft Windows 10 (64 bit) or later
- MacOS X 10.0 or later

4.1.2 Internet Connection

Download

- >10 MB/s download speed internet connection
- You can test your connection to our servers here: https://www.sophiagenetics.com/downloads/

Upload

- The recommended upload speed is 20 MB/s. If hard to achieve with wireless, we recommend an ethernet connection.
- The upload limit per run is 100 GB.
  Note: If any of the individually uploaded FASTQ files have a size of >2.5 GB, they will be downsampled to 2.5 GB per file.
- User should ensure they have a strong upload connection by using https://ca.sophiagenetics.com/html5/index.html

Firewall

Users should make sure to whitelist traffic from all SOPHiA sub-domains *.sophiagenetics.com on port 443 and 80.

Specifically:

ca.sophiagenetics.com	www.broadinstitute.org
ddm.sophiagenetics.com	exac.broadinstitute.org
dev.sophiagenetics.com	cancer.sanger.ac.uk/cosmic
uploaderstoragecha.blob.core.windows.net	www.internationalgenome.org
lma.sophiagenetics.com/sgml_sga_1/ddm	evs.gs.washington.edu/EVS
www.ncbi.nlm.nih.gov/SNP	127.0.0.1:21000
www.ncbi.nlm.nih.gov/SNP	127.0.0.1:60151

In case user is using a proxy as well, once they have downloaded DDM, they will need to create a file in c:\Users\userfolder\.dg-cache\proxyCfg.txt with their own proxy information, for example:

```
-Dhttps.proxyHost=10.165.3.76
-Dhttps.proxyPort=8080
-Dhttp.proxyHost=10.165.3.76
-Dhttp.proxyPort=8080
```

4.2 SOPHiA DDM™ Platform Installation Instructions

For instructions on SOPHiA DDM™, please visit https://www.sophiagenetics.com/docs/ and download the SOPHiA DDM™ User Manual.

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® NextSeq® 500/550 And NovaSeq™ 6000 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.

  Note: The SOPHiA DDM™ Platform merges multi-lane FASTQ files before analysis: For example, 8 NextSeq® 500/550 files (4 lanes, forward + reverse reads) will be merged into 2 files (forward + reverse).

- 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

After selecting the FASTQ files to be analyzed, the user is prompted to specify/confirm the kit that was used for data generation. Select “SOPHiA DDM™ GIInger Genomic Integrity Solution”.

Before finalizing the upload request, please ensure that the correct kit – “SOPHiA DDM™ GIInger Genomic Integrity Solution” – 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® NextSeq® 500/550 And NovaSeq™ 6000 Folder Configuration

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

FolderRun1	
FolderSample1	FolderSample2
SampleIdentifier1_S01_L001_R1_001.fastq.gz	SampleIdentifier2_S02_L001_R1_001.fastq.gz
SampleIdentifier1_S01_L001_R2_001.fastq.gz	SampleIdentifier2_S02_L001_R2_001.fastq.gz
SampleIdentifier1_S01_L002_R1_001.fastq.gz	SampleIdentifier2_S02_L002_R1_001.fastq.gz
SampleIdentifier1_S01_L002_R2_001.fastq.gz	SampleIdentifier2_S02_L002_R2_001.fastq.gz

If the total volume of data produced by the sequencing run exceeds 100 gigabytes (GB), FASTQ files should be organized in multiple batches, which must be uploaded independently. The following example illustrates how data from a single run (Run1) is organized in two batches (Run1_Batch1 and Run1_Batch2).

FolderRun1_Batch1	
FolderSample1	FolderSample2
SampleIdentifier1_S01_L001_R1_001.fastq.gz	SampleIdentifier2_S02_L001_R1_001.fastq.gz
SampleIdentifier1_S01_L001_R2_001.fastq.gz	SampleIdentifier2_S02_L001_R2_001.fastq.gz
SampleIdentifier1_S01_L002_R1_001.fastq.gz	SampleIdentifier2_S02_L002_R1_001.fastq.gz
SampleIdentifier1_S01_L002_R2_001.fastq.gz	SampleIdentifier2_S02_L002_R2_001.fastq.gz

FolderRun1_Batch2	
FolderSample3	FolderSample4
SampleIdentifier3_S03_L001_R1_001.fastq.gz	SampleIdentifier4_S04_L001_R1_001.fastq.gz
SampleIdentifier3_S03_L001_R2_001.fastq.gz	SampleIdentifier4_S04_L001_R2_001.fastq.gz
SampleIdentifier3_S03_L002_R1_001.fastq.gz	SampleIdentifier4_S04_L002_R1_001.fastq.gz
SampleIdentifier3_S03_L002_R2_001.fastq.gz	SampleIdentifier4_S04_L002_R2_001.fastq.gz

Data upload for large runs that require batching may be performed automatically via the CLI-based approach (see section *5.7 CLI-Based Data Upload (Optional)*).

5.6 Sample Upload Requirements

5.6.1 Illumina® NextSeq® 500/550

Data upload should be performed manually following the instructions provided in section 2.2 of the SOPHiA DDM™ User Manual and section *5 Data Upload* of this document.

5.6.2 Illumina® NovaSeq™ 6000

Data upload can be performed either manually following the instructions provided in section 2.2 of the SOPHiA DDM™ User Manual and section *5 Data Upload* of this document or via the CLI tool (see section *5.7 CLI-Based Data Upload (Optional)*).

The maximum run size supported for a single upload (i.e., the total size of all FASTQ files uploaded simultaneously) is 100 GB. Data from sequencing runs exceeding 100 GB should be split and uploaded in multiple batches.

5.7 CLI-Based Data Upload (Optional)

This command-line interface (CLI) utility takes advantage of the SOPHiA DDM™ Application Programming Interface (API), allowing the automatic upload of raw sequencing data. It enables scheduled upload directly after sequencing runs and a light integration with local Laboratory Information Management Systems (LIMS).

Setup of this functionality needs support from SOPHiA GENETICS' IT team and familiarity with command-line utilities. Please contact the support team at support@sophiagenetics.com.

This utility is a Java tool that is delivered separately, and the user guide available in the link below helps to get started with the tool. It has no graphical user interface and can only be used through the command-line.

For further instructions, see https://api-fr.sophiagenetics.com/uploader/cli/.

6 ANALYSIS

6.1 Analysis Workflow Description

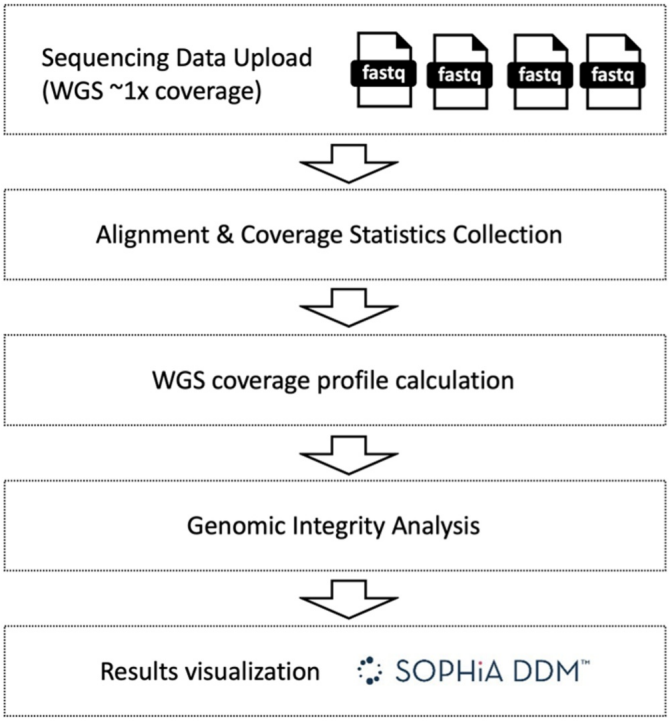


Figure 1: Analysis workflow overview for SOPHiA DDM™ GIInger Genomic Integrity Solution

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 provider's Reference Guide for the kit and Whole Genome Sequencing of these samples on Illumina® NextSeq® 500/550 or NovaSeq™ 6000 sequencers.

Sequencing should be performed in paired-end mode (2 x 150 bp cycles). Sequencing multiplexing should be performed to target 12 M clusters (i.e., DNA fragments) per sample, corresponding to a targeted coverage of ~1x. FASTQ files exceeding 2.5 GB (corresponding to approximately 24 M clusters) will be truncated by the pipeline before bioinformatics analysis.

Some data generation kits contain molecular barcodes (UMIs). These barcodes need to be trimmed from the FASTQ files before the upload. This can be done either by providing appropriate sample sheets to the sequencing facility or by pre-trimming the UMI corresponding region in the FASTQ files before uploading the data.

Library Pooling For Low-Pass Whole Genome Sequencing (lpWGS)

For data generation on Illumina® sequencers, the following recommendations for normalization and dilution of individual WGS libraries should be followed.

Materials

- Individual libraries
- IDTE
- DNA low-binding 1.5 ml tubes

Procedure

1. Determine the molarity of each library using the average size of the library (peak size in base pairs) and concentration (ng/µl) as follows:

$$\text{Library molarity (nM)} = \frac{\text{Library concentration (ng/µl)}}{\text{Average size in base pairs} \times 649.5} \times 10^6$$

2. Transfer 2 µl of each library individually into a new tube and dilute to 10 nM with IDTE.
3. Pool individual libraries at 10 nM by combining 5 µl from each library dilution to create a WGS pool.

Safe stopping point overnight at 4 °C or -20 °C for longer storage.

Pay attention to adapter dimers during quality checks of whole genome libraries using capillary electrophoresis systems, such as Agilent Fragment Analyzer™ or Bioanalyzer™. Adapter dimers manifest as small molecular weight peaks around 120–170 bp. Adapter dimers contain full-length sequencing adapters and are able to bind to Illumina® sequencing flow cells and generate clusters. Large proportions of adapter dimers may significantly reduce the number of sequenced library fragments, consequently reducing the number of mapped reads available for Genomic Integrity analysis. Note that patterned flow cells are more sensitive to adapter dimer presence. Illumina® recommends to sequence libraries with no more than 0.5% adapter dimers on patterned flow cells and no more than 5% adapter dimers on non-patterned flow cells.

6.2.2 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)
- NovaSeq™ 6000 sequencer (e.g., NovaSeq™ 6000 Sequencing Systems Guide, Document #1000000019358 v17)

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 is performed against the GRCh37 reference genome (the previous major release of the human reference genome, also referred to as hg19).

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

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

6.3.2 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.
- Down-sample the FASTQ files to a maximum file size of approximately 2.5 GB per sample (if applicable).

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.

6.3.3 Included Analysis Modules

The different results that this pipeline returns are: Genomic Integrity analysis

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 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 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: GI Status determination:

A 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:

- *Nb. WGS fragments:* Total number of DNA fragments available for the raw coverage WGS profile calculation. This QA metric is compared against a threshold of 4 million to determine the GI QA status.
- *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.

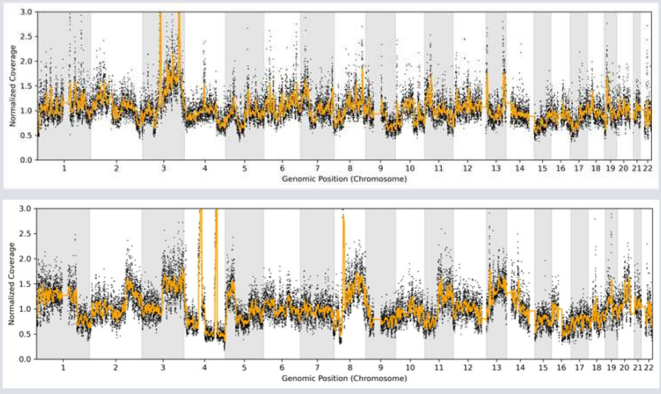
- *QA Status:* GI quality assessment status of a sample.
- *GI Index:* The GI index is a scalar value, ranging between −20 and 20, that reflects the level of genomic integrity. High GI indices reflect low levels of genomic integrity. Low GI indices reflect high levels of genomic integrity.
- *GI Status:* The genomic integrity status of a sample.

6.4 Considerations Regarding Negative*, Rejected, And Inconclusive Genomic Integrity Results

6.4.1 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 *Table 1* below to identify the underlying cause.

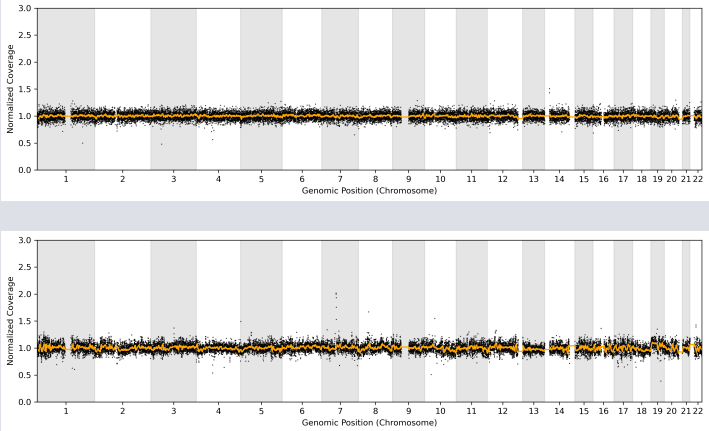
Table 1: Recommendations in case of GI Rejection

ROOT CAUSE	HOW TO IDENTIFY THE ROOT CAUSE	RECOMMENDATION
Insufficient coverage due to issues in 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 resequencing your samples, ensuring to quantify your libraries carefully and to adjust all samples to the same molarity before pooling.
Insufficient coverage due to a 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 it to the expected number of reads for the specific sequencing instrument used.	Repeat the sequencing run adjusting the final sequencing library loading concentration to avoid under- or over-clustering of the run.
Excess of residual noise and/or 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%. • Check the sample QA section of the 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 and/or proportion of coverage outliers can be a sign of low-quality input DNA and/or suboptimal execution of the NGS experiment. We recommend repeating the whole NGS experiment, if possible, starting with a new DNA sample of higher quality.

6.4.2 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 2* below to identify the underlying cause.

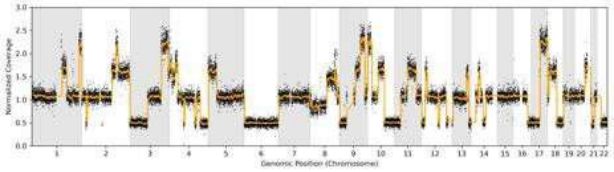
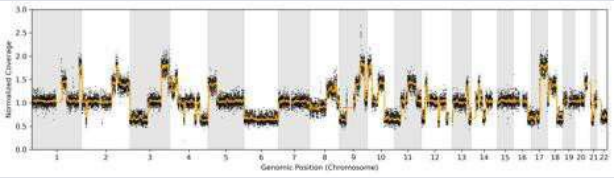
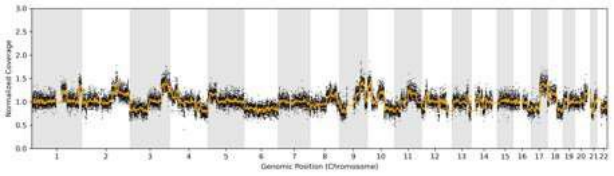
Table 2: Recommendations in case of GI Inconclusive status

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.

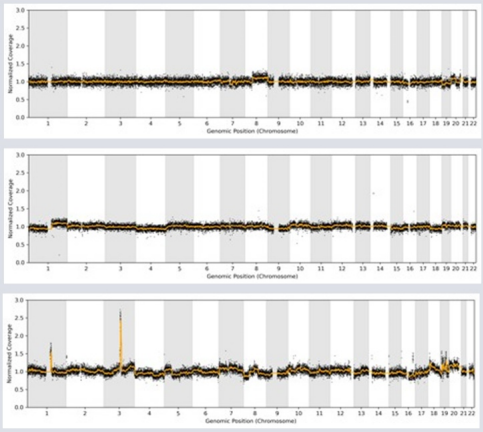
6.4.3 In Case Of GI Negative* Status

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 3* on the next page illustrates how insufficient tumor content impacts SNR.

Table 3: 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
		

7 PIPELINE DELIVERABLES

The following run-level output files are computed:

- **QA-report.pdf**: The pipeline QA report in PDF format.
- **GI_qc_table.txt**: A text file summarizing the GI quality assessment metrics.
- **GI_status_table.txt**: A text file summarizing the anonymized GI results.
- **Unanonymised_GI_status_table.txt**: A text file summarizing the un-anonymized GI results. This table contains the same results also provided in *GI_status_table.txt.*
- **GI-report.pdf**: The anonymized report for the genome integrity analysis in PDF format.

The following sample-level output files are computed:

- **<SAMPLE ID>- QA-patient.pdf**: The pipeline QA report in PDF format.
- **<SAMPLE ID>- GI- Report.pdf**: The un- anonymized report for the genomic integrity analysis in PDF format. This report contains the same results also provided in GI-report.pdf.

The SOPHiA DDM™ platform allows to download un-anonymized genomic integrity analysis reports for all samples included in the run via the request-level output files download tab.

7.1 Genomic Integrity Analysis Characteristics

7.1.1 Genomic Integrity Status Equivalence With Comparator Assay

The performance of the SOPHiA DDM™ GIInger Genomic Integrity Solution has been evaluated through equivalence studies performed against the SOPHiA DDM™ HRD Solution.

Library Preparation Via SOPHiA DDM™ HRD Solution

129 DNA samples (50 ng, DQN>3, tumor content>30%) extracted from FFPE Ovarian Cancer tissue have been processed using the library preparation kit provided as part of the SOPHiA DDM™ HRD Solution. Sequencing was performed on an Illumina® NovaSeq™ 6000 instrument. The resulting NGS data have been analyzed with the bioinformatics pipeline provided as part of the SOPHiA DDM™ HRD Solution (v2) to obtain Genomic Integrity Analysis results. In parallel, whole-genome sequencing (10 M clusters per sample) of the libraries obtained during this process was performed on Illumina® NovaSeq™ 6000 and processed with the SOPHiA DDM™ GIInger Genomic Integrity Solution.

See *Table 4* below for a comparison of the Genomic Integrity Statuses obtained with the two solutions.

Table 4: Comparison of Genomic Integrity Status assignment via SOPHiA DDM™ GIInger Genomic Integrity Solution and SOPHiA DDM™ HRD Solution

		SOPHIA DDM™ GIINGER GENOMIC INTEGRITY SOLUTION GENOMIC INTEGRITY STATUS				
		Positive	**Negative**	**Negative***	**Rejected**	**Inconclusive**
SOPHIA DDM™ HRD SOLUTION GENOMIC INTEGRITY STATUS	**Positive**	63	-	-	-	-
	Negative	-	58	-	-	-
	Negative*	-	2	4	-	-
	Rejected	-	1	-	1	-
	Inconclusive	-	-	-	-	-

Library Preparation Via SOPHiA GENETICS Universal Library Prep Kit

23 DNA samples (50 ng, DQN>3, tumor content>30%) extracted from FFPE Ovarian Cancer tissue have been processed using the SOPHiA DDM™ HRD Solution. The same samples were processed with the SOPHiA GENETICS Universal Library Preparation (ULP) kit, whole-genome sequenced (10 M clusters per sample), and analyzed with the SOPHiA DDM™ GIInger Genomic Integrity Solution. The equivalence was assessed by comparing the GI statuses obtained with the two solutions. **The observed GI statuses were concordant in all 23 samples.**

7.1.2 Limit Of Detection

The limit of detection for Genomic Integrity analysis was measured *in silico* using all WGS data described in section *7.1.1 Genomic Integrity Status Equivalence With Comparator Assay* for which:

i. Genomic Integrity Index obtained via the SOPHiA DDM™ GIInger Genomic Integrity Solution was >3 and

ii. sample tumor content could be confidently established.

The 36 samples fulfilling these criteria were serially diluted *in silico* using a normal control (to mimic sample tumor contents ranging between 0% and the original sample tumor content) and analyzed via the SOPHiA DDM™ GIInger Genomic Integrity Solution. A limit of detection of 25% tumor content was observed when measuring the lowest tumor content, at which 95% of the Genomic Integrity calls were Positive.

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.

8.1.2 Module-Specific Warnings

1. The Genomic Integrity Status only applies to Ovarian Cancer samples.
2. The performance of the SOPHiA DDM™ GIInger Genomic Integrity Solution has only been assessed on DNA (50 ng) extracted from Ovarian Cancer FFPE samples with DQN ≥3 (DQN defined as one-tenth of the percentage of DNA sample fragments >300 bp), measured using the Agilent Fragment Analyzer™ system), tumor content ≥30%, and prepared with:
    - The library preparation kit provided as part of the SOPHiA DDM™ HRD Solution and sequenced in WGS mode on Illumina® NovaSeq™ 6000 or
    - The SOPHiA GENETICS Universal Library Preparation (ULP) kit and sequenced in WGS mode on Illumina® NovaSeq™ 6000
3. The limit of detection of the SOPHiA DDM™ GIInger Genomic Integrity Solution has only been assessed on DNA (50 ng) extracted from Ovarian Cancer FFPE samples with DQN ≥3 (DQN defined as one-tenth of the percentage of DNA sample fragments >300 bp, measured using the Agilent Fragment Analyzer™ system), tumor content ≥30%, prepared with the library preparation kit provided as part of the SOPHiA DDM™ HRD Solution, and sequenced in WGS mode on Illumina® NovaSeq™ 6000.
4. Poor quality of raw NGS data can confound the data analysis and cause False Positives, False Negatives, and Inconclusive results.
5. The Genomic Integrity Status computed by SOPHiA DDM™ GIInger Genomic Integrity Solution does not take into consideration the mutational status of BRCA1 and BRCA2, as well as other genomic aberrations known to cause Homologous Recombination Deficiency. A non-positive Genomic Integrity Status does not exclude the presence of mutations in BRCA1 and BRCA2 (as well as in other genes).

8.2 Limitations

8.2.1 General Limitations

1. The product is designed to process, in a single request, the data volume produced by a single Illumina® NextSeq® 500/550 run (Mid Output and High Output kits).
2. The maximal request size cannot exceed 100 GB. Illumina® NovaSeq™ 6000 runs exceeding this limit cannot be processed via a single request and must be split into multiple batches.
3. The maximal file size per sample cannot exceed 10 GB. The pipeline merges multi-lane FASTQ files and performs random read subsampling to all compressed FASTQ files exceeding 2.5 GB of memory.
4. Even if sample multiplexing recommendations are followed, the total number of reads may be insufficient to provide conclusive results for various reasons, including poor sample quality, poor NGS data quality, and significant uneven read allocation between samples multiplexed in the same run.
5. The product does not detect nor report sample cross-contamination events.
6. The product does not measure nor report sample tumor content.
7. The limit of detection (LoD) for Genomic Integrity status determination is 30% tumor content. 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.
8. Samples with Negative* Genomic Integrity Status feature a higher risk of False Negative results.
9. Low-pass WGS data do not allow to reliably distinguish samples lacking large copy number alterations from samples having insufficient tumor content. Processing samples that do not feature large copy number alterations will lead to GI inconclusive results.

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.

Document Approvals

Approved Date: 29 Jun 2026

Approval Verdict: Approve	Technical Approval 26-Jun-2026 09:34:59 GMT+0000
QA Approval Verdict: Approve	Quality Assurance Approval 29-Jun-2026 07:38:57 GMT+0000