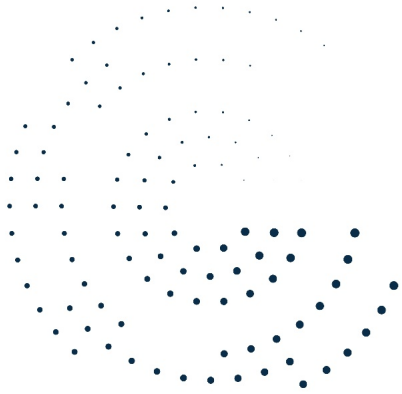
INSTRUCTIONS FOR USE

16, 32, 48 and 96 Samples

SOPHiA DDM™

Clinical Exome Solution v3



Using the SOPHiA GENETICS™ DNA Library Prep Kit I

SOPHiA GENETICS™

SUMMARY INFORMATION

Product Name	SOPHiA DDM™ Clinical Exome Solution v3
Product Type	Bundle Solution
Product Family	Molecular kit + analytics
Algorithm ID	ILL1XG1G5_CNV_exome_NextSeq
Gene Panel ID	CES_v3_hg38
Product Version	v1.0
Release Version	v5.10.3 - p5.5.45
Sample Type	Germline DNA isolated from peripheral blood
Sequencer	Illumina® NextSeq® 500/550
Document ID	SG-00164
Document Version	v3.2
Revision Date	Jan 2024

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

BS0102ILLRGLL01-016;
BS0102ILLRGLL01-032;
BS0102ILLRGLL01-048;
BS0102ILLRGLL01-096

RUO

PRODUCT CODES

	FULL PRODUCT CODE	BOX 1	BOX 2	LIBRARY PREPARATION KIT
REF	BS0102ILLRGLL01-016	B1.01.0002.R-16	B2.0000.R-16	700232
	BS0102ILLRGLL01-032	B1.01.0002.R-32	B2.0000.R-32	700232
	BS0102ILLRGLL01-048	B1.01.0002.R-48	B2.0000.R-48	700234
	BS0102ILLRGLL01-096	B1.01.0002.R-96	B2.0000.R-48	700234

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

AMPure® and Agencourt® are registered trademarks of Beckman Coulter, Inc.

Qubit® and Dynabeads® are registered trademarks of Thermo Fisher Scientific Inc.

KAPA™ is a trademark of Roche.

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-00164 – v3.2	Jan 2024	• Section *4.5 Sequencing*: Added instructions for MiSeq® • Section *5.1.2 Data Generation*: Added MiSeq® to the list of compatible sequencers • Section *7.2.2 Module-Specific Limitations - CNVs - General Recommendations/Limitations*: Added MiSeq®-specific sample number recommendations to list item (3.) • Section *8 Symbols*: Added missing symbols
SG-00164 – v3.1	Oct 2023	• Section 6.2 - Added description of new output file "Exon_coverage_stats_ complete.txt" and modified descriptions of "Exon_coverage_stats_v3.txt" and "Target_coverage_stats.txt". • Section 10.1.5 - Added clarifications regarding the orientation of the i5 index sequence depending on the used sequencer

SG-00164 – v3.0	Sep 2023	• Removed mention of third-party provider's intellectual property from sections *4.1.2 Kit Content*, *4.3.1 Library Pooling for Hybridization and Capture*, and *4.4.1 Hybridization*.
		• Sequencer field added to the Summary Information table.
		• Intended Use section title changed to Product Introduction.
		• Section 5.1.1 - Installation instructions referred to the SOPHiA DDM™ User Manual.
		• Formatting changes to the whole document.
		• Section 5.3.1 - Added information regarding pipeline parameters
		• Section 5.3.2
		  • SNVs/INDELs - Modified reference to variant filtering step; added information regarding homolog and pseudogene processing; added information regarding gene annotations
		  • CNVs - Modified table detailing genes/regions included in panel
		• Section 6.2 - Modified information regarding Exon_coverage_stats_v3.txt file and added reference to Target_coverage_stats.txt file
		• Section 6.3 - Added information regarding display of mitochondrial variants
		• Section 6.4 - Added information regarding custom filters
		• Section 6.5.2 - Added clarification regarding variants in genes with >1 possible annotation and regarding pseudogene/homolog warnings
		• Added new section 6.5.5 Homologs And Pseudogenes
		• Section 7.2.2

DOCUMENT ID / VERSION	DATE	DESCRIPTION OF CHANGE
		• SNVs/INDELs - Modified limitation 10. and added new limitation 11. • Mitochondrial Analysis - Modified limitation 3. and added new limitation 4.
SG-00164 – v2.0	Feb 2023	• Section 7.2 - text amended to include SNV/INDEL, CNV targets and target genes . • Section 7.2.2 - table amended to include the correct number of target regions and genes . • Section 7.3 - point 1, bullet point 3 - addition of low coverage regions at bullet point 3, point 1. • Section 7.3 - point 2 - additional information provided for exon_coverage_ stats_v3.txt and flagged_regions.txt, addition of flagged_low_coverage.txt as subpoint f. • Section 5.2 & 7.3 - all screenshots updated to reflect the latest user interface. • Section 7.3 - point 5, subpoint 1 - UNIPROT added to Gene level data. • Section 7.3 - point 5, subpoint 2 - table updated to include the last 3 rows: uniprot_acc, HGVS_gnomen and HGVS_pnomem.
ID-60101-48 – v1.1	Mar 2022	• Document ID - version and language added as a header throughout. • Section 1 Product Information - minor modifications. • Section 4.2.4 and 4.3.6 Procedure - note added before step 1 . • Section 4.2.5 Post-Ligation Clean Up (step 3 and 11). • Section 4.2.6 Dual Size Selection (step 3 and 6). • Section 4.2.8 Post-Amplification Clean Up (step 3 and 11). • Section 4.3.7 Post-Capture Amplification Clean Up (step 3 and 11) - "3 minutes" modified to "5 minutes". • Section 5.1 - FAQ link rectified. • Section 7.2 - note added on obtaining full list of targets. • Section 8.3.2 CNVs - general recommendations/limitations section elaborated. • Minor language changes for clarity, abbreviations expanded.
ID-60101-48 – v1.0	Jul 2021	• Initial release.

TABLE OF CONTENTS

1 PRODUCT INTRODUCTION

SOPHiA DDM™ Clinical Exome Solution v3 is a bundle solution including library preparation and capture kit, as well as the fully integrated, FASTQ-to-report bioinformatic workflow for the analysis, interpretation and reporting of sequencing data obtained on an Illumina® NextSeq® 500/550 instrument. The performance of the workflow has been optimized for use in the analysis and interpretation of sequencing data for the detection of germline variants in genes known to be linked to Mendelian and rare disorders.

The workflow consists of a secondary analysis pipeline and tertiary analytical tools and resources. The secondary analysis pipeline detects Single Nucleotide Variants (SNVs), Insertions/Deletions (INDELs), Copy Number Variants (CNVs) and mitochondrial variants in the mitochondrial genome, but also ~200 non coding pathogenic variants known to be associated with rare diseases. The product has been designed and built using the GRCh38/hg38 genome assembly. The tertiary analysis tools and resources aid in the interpretation of variants related to inherited and rare diseases, covering the majority of coding genes known to be associated with Mendelian diseases.

2 GENERAL STATEMENT OF THE TEST PRINCIPLES AND PROCEDURE

The SOPHiA DDM™ Clinical Exome Solution v3 is intended for preparation of DNA libraries from blood samples and the analysis and interpretation of germline genomic sequencing data.

3 PRODUCT COMPONENTS

The SOPHiA DDM™ Clinical Exome Solution v3 consists of three major components: DNA library preparation and capture kit, the analytical pipeline and the platform.

The purpose of the DNA library preparation and capture kit is to allow preparation of libraries from blood samples suitable for sequencing on an Illumina® sequencing platform. The purpose of the pipeline is to analyze the sequencing data obtained from an Illumina® NextSeq® 500/550 sequencer and return variant calls (i.e. secondary analysis), while the purpose of the platform is to host the pipeline and serve as the interface for the upload of sequencing data and tertiary analysis.

4 PROTOCOL

4.1 Materials

4.1.1 Initial Considerations

Upon receipt, please ensure that all tubes are physically intact and stored at recommended temperatures for optimum kit performance. Inappropriate handling and storage of the kit components under other conditions may adversely affect the kit's performance.

4.1.2 Kit Content (16, 32, 48 And 96 Samples)

Always briefly spin the tubes before use to collect all liquid.

Depending on the kit format, the following components are provided:

COMPONENT	NUMBER OF ITEMS DEPENDING ON KIT FORMAT			
	16 SAMPLE KIT	32 SAMPLE KIT	48 SAMPLE KIT	96 SAMPLE KIT
Box 1	1	1	1	2 (48 samples each)
Illumina®-compatible Dual Index Adapters (in 96-well plate format included in Box 1)	16	32	48	96 (Plate contained in one of the two Box 1)
Box 2	1	1	1	2 (48 samples each)
SOPHiA GENETICS™ DNA Library Prep Kit I	1	2 (16 samples each)	1	2 (48 samples each)

Box 1 (Store At -25°C To -15°C)

- Universal Blockers - TS Mix (12 µl)
- Human Cot DNA (25 µl)
- Clinical Exome Solution v3 Probes by SOPHiA GENETICS (20 µl)
- 2x Hybridization Buffer (50 µl)

- Hybridization Buffer Enhancer (20 µl)
- 2x Bead Wash Buffer (1250 µl)
- 10x Stringent Wash Buffer (200 µl)
- 10x Wash Buffer I (160 µl)
- 10x Wash Buffer II (110 µl)
- 10x Wash Buffer III (110 µl)
- Depending on the kit format: 16, 32, 48 or 96 Illumina®-compatible Dual Index Adapters in 96-well plate format (7 µl adapter per well). See Appendix I for adapter display and sequences.

Box 2 (Store At +2°C To +8°C)

- Dynabeads® M-270 Streptavidin (440 µl)
- Agencourt® AMPure® XP (3 x 1.5 ml for 16 samples, 8.7 ml for 32 samples and 11.6 ml for 48 samples, see Note below for 96 samples)
- IDTE Low TE Buffer (10 ml)
- Nuclease-free water (20 ml)

Note: For 96 samples, two times Box 2 of 48 samples are provided (see table on previous page).

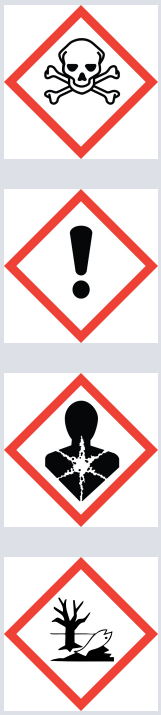
4.1.3 SOPHiA GENETICS™ DNA Library Prep Kit I (Store At -25 °C To -15 °C)*

REAGENT	KIT FORMAT	
	16 SAMPLES	48 SAMPLES
HiFi PCR Master Mix 2x (µl)	500	1560
Primer Mix Illumina® Library Amp (µl)	30	95
Fragmentation Enzyme Mix (µl)	200	625
Fragmentation Buffer 10x (µl)	100	315
Fragmentation Enhancer (µl)	100	315
Ligation Enzyme (µl)	200	625
Ligation Buffer 5x (µl)	400	1250

* SOPHiA GENETICS is the exclusive distributor of this Library Prep kit.

Refer to Warnings and Precautions below for additional details.

4.1.4 Warnings and Precautions

PRODUCT NAME	GHS PICTOGRAM	H&P STATEMENTS	SIGNAL WORD	HAZARDOUS COMPONENT
2x Hybridization Buffer		• H300 Fatal if swallowed. • H311 Toxic in contact with skin. • H315 Causes skin irritation. • H370 Causes damage to organs. • H370 Causes damage to organs (Central nervous system). • H411 Toxic to aquatic life with long-lasting effects. • P260 Do not breathe vapor/ spray. • P264 Wash contaminated skin thoroughly after handling. • P270 Do not eat, drink or smoke when using this product. • P273 Avoid release to the environment. • P280 Wear protective gloves/ protective clothing/ eye protection/face protection. • P301+P310 If swallowed: Immediately call a poison center/doctor. • P302+P352 If on skin: Wash with plenty of water. • P308+P311 If exposed or concerned: Call a poison center or doctor. • P321 Specific treatment (see medical advice on this label). • P330 Rinse mouth. • P332+P313 If skin irritation occurs: Get medical advice/ attention. • P362+P364 Take off contaminated clothing and wash it before reuse. • P391 Collect spillage. • P405 Store locked up. • P501 Dispose of contents/ container in accordance with national regulations.	Danger	Tetramethyl- ammonium chloride
Hybridization Buffer Enhancer		• H351 Suspected of causing cancer. • H360 May damage fertility or the unborn child. • H373 May cause damage to	Danger	Formamide

PRODUCT NAME	GHS PICTOGRAM	H&P STATEMENTS	SIGNAL WORD	HAZARDOUS COMPONENT
		organs through prolonged or repeated exposure. • P201 Obtain special instructions before use. • P202 Do not handle until all safety precautions have been read and understood. • P260 Do not breathe vapour/ spray. • P280 Wear protective gloves/ protective clothing/ eye protection/ face protection. • P308+P313 IF exposed or concerned: Get medical advice/ attention. • P314 Get medical advice/ attention if you feel unwell. • P405 Store locked up. • P501 Dispose of contents/ container in accordance with national regulations.		
10x Stringent Wash Buffer		• H302 Harmful if swallowed. • H315 Causes skin irritation. • H319 Causes serious eye irritation	Danger	Ethylenediaminetetraacetic acid disodium salt
10x Wash Buffer I		• H228 Flammable solid. • H302 Harmful if swallowed. • H315 Causes skin irritation. • H318 Causes serious eye damage. • H332 Harmful if inhaled. • H401 Toxic to aquatic life. • H402 Harmful to aquatic life. • H412 Harmful to aquatic life with long lasting effects. • P273 Avoid release to the environment. • P280 Wear protective gloves/ protective clothing/ eye protection/ face protection. • P305+P351+P338 If in eyes: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. • P310 Immediately call a poison	Danger	Sodium dodecyl sulfate

PRODUCT NAME	GHS PICTOGRAM	H&P STATEMENTS	SIGNAL WORD	HAZARDOUS COMPONENT
		center/ doctor. • P501 Dispose of contents/ container in accordance with national regulations.		
Fragmentation Buffer		• H302 Harmful if swallowed. • H370 Causes damage to organs. • H412 Harmful to aquatic life with long lasting effects. • P273 Avoid release to the environment. • P260 Do not breathe vapor. • P270 Do not eat, drink or smoke when using this product. • P264 Wash thoroughly after handling. • P308 + P311 IF exposed: Call a POISON CENTER or doctor. • P301 + P312, P330 IF SWALLOWED: Call a POISON CENTER or doctor if you feel unwell. Rinse mouth. • P405 Store locked up. • P501 Dispose of contents and container in accordance with all local, regional, national and international regulations.	Danger	Tetramethyl-ammonium chloride

Please use  and  as personal protective equipment.

4.1.5 Materials Required (Not Provided)

User-Supplied Materials (To Be Purchased Separately)

Lab-related plasticware

- DNA low binding 1.5 ml tubes
- 50 ml conical tubes
- Filter tips
- 1.5 ml tubes
- RNase/DNase-free 0.2 ml 8-tube strips

Reagents

- Ethanol (molecular biology grade)
- KAPA Library Amplification Kit KK2620 (Roche Cat. No: 07958978001)

Other

- Ice to ensure reagents and pre-mixes are kept at 4 °C
- Illumina® sequencing reagents

Laboratory Equipment

To avoid sample contamination, separate the workspace into:

- *Pre-PCR zone*
    - Fluorometric quantitation equipment and reagents (e.g., Qubit® or similar)
    - Magnetic separation rack (96-well type)
    - Multichannel pipettes (P10 or P20; P100; P200) with appropriate filter tips
    - Tabletop microcentrifuge (8-tube strips compatible)
    - Thermal cycler (programmable heated lid; 0.2 ml PCR tubes compatible)
    - Vortex mixer
- *Post-PCR zone*
    - Capillary electrophoresis system
    - DNA vacuum concentrator (e.g., SpeedVac™ or similar)
    - Fluorometric quantitation equipment and reagents (e.g., Qubit® or similar)
    - Magnetic separation rack (1.5 ml tube compatible)
    - Magnetic separation rack (96-well type)
    - Multichannel pipettes (P10 or P20; P100; P200) with appropriate filter tips
    - Tabletop microcentrifuge (8-tube strips compatible)
    - Thermal cycler (programmable heated lid; 0.2 ml PCR tubes compatible)
    - Thermoblock or water bath (1.5 ml tube compatible)
    - Vortex mixer

4.2 Library Preparation

4.2.1 Input Material Preparation

Materials

- Double-stranded high quality genomic DNA (gDNA)
- FX Enhancer
- IDTE
- RNase/DNase-free 0.2 ml 8-tube strips

Input Recommendations

DNA integrity, concentration, and purity are critical during this step. The purity of the DNA can be assessed using a UV spectrophotometer. Recommended absorbance ratios are between 1.8–2.0 for the 260/280 nm ratio and within 1.6–2.4 for 260/230 nm. We recommend confirmation of the sample integrity by capillary electrophoresis or an equivalent technique.

To avoid mistakes with DNA input, an initial dilution to obtain a concentration in the 50–100 ng/µl range is recommended. The DNA concentration should be confirmed by a fluorometric quantification method (e.g., Qubit®, Thermo Fisher), and the obtained value should be used to calculate the final dilution.

Preparation

Remove the FX Enhancer from -20 °C storage and thaw at room temperature. After thawing, mix the reagent by gently inverting the tube 5 times and briefly spin in a microcentrifuge.

Depending on the kit format, the number of DNA samples to be pooled per capture reaction will vary according to the following table. This must be taken into consideration before starting.

KIT FORMAT	16 SAMPLE KIT	32 SAMPLE KIT	48 SAMPLE KIT	96 SAMPLE KIT*
Number of individual libraries per capture	4	8	12	12

*For 96 samples, two 48 sample kits are provided which include 8 capture reactions.

Procedure

1. Prepare the following PCR strips according to the number of reactions:

NUMBER OF REACTIONS	4	8	12	16	24	32	48
PCR strip	**4-tube**	**4-tube**	**4-tube**	**8-tube**	**8-tube**	**8-tube**	**8-tube**
Number of strips	1	2	3	2	3	4	6

2. Prepare a dilution for each extracted DNA sample into the appropriate number of PCR strips, in the following manner:

DNA DILUTION	
DNA	200 ng
IDTE	Complete to 30 µl

3. Mix briefly by gently pipetting up and down 5 times followed by a brief spin in a microcentrifuge to collect all liquid.

Safe stopping point overnight at 4 °C.

4. Depending on the number of samples, proceed as follows:
    - If processing **4 samples**, add 5 µl of FX Enhancer to each tube of the 4-tube strip containing 30 µl of DNA samples (total of 35 µl in each tube of the 4-tube strip).
    - If processing **8 or more samples**, proceed as follows:
        a. To facilitate pipetting, create a reservoir of FX Enhancer by adding the following volumes to a new set of 4 or 8-tube strips, according to the following scheme:

NUMBER OF REACTIONS	8	12	16	24	32	48
PCR Strip (1 strip)	**4-tube**	**4-tube**	**8-tube**	**8-tube**	**8-tube**	**8-tube**
FX Enhancer (µl)	11.5	17.5	11.5	17.5	24	36

        b. Using a multichannel pipette, add 5 µl of FX Enhancer from the above tubes to 30 µl of DNA samples (total of 35 µl in each tube of the 4 or 8-tube strips).
    - Using a multichannel pipette set to 20 µl, mix gently by pipetting up and down 5 times and briefly spin in a microcentrifuge.
5. Keep on ice until Enzymatic Fragmentation reaction set-up.

4.2.2 Pre-Mixes and Reagents Preparation

Components And Reagents

- FX Enzyme Mix
- FX Buffer 10x
- Ligation Enzyme
- Ligation Buffer 5x
- HiFi PCR Master Mix 2x
- Primer Mix Illumina® Library Amp
- Nuclease-free water

- AMPure® XP beads
- Ethanol

Preparation

1. Remove the SOPHiA GENETICS™ DNA Library Prep Kit I components from -20 °C storage and thaw on ice.
2. Remove the Dual Index Adapters Plate from -20 °C storage and put it into 4 °C refrigerator for later use.
3. Remove the AMPure® XP beads from 2–8 °C storage and let them equilibrate at room temperature for at least 30 minutes. Vortex the AMPure® XP beads thoroughly, to ensure proper resuspension of the beads, prior to use in all subsequent steps.
4. Prepare fresh 80% Ethanol (volume according to the following scheme based on number of reactions):

80% ETHANOL VOLUMES							
Number of Reactions	**4**	**8**	**12**	**16**	**24**	**32**	**48**
80% Ethanol (ml)	10	20	30	30	40	50	70

5. Once the SOPHiA GENETICS™ DNA Library Prep Kit I components are thawed, mix the reagents by inverting the tube 5–10 times to ensure that there is no remaining precipitate and briefly spin in a microcentrifuge. Place on ice.

Pre-Mixes

1. Prepare the **FX pre-mix** as follows:

FRAGMENTATION PRE-MIX							
Number of Reactions	**4**	**8**	**12**	**16**	**24**	**32**	**48**
Fragmentation Buffer 10x (µl)	23.6	47.1	75	95	150	190	300
Fragmentation Enzyme (µl)	47.1	94.2	150	190	300	380	600

FX PRE-MIX							
Number of Reactions	**4**	**8**	**12**	**16**	**24**	**32**	**48**
FX Buffer 10x (µl)	23.6	47.1	75	95	150	190	300
FX Enzyme Mix (µl)	47.1	94.2	150	190	300	380	600

   - Mix thoroughly by pipetting up and down 10 times and spin briefly.
   - Keep on ice.

2. Prepare the **Ligation pre-mix** as follows:

LIGATION PRE-MIX							
Number of Reactions	**4**	**8**	**12**	**16**	**24**	**32**	**48**
Ligation Buffer 5x (µl)	95	190	300	380	600	760	1200
Ligation Enzyme (µl)	47.5	95	150	190	300	380	600
Nuclease-free water (µl)	71.3	142.5	225	285	450	570	900

- Mix thoroughly by pipetting up and down 10 times and spin briefly.
- Keep on ice.

The DNA Ligase Buffer 5x is highly viscous, pipette gently and make sure to obtain a homogeneous Ligation pre-mix.

3. Prepare the **PCR pre-mix** as follows:

PCR PRE-MIX							
Number of Reactions	**4**	**8**	**12**	**16**	**24**	**32**	**48**
HiFi PCR Master Mix 2x (µl)	115	230	345	460	690	920	1380
Primer Mix Illumina® Library Amp (µl)	6.9	13.8	20.7	27.6	41.4	55.2	82.8
Nuclease-free water (µl)	16.1	32.2	48.3	64.4	96.6	128.2	193.2

- Mix thoroughly by pipetting up and down 10 times and spin briefly.
- Keep on ice.

4.2.3 Enzymatic Fragmentation, End Repair and A-Tailing

Materials

- Diluted and conditioned double stranded DNA in 35 µl
- FX pre-mix
- RNase/DNase-free 0.2 ml 8-tube strips

Preparation

1. Program the thermal cycler for Fragmentation with the following settings:

	TEMPERATURE (°C)	TIME (MINUTES)
Lid	70	-
Step 1	4	1
Step 2	32	5
Step 3	65	30
Hold	4	∞

2. Start the Fragmentation program. When the block reaches Step 1 (4 °C), pause the program.

Procedure

Always keep the samples and pre-mix on ice before and after the incubation to block the enzymatic reaction.

1. Depending on the number of samples, proceed as follows:
    - If processing **4 samples**, proceed to step 2.
    - If processing **8 or more samples**, to facilitate pipetting, create a reservoir of FX pre-mix by adding the following volumes to a new set of 4 or 8-tube strips according to the following scheme:

NUMBER OF REACTIONS	8	12	16	24	32	48
PCR strip (1 strip)	**4-tube**	**4-tube**	**8-tube**	**8-tube**	**8-tube**	**8-tube**
FX pre-mix (µl)	33	52.5	33	52.5	66	105

2. Assemble the reaction as follows:
    - Using a multichannel pipette add 15 µl of Fragmentation pre-mix to each of the 35 µl of DNA samples (total of 50 µl in 4 or 8-tube strips).
    - Using a multichannel pipette set to 40 µl, mix thoroughly by pipetting up and down 5 times and briefly spin in a microcentrifuge.
3. Place in the thermal cycler and continue the Fragmentation program.

Proceed immediately to Ligation.

4.2.4 Ligation

Materials

- Fragmentation reaction products in 50 µl each
- Ligation pre-mix

- Dual Index Adapters
- RNase/DNase-free 0.2 ml 8-tube strips

Preparation

1. Remove the Dual Index Adapter plate from 4 °C (transferred from -20 °C to 4 °C earlier) and briefly spin the plate to collect all the liquid. Refer to *Appendix I* for the respective plate format.
2. During the Fragmentation, prepare new PCR strips with 5 µl of different Dual Index Adapters per tube according to the following scheme:

NUMBER OF REACTIONS	4	8	12	16	24	32	48
PCR strip (1 strip)	**4-tube**	**4-tube**	**4-tube**	**8-tube**	**8-tube**	**8-tube**	**8-tube**
Number of strips	1	2	3	2	3	4	6

3. Set up the thermal cycler at 20 °C (open lid).

Procedure

Always keep the samples and pre-mix on ice before and after the incubation to block the enzymatic reaction.

1. Depending on the number of samples, proceed as follows:
   - If processing **4 samples**, proceed to step 2.
   - If processing **8 or more samples**, to facilitate pipetting, create a reservoir of Ligation pre-mix by adding the following volumes to a new set of 4 or 8-tube strips according to the following scheme:

NUMBER OF REACTIONS	8	12	16	24	32	48
PCR strip (1 strip)	**4-tube**	**4-tube**	**8-tube**	**8-tube**	**8-tube**	**8-tube**
Ligation pre-mix (µl)	100	160	100	160	200	320

2. Using a multichannel pipette, transfer the 50 µl of each Fragmentation reaction product to the 4 or 8-tube strips containing 5 µl of Dual Index Adapters.
3. Mix thoroughly by pipetting up and down 10 times and spin briefly.
4. Using a multichannel pipette, add 45 µl Ligation pre-mix to each Fragmentation reaction product (55 µl in each tube of the 4 or 8-tube strip).
5. Mix thoroughly by pipetting up and down 10 times and spin briefly.
6. Incubate in the thermal cycler at 20 °C for 15 minutes (open lid).

Proceed to Post-Ligation Clean Up.

Do not place the strip(s) on ice at the end of the ligation as it might decrease the binding of the DNA to the beads.

4.2.5 Post-Ligation Clean Up

Materials

- Ligation reaction products in 100 µl each
- AMPure® XP beads equilibrated at room temperature
- Freshly prepared ethanol 80%
- Nuclease-free water
- RNase/DNase-free 0.2 ml 8-tube strips

Procedure

1. Using a multichannel pipette, add 80 µl of AMPure® XP beads to each of the 100 µl ligation reaction products. Mix thoroughly by pipetting up and down 10 times.
2. Incubate at room temperature for 5 minutes and spin briefly if required.
3. Place the 4 or 8-tube strip on a 96-well plate format magnetic rack for 5 minutes or until the liquid becomes clear.
4. Carefully discard 170 µl of supernatant using a multichannel pipette.

   Keep the tubes on the magnetic rack for the following steps.
5. Using a multichannel pipette, add 170 µl of 80% ethanol to the beads. Incubate for 30 seconds to 1 minute.
6. Carefully discard the ethanol using a multichannel pipette.
7. Repeat steps 5 and 6 once.
8. Remove the residual ethanol using a P10 or P20 multichannel pipette.
9. Air-dry the beads at room temperature for 5 minutes. Do not over-dry the beads because this could decrease the amount of recovered DNA.

   Remove the tubes from the magnetic rack.
10. Using a multichannel pipette, add 105 µl of nuclease-free water to the beads and wait for a few seconds. Mix thoroughly by pipetting up and down 10 times.
11. Incubate at room temperature for 5 minutes and spin briefly if required.
12. Place the 4 or 8-tube strip on a 96-well plate format magnetic rack for 5 minutes or until the liquid becomes clear.
13. Using a multichannel pipette, carefully transfer 100 µl of the supernatant to new, labelled 4 or 8-tube strip.

Proceed to Dual Size Selection.

4.2.6 Dual Size Selection

Materials

- Ligation reaction products in 100 µl each
- AMPure® XP beads equilibrated at room temperature
- Freshly prepared ethanol 80%
- IDTE
- RNase/DNase-free 0.2 ml 8-tube strips

Procedure

1. Using a multichannel pipette, add 60 µl of AMPure® XP beads to each of the 100 µl ligation reaction products. Mix thoroughly by pipetting up and down 10 times.
2. Incubate at room temperature for 5 minutes and spin briefly if required to collect all liquid.
3. Place the 4 or 8-tube strip on a 96-well plate format magnetic rack for 5 minutes or until the liquid becomes clear.
4. Using a multichannel pipette, carefully transfer 140 µl of the supernatant to a new, labelled 4 or 8-tube strip containing 20 µl of AMPure® XP beads. Mix thoroughly by pipetting up and down 10 times.
5. Incubate at room temperature for 5 minutes and spin briefly if required to collect all liquid.
6. Place the 4 or 8-tube strip on a 96-well plate format magnetic rack for 5 minutes or until the liquid becomes clear.
7. Carefully discard 150 µl of supernatant using a multichannel pipette.

   Keep the tubes on the magnetic rack for the following steps.
8. Using a multichannel pipette, add 170 µl of 80% ethanol to the beads. Incubate for 30 seconds to 1 minute.
9. Carefully discard the ethanol using a multichannel pipette.
10. Repeat steps 8 and 9 once.
11. Remove the residual ethanol using a P10 or P20 multichannel pipette.
12. Air-dry the beads at room temperature for 4 minutes. Do not over-dry the beads because this could decrease the amount of recovered DNA.

    Remove the tubes from the magnetic rack.
13. Using a multichannel pipette, add 20 µl of IDTE to the beads. Mix thoroughly by pipetting up and down 10 times and spin briefly.

Proceed to Library Amplification.

4.2.7 Library Amplification

Materials

- Dual size selected Ligation reaction products and beads resuspended in 20 µl IDTE each
- PCR pre-mix

Preparation

1. Program the thermal cycler for Library Amplification with the following settings:

	TEMPERATURE (°C)	TIME (SECONDS)	
Lid	99	-	8 cycles
Step 1: Initial Denaturation	98	120	
Step 2: Denaturation	98	20	
Step 3: Annealing	60	30	
Step 4: Extension	72	30	
Step 5: Final Extension	72	60	
Hold	10	∞	

Procedure

1. Depending on the number of samples, proceed as follows:
    - If processing **4 samples**, proceed to step 2.
    - If processing **8 or more samples**, to facilitate pipetting, create a reservoir of PCR pre-mix by adding the following volumes to a new set of 4 or 8-tube strips according to the following scheme:

NUMBER OF REACTIONS	8	12	16	24	32	48
PCR strip (1 strip)	**4-tube**	**4-tube**	**8-tube**	**8-tube**	**8-tube**	**8-tube**
PCR pre-mix (µl)	65	100	65	100	130	200

2. Assemble the reaction as follows:
    - Using a multichannel pipette, add 30 µl of PCR pre-mix to the dual size selected ligation products and beads (50 µl in 4 or 8-tube strips).
    - Mix thoroughly by pipetting up and down 10 times and spin briefly.
3. Place the tubes in the thermal cycler and run the Library Amplification program.

Safe stopping point overnight at 4 °C.

4.2.8 Post-Amplification Clean Up

Materials

- PCR reaction products in 50 µl each
- AMPure® XP beads equilibrated at room temperature
- Freshly prepared ethanol 80%
- Nuclease-free water
- DNA low-binding tubes for the storage of libraries

Procedure

1. Using a multichannel pipette, add 50 µl of AMPure® XP beads to each 50 µl of the PCR product. Mix thoroughly by pipetting up and down 10 times.
2. Incubate at room temperature for 5 minutes and spin briefly if required.
3. Place the 4 or 8-tube strips on a 96-well plate format magnetic rack for 5 minutes or until the liquid becomes clear.
4. Carefully discard 90 µl supernatant using a multichannel pipette.

   Keep the tubes on the magnetic rack for the following steps.
5. Using a multichannel pipette, add 170 µl of 80% ethanol to the beads. Let the tubes stand for 30 seconds to 1 minute.
6. Carefully discard the ethanol using a multichannel pipette.
7. Repeat steps 5 and 6 once.
8. Remove the residual ethanol using a P10 or P20 multichannel pipette.
9. Air-dry the beads at room temperature for 5 minutes. Do not over-dry the beads because this could decrease the amount of recovered DNA.

   Remove the tubes from the magnetic rack.
10. Using a multichannel pipette, add 20 µl of nuclease-free water to the beads. Mix thoroughly by pipetting up and down 10 times. Incubate at room temperature for 5 minutes and spin briefly if required.
11. Place the 4 or 8-tube strip on a 96-well plate format magnetic rack for 5 minutes or until liquid becomes clear.
12. Carefully transfer 18 µl (transferring two times 9 µl is recommended at this step) of the supernatant to a new and labeled library storage tube.

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

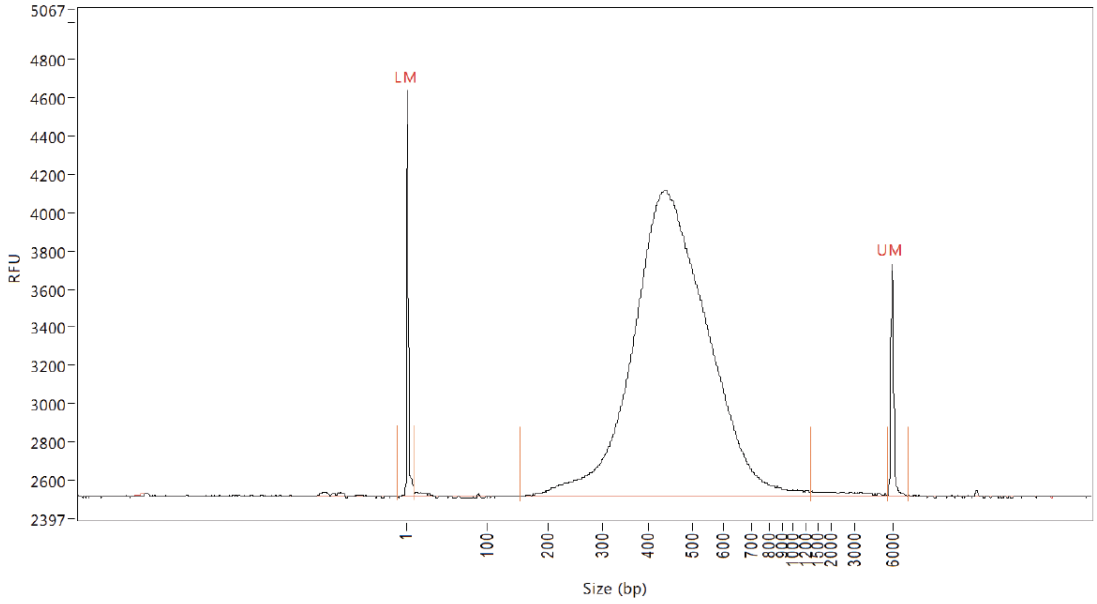
4.2.9 Individual Library Quantification and Quality Control

Materials

- Fluorometric quantitation equipment and reagents
- Capillary electrophoresis system
- Nuclease-free water
- RNase/DNase-free 0.2 ml 8-tube strips

Procedure

1. Prepare a 4-time dilution of each library with nuclease-free water (e.g., 2 µl of library in 6 µl nuclease-free water).
2. Quantify the libraries with a fluorometric method (e.g., Qubit HS quantification using 2 µl of the 4x library dilution prepared previously).
3. Quality control the libraries by analyzing their profile via capillary electrophoresis. Library DNA fragments should have a size distribution between 300 bp and 700 bp.



Example of a DNA library distribution obtained with the Agilent Fragment Analyzer™ capillary electrophoresis system. LM: Lower Marker, UM: Upper Marker.

4.3 Library Pooling

4.3.1 Library Pooling for Hybridization and Capture

Materials

- Individual sequencing libraries
- Human Cot DNA
- Universal Blockers – TS Mix
- DNA low-binding 1.5 ml tubes

Procedure

1. Prepare a library Capture pre-mix of the following in a DNA low-binding tube:

CAPTURE PRE-MIX				
Number of captures	**1**	**2**	**3**	**4**
Human Cot DNA (µl)	5	11	16.5	22
Universal Blockers – TS Mix (µl)	2	4.4	6.6	8.8

2. Prepare one DNA low-binding 1.5 ml tube per capture.
3. Pipette 7 µl of the above Capture pre-mix into individual DNA low-binding tubes for each capture.
4. To the individual tubes containing the above Capture pre-mix, add a pool of individual libraries according to the kit format:

KIT FORMAT	16 SAMPLES KIT	32 SAMPLES KIT	48 SAMPLES KIT	96 SAMPLES KIT
Number of individual libraries per capture	4	8	12	12
Amount of each library per capture	300 ng	200 ng	150 ng	150 ng
Total amount of libraries per capture	1200 ng	1600 ng	1800 ng	1800 ng

5. Mix thoroughly by pipetting up and down 10 times and spin briefly.
6. Dry each mix using a vacuum DNA concentrator until mix is completely lyophilized. Use mild heating (45–50 °C) to speed up the lyophilization.

Safe stopping point overnight at -20 °C.

4.4 Capture

4.4.1 Hybridization

Materials

- Lyophilized libraries
- 2x Hybridization Buffer
- Hybridization Buffer Enhancer
- Clinical Exome Solution v3 Probes
- Nuclease-free water
- RNase/DNase-free 0.2 ml 8-tube strips
- 1.5 ml Tubes
- 10x Wash Buffer I
- 10x Wash Buffer II
- 10x Wash Buffer III
- 10x Stringent Wash Buffer
- 2x Beads Wash Buffer

Preparation

1. Pre-warm the thermal cycler to 95 °C (set lid to 99 °C).
2. After the 2-minute denaturation, switch directly to 65 °C (set lid to 75 °C).

We recommend the use of different thermal cyclers for 95 °C and 65 °C incubations, if available.

Procedure

1. Prepare a Hybridization pre-mix according to the number of capture reactions:

HYBRIDIZATION PRE-MIX				
Number of captures	**1**	**2**	**3**	**4**
2x Hybridization Buffer (µl)	8.5	18.7	28.05	37.4
Hybridization Buffer Enhancer (µl)	3.4	7.48	11.22	14.96
Nuclease-free Water (µl)	1.1	2.42	3.63	4.84

2. Resuspend the lyophilized pellet in 13 µl of the Hybridization pre-mix.
3. Transfer the resuspended pellet to a PCR tube (one tube per capture reaction).
4. Add 4 µl of probes to the mix. Using a pipette set to 13 µl, mix thoroughly by pipetting up and down 5 times.
5. Incubate in the thermal cycler at 95 °C for 2 minutes.

Do not let the tube temperature drop below 65 °C from step 5 to 7 as this can lead to incorrect probe annealing.

6. Move the PCR tube from the 95 °C to 65 °C thermal cycler.
7. Incubate in the thermal cycler at 65 °C for 4 hours.
8. Prepare the 1x working solutions of different wash buffers in advance as described in the following sections to allow them to reach equilibrium during the hybridization reaction.

Wash Buffer Preparation For 1 Reaction

BUFFER	STOCK BUFFER (µl)	WATER (µl)	FINAL VOLUME 1X (µl)
10x Wash Buffer I	33	297	**330**
10x Wash Buffer II	22	198	**220**
10x Wash Buffer III	22	198	**220**
10x Stringent Wash Buffer	44	396	**440**
2x Bead Wash Buffer	275	275	**550**

Pre-warm 1x Stringent Buffer and an aliquot of 110 µl of 1x Wash Buffer I at 65 °C in a thermoblock or water bath for at least 2 hours. Keep the remaining Wash Buffer I at room temperature.

Wash Buffer Preparation For 2 Reactions

BUFFER	STOCK BUFFER (µl)	WATER (µl)	FINAL VOLUME 1X (µl)
10x Wash Buffer I	66	594	**660**
10x Wash Buffer II	44	396	**440**

BUFFER	STOCK BUFFER (µl)	WATER (µl)	FINAL VOLUME 1X (µl)
10x Wash Buffer III	44	396	**440**
10x Stringent Wash Buffer	88	792	**880**
2x Bead Wash Buffer	550	550	**1100**

Pre-warm 1x Stringent Buffer and an aliquot of 220 µl of 1x Wash Buffer I at 65 °C in a thermoblock or water bath for at least 2 hours. Keep the remaining Wash Buffer I at room temperature.

Wash Buffer Preparation For 3 Reactions

BUFFER	STOCK BUFFER (µl)	WATER (µl)	FINAL VOLUME 1X (µl)
10x Wash Buffer I	99	891	**990**
10x Wash Buffer II	66	594	**660**
10x Wash Buffer III	66	594	**660**
10x Stringent Wash Buffer	132	1188	**1320**
2x Bead Wash Buffer	825	825	**1650**

Pre-warm 1x Stringent Buffer and an aliquot of 330 µl of 1x Wash Buffer I at 65 °C in a thermoblock or water bath for at least 2 hours. Keep the remaining Wash Buffer I at room temperature.

Wash Buffer Preparation For 4 Reactions

BUFFER	STOCK BUFFER (µl)	WATER (µl)	FINAL VOLUME 1X (µl)
10x Wash Buffer I	132	1188	**1320**
10x Wash Buffer II	88	792	**880**
10x Wash Buffer III	88	792	**880**
10x Stringent Wash Buffer	176	1584	**1760**
2x Bead Wash Buffer	1100	1100	**2200**

Pre-warm 1x Stringent Buffer and an aliquot of 440 µl of 1x Wash Buffer I at 65 °C in a thermoblock or water bath for at least 2 hours. Keep the remaining Wash Buffer I at room temperature.

4.4.2 Streptavidin Beads Preparation

Materials

- Streptavidin beads equilibrated at room temperature
- 1x Bead Wash Buffer

- 1.5 ml tubes
- RNase/DNase-free 0.2 ml 8-tube strips

Procedure

Perform these steps just before the end of the hybridization incubation.

1. Mix the beads by vortexing them for 15 seconds.
2. Transfer 100 µl of beads per capture (200 µl for 2 reactions, 300 µl for 3 reactions, 400 µl for 4 reactions) to a single 1.5 ml tube.
3. Place the tube on a magnetic rack and let it stand until the solution becomes clear. Carefully remove and discard the supernatant.
4. Add 200 µl of 1x Bead Wash Buffer per capture (400 µl for 2 reactions, 600 µl for 3 reactions, 800 µl for 4 reactions) to the tube. Vortex for 10 seconds.
5. Place the tube on a magnetic rack and let it stand until the solution becomes clear. Carefully remove and discard the supernatant.
6. Repeat steps 4 and 5 once.
7. Add 100 µl of 1x Bead Wash Buffer per capture (200 µl for 2 reactions, 300 µl for 3 reactions, 400 µl for 4 reactions) to the tube. Vortex for 10 seconds.
8. Transfer 100 µl of cleaned beads to a new PCR tube (one tube per capture reaction).
9. Place tube(s) on a 96-well plate format magnetic rack and let it/them stand until the solution becomes clear. Carefully remove and discard the supernatant.

Do not allow the beads to dry.

Proceed immediately to Binding of Hybridized Targets to the Beads.

4.4.3 Binding of Hybridized Targets to the Beads

Materials

- Cleaned Streptavidin beads in PCR tube(s)
- Hybridization reaction(s)

Procedure

Work quickly to ensure that the temperature of the sample(s) remains close to 65 °C.

1. Remove the hybridization reaction(s) from the thermal cycler and briefly spin down the tube(s) and place them back on the thermocycler.

2. Place the washed Streptavidin bead tubes in the thermocycler (no more than two tubes at a time to avoid drying of beads).
3. For each hybridization reaction, transfer 17 µl of the hybridization reaction solution to one PCR tube containing cleaned beads. Resuspend the beads by pipetting up and down until the solution is homogeneous.
4. Bind the DNA to the beads by placing the tube(s) into a thermal cycler set at 65 °C (lid at 75 °C). Incubate for 45 minutes.
5. During the incubation, gently pipette up and down the tube(s) every 15 minutes to ensure that the beads remain in suspension.

Proceed immediately to Wash Streptavidin Beads to Remove Unbound DNA.

4.4.4 Wash Streptavidin Beads to Remove Unbound DNA

Materials

- Hybridized targets on beads
- RNase/DNase-free 0.2 ml 8-tube strips
- DNA low-binding 1.5 ml tubes
- 1x Wash Buffer I (⅓ at 65 °C and ⅔ at room temperature)
- 1x Wash Buffer II
- 1x Wash Buffer III
- 1x Stringent Wash Buffer (at 65 °C)
- Nuclease-free water
- IDTE

Procedure

Ensure that the temperature remains close to 65 °C for steps 1 to 7.

Note: If working with 2 or more capture tubes, work in a staggered manner from steps 2 to step 8, including the following:

1. When placing the first tube in thermoblock at 65 °C for the first incubation of 5 min (step 5), start a timer.
2. Begin processing the second tube.
3. When placing the second tube at 65 °C, note the time separating the tubes and ensure to respect this time gap along step 2 to step 8 to ensure each tube incubates exactly 5 min at 65 °C with the stringent wash.

1. Add 100 µl of 1x Wash Buffer I (at 65 °C) to each of the hybridized target/streptavidin beads tubes.
2. Working with one tube at a time, resuspend and transfer the mix one by one to a new DNA low-binding 1.5 ml tube. If working with two or more capture tubes, work in a staggered manner as indicated above.
3. Place tube on a magnetic rack and let it stand until the solution becomes clear. Carefully remove and discard the supernatant.
4. Add 200 µl of 1x Stringent Wash Buffer (at 65 °C) to the tube. Gently resuspend the beads by pipetting up and down.

Strong mixing of beads with the stringent wash buffer could decrease the quality of the capture.

5. Incubate at 65 °C for 5 minutes.
6. Place the tube on a magnetic rack and let it stand until the solution becomes clear. Carefully remove and discard the supernatant.
7. Repeat steps 4 to 6 once.

   Work at room temperature.

8. Add 200 µl of 1x Wash Buffer I (at room temperature) to your tube. Gently resuspend the beads by pipetting up and down.

   If working with 2 or more capture tubes; from this step on, process all the tubes at the same time.

9. Vortex for 2 minutes.
10. Place tube(s) on a magnetic rack and let it/them stand until the solution becomes clear. Carefully remove and discard the supernatant.
11. Add 200 µl of 1x Wash Buffer II to each tube(s). Vortex for 1 minute.
12. Place tube(s) on a magnetic rack and let it/them stand until the solution becomes clear. Carefully remove and discard the supernatant.
13. Add 200 µl of 1x Wash Buffer III to each tube(s). Vortex for 30 seconds. Spin briefly to collect all the liquid.
14. Place tube(s) on a magnetic rack and let them stand until the solution becomes clear. Carefully remove and discard the supernatant.
15. Add 200 µl of 1x IDTE to each tube(s). Resuspend the beads. Spin briefly to collect all the liquid.
16. Place tube(s) on a magnetic rack and let it/them stand until the solution becomes clear. Carefully remove and discard the supernatant.
17. Remove all the remaining liquid by using a P10 or P20 pipette.
18. Add 20 µl of nuclease-free water to each tube(s), resuspend and transfer the beads/water mix to a new PCR tube.

4.4.5 Post-Capture Amplification

Materials

- Streptavidin beads/nuclease-free water suspension (20 µl)
- KAPA HiFi HotStart ReadyMix 2x

- Library Amplification Primer Mix 10x
- Nuclease-free water

Preparation

1. Program the thermal cycler for Post-Capture Amplification using the following settings:

	TEMPERATURE (°C)	TIME (SECONDS)	
Lid	99	-	12 cycles
Step 1: Initial Denaturation	98	45	
Step 2: Denaturation	98	15	
Step 3: Annealing	60	30	
Step 4: Extension	72	30	
Step 5: Final Extension	72	60	
Hold	10	∞	

Procedure

In order to preserve enzyme activity, we recommend aliquoting the KAPA HiFi HotStart ReadyMix 2x so as not to exceed a maximum of four freeze thaw cycles. Please prepare aliquots of appropriate volume depending on the number of captures usually performed in the lab.

1. Prepare the PCR pre-mix as follows:

PCR PRE-MIX				
Number of Reaction(s)	**1**	**2**	**3**	**4**
KAPA HiFi HotStart ReadyMix 2x (µl)	25	55	82.5	110
Library Amplification Primer Mix 10x (µl)	2.5	5.5	8.25	11
Nuclease-free water (µl)	2.5	5.5	8.25	11

2. Add 30 µl of PCR pre-mix to each bead suspension. Mix thoroughly by pipetting up and down 10 times and spin briefly.
3. Place the tube(s) in the thermal cycler and run the Post-Capture Amplification program.

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

4.4.6 Post-Capture Amplification Clean Up

Materials

- PCR reaction products in 50 µl each
- AMPure® XP beads equilibrated at room temperature
- Freshly prepared ethanol 80%
- IDTE
- DNA low-binding tubes for library storage

Procedure

1. Add 50 µl of AMPure® XP beads to each of the 50 µl PCR reaction products. Mix thoroughly by pipetting up and down 10 times.
2. Incubate at room temperature for 5 minutes and spin briefly if required to remove the liquid from the tube walls.
3. Place tube(s) on a magnetic rack for 5 minutes or until the liquid becomes clear.
4. Carefully discard 90 µl supernatant using a multichannel pipette.

   Keep the tube(s) on the magnetic rack for the following steps.
5. Using a multichannel pipette, add 170 µl of 80% ethanol to the beads. Let the tube(s) stand for 30 seconds to 1 minute.
6. Carefully discard the ethanol.
7. Repeat steps 5 and 6 once.
8. Remove the residual ethanol using a P10 or P20 pipette.
9. Air-dry the beads at room temperature for 5 minutes. Do not over-dry the beads because this could decrease the amount of recovered DNA.

   Remove the tube(s) from the magnetic rack.
10. Add 20 µl of IDTE to the beads. Mix thoroughly by pipetting up and down 10 times. Incubate at room temperature for 5 minutes and spin briefly if required to remove the liquid from the tube walls.
11. Place tube(s) on a magnetic rack for 5 minutes or until liquid becomes clear.
12. Carefully transfer 18 µl of the supernatant (transferring two times 9 µl is recommended at this step) to a new, labeled library storage tube.

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

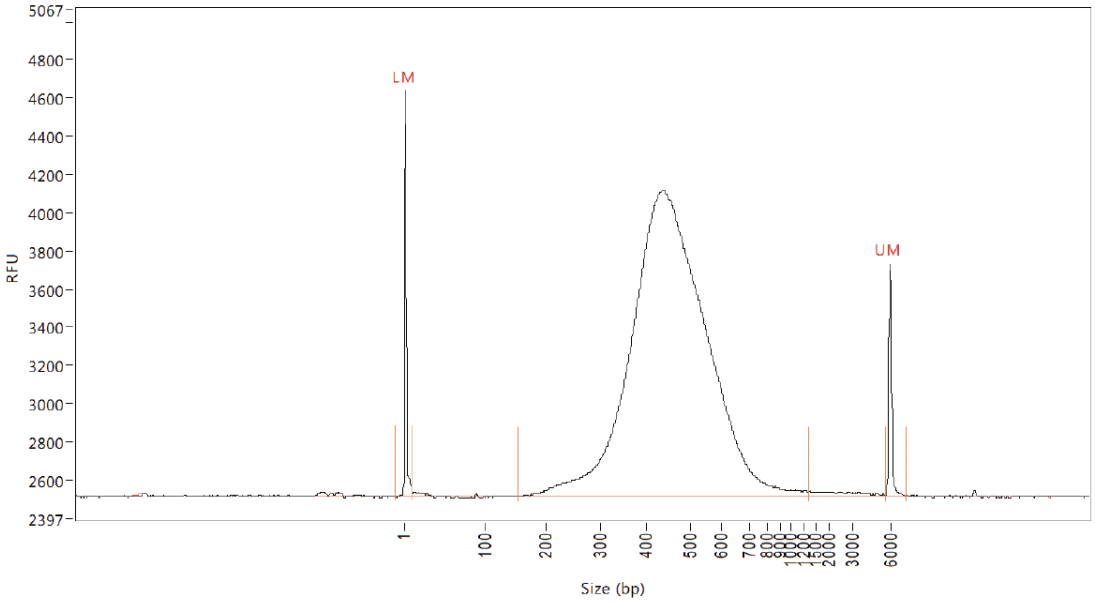
4.4.7 Final Library Quantification and Quality Control

Materials

- Fluorometric quantitation equipment and reagents
- Capillary electrophoresis system

Procedure

1. Quantify each captured library pool with a fluorometric method (e.g., Qubit® HS quantification using 2 µl of the library).
2. Control the quality of the captured pools of libraries by analyzing their profile via capillary electrophoresis. Library DNA fragments should have a size distribution between 300 bp and 700 bp.



Example of post-capture library pool size distribution obtained with the Agilent Fragment Analyzer™ capillary electrophoresis system. LM: Lower Marker, UM: Upper Marker.

4.5 Sequencing

4.5.1 Sequencing Preparations

Materials

- Illumina® MiSeq® v3 or NextSeq® 500/550 Sequencing Kit
- Final captured libraries
- EBT Buffer or similar

Procedure

1. Determine the molarity of each captured pool with the average size of the library (peak size in base pairs) and concentration (ng/µl) obtained during step *4.4.7 Final Library Quantification and Quality Control* 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. Dilute each captured pool to 4 nM.
3. If processing multiple sequencing pools, mix them in equal amounts (e.g., 5 µl of each) and use this dilution according to Illumina® standard denaturation recommendations.
4. Load a 1.5 pM dilution of the denatured libraries on NextSeq® 500/550 or a 10 pM dilution on MiSeq®.
5. For recommended reads per sample, see the table below:

READ LENGTH (BP)	RECOMMENDED TOTAL READS PER SAMPLE
150	16.25 million
300	12.42 million

5 ANALYSIS

5.1 Analysis Prerequisites

5.1.1 SOPHiA DDM™ Desktop App Installation Instructions

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

5.1.2 Data Generation

Sequencing data for analysis with the workflow should be generated starting with blood samples by following the instructions in section *4 Protocol* using the SOPHiA DDM™ Clinical Exome Solution v3 library preparation and capture kit, and sequencing these samples on Illumina® NextSeq® 500/550 sequencers.

Note: SOPHiA DDM™ Clinical Exome Solution v3 is compatible with sequencing data generated by Illumina® MiSeq® and NextSeq® 500/550 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 using the recommended number of reads and samples as described in sections *4.5.1 Sequencing Preparations* and *5.1.3 Run Composition*.

Using other systems can result in unexpectedly lower technical or analytical performance. As detailed in section *5.3 Analysis Description and Parameters*, appropriate down-sampling is applied to ensure good turn-around time without affecting results.

5.1.3 Run Composition

For high quality results based on SOPHiA DDM™ analytics, a minimum of 8.1 million fragments or read pairs (i.e., 16.25 million reads) per sample is recommended. Please note that certain factors such as uneven read allocation between samples, skewed read depth across the gene panel (e.g., consistently lower read depth in AT-rich regions), or excessive mitochondrial read coverage may lead to insufficient read depth – either across the entire gene panel or only in parts thereof – and may therefore negatively impact results for the affected samples. Consequently, the user is invited to consider that increasing the number of reads per sample is expected to provide more confident and sensitive variant calls, especially in regions with otherwise relatively low read depth.

5.1.4 Data Upload

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™ Desktop App 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™ Desktop App user that started the run.

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.

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.

Naming Convention

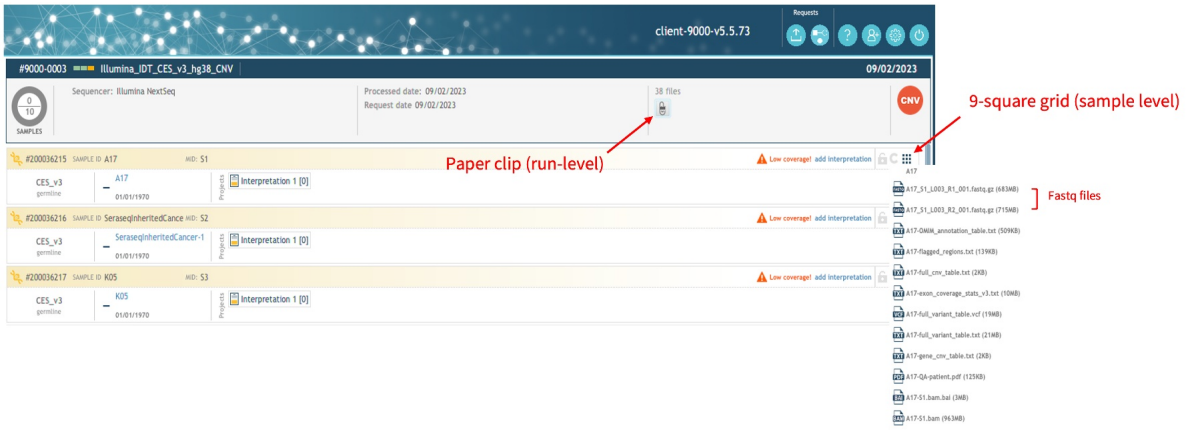
SOPHiA DDM™ desktop app uses the Illumina® naming convention and folder configuration to find the FASTQ files. Therefore, it is strongly advised **NOT** to reorganize or rename the FASTQ files after they are copied from the sequencer, otherwise the run might fail.

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.

  Note: The SOPHiA DDM™ desktop app 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.

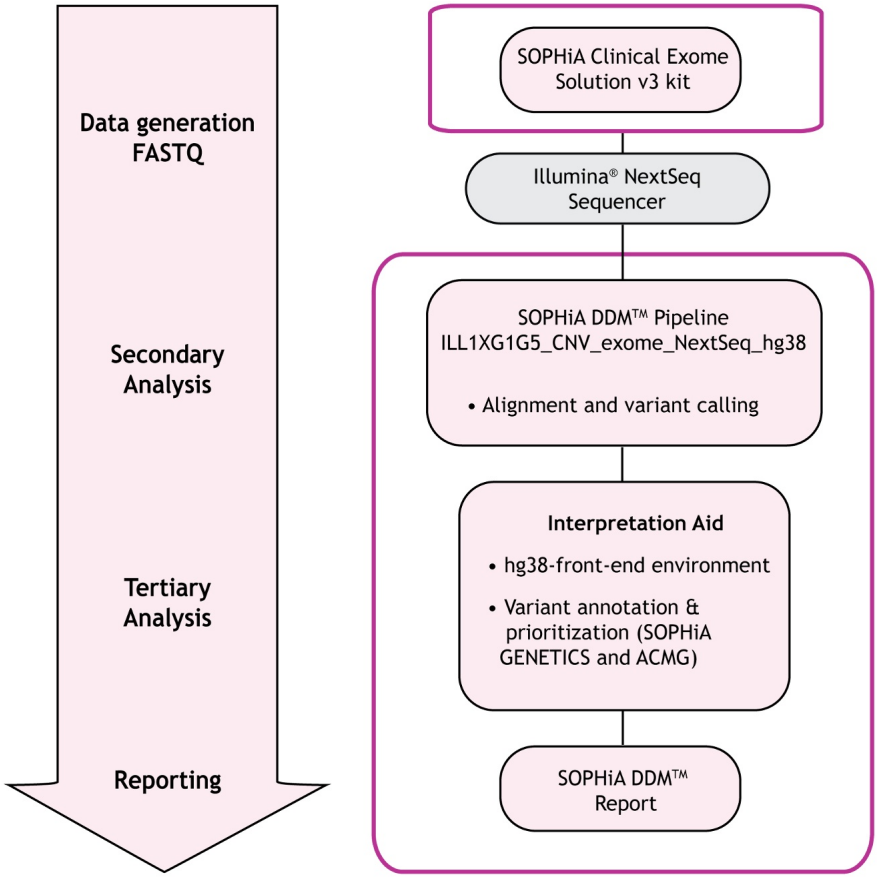


Illumina® Folder Configuration

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

5.2 Analysis Workflow Description



Analysis workflow description for the SOPHiA DDM™ Clinical Exome Solution v3

5.3 Analysis Description And Parameters

5.3.1 Resource Files

Alignment and variant calling are performed against the latest major release of the reference genome: GRCh38 (in this guide referred to as hg38). Variant coordinates are provided against GRCh37 – the previous major release of the human reference genome (referred to as hg19) – via a conversion operation termed liftOver.

A file containing pipeline parameters is available on request via support at support@sophiagenetics.com.

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.

> **Note:** This assembly contains the human decoy sequence from hs38d1 (GCA_000786075.2) in addition to all the sequences present in the no_alt_analysis set (mitochondrial genome, unlocalized scaffolds, and unplaced scaffolds). Importantly, following alignment, only reads mapping to the main chromosomes (1-22, X and Y) are retained for variant calling.

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

5.3.2 Included Analysis Modules

The different results that this pipeline returns are: SNVs/INDELs, CNVs, Familial Variant Analysis and Mitochondrial Analysis. **For a full list of the targets of this solution, (including SNVs/INDELs targets, CNV targets and target genes possessing pseudogenes/homologs), please contact support at support@sophiagenetics.com.**

SNVs/INDELs

Preprocessing

1. Collect quality metrics based on the raw FASTQ files.
2. Truncate the FASTQ files to a maximum size of 6,500 MB.

Alignment

1. Cut adapters and trim low quality ends from reads (base quality below 20).
2. Align reads to the hg38 human reference genome in paired-end mode.

3. Compute alignment statistics and coverage metrics on the raw alignment files.

Variant Calling

1. Main chromosomes
    a. Assign reads to PCR duplicate groups based on start and end positions.
    b. For SNVs, minimum base quality of Q20; minimum variant read coverage of 5; minimum variant fraction of 10%.
    c. For INDELs, minimum variant read coverage of 5; minimum variant fraction of 5%.
    d. Merge variants together if they are on the same allele (phasing).
    e. Re-quantify the variant fraction considering all the haplotypes in the neighboring region.
2. Mitochondria
    a. Assign reads to PCR duplicate groups based on start and end positions.
    b. For SNVs, minimum base quality of 20, minimum variant read coverage of 5, minimum variant fraction of 0.9%.
    c. For INDELs, minimum variant read coverage of 5, minimum variant fraction of 0.9%.
    d. Merge variants together if they are on the same allele (phasing).
    e. Re-quantify the variant fraction considering all the haplotypes in the neighboring region.

Variant Filtering

1. Main chromosomes
    a. Filter variants with variant percent below expected thresholds: 25% for SNVs, 20% for INDELs of length below 25bp, and 5% for INDELs of 25bp or longer (filter = low_variant_ fraction).
    b. Filter variants with read coverage below 20 (filter = low_coverage).
    c. Filter variants outside of the target regions (filter = off_target).
    d. Filter variants overlapping problematic regions, which are intronic or intergenic regions with systematic background noise (filter = problematic_region).
    e. Filter INDELs in homopolymers of length equal to or greater than 10 bp (filter = homopolymer_region).
    f. Filter variants where variant fractions are significantly different between the two strands and where the supporting strand is detected to be noise (filter = strand_bias).
2. Mitochondria
    a. Filter variants with variant percent below expected thresholds: 5% for SNVs, 5% for INDELs of length below 25bp, and 5% for INDELs of 25bp or longer (filter = low_variant_ fraction).
    b. Filter variants with read coverage below 20 (filter = low_coverage).
    c. Filter INDELs in homopolymers of length equal to or greater than 10 bp (filter = homopolymer_region).
    d. Filter variants where variant fractions are significantly different between the two strands and where the supporting strand is detected to be noise (filter = strand_bias).

Variant Annotation

1. Extract low coverage regions based on a threshold of 20 reads.
2. Annotate variants with dbSNP identifier, allele frequency from the 1000 genomes project, ExAC, ESP, GnomAD, normalized prediction scores from dbNSFP, which include SIFT, PolyPhen2, MutationTaster, LRT, GERP and clinical significance from ClinVar.
3. Calculate transcript dependent variant annotation such as cDNA and protein notation, exon, rank and distance to exon. For variants in a region where more than one gene annotation is possible, the default annotation will be given for a gene with the highest HPO score (i.e., the number of counts of HPO terms associated with a gene).
4. Annotate variants with OMIM.
5. For variants present in genes that possess pseudogenes or homologs, a warning will be displayed in SOPHiA DDM™. The warning does not indicate that the variant call is low confidence.

CNVs

Copy-number variations (CNVs) are structural changes in the DNA associated with variations in the number of copies of the affected DNA sections.

CNV analysis for this gene panel is based on the hg38 reference genome and is performed by the following steps:

1. Coverage levels of the target regions across samples within the same batch are analyzed (a target region in the context of CNV detection is defined as a continuous region covered by probes in the capture-based amplification technology or an amplicon region in the amplicon-based technology). For each sample, the algorithm automatically selects a set of reference samples from the same run based on the similarity of coverage patterns.
2. Using the reference samples, the coverage is normalized by sample and target region, and CNV calling is performed using a hidden-Markov-model algorithm. As a result, the most probable copy number for each target region is determined. Additionally, confidence grading is calculated both at the sample and target-region level.
3. Samples are classified into rejected, medium-noise, and low-noise based on the residual coverage noise after normalization and CNV calling. No CNV results are reported for the rejected samples.
4. Individual target regions (in non-rejected samples) are also classified into three categories: high-confidence, medium-confidence, and undetermined.

Total number of regions in the panel	65736
Number of regions retained in the CNV module	60950 (92.7% of total regions)
Total number of genes in the panel	6380 (all genes overlapping with target regions, including 5111 protein-coding genes)
Number of genes retained in the CNV module	4746 (out of 5731 genes overlapping regions eligible for CNV detection; only a single gene is chosen for annotation in the CNV report for regions with multiple overlapping genes)

Familial Variant Analysis

1. Based on alt and ref counts associated with each variant in each of the family members, compute the p-values characterizing the consistency of sequencing data with the hypothesis that the genotype of the family member:

   a. Contains the variant (pval_present)
   b. Does not contain the variant (pval_absent)
   c. Contains exactly one copy of the variant allele (pval_ heterozygous) and
   d. Is homozygous variant (pval_homozygous).

   The computed p-values are used by SOPHiA DDM™ to call the individual genotypes.

2. Compute the p-values characterizing the relevance of each variant for the observed clinical conditions of the patient under the assumption that the disease follows either:
   a. Autosomal recessive inheritance pattern (pval_autosomal_recessive)
   b. Autosomal dominant pattern (pval_autosomal_dominat)
   c. X-linked pattern (pval_x_linked), or
   d. Is caused by a *de novo* mutation (pval_de_novo).

   The computed p-values are used by SOPHiA DDM™ for filtering out the irrelevant variants once the suspected inheritance pattern is specified by the user.

3. Identify possible “compound heterozygous” groups of variants. Those groups are formed by variants that are all located within the same gene that are heterozygous in exactly one of the parents, absent in another, and heterozygous in the proband. For each of the identified “compound heterozygous groups”, all the possible pairs of variants are tested for being a possible cause of the disease resulting from the proband combining the variants that were separately present in parents. The procedure results in the pval_compound_het p-value for each variant identified as a member of a “compound heterozygous” group. The computed p-values are used by SOPHiA DDM™ for filtering out the irrelevant variants once the compound heterozygous inheritance pattern is specified by the user.
4. The Benjamini-Hochberg procedure is applied to the computed p-values column by column.

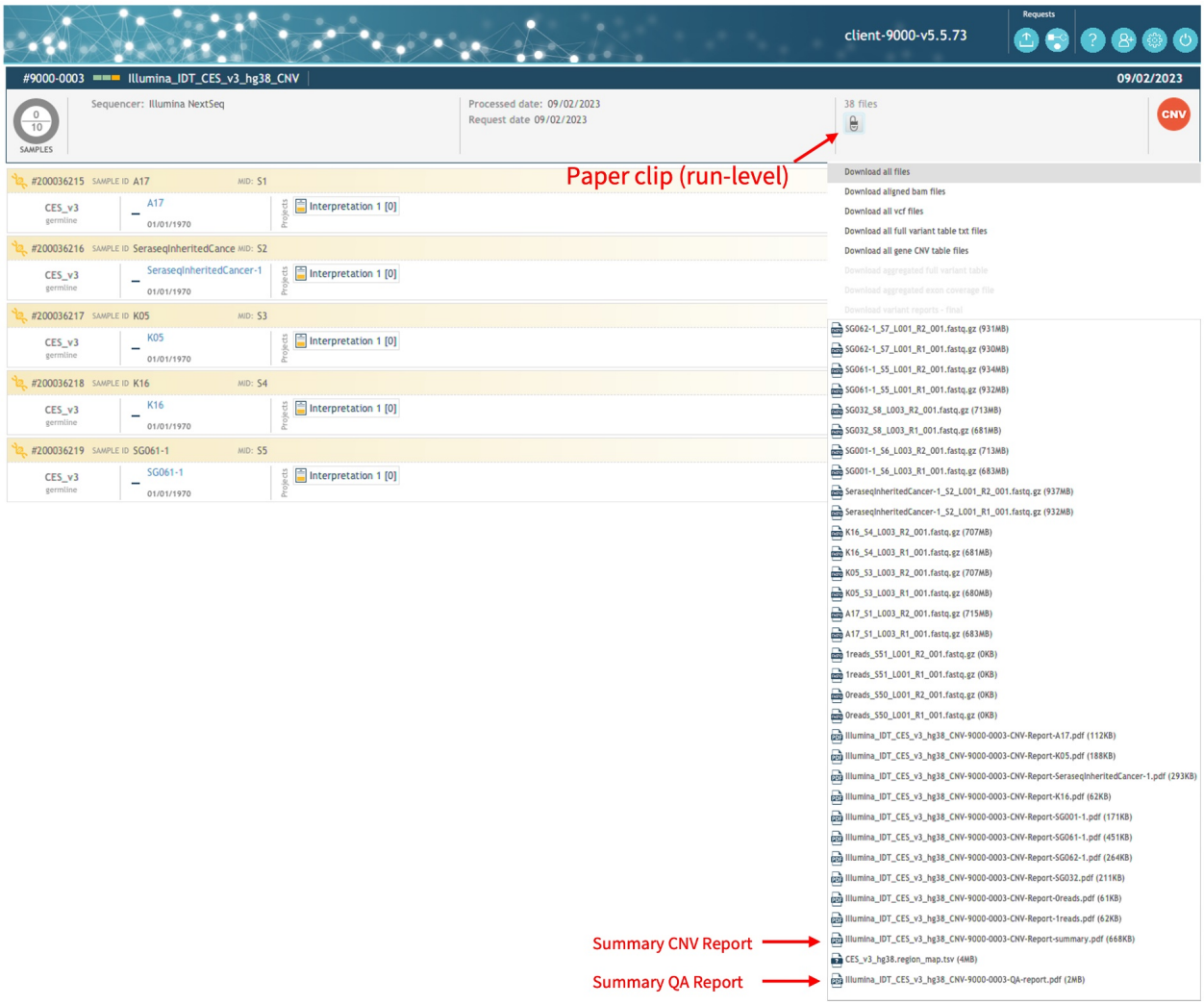
6 PIPELINE DELIVERABLES

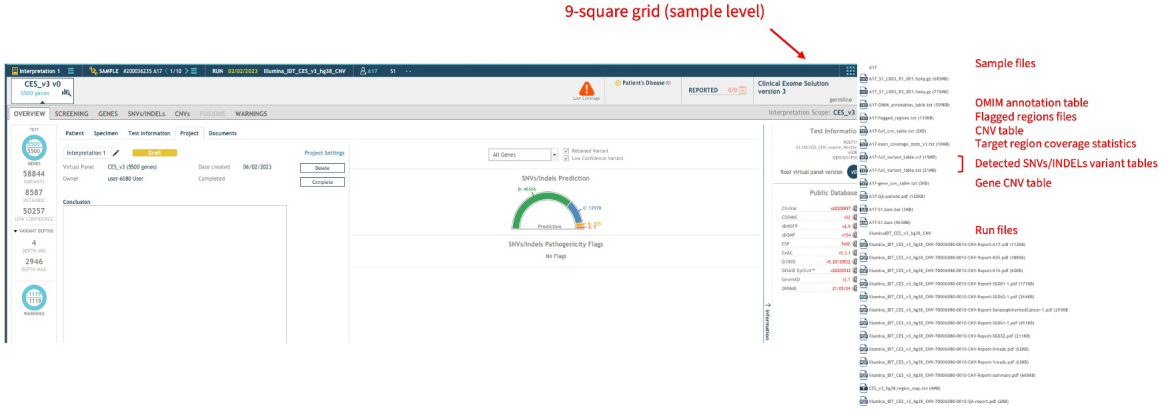
6.1 Results Location

Output files from SOPHiA DDM™ Desktop App can be found at the run level (paper clip) or sample level (9-grid square).

- At the run level, a group of files can be downloaded – all files (including FASTQ), all vcf, all aligned bam files, aggregated full variant table, aggregated exon coverage, and final variant reports.
- The summary pipeline QA report can be downloaded here, as well as other summary reports (e.g., CNV summary report).
- At the sample level, sample-specific files can be downloaded – FASTQ, aligned bam, target region coverage statistics, flagged regions, full variant table, sample-specific pipeline QA report and run-level summary reports (e.g., QA report, CNV summary report).

An example is shown in the schematic below:





6.2 File Composition

- **Fastq.gz files:** The input FASTQ files zipped.

- **Exon_coverage_stats_v3.txt** Coverage statistics for each exon of each transcript in the gene panel taking into account only the coding sequence (CDS) (% targeted, min/average/max coverage, % target region covered at different coverage levels (10x, 15x, 20x, 25x, 30x, 50x, 100x, 200x, 500x, 1000x). The file provides coverage information for each exon of each transcript in the gene panel and coverage information for each transcript overall. The file enables users to determine the coverage of a coding portion of a particular gene (or set of genes) for a given sample. Note that exons/transcripts are ordered first by the chromosome they fall on (ordered 1–22, X, Y, MT) and then their starting position.

- **Exon_coverage_stats_complete.txt** Coverage statistics for each exon of each transcript in the gene panel, taking into account both the coding and non-coding sequences (e.g., 5' UTR, 3' UTR), as well as coverage information for the entire transcript overall (indicated by lines with exon_rank=0). The statistics calculated will include: % of exon/transcript targeted by the target regions, min/average/max coverage, % of the exon/transcript covered at different coverage levels (10x, 15x, 20x, 25x, 30x, 50x, 100x, 200x, 500x, 1000x). The file enables users to determine the coverage of a particular gene (or set of genes) for a given sample. Note that exons/transcripts are first ordered by the chromosome they fall on (ordered 1–22, X, Y, MT) and then their starting position. Since coverage statistics are calculated for both coding and non-coding regions, this file will include coverage for completely non-coding transcripts, which will not be included in exon_coverage_stats_v3.txt

- **Target_coverage_stats.txt** Coverage statistics for each target region in the gene panel. The statistics calculated will include the average coverage in the target region. Note that this file will include information for any target region, including those not associated with a coding or non-coding transcript (i.e., intergenic hotspots), and helps determine coverage that is otherwise not included in exon_coverage_stats_v3.txt or exon_coverage_stats_complete.txt.

    - For information related to the customized exports Coverage Calculator, please refer to section 4.8.5 “Coverage Calculator” of the user manual.

- **QA-report.pdf:** The pipeline QA report in PDF format at the sample and run level (e.g., % mapped reads, on-target rate, molecule coverage, etc.)(e.g. % mapped reads, on-target rate, molecule coverage with coverage metrics statistics split between nuclear and mitochondrial genomes, etc.).

- **CNV-report.pdf:** Summary (run-level) CNV results, depending on the available run and pipeline.

    - For more information regarding CNV interpretation, please refer to the CNV report.

- **OMIM_annotation_table.txt** Summarizes OMIM annotation at gene level for each gene containing variants. Lists the inheritance mode of the disease and provides the appropriate DOID.

- **flagged_low_coverage.txt** List of target regions insufficiently covered (i.e., < 20 x) to ensure reliable variant calling.

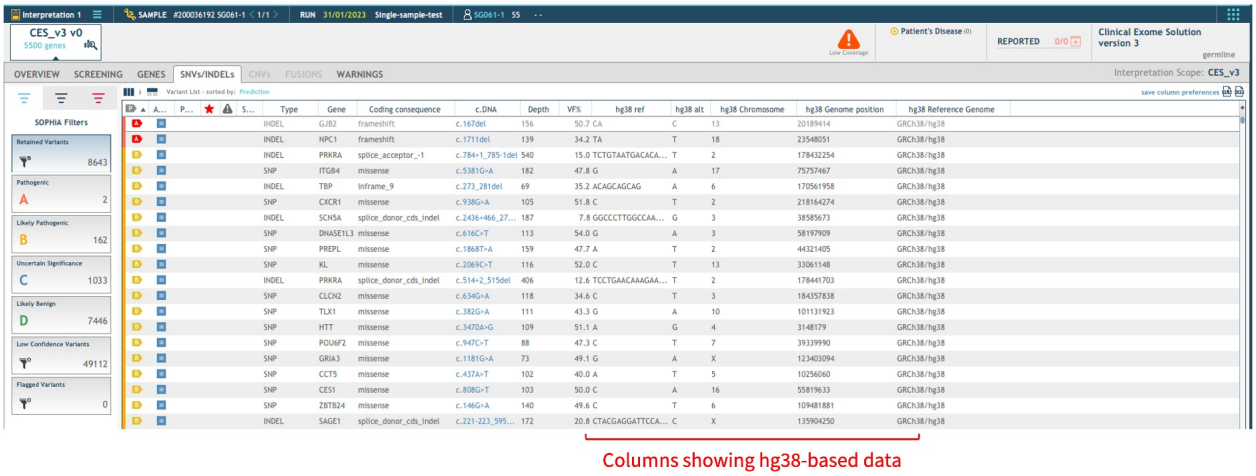
- **flagged_regions.txt** Conveys warnings and information about particular regions of the genome for a given patient sample. Importantly, there are flagged regions that are specific to a given patient sample within the file (i.e. their specificity is “sample”), and flagged regions that are common to all samples for a given product (i.e. their specificity is “design”). This is because some regions will vary from patient to patient; for example, sequencing coverage will vary across experiments, and therefore “low_coverage” flagged regions will be specific to a patient sample. Alternatively, some flagged regions like “low_complexity_region” are a property of the genome and will be the same across patient samples.

To download the flagged_regions.txt file, begin within an Interpretation for a patient, then navigate to Overview Tab > Documents Tab > Analysis Files Tab, and click on the icon labeled flagged_regions.txt. Alternatively, when working within an Interpretation for a patient, you can click the nine-square grid icon in the upper right corner and download the file.

6.3 Results Visualization

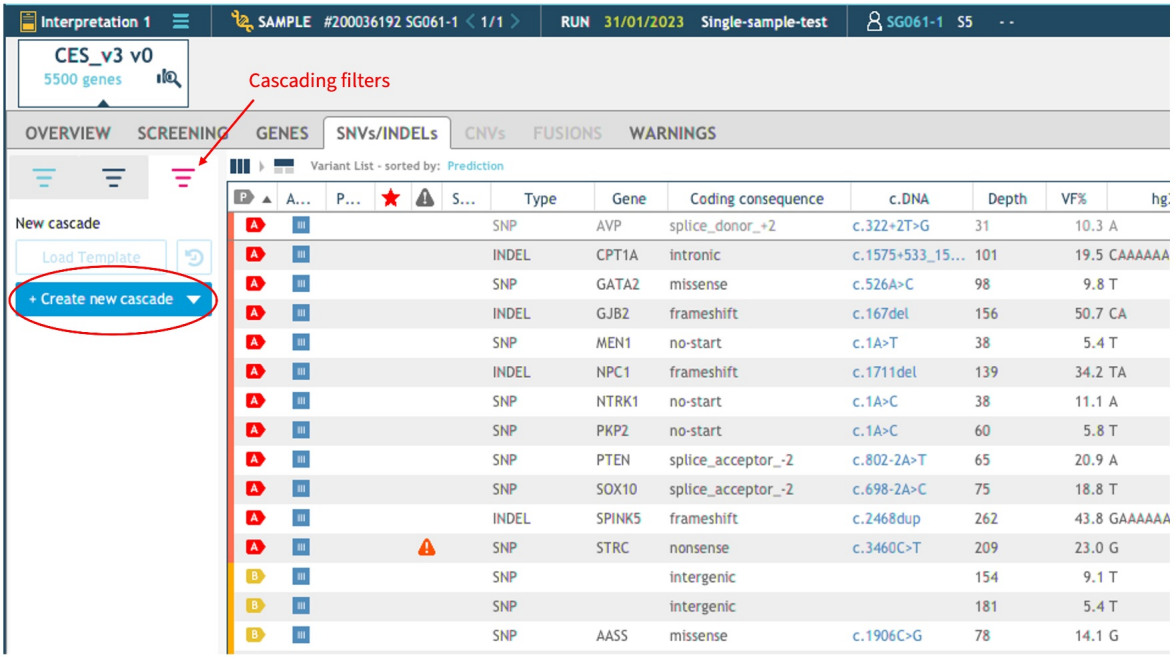
For information related to accessing analysis results (Variant display, QA report), please refer to Section 4 “Data Analysis” of the SOPHiA DDM™ Desktop App User Manual. Note that as certain modules are only active for selected pipelines, sections of the platform will be greyed out or bolded.

For example, the “SNVs and INDELs” tab is activated, while the “Fusions” tab is inactive in this pipeline.

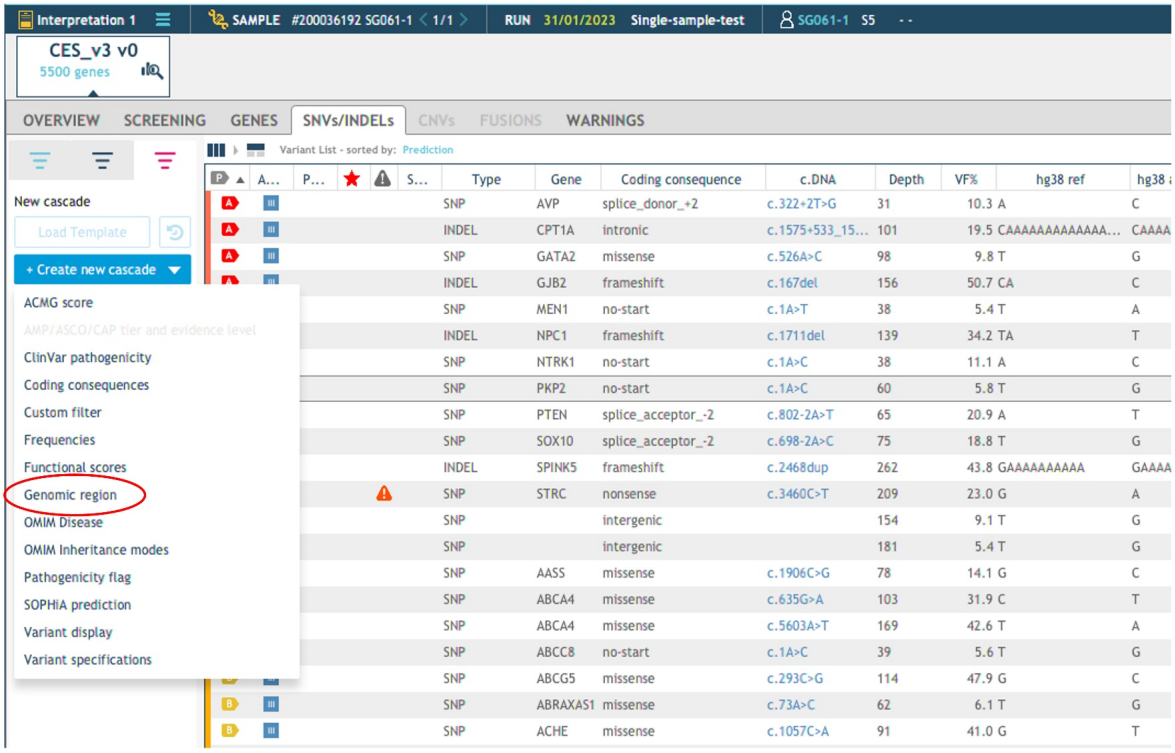


In order to visualize mitochondrial variants, cascading filtering may be applied to select for 'Chromosome MT' in the following manner:

1. Select cascading filters > Create new cascade

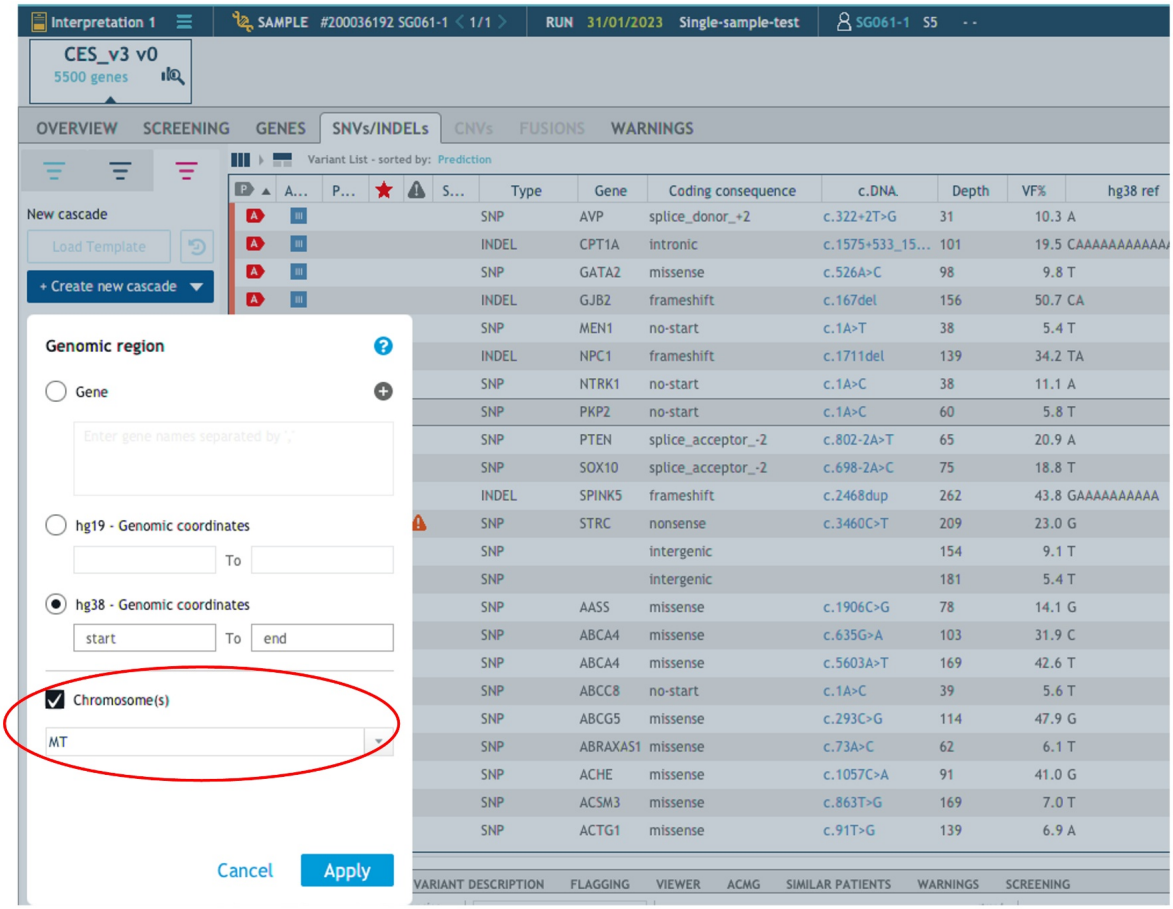


2. Select Genomic region

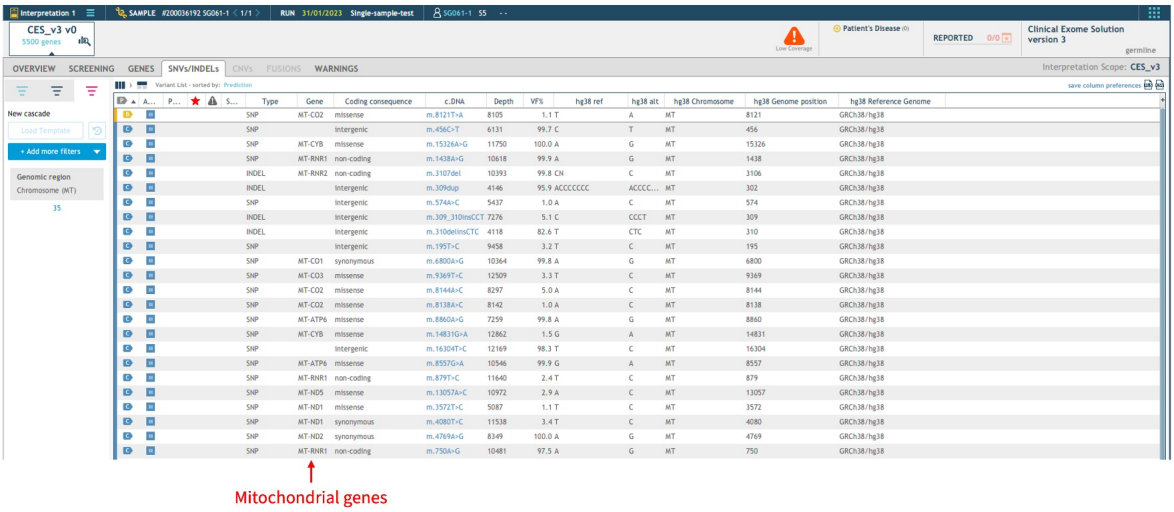


A...	P...	S...	Type	Gene	Coding consequence	c.DNA	Depth	VF%	hg38 ref	hg38
			SNP	AVP	splice_donor_+2	c.322+2T>G	31	10.3	A	C
			INDEL	CPT1A	intronic	c.1575+533_15...	101	19.5	CAAAAAAAAAAAA...	CAAAA
			SNP	GATA2	missense	c.526A>C	98	9.8	T	G
			INDEL	GJB2	frameshift	c.167del	156	50.7	CA	C
			SNP	MEN1	no-start	c.1A>T	38	5.4	T	A
			INDEL	NPC1	frameshift	c.1711del	139	34.2	TA	T
			SNP	NTRK1	no-start	c.1A>C	38	11.1	A	C
			SNP	PKP2	no-start	c.1A>C	60	5.8	T	G
			SNP	PTEN	splice_acceptor_-2	c.802-2A>T	65	20.9	A	T
			SNP	SOX10	splice_acceptor_-2	c.698-2A>C	75	18.8	T	G
			INDEL	SPINK5	frameshift	c.2468dup	262	43.8	GAAAAAAAAA	GAAAA
			SNP	STRC	nonsense	c.3460C>T	209	23.0	G	A
			SNP		intergenic		154	9.1	T	G
			SNP		intergenic		181	5.4	T	G
			SNP	AASS	missense	c.1906C>G	78	14.1	G	C
			SNP	ABCA4	missense	c.635G>A	103	31.9	C	T
			SNP	ABCA4	missense	c.5603A>T	169	42.6	T	A
			SNP	ABCC8	no-start	c.1A>C	39	5.6	T	G
			SNP	ABCG5	missense	c.293C>G	114	47.9	G	C
			SNP	ABRAXAS1	missense	c.73A>C	62	6.1	T	G
			SNP	ACHE	missense	c.1057C>A	91	41.0	G	T

3. Check Chromosome > Select MT from the dropdown menu > Click on Apply



4. All the mitochondrial variants will be displayed. Large mitochondrial deletions of several kilobases in size are also displayed in the SNV/indel tab rather than the CNV tab.



6.4 Variant Filtering

A list of reasons for variants being in the low confidence (i.e., being considered as low-confidence) tab is given in the table below.

FILTER TYPE	DESCRIPTION
off_target	The variant is placed outside of the panel target region.
low_variant_fraction	The variant has a variant fraction value below cutoff (cutoff can be different for different variant types: SNVs/INDEL, INDELs in homopolymers, etc.).
homopolymer_region	The variant is in a homopolymer region with a length higher than the cutoff.
problematic_region	The variant position overlaps with predefined problematic regions (e.g., low complexity region, GC-rich regions) or variant matches common sequencing artifacts.
low_coverage	The variant is supported by insufficient number of high quality read position (phred score >20).
strand_bias	The variant shows a significant difference of variant fraction between the two strands and the supporting strand shows higher than normal noise levels.

Please refer to section *5.3.2 Included Analysis Modules - SNVs/INDELS* for more details.

Users can create a custom filter based on criteria from any column (e.g., to retain off-target variants) to display chosen variants. Such a filtered list can then be used as input to the cascading filter.

6.5 Variant Annotation

6.5.1 Database Information And Location Of This Information

1. When a catalog in SOPHiA DDM™ is updated, the clients get notified via the release notes and the catalog version changes on the front end of the software.
2. Variant (SNVs/INDELs) catalogs:
    a. Clinvar
    b. dbNSFP
    c. dbSNP
    d. ESP
    e. ExAC
    f. G1000
    g. GnomAD
3. CNV catalogs
    a. GnomAD_SV
    b. dbVAR
4. Gene level data
    a. RefSeq
    b. HUGO HGNC
    c. UNIPROT

6.5.2 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. Please note that the examples mentioned in the table in italics are not from the same variant.

For variants in a region where more than one gene annotation is possible, the default annotation will be given for a gene with the highest HPO score (i.e., the number of counts of HPO terms associated with a gene). A warning will be displayed for such a variant, informing the user about the presence of an overlapping gene.

A warning will be displayed in SOPHiA DDM™ for variants in genes possessing pseudogenes or homologs. The warning does not indicate that the variant call is low confidence.

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

FIELD	DESCRIPTION	EXAMPLE
type	variant type	SNP / INDEL
codingConsequence	consequence on protein	5'UTR
refGenome	hg19 reference genome for hg19 panels (liftover to hg19 for hg38 panels)	GRCh37/hg19
chromosome	hg19 “chromosome” for hg19 panels (liftover coordinates to hg19 for hg38 panels)	13
genome_position	hg19 “genomic position”as generated by the variant caller) for hg19 panels (liftover to hg19 coordinates for hg38 panels)	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	*p.(Ser309Glnfs*26)*
ref	hg19 “ref” in reference genome for hg19 panels and liftover to hg19 for hg38 panels	G
alt	hg19 “alternative allele” for hg19 panels and liftover to hg19 for hg38 panels	A
refNum	depth reference allele	6178
altNum	depth alternative allele	6029
refSeq	ref codon	*GGC*
altSeq	alt codon	*GGG*
refAA	ref amino-acid	*G*
altAA	alt amino-acid	*G*
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
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	.

FIELD	DESCRIPTION	EXAMPLE
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 which has been normalised	
LJB_PolyPhen2	dbNSFP's precomputed PolyPhen2 score which has been normalised	
LJB_PolyPhen2_HumDiv	dbNSFP's precomputed PolyPhen2_HumDiv score which has been normalised	
LJB_LRT	dbNSFP's precomputed LRT score which has been normalised	
LJB_MutationTaster	dbNSFP's precomputed MutationTaster score which has been normalised	
LJB_GERP	dbNSFP's precomputed GERP score	
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	The flagged region id is used to link a variant to a specific warning region from the warnings tab	
depth_uniq	Not applicable	Not applicable
refNum_uniq	Not applicable	Not applicable
altNum_uniq	Not applicable	Not applicable

FIELD	DESCRIPTION	EXAMPLE
matchStatus		exact
fva_compound_ heterozygous_pop_freq	***Only in familial variant analysis*** g1000 allele frequency limiting the appearance of the compound heterozygous group (if the variant was identified as a possible member of a group)	
fva_compound_ heterozygous_group	***Only in familial variant analysis*** id of compound heterozygous group to which the variant belongs (if any)	
fva_compound_ heterozygous_pval	***Only in familial variant analysis*** p-value (in log scale with with Benjamini–Hochberg correction) corresponding to the hypothesis that the disease is caused by the compound heterozygous genotype of the proband (one allele is the given variant and the second allele is another variant from the same group)	-0.1256
fva_pval_present	***Only in familial variant analysis*** p-value (in log scale and with Benjamini–Hochberg correction) corresponding to the hypothesis that the variant allele is present in the genotype	-0.1256
fva_pval_absent	***Only in familial variant analysis*** p-value (in log scale with Benjamini–Hochberg correction) corresponding to the hypothesis that the variant allele is absent in the genotype	-0.1256
fva_pval_homozygous	***Only in familial variant analysis*** p-value (in log scale with Benjamini–Hochberg correction) corresponding to the hypothesis that the genotype is homozygous variant	-0.1256
fva_pval_heterozygous	***Only in familial variant analysis*** p-value (in log scale with Benjamini–Hochber correction) corresponding to the hypothesis that the genotype is heterozygous variant	-0.1256
fva_pval_autosomal_ recessive	***Only in familial variant analysis*** p-value (in log scale with Benjamini–Hochberg correction) corresponding to the hypothesis that the variant is present and is a cause of an autosomal recessive disease	-0.1256
fva_pval_autosomal_ dominant	***Only in familial variant analysis*** p-value (in log scale with Benjamini–Hochberg correction) corresponding to the hypothesis that the variant is present and is a cause of an autosomal	-0.1256

FIELD	DESCRIPTION	EXAMPLE
	dominant disease	
fva_pval_de_novo	***Only in familial variant analysis*** p-value (in log scale with Benjamini–Hochberg correction) corresponding to the hypothesis that the variant is a de novo mutation causing the disease	-0.1256
fva_pval_x_linked	***Only in familial variant analysis*** p-value (in log scale with Benjamini–Hochberg correction) corresponding to the hypothesis that the variant is present and is a cause of an x-linked disease	-0.1256
OMIM	mim2gene identifier	600185
hg38_chrom	hg38 “chromosome” if panel is in hg38 (otherwise if panel is in hg19 it is the lifted hg38 equivalent)	13
hg38_pos	hg38 “genome_position” if panel is in hg38 (otherwise if panel is in hg19 it is the lifted hg38 equivalent)	32890572
hg38_ref	hg38 “ref” if panel is in hg38 (otherwise if panel is in hg19 it is the lifted hg38 equivalent)	G
hg38_alt	hg38 “alt” if panel is in hg38 (otherwise if panel is in hg19 it is the lifted hg38 equivalent)	A
lift_diagnostic	Tool used for liftover if carried out (if panel is in hg19)	PICARD
hg38_refGenome	“refGenome” if panel is in hg38 (otherwise if panel is in hg19 it is the lifted hg38 equivalent)	GRCh38/hg38
sgid	md5sum representation of the annotation_id. If panel is in hg38 this is equal to hg38_sgid (otherwise if panel is in hg19 it is the lifted hg19 sgid equivalent)	00010005c3cfe ba87b09a0c35c 99603e893aff80
hg38_sgid	md5sum representation of the annotation_id for hg38	0001010185b1b a2fd14b7152885 1d8058b589bf5
uniprot_acc	UNIPROT ACCESS code. Applicable only to coding mitochondrial proteins.	P03897
HGVS_gnomen	HGVS compliant representation of the genomic position of a variant.	NC_012920.1: g.7579472G>C
HGVS_pnomen	HGVS compliant representation of the impact of a variant at protein level.	P03897: p.(Tyr641Lys)

6.5.3 LiftOver

Hg19 coordinates are obtained for all called SNVs/INDELs through LiftOver. The hg19 lifted values are found in the full variant table in the regular coordinate columns. The called coordinates are placed in the columns with names starting with hg38_ prefix. The custom LiftOver procedure employs the Picard tool (see below for details).

Picard settings:

Version 2.23.8 of Picard is being used to lift-over variants with the following settings:

```
LiftoverVcf \
CHAIN=chain_file\
DISABLE_SORT=true \
ALLOW_MISSING_FIELDS_IN_HEADER=true \
WRITE_ORIGINAL_POSITION=true \
WRITE_ORIGINAL_ALLELES=true \
WARN_ON_MISSING_CONTIG=true;
```

The chain file for the lift- over from GRCh37 to GRCh38 was obtained from the ENSEMBL website: ftp://ftp.ensembl.org/pub/assembly_mapping/homo_sapiens/GRCh37_to_GRCh38.chain.gz (January 2022)

6.5.4 Non-lifted Variants And Rescuing Procedure

To keep the integrity between lifted outputs and the original VCF input file, it is necessary to parse the variants rejected for analysis by Picard to properly label the root cause of the lifting failure. Accordingly, Picard explicitly excludes variants for the four following failure cases:

1. **cannotLiftOver:** The variant requires reverse-complementing that fails for unclear reasons. The variant cannot be lifted.
2. **noTarget:** The variant cannot be lifted among genomes owing to null or insufficient overlap with the chain file.
3. **IndelStraddlesMultipleIntevals:** The variant is an INDEL that overlaps multiple intervals in the chain file.
4. **MismatchedRefAllele:** The variant coordinates are lifted over, but the reference allele in hg19 does not match that of hg38 at the corresponding location. This failure case can be handled using a dedicated rescuing procedure described below.

Rescuing Procedure

RevComp	The chain file specifies a change of strand for the locus, therefore the variant has been reverse complemented.	A/G	→	T/C
hg19var	The alt allele in hg19 is not different from the ref allele in hg38, i.e., the variant can be called only against hg19 and is not defined in hg38.	A/G	→	G/G
MRA	The reference allele in hg38 has changed from the reference allele in hg19, but the alt allele still differs from the reference.	A/G	→	C/G

RevComp_hg19var	The variant has been reverse complemented, but the complement does not differ from the reference allele.	A/G	→	C/C
RevComp_MRA	The variant has been reverse complemented, but the complement still differs from the reference allele, which has changed between assemblies.	A/G	→	G/C

6.5.5 Homologs And Pseudogenes

Variants in genes that possess homologs or pseudogenes will be labeled with an appropriate flag in SOPHiA DDM™. The flag is displayed on a per-gene basis. This annotation is knowledge-based, using multiple data sources. The flag does not indicate that the variant call is low confidence.

7 WARNINGS, LIMITATIONS, and PRECAUTIONS

7.1 Warnings

7.1.1 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 SOPHiA DDM™ Desktop App User Manual.
3. If any part of the handling, protocol, sequencer, multiplexing etc. is changed, the analyses are not covered by the described IFU parameters.

7.2 Limitations

7.2.1 General Limitations

1. While great care has been taken to ensure the highest performance and accuracy of this product, variant calling artifacts including False Positive and False Negative variant calls may occur, e.g. due to, but not limited to, factors associated with sample quality, sample preparation, probe design, DNA capture and amplification, sample multiplexing, sequencing, read alignment, read depth, variant calling, variant annotation, or variant cross-talk between samples.
2. 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.
3. All analyses depend on alignment of the sequencing reads to the hg38 human reference genome sequence and may therefore be adversely affected by limitations of that genome assembly like incorrect or incomplete reference genome representations. Furthermore, the analyses, including variant detection and annotation, are optimized for the main chromosomal sequences of the reference genome and do not support short alternate genome sequence contigs (e.g., "chr1_ KI270762v1_alt").
4. Variant cross-talk between samples may occur e.g. due to index hopping or low-level index cross-contamination. Such variant cross-talk 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.
5. Even when sample multiplexing recommendations are followed and a recommended AVERAGE number of reads per sample is achieved for a given sequencing run, (local) read depth in a given sample may turn out to be too low and insufficient to provide sensitive and accurate results for various reasons, for example:
    - Significantly uneven read allocation between samples.
    - Skewed read depth across the gene panel (e.g., consistently lower read depth in AT-rich regions).
    - Excessive mitochondrial read coverage in a given sample.
    - Poor sample quality.
    - Poor data quality.

7.2.2 Module-Specific Limitations

SNVs/INDELs

1. Problematic regions that may cause high levels of noise in exonic, intergenic or intronic locations have been identified. SNVs/INDELs in these regions can be artefacts and are therefore either flagged with a corresponding warning or are filtered and considered low confidence in SOPHiA DDM™.
2. 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) have an elevated risk of being subject to False Negative or False Positive SNV/INDEL variant calls.

3. For stable performance of the algorithms, we recommend read coverage of at least 20x and higher (e.g. 50x) if possible. In particular, lower read coverage significantly increases the risk of False Negative variant calls.
4. SNVs/INDELs in homopolymers of length ten or higher cannot be called confidently as their detection is confounded by high background noise for homopolymers of this length.
5. SNVs/INDELs in long repeat or low complexity nucleotide regions may be missed due to high levels of noise.
6. In rare cases, a variant may be represented in multiple forms in a given region. If one portion of the reads reports the variant in a different form or does not capture the variant, variant fraction may be under-estimated and the corresponding variant may be missed due to low variant fraction.
7. Variants may be missed or wrongly called due to limitations in the probe design or in the performance of kit components.
8. DNA fragments containing certain types and combinations of variants may not be effectively captured by the probes, which may prevent these variants from being correctly detected and reported, for example:
    - Large insertions or duplications that are longer than ~1/3 of read length.
    - DNA fragments that contain multiple variants, e.g. a variant of clinical interest that co-occurs together with other variants (e.g. common polymorphisms).
9. Large deletions in the nuclear genome that are longer than approximately 10,000 bp can only be detected with the CNV module. Large mitochondrial deletions of several kilobases in size are also displayed in the SNV/indel tab rather than the CNV tab.
10. Poor data quality due to issues in sample preparation or sequencing can confound the data analysis and cause False Positive or False Negative variant calls.
11. The same variant may have varying consequences when annotated with different transcripts, particularly transcripts of distinct genes.

CNVs

General Recommendations/Limitations

1. In order to provide the optimal performance of the CNV module, it is essential that all the samples in the batch are processed under the same laboratory condition. For processing involving PCR amplification, it is strongly recommended that only samples processed together at the PCR amplification step(s) are combined in one batch.
2. At the PCR step(s) of library preparation, temperature gradients should be avoided as much as possible (e.g., by not using the outer rows of PCR plates). Temperature gradients may produce false positive CNV calls.
3. A sufficient number of reference samples is needed for a reliable estimate of the reference coverage level. For good performance, we typically recommend having 20 samples in a batch (and 8 samples as an absolute minimum), not counting possible relatives and replicates.

    For MiSeq® v3 (300 bp reads), we recommend sequencing 8 independent samples in two runs (4 samples per run) to meet the minimum requirement of 8 samples for CNV analysis.

    For capture amplification technology, this number of samples should be maintained in each capture pool. These general recommendations may be overridden in the gene-panel-specific section below.
4. The CNV detection algorithm is based on the statistical analysis of coverage levels. Therefore, statistical fluctuations could result in false positive or false negative calls for borderline cases (e.g., short CNVs in noisy regions).

5. The CNV module is optimized to reduce the occurrence of false negatives (missed CNVs). As a consequence, one should expect a certain number of false positives and rejected regions.
6. Samples may get rejected for two reasons: either because of quality problems in the library preparation (e.g., failed PCR) or because of a different coverage pattern from the rest of the samples (for any reason). If the total number of non-rejected samples in the batch is below 4, the whole batch is rejected.
7. The algorithm is based on the statistical inference of CNVs from comparing samples to each other. Therefore, it assumes that, for each region, the CNV are only present in a small fraction of samples. It may fail to detect a CNV in batches with a large fraction of samples having CNVs at the same position.
8. The performance of the CNV module is generally higher for longer CNVs than for shorter CNVs and also higher for deletions than for duplications (insertions).
9. High AT/GC regions have higher coverage fluctuations which may lead to lower performance of CNV algorithm in those regions (e.g. first exons of genes).
10. If a processed pseudogene is present for a region, possible actual CNVs in the corresponding gene may be missed (masked by the processed pseudogene). Duplications in these genes may correspond to processed pseudogenes.
11. The algorithm is designed to detect integer copy number changes in the germline application; hence it is not designed to detect mosaic variants.
12. The maximal copy number considered in the algorithm is copy number 6.
13. In case of a target region covered by only one probe, a deletion may be called in case of variants in the probe region, if they are sufficiently significant to suppress probe binding.
14. The CNV detection algorithm only reliably detects CNVs with whole (or almost whole) target regions deleted or duplicated. In case of CNVs with breakpoints inside target regions, the boundaries of such CNVs may be determined incorrectly.
15. For protocols with samples pooled for capture, CNV detection performs best when captures are processed separately. Combining several capture pools in one batch is possible, if their number remains small. Typically, the number of capture pools combined for CNV processing should not exceed four. This recommendation may be overridden in the gene-panel specific section below.

Gene Panel-specific Limitations

1. In order to improve the stability of the CNV-detection algorithm, some regions are excluded from CNV detection. Regions may be excluded for different reasons, such as pseudogene issues, polymorphic copy-number variations, large coverage fluctuations, regions from the Y chromosome. Out of 65,736 regions included in the panel, 5,698 regions are excluded and 60,038 regions are retained. The full list of excluded and retained regions is available separately.
2. In this large panel, the resolution of the CNV detection adjusts to the sequencing depth. As a guideline, at 160M pair-ended reads (80M fragments) per sample, we expect the CNV resolution limit to be around 1 region, at 40M pair-ended reads (20M fragments) around 1-2 consecutive regions, at 20M pair-ended reads (10M fragments) around 2-3 consecutive regions, at 10M pair- ended reads (5M fragments) around 4-5 consecutive regions. The resolution limit is understood here as the minimal size of CNV which is detected with a sensitivity higher than 50%; the limit for a confident CNV detection may be slightly higher.
3. For detecting CNVs in the X chromosome, the sex of the subject is determined automatically from the ratio of the coverage of the Y chromosome to that of the autosomes (otherwise, regions in the Y chromosome are not included in the CNV detection). This method may produce inaccurate results, and it is recommended to check the automatically determined sex against the patient records.

4. In spite of excluding many known polymorphic and pseudogene-affected regions, some of them may still remain in less known genes. CNV results may be inaccurate in such regions, in case of frequent polymorphic CNVs or of read misalignment between homologous regions.
5. CNV analysis for this gene panel is based on the hg38 reference genome.
6. In this panel, samples are additionally rejected if they have an excessive number of CNV calls (more than 100 stretches).

Familial Variant Analysis

1. As any probabilistic analysis based on p-values, the familial variant analysis is a "rejection scheme": p-values associated to variants are used (by SOPHiA DDM™; not in the pipeline) to filter out the possibilities that are inconsistent with the observed data (sequencing data or the clinical status). The possibilities that are characterised by p-values smaller than the predefined confidence level are rejected while the others are accepted. Only small p-values (around the confidence level) carry sensible information. The advantage of such a rejection scheme is the controlled (and low) rate for the "false positive" errors, i.e. the situations when a variant that was actually responsible for the phenotype was filtered out and was not presented to the user as a possible cause of the phenotype.
2. The p-values computed in the familial variant analysis are used by SOPHiA DDM™ to call the genotype of the family members (wild type, heterozygous variant or homozygous variant). In some cases the available data may not be enough for a reliable identification of one of those three possibilities. In particular, as a result of e.g. statistical noise at the given locus being slightly different from the one assumed by the system or other reasons, the observed data can look consistent (number of alt counts around 75%) with both homozygous variant AND heterozygous variant genotypes. Alternatively, the number of observed alt counts can be a bit too small to correspond to the heterozygous variant genotype but still too large to be attributed to the pure noise. In BOTH of these cases SOPHiA DDM™ will call the “genotype 4” with the meaning that variant allele is (almost definitely) present but the precise genotype cannot be determined.

Mitochondrial Analysis

1. The same limitations for calling SNVs and INDELs in the main chromosomes listed above also apply for calling these variants in the mitochondria. In contrast to germline applications where the variant fractions expected are 50 or 100%, mitochondrial variants can be heteroplasmic and therefore show a continuous range of variant fractions for different variants. For mitochondrial variants present at low variant fractions, any variant detected below a variant fraction of 5% will be filtered (marked as low_variant_fraction), and will not be detected below a variant fraction of 0.9%. Therefore, variants at very low levels of heteroplasmy may be present in the sample but will not be addressed by this method.
2. Additionally, it is important to note that coverage of mitochondrial variants is in general higher than variants on the main chromosomes. This is due to the fact that mitochondrial chromosomes exist in greater copy numbers in the cell than the nuclear chromosomes. Therefore, differences in depth of coverage between genes located in the mitochondria and genes on the nuclear chromosomes is to be expected.
3. Low coverage related statistics for the whole panel (main chromosomes and mitochondria) in the QA report and in the flagged regions are based on the threshold of 20x for the main chromosomes and 1000x for the mitochondria.
4. Large mitochondrial deletions of several kilobases in size are also displayed in the SNV/indel tab rather than the CNV tab.

Annotation

1. Annotation catalogs may be of varying quality and subject to change with time. Those facts can expose the annotation system to exceptions that find their origin into the original data providers. As a consequence, the external data served through the annotation system cannot be fully guaranteed against regressions; although the complete data processing (import, reformatting and database storage) is performed under tight control and quality checks, so as to offer the best guaranties on data fidelity and integrity.
2. Catalog updates may impact on annotations and cause changes in variant priorizations. The ABCD and ACMG scores are notably subject to such variations, notably when updating ClinVar. Validations are performed at update time to assess and mitigate interpretation risks. Nevertheless, the user is invited to consider annotations as mutable data and therefore subject to change with time.
3. LiftOver operations: coordinate conversions among reference genomes relies on external algorithms (Picard) and pre-computed resources (chain files). The present product limits the use of LiftOver operations to functionalities enabling the cross talk of variant flags and frequencies among hg19 and hg38 products. The product relies on several quality checks and emits diagnostic flags allowing the user to gauge the LiftOver result. Nevertheless, the user is invited to consider that information transfers between hg19 and hg38 products are provided within the limitations of the LiftOver approach.
4. Transcript selection: for variant in any gene with a known MANE Select transcript, this transcript will be chosen for variant annotation, unless the MANE Select transcript does not overlap with the target list. For variants in genes without available MANE Select transcript, RefSeq transcript with a worst consequence will be chosen for annotation.

8 SYMBOLS

SYMBOL	TITLE
	Consult instructions for use
	Catalog number
	Batch code (Lot Number)
	Caution
	Manufacturer
	Date of manufacture
	Temperature Limit
	Use-by date
	Research Use Only
	Contains sufficient for <n> tests
	Refer to Warnings and Precautions section.
	Refer to Warnings and Precautions section.
	Refer to Warnings and Precautions section.
	Refer to Warnings and Precautions section.

SYMBOL	TITLE
	Refer to Warnings and Precautions section.
	Box 1
	Box 2

9 SUPPORT

In case of difficulty using the SOPHiA DDM™ Desktop App, please consult the troubleshooting section of the "General information about usage of SOPHiA DDM™" document or contact our support line by telephone at +41 21 694 10 60 or e-mail support@sophiagenetics.com. Please visit www.sophiagenetics.com for further details.

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 support@sophiagenetics.com if there are any concerns with the kits.

10 APPENDICES

10.1 Appendix I: Dual Index Adapter Plates

10.1.1 16 Illumina®-compatible Dual Index Adapters In 96-well Plate Format (7 µl Each)

	1	2	3	4	5	6	7	...	12
A	701-501	701-502	701-503	701-504					
B	702-501	702-502	702-503	702-504					
C	703-501	703-502	703-503	703-504					
D	704-501	704-502	704-503	704-504					
E									
F									
G									
H									

10.1.2 32 Illumina®-compatible Dual Index Adapters In 96-well Plate Format (7 µl Each)

	1	2	3	4	5	6	7	...	12
A	701-501	701-502	701-503	701-504					
B	702-501	702-502	702-503	702-504					
C	703-501	703-502	703-503	703-504					
D	704-501	704-502	704-503	704-504					
E	705-501	705-502	705-503	705-504					
F	706-501	706-502	706-503	706-504					
G	707-501	707-502	707-503	707-504					
H	708-501	708-502	708-503	708-504					

10.1.3 48 Illumina®-compatible Dual Index Adapters In 96-well Plate Format (7 µl Each)

	1	2	3	4	5	6	7	...	12
A	701-501	703-502	705-503	707-501	709-502	711-503			
B	702-501	704-502	706-503	708-501	710-502	712-503			
C	703-501	705-502	701-504	709-501	711-502	707-504			
D	704-501	706-502	702-504	710-501	712-502	708-504			
E	705-501	701-503	703-504	711-501	707-503	709-504			
F	706-501	702-503	704-504	712-501	708-503	710-504			
G	701-502	703-503	705-504	707-502	709-503	711-504			
H	702-502	704-503	706-504	708-502	710-503	712-504			

10.1.4 96 Illumina®-compatible Dual Index Adapters In 96-well Plate Format (7 µl Each)

	1	2	3	4	5	6	7	8	9	10	11	12
A	701-501	702-501	703-501	704-501	705-501	706-501	707-501	708-501	709-501	710-501	711-501	712-501
B	701-502	702-502	703-502	704-502	705-502	706-502	707-502	708-502	709-502	710-502	711-502	712-502
C	701-503	702-503	703-503	704-503	705-503	706-503	707-503	708-503	709-503	710-503	711-503	712-503
D	701-504	702-504	703-504	704-504	705-504	706-504	707-504	708-504	709-504	710-504	711-504	712-504
E	701-505	702-505	703-505	704-505	705-505	706-505	707-505	708-505	709-505	710-505	711-505	712-505
F	701-506	702-506	703-506	704-506	705-506	706-506	707-506	708-506	709-506	710-506	711-506	712-506
G	701-507	702-507	703-507	704-507	705-507	706-507	707-507	708-507	709-507	710-507	711-507	712-507
H	701-508	702-508	703-508	704-508	705-508	706-508	707-508	708-508	709-508	710-508	711-508	712-508

10.1.5 Index Sequences For The Illumina®-compatible Dual Index Adapters

The table below lists the index sequences. Please note that when using manual Sample Sheets, the following Illumina® instruments require the **reverse orientation** of the i5 sequence:

- NextSeq® 500/550
- NextSeq® 1000/2000 in Standalone mode (Sample Sheet v1)
- HiSeq® 3000/4000/X
- NovaSeq™ 6000 with v1.5 reagent kits and NovaSeq™ X/X Plus
- MiniSeq™ (except when used with Rapid Reagent Kits)
- iSeq™ 100

All other Illumina® instruments (e.g., MiSeq®) and instrument configurations require the **forward** i5 orientation.

This information is subject to change. Please always refer to the Illumina® support pages for the latest information on which sequencing systems require the i5 index sequence in forward or reverse (complement) orientation:

https://support-docs.illumina.com/SHARE/AdapterSequences/Content/SHARE/AdapterSeq/Overview.htm

i5	i5 SEQUENCE
D501	AGGCTATA
D502	GCCTCTAT
D503	AGGATAGG
D504	TCAGAGCC
D505	CTTCGCCT
D506	TAAGATTA
D507	ACGTCCTG
D508	GTCAGTAC

i7	i7 SEQUENCE
D701	ATTACTCG
D702	TCCGGAGA
D703	CGCTCATT
D704	GAGATTCC
D705	ATTCAGAA
D706	GAATTCGT
D707	CTGAAGCT
D708	TAATGCGC
D709	CGGCTATG
D710	TCCGCGAA
D711	TCTCGCGC
D712	AGCGATAG

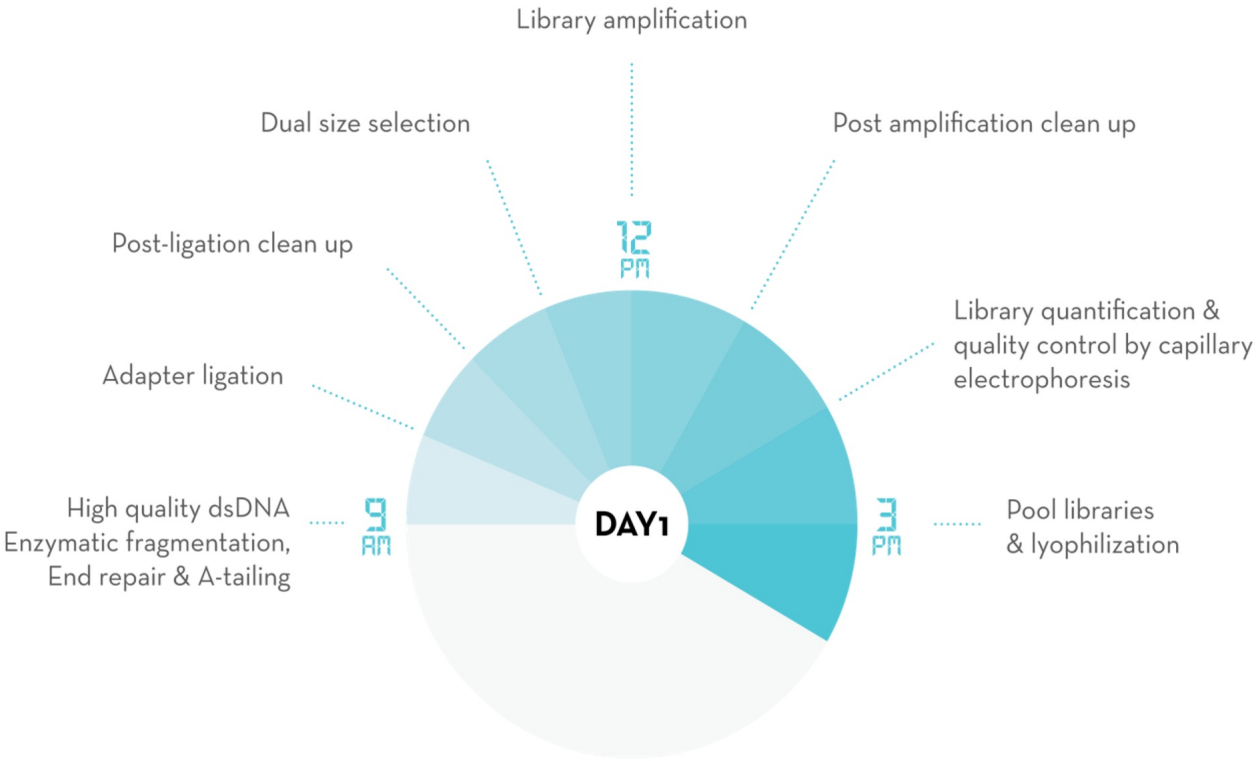
10.2 Appendix II: Laboratory Equipment Used in SOPHiA GENETICS Laboratory

USER-SUPPLIED MATERIALS	SUPPLIER	PRODUCT NAME	CATALOG N°
RNase/DNase-free 8-tube strips (0.2 ml)	Thermo Fisher Scientific	EasyStrip Snap Tubes	AB-2000
DNA low binding tubes (1.5 ml)	Axygen	MaxyClear Microcentrifuges Tubes	MCT-175-C
Tubes (1.5 ml)	Eppendorf	Eppendorf Tubes	3810X
Conical tubes (15 ml and 50 ml)	Falcon	15 ml & 50 ml Conical Centrifuge Tubes	352096 & 352070
Filter tips	Starlab	TipOne RPT	S1180-3710, S1183-1740, S1180-8710, S1180-9710, S1182-1730
Ethanol (molecular biology grade)	Merck	Ethanol Absolute	1.00983.1000

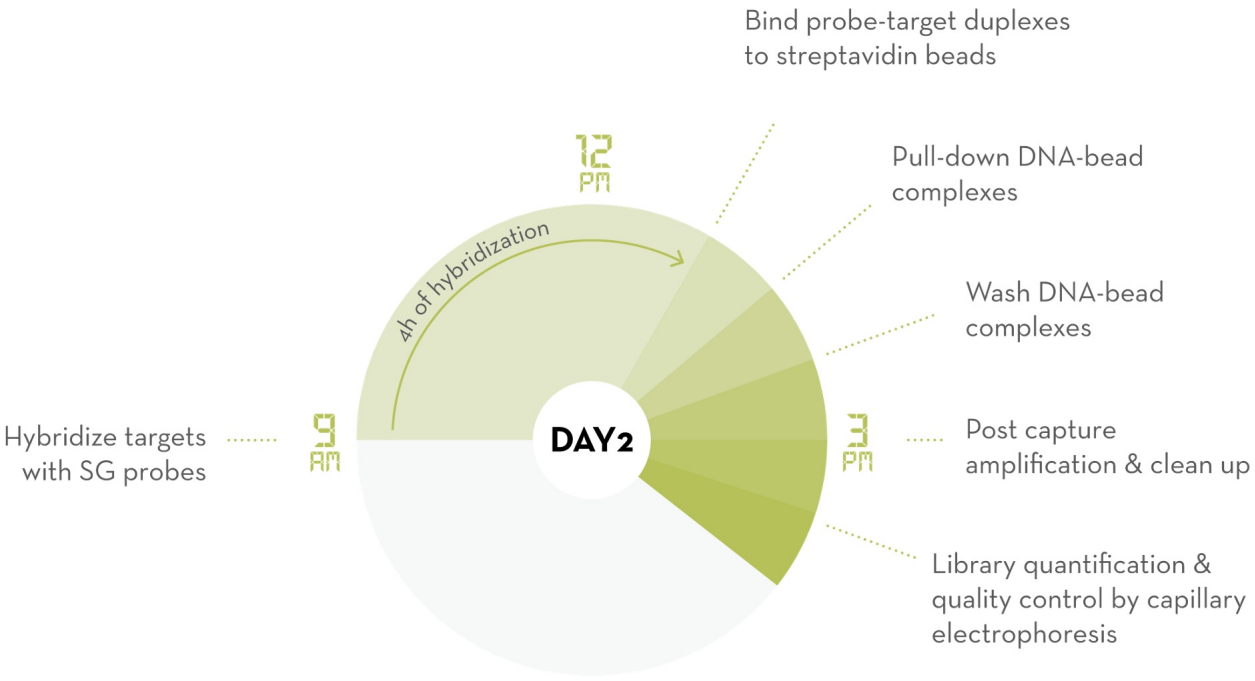
PRE-PCR ZONE	SUPPLIER	PRODUCT NAME	CATALOG N°
Vortex mixer	Scientific Industries	Vortex Genie 2	SI-0236
Table top microcentrifuge (8-tube strips compatible)	Starlab	Mini Centrifuge	N2631-0007
Magnetic separation rack 96-well type	Alpaqua	96S Super Magnet Plate	A001322
Magnetic separation rack 96-well type	Thermo Fisher Scientific	DynaMag-96 Side Magnet	12331D
Multichannel pipettes (P10; P100; P300)	StarLab	ErgoOne	S7108-0510, S7108-1100, S7108-3300
Thermal cycler with programmable heated lid	Biometra	TAdvanced 96	
Fluorometric quantitation equipment and reagents	Thermo Fisher Scientific	Qubit 3.0 Fluorometer & Qubit dsDNA HS Assay kit	Q33216 & Q32854
Single channel pipettes (P10; P100; P200; P1000)	StarLab	ErgoOne	S7100-0510, S7100-1100, S7100-2200, S7100-1000

POST-PCR ZONE	SUPPLIER	PRODUCT NAME	CATALOG N°
Thermal cycler with programmable heated lid	Biometra	TAdvanced 96	
Capillary electrophoresis system	Advanced Analytical	Agilent Fragment Analyzer	
Vacuum concentrator (SpeedVacTM or similar)	Thermo Fisher Scientific	Savant DNA120-230	
Dry block heater or water bath(1.5 ml tube compatible)	Techne	Dri-Block DB-1	
Magnetic separation rack (1.5 ml tube compatible)	Thermo Fisher Scientific	MagJET Separation Rack, 12 x 1.5 mL tube	MR02
Magnetic separation rack (96-well type)	Alpaqua	96S Super Magnet Plate	A001322
Magnetic separation rack 96-well type	Thermo Fisher Scientific	DynaMag-96 Side Magnet	12331D
Vortex mixer	Grant instrument	Multi-tube Vortex Mixer, V32	
Vortex mixer	Scientific Industries	Vortex Genie 2	SI-0236
Table top microcentrifuge (8- tube strips compatible)	StarLab	Mini Centrifuge	N2631-0007
Multichannel pipettes (P10; P100; P300)	StarLab	ErgoOne	S7108-0510, S7108- 1100, S7108-3300
Fluorometric quantitation equipment and reagent	Thermo Fisher Scientific	Qubit 3.0 Fluorometer & Qubit dsDNA HS Assay kit	Q33216 & Q32854
Single channel pipettes (P10; P100; P200; P1000)	StarLab	ErgoOne	S7100-0510, S7100- 1100, S7100-2200, S7100-1000

10.3 Appendix III: General Workflow SOPHiA DDM™ Capture Solution



Library Preparation with SOPHiA GENETICS™ DNA Library Prep Kit I



CAPTURE
EASY WORKFLOW
- ONLY 1-4 TUBES TO HANDLE (MULTIPLEX POOLED LIBRARIES)
- ONLY 3 HOURS HANDS-ON TIME

Document Approvals

Approved Date: 08 Feb 2024

Approval Verdict: Approve	Martin Fritzsche, (mfritzsche@sophiagenetics.com) Technical Approval 01-Feb-2024 19:32:37 GMT+0000

Approval Verdict: Approve	Paolo Florent, (pflorent@sophiagenetics.com) Regulatory Approval 08-Feb-2024 12:44:43 GMT+0000

QA Approval Verdict: Approve	Denisa Dzulova, (ddzulova@sophiagenetics.com) Quality Assurance Approval 08-Feb-2024 13:51:00 GMT+0000