# Analysis Quick Start: Annotated Rare Variant Calling

DOCUMENT NUMBER: CG-30375

DOCUMENT REVISION: B

Effective Date: 04/16/2024

For Research Use Only. Not for use in diagnostic procedures.

#### **Table of Contents**

Revision History	3
Rare Variation Detection - Introduction	4
Workflow Overview	5
Accessing Bionano Data and Filtering with Recommended Settings	6
Goal of the Tutorial	6
Entering Bionano Access	6
Variant Annotation Pipeline	8
Circos Plot View	10
Visual Comparison of Circos Plot (optional)	11
Identify Somatic Variations in a Tumor - Normal Comparison	13
Practice Session – Identifying Potentially Pathogenic SVs, VAP Filter	18
Filter Based on Gene Annotations	18
Adding BED Files for Filtering	19
Report Generation, Exporting Data, and Creating Images	26
Report Generation and Data Export	26
Creating Images	28
Resources and Information	28
Technical Assistance	30
Legal Notice	31
Patents	31
Trademarks	31

### **Revision History**

REVISION	NOTES
A	Initial release.
В	Update screenshots and candidate gene analysis description.

#### **Rare Variation Detection - Introduction**

Bionano Access supports different workflows for the detection of structural variants (SV) in the human genome (**Figure 1**). As per Bionano technology, a rare variant is defined as one that is present in low abundance in the sample and is not present in the reference molecule. To identify such variants effectively, a dedicated bioinformatics pipeline is used that locally aligns molecules to a reference, assembles them with putative differences into consensus maps, and identifies the structural variations with less computational burden. For more information on each of these workflows, please see *Bionano Solve Theory of Operation: Structural Variant Calling* (CG-30110), *Bionano Solve Theory of Operation: Variant Annotation Pipeline* (CG-30190), and *Bionano Solve Theory of Operation: EnFocus FSHD Analysis* (CG-30321). To obtain information on data coverage targets, which can vary depending on the analysis, see Data Collection Guidelines (CG-30173). To obtain information on how Bionano determines quality control of raw data, please refer to *Bionano Access Dashboard and Chip Metrics Guidelines* (CG-30304) and *Bionano Access Molecule Quality Report Guidelines* (CG-30223).



**Figure 1**. Schematic representation of the Bionano structural variant detection workflows. Variant Analysis Pipeline - includes first constructing a *de novo* map and then comparing the resulting map with a reference which is highly sensitive in detecting SVs.

#### **Workflow Overview**

This workflow, shown in **Figure 2**, discusses how to filter and narrow a search for high confidence SVs that were identified using the Rare Variant Analysis pipeline (RVA). Once DNA is isolated and labeled, Bionano chips are loaded and run on the Saphyr instrument. Molecules and labels imaged on the Saphyr instrument are digitized and saved as in a BNX format (raw data). This serves as input for downstream analyses. RVA detects SVs by the "split-read" analysis, where initial molecule alignment and molecule extension refinement allow identification and detection of somatic variants and identifies variants that are 5,000 bp or longer in size. For more information on the Variant Annotation Pipeline and relevant file formats, please refer to *Bionano Solve Theory of Operation: Variant Annotation Pipeline* (CG-30190) and *OGM File Format Specification Sheet* (CG-00045)..

To access the results from an RVA analysis, the user begins by logging into the web server location which contains user data. Once logged in, the user navigates to the project which contains the data for the RVA experiment of interest. RVA can be used for a single sample, duo (matched tumor-normal, for example) or trio (parents-child, for example) analysis. For any individual experiment, clicking on the sample data will take the user to a Circos plot which has a global overview of the mapped genome.

The user may then perform filtering of the rare variants and identify genes of interest that might have been affected by a rare variant event. Additional filtering and masking of low confidence calls may be performed using a BED file. Once the dataset has been filtered appropriately the user can visually explore the structural variants, create a report, and then export the data for further bioinformatic analysis. For more details on thresholds and experimental designs refer to *Bionano Solve Theory of Operations: Variant Annotation Pipeline* (CG-30190).



Figure 2. Bioinformatic workflow for post RVA data analysis in Bionano Access

# Accessing Bionano Data and Filtering with Recommended Settings

#### Goal of the Tutorial

This quick start guide directs users on how to view annotated rare variants, which are detected from Bionano's genome mapping data. It can be used to evaluate a user's own data or to familiarize users with the software using a pre-loaded demo dataset. The demo dataset is provided by Bionano to help users get familiarized with the Bionano Access software interface that enables analysis, visualization, and data export. To walk through the analysis, follow the numerical prompts in this guide. Text outside of the numerical steps is provided to facilitate learning. This guide assumes all necessary analyses pipelines have been run on user samples, and so this tutorial will guide users through selecting the proper data, and performing filtering based on current Bionano guidance before exporting the data for presentations or further analysis outside of Bionano Access. The filter sets used here represent current recommended settings but may be adjusted to fit the needs of a project. In the hands-on section of the tutorial, users will be walked through the analysis of tumor (HCC2281) and non-tumor (HCC2281BL) cell lines established from the same patient with breast cancer. The goal is to identify rare somatic variants in the HCC2281 breast carcinoma cell line when compared to its matched normal cell lines.

To obtain information on data coverage targets, which can vary depending on the analysis, see *Data Collection Guidelines* (CG-30173). To obtain information on how Bionano determines quality control of raw data, please refer to *Bionano Access Dashboard and Chip Metrics Guidelines* (CG-30304) and *Bionano Access Molecule Quality Report Guidelines* (CG-30223). To obtain information on ways by which bioinformatic analysis of annotated structural variants are recommended, see *Bionano Solve Theory of Operation: Variant Annotation Pipeline* (CG-30190).

#### **Entering Bionano Access**

 Using a web browser (Google Chrome is recommended), navigate to the web address of the user's Bionano Access data (e.g., <u>https://us.bionanoaccess.com</u>). Users with an installed Saphyr<sup>™</sup> System will utilize an Access Installation hosted on their local Bionano Access Server. Enter the appropriate User Name and Password, then press Login (Figure 3). To access the demo dataset or user's sample data generated by the Bionano Commercial Services Lab, a user will need login credentials as provided by a Bionano representative.

User Name			_
your login			
Password			

Figure 3. Login to Bionano Access

2. To open user sample data or demo project data, click on **Projects** on the Home page (**Figure 4**). A Project is the central repository for all objects such as experiments and any bioinformatic analysis that is performed.



Figure 4. Bionano Access login screen.

 In the Projects view, click on the hyperlinked project name corresponding to the submitted project, or BionanoAccess\_somatic to follow along with this demo dataset (Figure 5). This will open the list of data types (Objects) within that Project (Figure 6).

Project 🏹	Key 🍸	Description	Date Created 🛛 🕇	Export	Import	Setting	Delete	Edit	Access
BionanoAccess1.5_FSHD			06/01/2020 06:34 PM	1	*	\$	Û		
BionanoAccess_constitutional			06/01/2020 06:34 PM	1	*	•	Û		
BionanoAccess_somatic			06/01/2020 06:35 PM	1	*	•	Û		
Demo_BionanoAcces BionanoAccess_somatic			08/08/2021 01:19 AM	1	*	•	Û		
SV_Catalog_demo_BionanoAccess1.6			08/04/2021 12:10 PM	1	*	\$	Û		
SV_Catalog_demo_BionanoAccess1.6_molecul- es			08/20/2021 09:42 AM	<b>1</b>	*	•	Û		
Page 1 of 1	$\mathbf{D}$							1 - 6 c	of 6 items



The **Project Data** page (**Figure 6**) will list the various data types for a given sample or analysis. For rare variants with the variant annotation pipeline, a project may contain some or all the following file **Types**.

- a) **Molecules:** the raw, unassembled, molecules data (also known as a .bnx filetype). A Molecules file is generated by the image analysis software. It contains single-molecule optical mapping data generated on a Bionano Saphyr instrument and is the first data type produced. These may or may not be present and are not essential for visualization.
- b) De novo Assembly: the assembled consensus genome maps (also known as a .cmap filetype). The de novo Assembly file contains the assembled consensus genome maps which are then compared against a reference assembly (e.g., hg19 or hg38) to produce un-annotated SV calls. These may or may not be present.
- c) **Rare Variant Analysis:** The results from the rare variant annotation pipeline are output to a text file with/without annotation for each SV detected for each sample.
- d) Annotated Rare Variant Analysis: This file is unique to a duo analysis. The important thing about this filetype is that it is 1) annotated 2) compared to the internal control database. For duo analysis, the calls are then compared to each other and contain the rare variants identified in the cancer sample compared to its matched control samples.

On highlighting each file type, different operations become visible under the **Options** pane on the right.



Figure 6. Project Data page. This contains a list of objects and a list of operations (options) that can be performed with the selected object.

#### Variant Annotation Pipeline

In this tutorial, users will find the Project contains at least three different files, as seen in Figure 7:

- 1. HCC2218 Rare Variant Analysis (Solve3.6).
- 2. HCC2218BL Rare Variant Analysis (Solve3.6).
- 3. HCC2218BL Variant Annotation Pipeline

The rare variant calls from each sample (HCC2218 and HCC2218BL) are saved as separate files. The third file includes the dual annotation file which includes all rare variants identified in the samples compared to an internal database of more than 200 normal, apparently healthy individuals. In this tutorial, the dual analysis file is used to identify rare somatic variants.

Given sample name	Object type	Auto-generated sample name	Curation status	Sample description	Date of creation
Sample 💙	Туре 🕇	Name Y	Curation <b>Y</b>	Tag 🕇	Cre 🝸
HCC2218BL	Annotated Rar	HCC2218BL - Rare Variant Analysis	Curation	demo-data-1.7	12/03/202
HCC2218BL	Variant Annota	HCC2218BL - Variant Annotation Pipeline 1.6	Curation		12/02/202
HCC2218BL	Variant Annota	HCC2218BL - Variant Annotation Pipeline	Curation		11/20/202
HCC2218BL	Annotated Rar	HCC2218BL - Rare Variant Analysis - solve3.6	Curation	demo-data-1.6	11/18/202
HCC2218BL	Molecules	HCC2218BL_mol			11/18/202
HCC2218BL	Rare Variant A	HCC2218BL_rareVariant_import			06/16/202
(I) Page	1 of 1 (	$(\mathbf{F})$		1 -	6 of 6 items

Figure 7. Selecting the Variant Annotation Pipeline data type (file type). Each column lists the auto-generated sample name for a given sample

To view a sample, click on the sample name. Depending upon the experimental design, the content may be different. To identify the files used to generate the dual analysis, **click**, and highlight **HCC2218BL-Variant Annotation Pipeline** as shown in **Figure 8**. Check the sample name under "Content." For more information see *Bionano Solve Theory of Operation: Variant Annotation Pipeline* (CG-30190).

Sample 🍸	Туре 🕇	Name T	Cu 🝸	Tag 🕇	Cr <b>T</b>	Sample Details	6		
HCC2218BL	Annotated	HCC2218BL - Rare Variant Analysis	Curation	demo-data-1.7	12/03/20	Name	HCC2218BL		
HCC2218BL	Variant Ann	HCC2218BL - Variant Annotation Pipeline 1.6	Curation		12/02/20	Sample UID Created	17a0af58-b0 06/16/2020 <sup>-</sup>	1c-11ea-9 16:55 PM	9da8-0a03396e7481
HCC2218BL	Variant Ann	HCC2218BL - Variant Annotation Pipeline	Curation		11/20/20	Variant Annota	tion Pipeline	Details	
HCC2218BL	Annotated	HCC2218BL - Rare Variant Analysis - solve3.6	Curation	demo-data-1.6	11/18/20	Name	HCC2218BL	Variant	Annotation Pipeline
HCC2218BL	Molecules	HCC2218BL_mol			11/18/20	Poforonao	1.6		
HCC2218BL	Rare Varian	HCC2218BL_rareVariant_import			06/16/20	Description	ngso_DLE I		bers.cmap
(H) Pag	e 1 of	f1 ())		1 - 6	of 6 items	Created	12/02/2020 (	01:44 PM	volino
$\bigcirc \bigcirc$						Status	Complete	nation Fip	Jenne
						User	Dipa Roycho	oudhury	
						Job ID Command	★ 2556 Q Show Cor	mmand	
						Content			
						Joh Tune	Sample	Input	Innut Nome
						Annotated Rare	Name	Туре	HCC2218BL - Rare Variant
						Variant Analysis	HCC2218BL	Case	Analysis - solve3.6
						Variant Analysis	HCC2218	Control	Analysis - solve3.6

Figure 8. Dual Analysis File-The content lists the samples compared in this analysis. The first operation under **Options** is **View variant annotation results**. This launches the Circos plot.

4. Click on **View Results** to launch the Circos plot.

#### **Circos Plot View**

Users can view one sample or file type at a time in the Circos plot view. In a duo analysis, all SVs detected in normal versus tumor samples that is compared against the control database can be viewed on a Circos plot. Using the filer settings as discussed below, users can display SVs unique to tumor or normal.

In Bionano Access, the Circos plot (**Figure 9**) is displayed by default for human assemblies and variant annotations. The Circos plot is displayed as the circular plot on the left, with **Filter Criteria**, **SV Filter Confidence** scores, **BED Files**, and **Samples** on the right. To zoom on the Circos plot, rotate the mouse-wheel forward (scroll down). The user can also click and drag the plot around using the mouse. **NOTE**: For more mouse and keypad short cuts used to maneuver the Circos plot, click on the **?** icon, which is the last icon on the top panel.



Figure 9. Circos plot view includes a top panel with filtering options (E), options to add BED files (G), Sample Name (H), summary SV counts (I), and the central Circos plot with the different tracks (A-D).

In a Circos plot, each chromosome is positioned end-to-end in a large circle. Each track represents a different type of SV or annotation result. Clicking on the cytoband, SV, or copy number track will open the genome browser, and show the user a detailed view of a Bionano consensus genome map for that locus. The user can navigate back to this Circos plot from other views by clicking on the circle icon on the Top Panel (E) toolbar. The rings of the Circos plot are described from the outermost track to the inner most track.

- The outermost numerical track (A) corresponds to the chromosome number, with cytoband information shown in the black-and-white banding pattern. Currently, chromosome 23 corresponds to X and chromosome 24 to Y.
- The next track (B) includes the detected SVs as represented by color-coded dots. A user can hover over the dot to view information about a given SV call.
- The next track (C) shows copy number calls. For the copy number track, the baseline state of autosomes in a diploid genome is a copy number of 2. If a region shows a copy number gain, the line will move outward from 2 (colored blue), reflecting an increase in the local copy number state. If the region has a loss, it will move inward (colored red).

- The innermost track (D) contains information about translocations, both inter-chromosomal and intrachromosomal (intra-chromosomal translocations are defined by an abnormal fusion within one chromosome with breakpoints at least 5 Mbps apart). The lines (colored magenta) going from one chromosome to another are drawn to show chromosomal rearrangements.
- The top panel (E) of Figure 9 includes Anchor which corresponds to the chromosome number.

**NOTE: Range** (which corresponds to coordinate ranges on the genome) allows the user to view SVs of interest by specifying genomic coordinates.

By default, when using Access for the first time, all SVs detected can be visualized in the Circos plot and the **Filter Criteria** (F) lists the default settings used to detect rare variants in the analysis pipeline. Once changed, they will remain at that setting going forward until they are changed again or reset to default.

**NOTE**: In Bionano Access, "masking" refers to a genomic sequence that has been scanned for either N-base gaps, sub-centromeric and sub-telomeric regions that are prone to generating putative false positive translocation breakpoint calls. Apart from these, regions with putative SVs such as translocation breakpoints can be masked during the alignment step in Solve. Access allows users to view these annotations and regions during the SV filtering steps.

In this tutorial, the default settings including the SV Masking option were used.

#### Visual Comparison of Circos Plot (optional)

- 5. To compare the SVs detected in each sample, highlight the cancer sample HCC2218 (Figure 10).
- 6. Under the **Options** section on the right, click on the second option **View Results**. This will launch the Circos plot.



Figure 10. Launching the Circos plot for HCC2218.

7. Similarly, to view the Circos plot for the normal sample (HCC2218BL), first click on the back arrow icon on the top panel. This will navigate the user back to the **Project** page.



Figure 11. Launching the Circos plot for HCC2218BL

8. Click on the sample HCC2218BL to highlight it. From the **Options** section, click on **View Results** to launch the Circos plot as seen in **Figure 11**.



Figure 12. Comparison of the Circos plots from two samples with Tumor cell line (A) and Normal cell line (B)

For ease of comparison, two Circos plots are shown side by side in **Figure 12**. Note the difference in the nature of the distribution of rearrangements between the two cell lines depicted. The whole genome profile on the left (A) has 147 translocation rearrangements that are heavily clustered on Chromosome 17. In contrast, no such rearrangements were found in the normal (B) cell line present in the profile on the right-hand side.

**NOTE**: Users can customize the enzyme colors, SV colors, map (reference, genome) colors, minimum and maximum map height value, molecule and copy number height value, and others.

 To change the size of the SVs detected in the Circos plot, click on the Settings icon in the top panel as shown in Figure 13. Scroll down to find Circos plot SV size and change Circos plot translocation width size to 1 (Figure 14). This increases the size of variants in the SV track and the translocations in the innermost track. Click Save.



Figure 13. The Settings icon in the red box customization of the Circos plot includes Anchor and Range to navigate across the chromosomes and Molecules, Confidence and Find Map to customize the visualization area. The Settings icon in the red box customization of the Circos plot.

**NOTE**: Changed settings parameters are saved under the username and across datasets, then retained when Bionano Access is opened later. Select **Reset Options** to return to original default filters.



Figure 14. Viewing SV events. Customizable options that allow better visualization of the tracks are listed. Select Reset Options to return to default settings.

#### Identify Somatic Variations in a Tumor - Normal Comparison

In general, SV filters can be customized to identify structural variants that are *de novo* or proband-specific in family studies or somatic SVs in cancer datasets. They can also be used to identify rare variants found infrequently in the control database within single individuals, or variants that overlap with annotated genes. In this

tutorial, cancer-specific somatic variants are filtered for. The dual analysis for RVA filters only those cancer sample specific variations. To remove germline calls, Access filters can be adjusted to identify tumor-specific somatic variations from a dual annotation file called "HCC2218\_rareVariant\_import\_annotationDual file".

1. Click on either the back arrow or the home page to navigate back to the Project Browser (Figure 15).



Figure 15. Navigate back to the Project Data page to find the container file "HCC2218\_rareVariant\_import\_annotationDual file".

2. Follow instructions as described earlier to launch the Circos plot page for "HCC2218BL-Variant Annotation Pipeline."

**TIP:** To ensure none of the additional filtering options are active, check the filter criteria listed in **Figure 16** and click on **Reset Filters** before the recommended filters are chosen.

- 3. To choose filtering options, click the SV Filter funnel icon, shown in Figure 17.
- 4. The filter dialog window opens showing the Filter by SV Type options (Figure 18).



Figure 16. Filter Criteria for SVs are filtered on user defined criteria. By default, stringent confidence score settings are used to detect high confidence SVs.



Figure 17. Filter icon for SVs

In general, this wizard allows the user to set rules for identifying variants relevant to their data based on confidence of a call and the length of the detected SV. The confidence score of a detected SV is defined as the probability of the SV being a true positive call.

The filter dialog pop-up window contains four filter tabs:

- The **Filter by SV Type** tab allows basic filtering of type of SV using size cutoff and "recommended" or "all" confidence cutoffs. "Recommended" confidence cutoffs balance false positives and false negatives when calling SVs.
- The General SV Filters utilize the masking BED file, allowing users to select which chromosomes to display, the precision of the BED file overlap with called SVs and how the desired masking should occur (either by showing the areas that are masked or unmasked).
- The third tab is for controlling the Variant Annotation filtering and allows users to search for SVs present in a given percentage of the samples in the control database as well as determining whether SVs that are present in the control sample should be shown.
- The fourth tab is for filtering copy number variants.
- 5. For this demo dataset, choose the default or **Recommended** setting for SV Confidence and leave blank for SV Minimum Size (bp). In short, make no changes for the **Filter by SV Type** tab. Then click on **General SV Filters**.

r by S	SV Type General SV F	Filters Variant Annotation Filters	Copy Number Variant Filters
now	SV Type	SV Confidence 3	SV Minimum Size (bp)
2	Insertion	Recommended	~
-	Deletion	Recommended	<b>~</b>
2	Inversion	Recommended	•
2	Duplication	Recommended	~
2	Intra-Translocation	Recommended	~
2	Inter-Translocation	Recommended	~

Figure 18. Filter Settings include four different tabs.

6. In the General SV Filters tab (Figure 19), select the All chromosomes radio button. Leave SV BED Overlap Precision to 12 kbp to identify SVs overlapping a region of interest when using a BED annotation file to filter data. Select Non-Masked Structural Variants Only from the dropdown box under the SV Masking Filter.

er Settings			
Filter by SV Type	General SV Filters	Variant Annotation Filters	Copy Number Variant Filters
Chromosomes to D	isplay on Circos Plot:		
All chromosome	S		
Only chromosom	nes that have structural	variants	
Only chromosom	nes from this range:		
SV BED Overlap Pr	ecision (KDp):		
12			
SV Masking Filter:			
Non-Masked Struct	ural Variants Only		~

Figure 19. General SV Filters.

- 7. To filter for somatic variations unique to the cancer sample make the following changes in the **Variant Annotation Filters** tab, seen in **Figure 20**:
- SV in less than this percentage of the Bionano control samples: 100
- SV in less than this percentage of the Bionano control samples with the same enzyme: 100
- SV chimeric score filter: Show Not Failing Chimeric Score
- SV control sample assembly check: SV not found in control assemblies.
- SV control sample molecule check: SV not found in control molecules.
- SV self molecule check: SV found in self molecules.
- SV overlapping genes filter: SV with overlapping genes.
- Self Molecule Count: 5

ter Settings				
Filter by SV Type	General SV Filters	Variant Annotation Filters	Copy Number Variant Filters	
SV in less than this	% of the control sample	es with the same enzyme:		
100				
SV self molecule ch	eck:			
SV found in self me	olecules		~	
Self Molecule Count				
5				
SV in less than this	% of the control sample	PS:		
100				
SV chimeric score fi	ilter:			
Show Not Failing (	Chimeric Score		~	
SV overlapping gen	es filter:			
SV with overlappin	ng genes		~	
SV control sample a	assembly check:			
SV not found in co	ontrol assemblies		~	
SV control sample r	nolecule check:			
SV not found in co	ontrol molecules		~	

Figure 20. RVA Variant annotation filtering options. The filters in the variant annotation tab have eight different options that can be grouped in six different categories. (A) Check against somatic variants detected in control database of healthy individuals, (B) Check for Spurious alignment, (C) Check for against one control sample (for duo analysis) or two parental control samples (for trio analysis. Not shown), (D) Check for sample-specific assembly or molecule, (E) Check for overlapping genes.

The first option filters for somatic variants detected in a control database of healthy individuals. By setting these values to "100", the software will display structural variants which are present in any of the individuals in the control database. The control database is not exhaustive and does not include all variants. But the control database may help to identify pathogenic SVs that could be present in a control sample, those of which may not be "healthy" at the time of sample collection. The SV chimeric score filter ensures that regions on the Bionano maps from which SVs are detected have sufficient molecule support to create a confident consensus genome map. The next set of options shows filters to remove any variants detected in the matched normal sample. This is only applicable for a paired sample or a trio analysis. The SV **self molecule check** and **Self Molecule Count** options show filters to include either molecules that align only to the cancer sample (HCC2218) or limit the number of supporting molecules, respectively. The higher number of molecules supporting a call, the greater the confidence of that SV call. The option **SV overlapping genes** filter helps identify the biological importance of the calls by identifying genic information. For more details on any of the settings, please see *Bionano Solve Theory of Operation: Structural Variant Calling* (CG-30110). For more details, see the *Bionano Solve Theory of Operation: Variant Annotation Pipeline* (CG-30190) for guidance on setting these parameters.

After clicking Apply the user will be returned to the Circos Plot. Confirm that the SV counts lists three insertions, 36 deletions, 9 inversions, 9 duplications, 64 Intra-Translocation, 17 Inter-Translocation, 55 copy number gains and 25 copy number losses (Figure 21). These are represented by different colors on the SV track of the Circos plot. NOTE: the SVs detected after the filter is applied is lower than the SVs initially detected in sample HCC2218 (Figure 12A).

CC2218 - Variant Annotation Pipeline_sv	Cou
Insertion	3
Deletion	36
Inversion	9
Duplication	9
Intra-Translocation	64
Inter-Translocation	17
CNV Gain Segment	55
CNV Loss Segment	25

Figure 21. Structural Variants. Color coded by the different types.

There are many ways to identify a set of rare variants that might be of interest to a user. Detailed study is therefore needed to identify rare variants that are biologically important. One way to make sense of the data is to identify relevant, publicly available information for the region, type of SV or gene of interest. The next section describes how to apply a BED file to identify conserved genic regions in the overlapping genomic coordinates.

# Practice Session – Identifying Potentially Pathogenic SVs, VAP Filter

This section can be used to facilitate learning about how to find a specific SV of interest. Basic filtering to identify potential somatic variants that are rare among phenotypically "normal" individuals can be applied. Then a masking BED file to identify SVs specific to a gene of interest can be used. ATCC describes HCC2218 and HCC2218BL cell lines as patient derived lymphoblastoid cell lines that are highly positive for expression of Her2/neu and p53. HCC2218BL is an EBV transformed cell line from the same patient. The <u>Dependency Map (DepMap)</u> portal lists large numbers of mutations and gene fusion events for HCC2218 and also reported around <u>500 SNPs and 36</u> inter and intra-chromosomal translocation events. Of all the genes known to have a translocation, INTS2 genes was reported to have six different intra-chromosomal rearrangements. The cancer cell line HCC2218 can be checked to identify <u>INTS2</u> CNV gain as reported by DepMap.

As a proof of principle, simple filtering criteria can be applied to identify the ZNF652-INTS2 translocation in the HCC2218 cancer sample. To identify the genomic coordinates of the SVs impacting them, annotations to the SV calls should be added by providing overlapping gene information. The same filtering options can be applied to identify other variants rare among phenotypically "normal" individuals that may overlap with annotated genes or are potential false positive calls.

#### Filter Based on Gene Annotations

To visualize genic regions on a chromosome, users can import annotations from public resources such as the UCSC Browser or use a pre-loaded BED file available in the software. A BED file (.bed) is a tab-delimited text file that defines a feature track. Canonical, or known genes, datasets are constructed by identifying a single transcript or isoform for a gene that has many reported isoforms. For ease of analysis, only the longest isoform is chosen for a gene and its chromosomal coordinates are included in the BED file. Because the chromosomal coordinates are reported in a BED file, genes can be visualized as tracks in a Circos plot or genome map. Click <u>here</u> to read on how canonical genes are identified. In this tutorial, known genes corresponding to human build hg38 ("hg38 known canonical.BED") is used for analysis since raw data was aligned using the hg38 genome. This BED can be easily customized and has been manually curated to include recent genes.

#### Adding BED Files for Filtering

 To add a BED file, click the Add BED button on the Circos plot display page as shown in Figure 22; this will prompt the user to select from the numerous reference genome-specific BED files that can be easily added to any analysis. "hg38 known Canonical" will also be used to filter structural variants to just those that overlap known genes and view them in the Circos plot or the Genome Browser.

RED SV Overlap Precision	(Khp):	12	
RED CNIV Overlap Precision	(KOD):	500	
SV Masking Filter:	n (Kop).	non n	nasked
Copy Number Type:		all	in a since a
Copy Number Confidence		0.99	
Copy Number Min Size (b	o);	5000	00
Self Molecule Count:		5	
% in Control:		100	
% in Control for Enzyme:		100	
SV Chimeric Score:		all	
Found in Control Sample	Assembly:	no	
Found in Control Sample I	Molecules:	no	
Found in Self Molecules:		yes	
Overlap Genes:		yes	
SV Filter	Confidence	Min S	ize (bp)
Insertion	0		
Deletion	0		
Inversion	0.7		
<ul> <li>Duplication</li> </ul>	-1		
Intra-Translocation	0.3		
Inter-Translocation	0.65		
BED File	SV	CNV	Acti
hg38 cnv masks 🦰	۲	۲	ť
Add BED			
Camples			
Samples Reference: hq38 DLE1	Okb Olabels		
Samples Reference: hg38_DLE1_	0kb_0labels		
Samples Reference: hg38_DLE1_ HCC2218 - Variant Annotatio	Okb_Olabels		Cour
Samples Reference: hg38_DLE1_ HCC2218 - Variant Annotatio	Okb_Olabels		Cour 3
Samples Reference: hg38_DLE1_ HCC2218 - Variant Annotatio Insertion Deletion	Okb_Olabels		Cour 3 36
Samples Reference: hg38_DLE1_ HCC2218 - Variant Annotatio Insertion Deletion Inversion	Okb_Olabels		Cour 3 36 9
Samples Reference: hg38_DLE1_ HCC2218 - Variant Annotation Deletion Inversion Duplication	Okb_Olabels		Cour 3 36 9 9
Samples Reference: hg38_DLE1_ HCC2218 - Variant Annotatio Deletion Inversion Duplication Intra-Translocation	Okb_Olabels		Cour 3 36 9 9 64
Samples Reference: hg38_DLE1_ HCC2218 - Variant Annotation Deletion Inversion Duplication Intra-Translocation Oth/C are Secreted	Okb_Olabels		Cour 3 36 9 9 64 17

Figure 22. Filtering settings and BED (masking) file.

2. In the BED Configuration window, select the **hg38 Known Canonical** file from the dropdown as shown in **Figure 23**.

BED Configuration	
PED Selection	
BED Selection:	
	~
D474 with controls	A
D4Z4 What controls	
Rare Variant hg19 DLE-1 SV Mask	
Rare Variant hg38 DLE-1 SV Mask	
hg19 ACMG genes 🦲	
hg19 BspQI SV Mask 🦲	
hg19 BssSI SV Mask 🦰	
hg19 DLE-1 SV Mask 🦰	
hg19 Known Canonical 🦰	
hg19 cnv masks 🦰	
hg38 ACMG genes 🦰	
hg38 BspQI SV Mask	
hg38 BssSI SV Mask	
hg38 DLE-I SV Mask	
hg38 cnv masks	
genomicSuperDups_hg19	
genomicSuperDups_hg38_Blue10kbp.bed	
genomicSuperDups_hg38_chr.bed	-

Figure 23. Selecting hg38 Known Canonical

3. Choose the button for **Show structural variants that overlap BED Regions** as highlighted (**Figure 24**) and then press **Ok**.

D Selection:		
hg38 Known Canonical Mapped 角	~	
Structural Variant Overlap Precision (Kbp):		
12		
Structural Variant Filter Setting:		
Show all structural variants = Image Im		
$\bigcirc$ Show structural variants that overlap BED Regions = $\Im$		
O Hide structural variants that overlap BED Regions = S		
CNV Segment Overlap Precision (Kbp):		
500		
CNV Segment Filter Setting:		
Show all CNV segments = (9)		
○ Show CNV segments that overlap BED Regions =		
Hide CNV segments that overlap BED Regions = S		

Figure 24. Selecting a BED file. Preloaded bed files can be used filter out structural variation in non-specific regions.

Since the visualization is interactive, the Circos plot will now show an extra annotation track. **Figure 25** now shows the known canonical genes on the second track right next to the cytobands. In this track, each vertical line represents a gene.



Figure 25. Rare variants overlapping known genes - Circos plot with BED track.

The dataset has now been filtered according to Bionano's recommended filter settings and a BED masking file has been applied to the dataset so that only SVs that overlap with known genes will be visible. Confirm the SV counts as shown in **Figure 25**. These are represented by different colors on the SV track of the Circos plot. **NOTE** Upon implementing the filters, the lowest number of rare variants is identified.

Now the user can confirm whether the SVs can be found on chromosomes 17 in the regions specified by the DepMap database.

4. Zoom in on the Circos plot region for chromosome 17 using the mouse wheel or track pad (Figure 26).



Figure 26. Circos plot with Variant Annotation filters applied. Zoom in and mouse over SVs on chromosome 17.

- 5. Hover the mouse cursor over the SV track for chromosome 17, which also has a large copy number change associated with that region.
- 6. Click over the deletion on chr 17. Confirm that the position is roughly 41,109,560-67,483,840 according to the Bionano data.
- 7. Click on the intrachromosomal fusion to open the Genome Map view, seen in Figure 28. Alternatively, choose Genome Browser and Chr 17 for Anchor from the dropdown as shown in Figure 27. This will launch the Genome browser view. This genome map view shows an intrachromosomal fusion event on chromosome 17.



Figure 27. Dropdown from the top panel allows navigation to Genome Browser view, Anchor represents chromosome number.



Figure 28. Genome Map View visualization

This zoomed out genome map view shows chromosome 17. The top panel of the browser is consistent with the Circos plot view. The cytoband is the top track. The **Anchor** and **Range** show the precise coordinates on top reference that is being viewed. The SV can be viewed in more detail in the central region of the map.

In **Figure 28**, the small red box in this track indicates the location of the intrachromosomal fusion. The second track includes the SVs detected (red box) which in this example is an intrachromosomal fusion. Known genes from hg38 are added in the view as a reference. The vertical lines on the reference map and assembled genome map are the recognition sequence of the labeling enzyme and the connections between maps (grey lines) are alignments. If the labels on the assembled genome maps match with the reference map, then no structural variations are reported in a region of interest. In this example, by comparing the enzyme marks on the Bionano map (blue) and reference genome locations (green), one can see there are alignment marks from the assembled genome that matches marks of chromosome 17. The abnormal alignment marks (pink lines) show the region of the map that skips contiguous alignment to the reference map suggesting a translocation event. Using the mouse, draw a box on the cytoband track around 41 Mb to 44 M region to zoom in. In the tabular section of the view, click on the header of each column to sort it. Use the mouse to scroll zoom in or zoom out. Please refer to *OGM File Format Specification Sheet* (CG-00045). for more details about each column.

- 8. Use the scroll bar at the base of the table (**Figure 29**) to navigate to the **Overlap Genes** column. This lists the overlapping genes and nearest non-overlap genes. The first few rows in the table lists the SVs detected in this region.
- 9. Click on the header column to sort the overlapping gene names. This will list all the SV events for a fused gene called ZNF652 INTS2.

Geno	ne Bro HCC2	owser 🗸 Ar 2218 - Variant Ar	chor 17 🗸 R notation Pipelir	ange [map17: 48,589]	127-50,179,028	Molecules Pack	<b>∨ !</b> , c	onfidence	✓   Find Ma	ap	<b>⊳</b> •@ 2	T 🖸 👁 🕸	♠ ?		
	0	I 3.2 M 6	1 I 4M 9.6M	I I 12.8 M 16 M 19	1 I I 2 M 22.4 M 25.6	M 28.8 M 32 M	I I 35.2 M 38.4	I M 41.6 M 44	I I 8 M 48 M	I I 51.2M 54.4M	I 57.6 M 60	I I 1.8 M 64 M	I I I 67.2 M 70.4 M 73.6 M	I I 76.8 M 80 M	83:
CN	0														
SV Ani	otatio	n Match Co	py Number B	BED Selected SV											
		Sample	Туре	Algorithm	RefcontigID1	RefcontigID2	RefStartPos	RefEndPos	Size	Zygosity	Score	Smapld	Present in % o	f control samples	
•															<b>^</b>
		HCC2218R_	deletion	assembly_comparison	17	17	35153931	39467783	3993854.1		0.99	4657		0.6	
Đ		HCC2218R	deletion	assembly_comparison	17	17	48631646	49036970	402156		7	4853		0	
Ð	0	HCC2218R_	deletion	assembly_comparison	17	17	48631646	49036970	402188		1	4831		0	
Đ	0	HCC2218R_	deletion	assembly_comparison	17	17	48706136	49946060	1224900.8		0.99	4837		0	
•		HCC2218R_	deletion	assembly_comparison	17	17	49121624	49437826	316202		1	4868		0	
	0	HCC2218R_	deletion	assembly_comparison	17	17	50758219	51547942.5	746882.8		0.99	4929		0	÷

Figure 29. Table at the bottom of the Genome Map lists the SVs for genes in the region.

10. Click on the first row to automatically navigate into this region with the SV overlapping the gene ZNF652 or INTS2. The table at the lower half lists all the SVs, CNVs and their associated annotations. The highlighted row lists "translocation\_intrachromosomal" under the **Type** column. The next few columns provide information about the algorithm used to call the SV, contig information and breakpoint coordinate information for the reference map, among others.

4						<u> </u>			, n							~			80 M
			н												H H	, <u>1</u>	·		
38 Know		10					1		6		in tra	ini i							1
	2 <sup>214</sup>	an	M	.02 <sup>10</sup>	ng AM		,W	W	- AN		.15 <sup>10</sup> .19 <sup>10</sup>		2M chaM	1514	Man Man	1-11-11-11-11-11-11-11-11-11-11-11-11-1	1M _ 15M	-n.e.M	
17	1	· · · · · · · · · · · · · · · · · · ·	<sup>1</sup>	ľ	<b>1</b>			<b>U</b> ÍT.	Ň	1	i na lina	<b>N CANAN</b>	1				<b>T</b>	<b>F</b>	,
			-14	-					17			1							
2218 -				3	-														
Annotat	en Match Copy I	Number BED Selected SV	_		_			11 - 17	K I	2	XX	8785 - 68 I.L		113	4000		<u>~_и</u>		-
	Sample	Type AL.	Refco. Re	afcon. R	ofStartPos R	RefEndPos	Size Zyg	. Score	Sma	P. 1	- Fail assemb.	Overlap Genes	Noares_ Noa	rest No	Putative Gene Fusion Self	Molecule _ Contro	Molec. Numb	er OverL. For	and in
	HOC2288R.	translocation_intrachr ass	17	17	49323917	61990098	-1	0.19489499	4882	0	0 pass	ZNF652/NTS2	Y_RNA	15917	2NF652-INT52	135	0	2	
		1100 N. 17 AVA 10	17	17	49.523917	63290098	-1	0.47824752_	-4/227	0	0 pass	2NF652/NT52	Y_RNA	15917	2145652-81752	100	0	-1	
	HCC278]R_	translocation_intrachr ass																	
	HCC2288_R.	translocation_intrachr ass	17	17	49323917	61990098	-1	0.929277702_	4884	0	0 toi	ZNF652JNTS2	Y_RN64	15917	ZNF652-INTS2	143	0	-7	
	HCC228R. HCC228R. HCC228R.	translocation_intrachr ass. translocation_intrachr ass.	17	17	49323917	61890/298 75725545	-1	0.929271702	4984	0 0	0 toi 0 toi	ZNF652JNT52 AC09880.5/TGB4	Y_RNA	15917 13293	ZNF652-WTS2 AC091180.5-ITGB4	143	0	9	
	HCC2288_R. HCC2288_R. HCC2288_R. HCC2288_R.	translocation_intrachr ass. translocation_intrachr ass. translocation_intrachr ass. translocation_intrachr ass.	17 17 17	17 17 17	49323917 49444568 49444568	68990098 75725545 75730132	-1 -1 -1	0.929277702. 0.97583083. 0.59642820.	4884 4902 4885	0 0 0	0 £20 0 £20 0 pass	ZNF652JNT52 ACO3NB0.5JTG84 ACO3NB0.5JTG84	Y_RNA LINCO2. LINCO2.	15917 13293 13293	2NF652-INT52 ACD91180.5-17G84 ACD91180.5-17G84	143 156 117	0 0 1	9 -4 -4	
	HCC228_R. HCC228_R. HCC228_R. HCC228_R.	translocation_intrache ass. translocation_intrache ass. translocation_intrache ass. translocation_intrache ass.	17 17 17 17	17 17 17	49323917 49444568 49444568 50595912	6890098 75725545 7573002 70454031	-1 -1 -1 -1	0.929277702. 0.97583083. 0.59642820. 0.48802552.	4984 4902 4985 4924	0 0 0	0 fail 0 fail 0 pass 0 pass	ZNF652INTS2 AC09080.5JTGB4 AC09080.5JTGB4 CACNAIG	Y_RNA LINCO2. LINCO2. ACOD4.	15917 13293 13293 13293 31120	2NF652-INT52 AC09180.5-ITG84 AC09180.5-ITG84	143 156 117 61	0 0 1 0	-1 -1 -1 -1	
	HCC228R_ HCC228R_ HCC228R_ HCC228R_ HCC228R_ HCC228R_	translocation_intractir auc. translocation_intractir ass. translocation_intractir ass. translocation_intractir ass. translocation_intractir ass. translocation_intractir ass.	17 17 17 17 17	17 17 17 17 17	49323917 49444568 49444568 50595902 50595902	68990098 75725545 75730132 70454031 70454031	-1 -1 -1 -1 -1	0.929277702. 0.975853083. 0.59642820. 0.48802562. 0.49578954.	4984 4902 4885 4924 4917	0 0 0 0	0 toi 0 toi 0 pass 0 pass 0 pass	ZNF652/NT52 ACD91180.5/TGB4 ACD91180.5/TGB4 CACNAIG CACNAIG	Y_RNA LINCO2 LINCO2 ACO04	15917 13293 13293 3120 3120	2NF652-WT52 ACD9880.5-ITG84 ACD9880.5-ITG84	143 156 117 61 108	0 0 1 0	-1 -1 -1 -1 -1 -1	
	HCC228R. HCC228R. HCC228R. HCC228R. HCC228R. HCC228R.	translocalizn_intrache auc. translocalizn_intrache auc. translocalizn_intrache auc. translocalizn_intrache auc. translocalizn_intrache auc.	17 17 17 17 17	17 17 17 17 17	49323917 49444568 49444568 50586912 50586912	68990098 75725545 75730132 70454033 70454033	-1 -1 -1 -1 -1 -1	0.929277702. 0.975832083. 0.59642820. 0.49802552. 0.49878954.	4884 4902 4885 4924 4917	0 0 0 0	0 toi 0 toi 0 pass 0 pass 0 pass	2NF662N152 AC091180.5JTG84 AC091180.5JTG84 CACNAIG CACNAIG	Y_RNA LINC02_ LINC02_ AC004_ AC004_	15917 13293 13293 31020 31020	2NF652-WT52 AC09880.5-IT684 AC09880.5-IT684	143 156 117 61 108	0 1 0	-1 -1 -1 -1 -1	
	HCC228_R. HCC228_R. HCC228_R. HCC228_R. HCC228_R. HCC228_R.	translocalizn_intrache auc. translocalizn_intrache auc. translocalizn_intrache auc. translocalizn_intrache auc. translocalizn_intrache auc.	17 17 17 17 17	17 17 17 17	49323917 49444568 49444568 50586812 50586812	68900398 25725545 25730332 20454033 20454033	-1 -1 -1 -1 -1	0.929277702. 0.97583083. 0.59642820. 0.49802562. 0.49578954.	4884 4902 4885 4924 4927	0 0 0	0 fai 0 fai 0 pass 0 pass 0 pass	ZNF662JNIS2 AC09180.5JTGB4 AC09180.5JTGB4 CACNAIG CACNAIG	У_ВНА LINCO2. LINCO2. ACOO4. ACOO4.	15917 13293 13293 31120 31120	2NF652-N1152 ACCOMBO.5-ITGBA ACCOMBO.5-ITGBA	143 156 117 61 208	0 1 0 0	9 -0 -0 -0 -0 -0 -0 -0 -0 -0 -0 -0 -0 -0	

Figure 30. ZNF652 gene showing several structural variations in this region.

11. Zoom in to the region map17: 49,100,615-49,570,725 to view the ZNF652 gene. This region shows several structural variants including translocations events in the region as well as copy number change. When regions located on the same chromosome fuse, they are called an intrachromosomal

rearrangement. These intrachromosomal rearrangements are well known characteristic of malignancies. For more details on translocation calls detected using Bionano optical mapping techniques, refer to *Bionano Solve Theory of Operations: Variant Annotation Pipeline* (CG-30190). Additionally, SV track and the table on the lower half also suggest two large deletion events spanning the region.

12. Scroll to the right to find the column **Self Molecule Count**. For the list of translocation events the selfmolecule count is listed in the range of 135 to 143, suggesting that the breakpoints for this translocation event was supported by molecules that line up to the breakpoint coordinate. While hovering over the table, users can press the left and right arrow keys to navigate to the additional cells or use the trackpad to drag the table left or right (**Figure 31**).

Br	owser 👻	Anchor 17 🗸	Range map17	7: 49,100	0.615-49,5	70,725	Molecules	Pack	<ul> <li>▶ 17</li> </ul>	Confid	lence 🛛 🗸	Find	і Мар				. 00	<b>n</b> # ?	HCC2218 -	Variant Annotation	Pipeline		
0	3.2 M	6.4 M	9.6 M 12.8	м	16 M	19.2 M	22.4 M	25.6 M	28.8 M	32 M	35.2 M	38.4 M	4	1.6 M 44.	8 M 48	51.2	M 54.4	M 57.6 M	60.8 M	64 M 67.2 M	70.4 M 7	3.6 M 76.8 M	80 1
																			1				
	-																						
w		B40	BALNT2			AB	13	,	0		ı	z	NF65:	2	н	н	н 1 (- Р	HB AC091180	5	н		H AC0064	AC01 87.1
			49.15M		1 1 1	1.2 M		49.25 M	<u>,</u>		49.3 <sup>M</sup>			64	35 M		49.4 14		49.45	A	49.5M		A9.5
17																							
-[]													_										
atio	n Match	Copy Numbe	r BED Selec	ted SV									_										
	Samp	ile L tr	Type ranslocation_intrach	e AL	Refco	Refcon	RefStartPos 49323917	RefEndPos 61890096	Size	Zyg	Score 0.19489499	Sma 4882	P_	P. Fail assen	nb C	Overlap Genes ZNF652;INTS2	Neares_ Y_RNA	Nearest No 15917	Putative Gene Fu	ion Self Molecule 152 135	Control Molec	Number OverL.	Found in
	HCC2218R		ranslocation_intrach	ass	17	17	49323917	61890096	3 -1		0.478247152	4891	0	0 1	Nass Z	NF652:INTS2	Y_RNA	15917	ZNF652-IN	TS2 108	c	-1	
	HCC2218R	L. U	anslocation_intrach	ass	17	17	49323917	61890096	-1		0.929271702	4884	0	0	fail 2	ZNF652;INTS2	Y_RNA	15917	ZNF652-IN	TS2 143	C	-1	
	HCC2218R	L tr	ranslocation_intrach	ass	17	17	49444568	75725545	-i		0.97583083.	4902	0	0	fail ACO	91180.5;1TGB4	LINCO2	13293	AC091180.5-17	584 156	c	-1	
	HCC2218R	L tr	anslocation_intrach	ar ass	17	17	49444568	75730132	2 -1		0.59642820	4885	0	0 1	Nass ACO.	91180.5;1TGB4	LINC02	13293	AC091180.5-17	584 117		-1	
	HCC2218R	t. tr	anslocation_intrach	ar ass	17	17	50595912	7045403	-1		0.48102552	4924	0	0 1	Nass	CACNAIG	AC004	31120		- 61	G	-1	
	HCC2218R	t. tr	anslocation_intrach	ar ass	17	17	50595912	7045403	-1		0.49578954	4917	0	0 ;	Nass	CACNAIG	AC004	31120		- 108	G	-1	

Figure 31. Genome view with the copy number track at the top.

- 13. Click on the "+" sign to the left of the row that has the intrachromosomal translocation SV of interest. This will add it to the "Selected SV's list, and the SV report to be printed out later.
- 14. Right-click on the assembled genome map (light blue) to list options to show the molecule pileup (**Show Molecules**) used to find the variants (**Figure 32**).



Figure 32. Right click options to add molecules or ruler to the genome map view

Thus, optical mapping confirms the translocation call. In addition, optical mapping also reports a large deletion (color coded orange) and copy number increase (in the copy number track) as well as other translocations in this region. To view the any other SV call, click on the variant and it will launch the optical map supporting the call and the molecules supporting. Similar filtering criteria can also be applied to a user's own dataset. The Circos plot and genome map navigation will be applicable to all datasets analyzed in Access. In this figure, the CNV track is shown with the SV track, gene information track and the assembled Bionano map with the reference map. To find more information about copy number calls, click on the **Copy Number** tab. For more information on how CNVs are calculated see *Bionano Solve Theory of Operations: Structural Variant Calling* (CG-30110).

15. To view the whole genome copy number view, use the dropdown on the top panel to navigate to the WGCNV view as shown in the **Figure 33**. Optical mapping also shows an uploidy and dramatic copy number changes in the genome.



**Figure 33**. Whole Genome Copy Number View. The individual label copy number data points are plotted in blue or red. The segmented copy number line is green, and the copy number variant segments identified are shaded according to the variant type. Aneuploidy information is also marked on the bottom. Zoom in/out to see the scale of chromosome.

#### **Report Generation, Exporting Data, and Creating Images**

#### **Report Generation and Data Export**

Data export can be useful for record keeping or for tertiary analysis. To learn more about the different data files, read the *OGM File Format Specification Sheet* (CG-00045). Results from the rare variant annotation pipeline are output to a text file with annotation for each SV detected for each sample. Exporting data from Access allows the user to share their findings with others or for record keeping. Access users can create a pdf report with analysis details, as well as scalable vector graphics (SVG) or screen captures for the genome map views or Circos plots. These figures are also a part of the pdf generated from Access. The image in the report will reflect what is projected on the screen when the SV is selected in the data grid by default.

1. To generate a PDF SV annotation report containing the SVs that have been added, make sure the Circos plot view is selected and then select the **Generate Report** icon (pdf icon and the pdf dialog box will pop up (**Figure 34**).

### 🖪 Q C / T 🔿 👁 🗱 🐂 🏦 ?

Figure 34. Generating a pdf report. The icon in the red box launches the PDF generating dialog box.

2. To download all the filtered variants, check the box according to the file format desired (or check all of them), ensure **Filtered Variants** is selected in the dropdown box and then click **Ok** (**Figure 35**). This command will save a zipped file in the default folder (downloads or desktop).

Additional Files to Downly	bed
Additional Files to Downie	Jau
Generate and download t	he following files as
Filtered Variants	~
Copy Number Varian	ts (CSV Format)
Structural Variants (S	MAP Format)
Annotated Structural	Variants (SMAP Format)
Aneuploidy File	
Structural Variants (V	'CF Format')
VCF Input Values	
Reference Accession:	hg38_DLE1_0kb_0labels
Experiment Id:	
Note: VCF export will autom	natically omit all SVs that were masked by
the SV calling pipeline as a regardless of whether those	result of unresolved CMPR regions,
SMAP does not automatical	ly omit these.

Figure 35. Generating a pdf report.

This file is called "HCC2218\_rareVariant\_import\_annotationDual\_879\_6\_30\_2020\_2\_33\_18.zip." This folder will contain a PDF report, a VCF file and two CSV files of the selected variant (**Figure 36**).

**TIP**: To download the variants detected in each sample, choose either "HCC2218" or "HCC2218BL," listed as **Rare Variant Analysis** under the column **Type**. For each of these files, click on **Download Rare Variant Analysis**.

**NOTE:** To generate a close-up of the Circos plot, zoom in to the desired location before selecting **Generate Report**.

Name	Туре	Compressed size	Password protected S	Size
HCC2218_rareVariant_import_annotationDual_879_6_30_2020_2	PDF File	953 KB	No	953 KB
HCC2218_rareVariant_import_annotationDual_879_6_30_2020_2	SMAP File	65 KB	No	65 KB
HCC2218_rareVariant_import_annotationDual_879_CNV_6_30_20	Microsoft Excel Comma Separated Values File	1 KB	No	1 KB
HCC2218_rareVariant_import_annotationDual_879_CSV_SV_6_30	Microsoft Excel Comma Separated Values File	60 KB	No	60 KB

Figure 36. Rare variant downloaded folder showing the different files downloaded as a part of the report

3. To view the files, Unzip the folder. Double-click on PDF to read the file.

This report includes a screen shot of the Circos plot complete with the legend, a screen shot of each structural variant, details on each structural variant, and general information about the dataset and any comments that were included during the analysis. To obtain more information about the SMAP file, refer to *OGM File Format Specification Sheet* (CG-00045). **NOTE:** Another option to export files or the whole project for the sample analyzed is to download the SV annotation files. To do this, first click on **Home** or the **Back** icon to go back to the project data page (**Figure 37**).





Figure 37. Relaunching the project. The back icon in the red box relaunched the project browser

#### **Creating Images**

Bionano Access is equipped with a tool to generate publication quality images in Jpeg file format in Access version 1.6. Preserving graphical quality is essential in publishing a reported finding.

- 4. To capture an image of a structural variant, navigate to the SV of interest.
- 5. Click on the camera icon displayed in the genome browser (**Figure 38**). A preview icon is also available to allow the user to view the image that was captured.



Figure 38. Generating images of regions of interest. Inset shows camera icon that launches the SVG export wizard.

6. This launches an interactive "AccessViewer.svg" file from the local download. This image can be further edited where it can scale to any dimension without losing quality.

#### **Resources and Information**

For more information about software and data analysis, click on the **Help** button (question mark icon) on the **Home** page. This opens the **Support** page.

bionan	0	Access			
	E Software	🔝 Tutorial Videos			
	Tool Version Bionano Access 1.8.1.1	Our video tutoriais have been centralized on our corporate support site so that we can continuously improve and evolve our help content and reduce our installation file size. For a fi index of tutorials available click the link below.	i at		
	Software License Agreement	Bionano Access Video Tutorials			
	Compute On Deman	ed 🔒 Security			
	Tool Version Bionano Solve Solve3.8.1.1	We recommend that you apply security patches to your linus systems with each software update according to our	e.		
	Terms and Conditions	frequent security patching status and to subscribe to security updates.	đ		
		Security Home			
		For technical assistance, contact Bionano Technical Suppor You can retrieve documentation on Bionano products, Safet Data Sheets, certificates of analysis, frequently asked questions, and other related documents from the Support page or by request.	L. X		

Click on the video tutorials for short videos to learn about specific analysis steps: http://www.bnxinstall.com/Videos/TutorialIndex.html

- For Structural Variants, Copy Number Variations and Rare Variant Analysis, refer to: Bionano Solve Theory of Operation: Structural Variant Calling
- For Variant Annotation Pipeline, refer to:

Bionano Solve Theory of Operation: Variant Annotation Pipeline

• For hybrid scaffolding pipeline details, refer to:

Bionano Solve Theory of Operation: Hybrid scaffold

• When using command line, refer to:

Guidelines for Running Bionano Solve Pipeline on the Command Line

You can also find the latest versions of these documents on Bionano's Data Analysis Documentation page (https://bionano.com/software-and-data-analysis-support-materials/).

#### **Technical Assistance**

For technical assistance, contact Bionano Technical Support.

You can retrieve documentation on Bionano products, SDS's, certificates of analysis, frequently asked questions, and other related documents from the Support website or by request through e-mail and telephone.

ТҮРЕ	CONTACT
Email	support@bionano.com
Phone	Hours of Operation: Monday through Friday, 9:00 a.m. to 5:00 p.m., PST US: +1 (858) 888-7663 Monday through Friday, 9:00 a.m. to 5:00 p.m., CET UK: +44 115 654 8660 France: +33 5 37 10 00 77 Belgium: +32 10 39 71 00
Website	www.bionano.com/support
Address	Bionano, Inc. 9540 Towne Centre Drive, Suite 100 San Diego, CA 92121

#### **Legal Notice**

#### For Research Use Only. Not for use in diagnostic procedures.

This material is protected by United States Copyright Law and International Treaties. Unauthorized use of this material is prohibited. No part of the publication may be copied, reproduced, distributed, translated, reverseengineered or transmitted in any form or by any media, or by any means, whether now known or unknown, without the express prior permission in writing from Bionano Genomics, Inc. Copying, under the law, includes translating into another language or format. The technical data contained herein is intended for ultimate destinations permitted by U.S. law. Diversion contrary to U. S. law prohibited. This publication represents the latest information available at the time of release. Due to continuous efforts to improve the product, technical changes may occur that are not reflected in this document. Bionano Genomics, Inc. reserves the right to make changes in specifications and other information contained in this publication at any time and without prior notice. Please contact Bionano Genomics, Inc. Customer Support for the latest information.

BIONANO GENOMICS, INC. DISCLAIMS ALL WARRANTIES WITH RESPECT TO THIS DOCUMENT, EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO THOSE OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. TO THE FULLEST EXTENT ALLOWED BY LAW, IN NO EVENT SHALL BIONANO GENOMICS, INC. BE LIABLE, WHETHER IN CONTRACT, TORT, WARRANTY, OR UNDER ANY STATUTE OR ON ANY OTHER BASIS FOR SPECIAL, INCIDENTAL, INDIRECT, PUNITIVE, MULTIPLE OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH OR ARISING FROM THIS DOCUMENT, INCLUDING BUT NOT LIMITED TO THE USE THEREOF, WHETHER OR NOT FORESEEABLE AND WHETHER OR NOT BIONANO GENOMICS, INC. IS ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

#### Patents

Products of Bionano Genomics<sup>®</sup> may be covered by one or more U.S. or foreign patents.

#### Trademarks

The Bionano logo and names of Bionano products or services are registered trademarks or trademarks owned by Bionano Genomics, Inc. ("Bionano") in the United States and certain other countries.

Bionano<sup>™</sup>, Bionano Genomics<sup>®</sup>, Saphyr<sup>®</sup>, Saphyr Chip<sup>®</sup>, Bionano Access<sup>™</sup>, and Bionano EnFocus<sup>™</sup> are trademarks of Bionano Genomics, Inc. All other trademarks are the sole property of their respective owners.

No license to use any trademarks of Bionano is given or implied. Users are not permitted to use these trademarks without the prior written consent of Bionano. The use of these trademarks or any other materials, except as permitted herein, is expressly prohibited and may be in violation of federal or other applicable laws.

© Copyright 2024 Bionano Genomics, Inc. All rights reserved.