Redefining Cell Characterization and QC for Genomic Integrity and Off-Target Monitoring

DOCUMENT NUMBER:
LIT-00017

DOCUMENT REVISION:
A

Effective Date:
08/21/2023

bionano

Table of Contents

Revision History

REVISION	NOTES
A	Initial release of document.

Introduction

The cell and gene therapy market and the respective developer pipelines are rapidly advancing. Significant growth has been spurred by a substantial increase in recent publications, and the availability of high-quality tools and funding in this exciting segment. However, the success rate for phase 2 approvals is still very low[1].Developers face important technological, economic, and regulatory challenges to ensure safety and efficacy, unobstructed time-to-clinic, and cost control. The FDA has released draft guidance[2] laying out the assessment data required over the course of product development and manufacturing for submissions. Substantial emphasis is placed on comprehensive characterization, quality control of gene-editing events and monitoring genomic integrity of cell lines – especially those that undergo prolonged cell culture and multiple passages, such as induced pluripotent stem cells (iPSCs)[3].

Typically, a constellation of cytogenetic and molecular methods are used to detect genome-wide variation across the spectrum of resolution and classes. This is inevitably time-consuming, expensive, and complex as development and manufacturing teams struggle to determine the most effective approach for their pipelines.

The Bionano Saphyr® system is redefining cell genetic characterization and off-target effect evaluation. Optical genome mapping (OGM) is an unbiased, genome-wide, high-resolution approach that can be used as a single-assay alternative to multiple cytogenetic techniques and importantly bridges the resolution gap between traditional cytogenetics and sequencing. The accelerated sample-to-answer workflow is straightforward to implement in-house as a superior one-stop alternative to karyotyping, fluorescent *in situ* hybridization (FISH) and chromosomal microarray (CMA) combined. Furthermore, the digital OGM readout does not reveal proprietary sequence information.

This white paper is intended to provide guidance to development and manufacturing teams in biopharma companies and CDMOs/CROs to enable efficient evaluation and implementation of Bionano Saphyr® and OGM workflows.

[1] Home | ClinicalTrials.gov

[2] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/human-gene-therapy-products-incorporating-human-genome-editing

[3] Jennifer L Dashnau 1, Qiong Xue 2, Monica Nelson 3, Eric Law 3, Lan Cao 2, Derek Hei 4, A risk-based approach for cell line development, manufacturing and characterization of genetically engineered, induced pluripotent stem cell-derived allogeneic cell therapies. Cytotherapy 2023 Jan;25(1):1-13.doi: 10.1016/j.jcyt.2022.08.001. Epub 2022 Sep 13.

Introduction to Optical Genome Mapping: Genome-Wide, Unbiased, High Resolution SV Detection

Detection of structural variants (SVs) is essential in genomics. Classical cytogenetic approaches such as karyotyping and FISH are the gold standard in SV detection. However, karyotyping has limited resolution, only detecting events 5Mbp or larger, and FISH is a targeted assay, not providing whole genome coverage. Even though chromosomal microarray (CMA) can detect copy number variation (CNV) genome-wide, it cannot detect balanced translocations or inversions, and has low sensitivity for low-level mosaicism. On the other end of the variant size spectrum, while short-read and long-read next generation sequencing (NGS) can detect nucleotide variants, they miss a significant portion of SVs[4] thus missing a significant amount of information critical to understanding disease etiology.

OGM can detect all classes of SVs in a single, genome-wide platform that bridges the gap in the variant size spectrum (see **Figure 1**). OGM can detect aneuploidy, deletions, duplications, inversions, translocations, and repeat expansions or contractions. Importantly, its workflow is designed to scale, and can generate high-coverage data to detect low allele fraction. Current tools, such as CMA, NGS, and karyotyping can detect some of these variants but have certain limitations, whereas OGM meets these needs with higher resolution of SVs than all other methods combined, as shown in **Table 1**. OGM provides the most comprehensive structural assessment of the genome, affording a high resolution, digital karyotype that can transform operations in a cytogenetic lab, while also finding or resolving even more actionable variation than traditional techniques.

[4]Chaisson, M.J.P., Sanders, A.D., Zhao, X. et al. Multi-platform discovery of haplotype-resolved structural variation in human genomes. Nat Commun **10**, 1784 (2019). https://doi.org/10.1038/s41467-018-08148-z

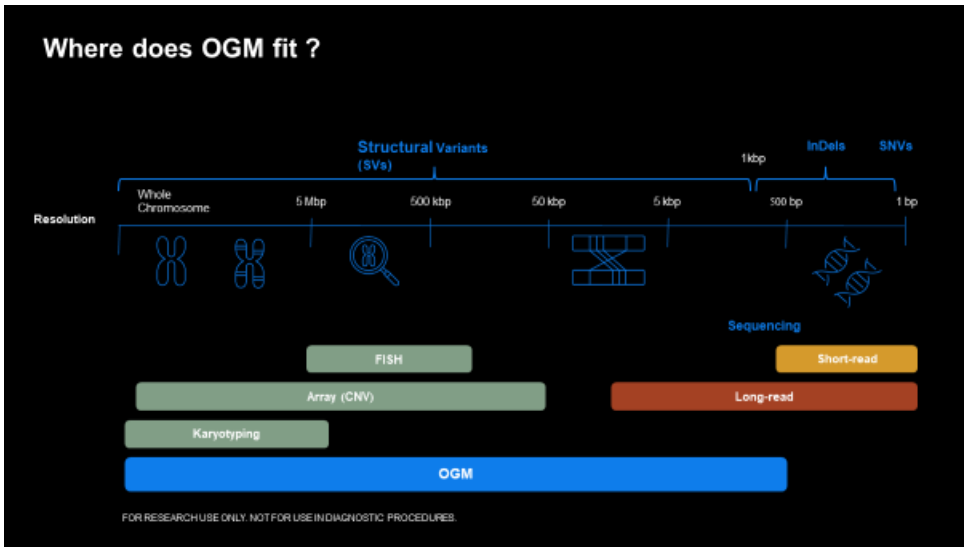

Figure 1. Size distribution of variants and technologies

Table 1. Comparison of OGM Capabilities across Variant Type versus Traditional Techniques

How does OGM measure up?

Higher structural variant coverage **than all other methods combined**

Variant Type	Karyotype	FISH	Microarray	NGS	OGM
Aneuploidy	✓	✓ Targeted	✓	✓	✓
Deletion	✓	✓ Targeted	✓	✓	✓
Duplication	✓ >5–10Mbp	✓ Targeted	✓	✓	✓
Translocation	✓ >5–10Mbp	✓ Targeted	✗	✓ Restricted Targets	✓
Inversion	✓ >5–10Mbp	✓ Targeted	✗	✗	✓
AOH	✗	✗	✓	✓	✓
Repeat Expansion	✗	✗	✗	✓ Limited to Small Repeats	✓ Limited to Large Repeats
Repeat Contraction	✗	✗	✗	✓ Limited to Small Repeats	✓
SNV	✗	✗	✗	✓	✗

bionano FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES. 18

A Tool for Screening Structural Variants Linked to Off target Events.

Quantitation of acquired SVs in induced pluripotent stem cells (iPSC) during culture is a good example of quality control (QC) analysis where SV gain and/or loss can be evaluated (see **Figure 2**), according to[5] Cell line QC is improved, and operations are streamlined with OGM.

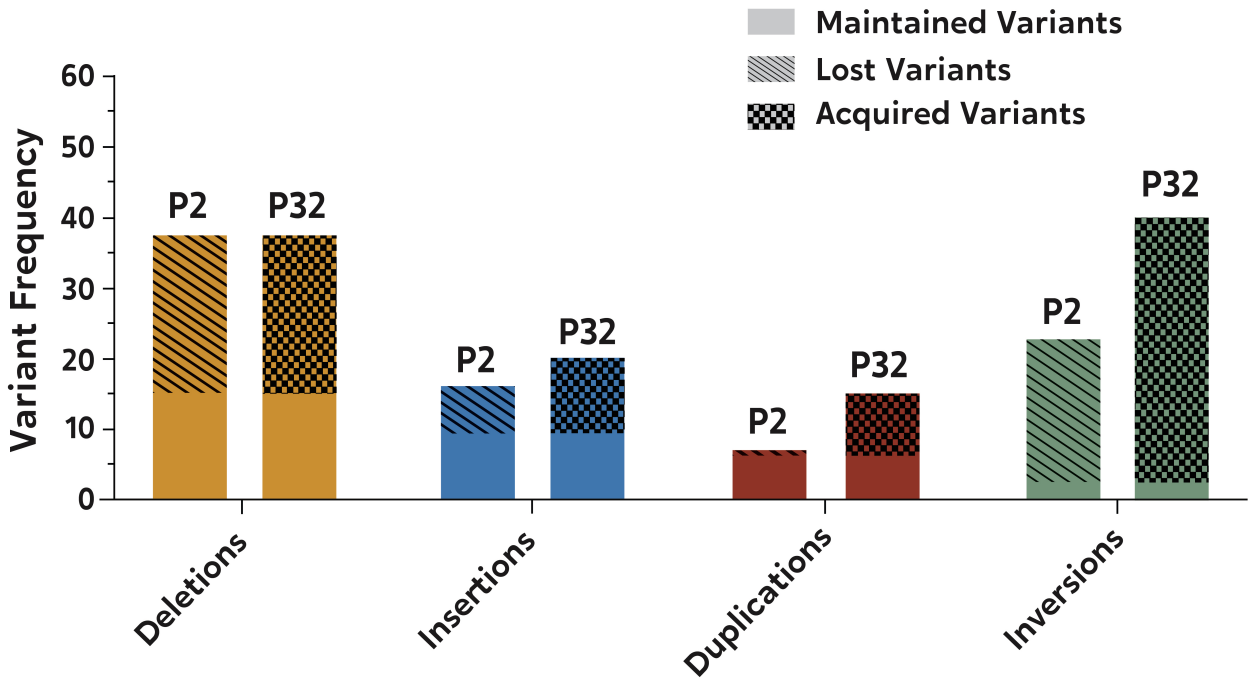

Figure 2. SV Gain/loss

[5]DuBose CO, Daum JR, Sansam CL, Gorbsky GJ. Dynamic Features of Chromosomal Instability during Culture of Induced Pluripotent Stem Cells. Genes (Basel). 2022 Jun 27;13(7):1157. doi: 10.3390/genes13071157. PMID: 35885940; PMCID: PMC9318709.

OGM MEETS QUALITY CONTROL ANALYSIS NEEDS IN A SINGLE PLATFORM

- High resolution for detection of structural and numerical variants.
- Can verify position of the transgene or insertion.
- Unbiased genome-wide detection of all classes of balanced and unbalanced SVs.
- Efficient, scalable tool.
- Intuitive, automated variant calling.
- Time to result is four-five days.

QC needs for cell bioprocessing include genome integrity assessment and gene editing validation. OGM meets both in a single platform.

Results

OGM is an emerging cytogenomic technology that offers an effective and genome-wide assessment of pathogenic variants in postnatal constitutional disorders[6] and in prenatal samples[7]. Homozygous and heterozygous germline variants can be detected comprehensively by running Bionano's *de novo* assembly on 400 Gbp of data (**Table 2**). OGM has also been used for hematological malignancies whole genome analysis[7]. To detect somatic variants down to about 5% variant allele fraction (VAF), one can collect 1.5 Tbp of data and run the rare variant pipeline (RVP).

To detect the even lower VAF events needed for cell characterization and quality control, a deeper coverage to 5.0 Tbp of data is required. We estimated the limit of detection at 5.0 Tbp using *in-silico* dilution of multiple cell lines with known SVs with GM24385 to achieve the desired VAF. Each line was analyzed independently with the rare variant pipeline, and each dilution was performed in triplicate (**Table 3**). This experimental, deep coverage, high sensitivity workflow is not fully validated, but preliminary results indicate that it can detect rare subclones at a VAF of 1-2 % (**Table 2**).

---

[6]Iqbal MA, Broeckel U, Levy B, Skinner S, Sahajpal NS, Rodriguez V, Stence A, Awayda K, Scharer G, Skinner C, Stevenson R, Bossler A, Nagy PL, Kolhe R. Multisite Assessment of Optical Genome Mapping for Analysis of Structural Variants in Constitutional Postnatal Cases. J Mol Diagn. 2023 Mar;25(3):175-188. doi: 10.1016/j.jmoldx.2022.12.005. PMID: 36828597.

[7]Levy B, Baughn LB, Akkari Y, Chartrand S, LaBarge B, Claxton D, Lennon PA, Cujar C, Kolhe R, Kroeger K, Pitel B, Sahajpal N, Sathanoori M, Vlad G, Zhang L, Fang M, Kanagal-Shamanna R, Broach JR. Optical genome mapping in acute myeloid leukemia: a multicenter evaluation. Blood Adv. 2023 Apr 11;7(7):1297-1307. doi: 10.1182/bloodadvances.2022007583. PMID: 36417763; PMCID: PMC10119592.

Table 2. Summary of performance at three different coverage levels for different OGM applications. **NOTE**: The data on the germ line workflow is based on the *de novo* assembly pipeline, while the somatic and ultra-rare variants workflows are based on the RVP.

OGM Workflows	**Germ line 50% VAF**	**Somatic 5% VAF**	**Ultra-rare variants – suggested thresholds based on preliminary data**
Target genome coverage	400 Gbp **80x effective coverage**	1.5 Tbp **300x effective coverage**	5.0 Tbp **1200x effective coverage**
# of flowcells	1	1	2-3
Target VAF %	50% VAF	5% VAF	1-2% VAF
Deletion	> 700 bp	> 7 kbp	> 50 kbp
Insertion	> 500 bp	> 5 kbp	> 20 kbp
Duplication	> 50 kbp	> 140 kbp	> ~ 100 kbp
Inversion	> 50 kbp	> 70 kbp	> ~ 100 kbp
Translocation	> 70 kbp	> 70 kbp	> ~ 100 kbp
Aneusomy	5%-15% VAF (depends on chromosome size)		
Notes	Sample with stable genome, e.g., expanded clone, verify major clone is correct, detects smallest variant (500 bp)	Used extensively for oncology, robust for all SV classes down to 5% VAF	Highest sensitivity for minor clone variants, down to 1% VAF

Table 3. SV calling performance by SV type with 5.0 Tbp mixtures at different proportions.

	>10% VAF	5% VAF	2.5% VAF	1% VAF	0.5% VAF	0% VAF
Translocations (n=14)	100 %	100 %	100 %	86 %	35 %	0 %
Large Duplication (2 Mbp, n=1)	100%	100%	100%	0%	0%	0%
Large Deletion (1 Mbp, n=1)	100 %	100 %	100 %	100 %	ND	0 %
Insertions (5-20kbp, n=10)	100 %	100 %	ND	70 %	ND	0 %
Deletions (>20kbp, n=6)	100 %	83 %	ND	33 %	ND	0 %
Sum	**100%**	**97%**	**100%**	**69%**	**26%**	**0 %**

Methods

Deep Coverage Data Collection

Figure 3 describes the workflow required to collect 5.0 Tbp of data needed to identify SVs as low as 1-2% VAF. Briefly, ultra-high molecular weight genomic DNA is extracted, labeled using twice the amount of standard Direct Label and Stain reagent, linearized, and uniformly stretched in high density nanochannel arrays on two to three flowcells. The samples are loaded on the Saphyr® chip, and the Saphyr® instrument runs the samples sequentially until 5.0 Tbp of data is obtained (**Figure 3)**. Down sampling occurs based on the scan count if the scan fails to meet QC. The run index can be used to exclude later scans which do not pass QC during long runs. Additionally, less merging is favored when generating the dataset; if 5.0 Tbp of data is realized with two flow cells, the third flow cell will not need to be used.

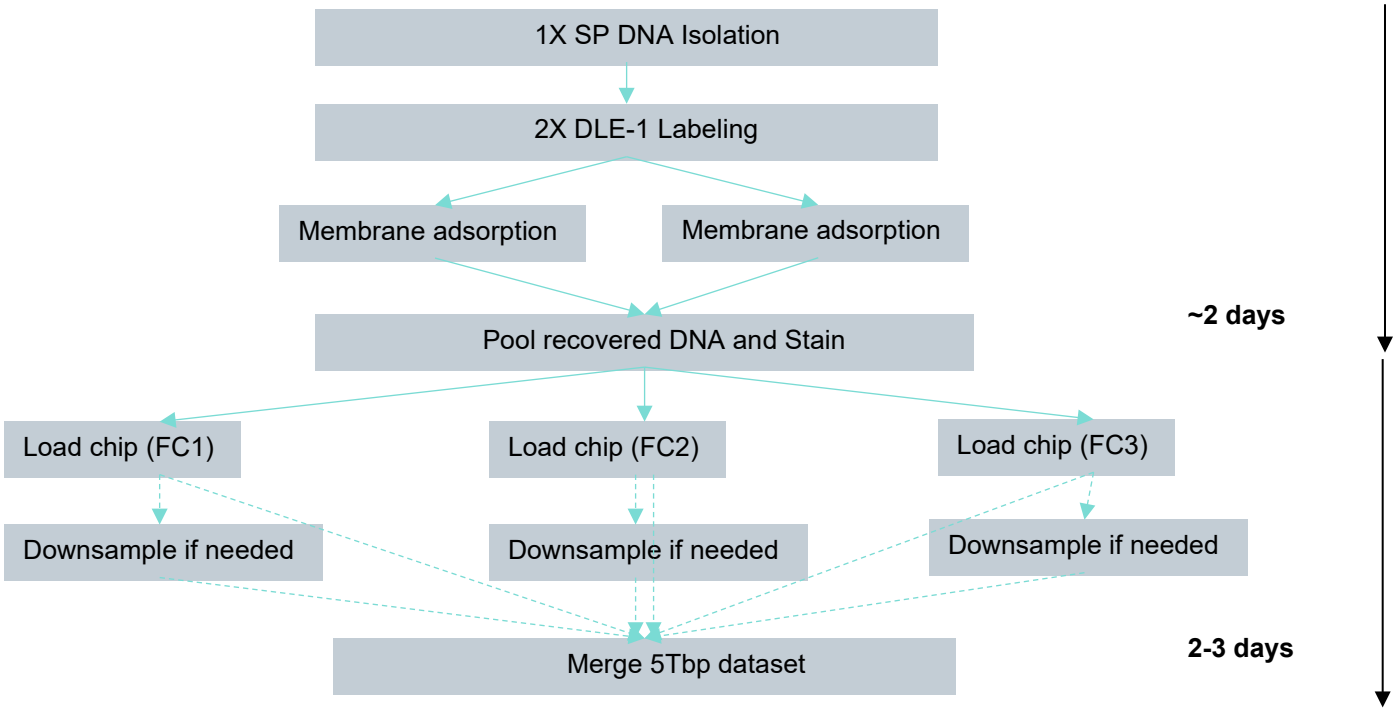

Figure 2. Deep coverage sample prep and data collection workflow

Experiment design and computational analysis pipelines

Figure 4 shows the experimental design of using OGM for quality assessment for cell bioprocessing. Both a test (treated) and a control (untreated) sample are required to assess the off and on target effects. A total of 5.0 Tbp of data is collected for each, and the SV in each sample is detected by rare variant pipeline. The dual variant annotation pipeline then compares the treated sample's SV profile against the control to identify treatment-specific variants, and it also utilizes a population control database that encompasses data from more than 200 ethnically diverse individuals to remove polymorphic variants. This comparison effectively reduces the number of structural variants to those that are relevant to both off and on-target effects.

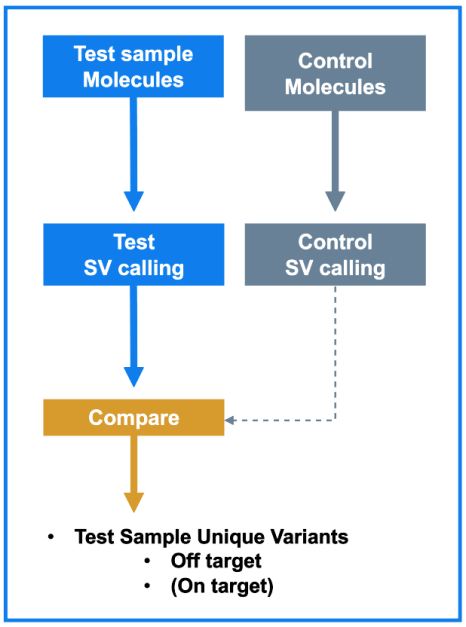

Figure 3. Computational analysis workflow. First, the Bionano Solve 3.7.2 rare variant pipeline was deployed on the "Test/Case" and "Control" samples to detect SVs throughout the genome. During the rare variant pipeline, molecules were aligned with the GRCh38 reference genome, and clusters of at least three molecules indicating SVs were assembled locally. These assemblies were used to confirm SV calls by comparing them to the reference genome. The dual variant annotation pipeline (brown box) using both the "Test" and "Control" sample was run to identify unique calls absent in the control. The pipeline also identified rare SVs by comparing them with SVs in the OGM population database. On and off target SVs unique to the test sample were then manually assessed using Bionano Access™ v1.7.2.

Candidate SVs filtering

Within Bionano Access, all SVs can be displayed in an interactive circos plot, presenting SV calls and CNV calls in a unified view. **Table 4** illustrates the recommended filtering criteria to evaluate test sample calls before and after comparison with the control sample. The **Initial Calls** column shows how to identify large, high confidence and non-polymorphic SVs, while the **Treatment Unique** column shows how to further select for SVs that may be treatment-induced.

Table 4. Recommended filtering criteria to identify, in the test sample, large and non-polymorphic SVs (Initial Calls), and treatment-induced on- and off-target SVs (Treatment Unique).

	Initial Calls	Treatment Unique
SV sizes (Insertions and Deletions)	>30 kbp	>30 kbp
SV sizes (Duplications and Inversions)	>100 kbp	>100 kbp
Confidence score	Default	Default
SV and CNV masks	On	On
SVs in self-molecules	Present	Present
Control database frequency	0%	0%
SVs in control SV calls	All SVs	Not in control
SVs in control molecules	All SVs	Not in control

Manual curation of variants

To rule out any artifacts that may arise, it is critical that the remaining post-filtered SV maps from the variant annotation pipeline (**Figure 5**) are evaluated manually within Bionano Access. There are two types of artifacts. The first type is an artificial SV call due to misalignments and the second is an SV being incorrectly classified as treatment specific. **Figure 6** illustrates the first type of artifacts and users should examine the alignments with other maps in proximity. Such problem can occur occasionally in ultra-deep coverage data sets where a few short, misaligned molecules can lead to false calls.

The second type of artifact requires a manual examination of molecule pileup to discern whether an SV exists in low frequency in the pre-treatment control. Molecule pileups can be displayed for either the case or control sample by right clicking on the treatment unique SV map and selecting the group of interest within the “Show Molecules for” dropdown selection. When evaluating the molecule pileups from the treatment unique SV maps, change the Molecules ordering selection from **Pack** to **Start**. To propel all the selected molecules that span the labels of interest to the top of the pileup, hover the cursor over the selected label and press the **c** button. Doing so will shift the selected molecules upwards within the pileup. One then would look for supporting molecules spanning both sides of the SV breakpoint. If there are at least one or two control molecules that can align by five labels on both sides of an SV junction, then the variant is considered to exist in the control sample, thus not specific in the test sample; if not, the variant is deemed to be truly treatment-specific and seen only in the test

sample. Some typical SV examples are shown in **Figures 5** through **9**, where molecule coverage plots of SVs are seen upon manual curation**.**

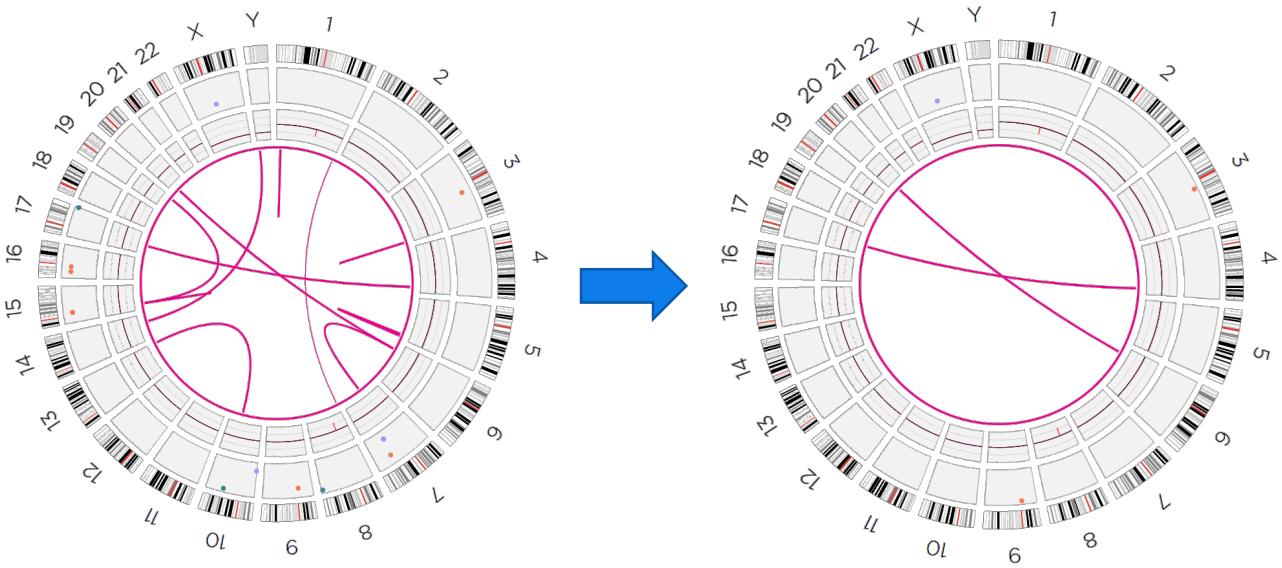

Figure 4. Deep coverage circos plots of SVs seen in the treated sample before (left) and after (right) auto-filtering against the untreated sample.

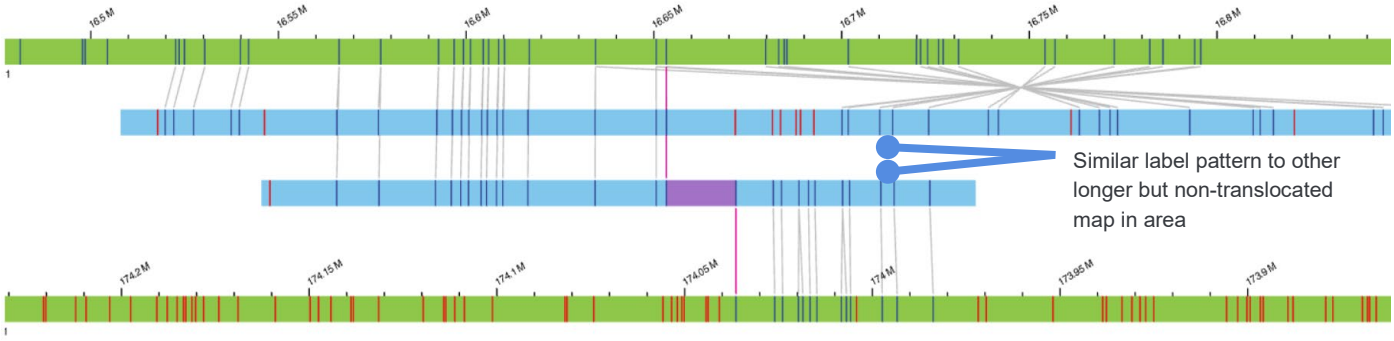

Figure 5. Artifact translocation due to similar patterns on two chromosomes. The low confidence or false translocation calls can be identified by the presence of similar maps with shorter length and similar label pattern to another chromosome. In this case, the longer blue map has the correct alignment (in inverted orientation with the top reference) rather than the shorter blue map that carries the translocation to the bottom reference. Therefore, the call itself is an alignment artifact.

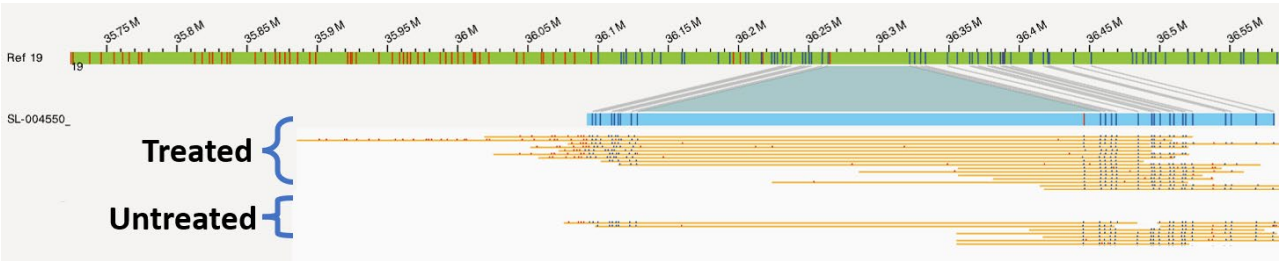

Figure 6. Not-treatment specific artifact – A 271.8kbp insertion has molecule support from both treated and untreated samples.

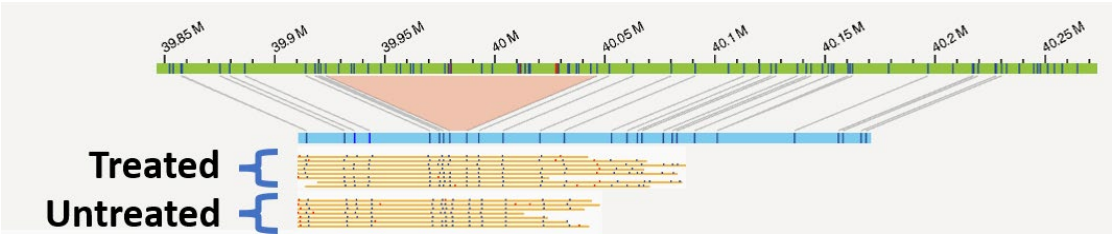

Figure 7. Not-treatment specific artifact – A 115.5 kbp deletion call that contains molecule support for the SV in both the treated and untreated sample.

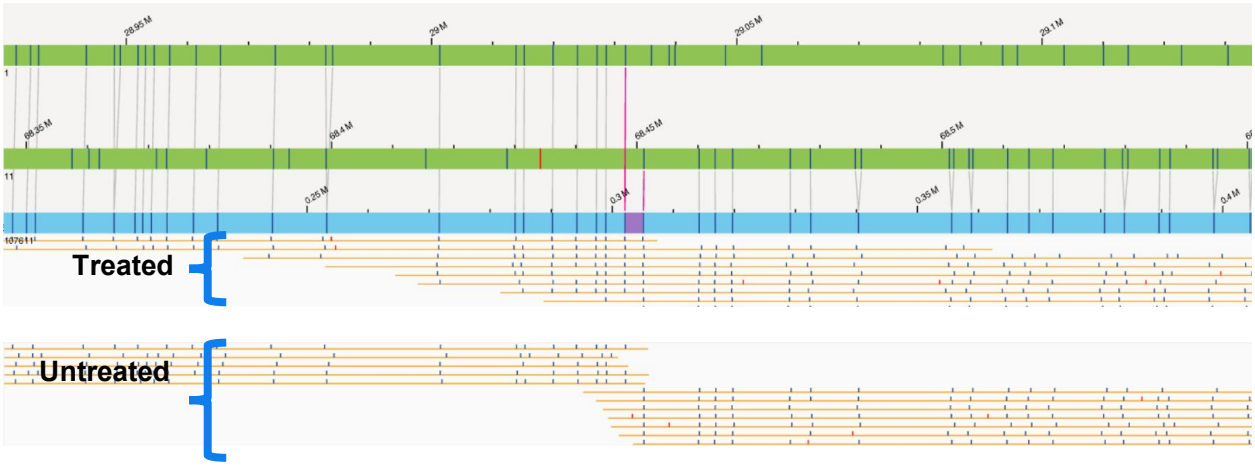

Figure 8. A true treatment specific variant – A translocation, ogm[GRCh38] t(1;11), seen only in the treated sample's molecules.

Finally, it is recommended to manually examine regions that contain the expected genetic edits to validate the integrity of the on-target regions. Bionano's deep coverage workflow, complemented by SV calls and molecule pileup views, presents a powerful toolset for identifying SVs, furthering the quality and safety of bioprocessing and research applications. It is important to highlight that unique sample-specific structural variant maps must be confirmed through detailed interpretation of the data via molecule pileups as this workflow and the presented results are based on preliminary data. By leveraging Bionano's deep coverage protocol and workflow, transgene insertions, mosaicisms, and concatemers can be detected with high sensitivity.

References

1. ClinicalTrials.gov

2. www.fda.gov/regulatory-information/search-fda-guidance-documents/human-gene-therapy-products-incorporating-human-genome-editing

3. Jennifer L Dashnau 1, Qiong Xue 2, Monica Nelson 3, Eric Law 3, Lan Cao 2, Derek Hei 4Cytotherapy 2023 Jan;25(1):1-13.doi: 10.1016/j.jcyt.2022.08.001. Epub 2022 Sep 13.

4. Chaisson, M.J.P., Sanders, A.D., Zhao, X. et al. Multi-platform discovery of haplotype-resolved structural variation in human genomes. Nat Commun 10, 1784 (2019). https://doi.org/10.1038/s41467-018-08148-z

5. DuBose CO, Daum JR, Sansam CL, Gorbsky GJ. Dynamic Features of Chromosomal Instability during Culture of Induced Pluripotent Stem Cells. Genes (Basel). 2022 Jun 27;13(7):1157. doi: 10.3390/genes13071157. PMID: 35885940; PMCID: PMC9318709.

6. Iqbal MA, Broeckel U, Levy B, Skinner S, Sahajpal NS, Rodriguez V, Stence A, Awayda K, Scharer G, Skinner C, Stevenson R, Bossler A, Nagy PL, Kolhe R. Multisite Assessment of Optical Genome Mapping for Analysis of Structural Variants in Constitutional Postnatal Cases. J Mol Diagn. 2023 Mar;25(3):175-188. doi: 10.1016/j.jmoldx.2022.12.005. PMID: 36828597.

7. Levy B, Baughn LB, Akkari Y, Chartrand S, LaBarge B, Claxton D, Lennon PA, Cujar C, Kolhe R, Kroeger K, Pitel B, Sahajpal N, Sathanoori M, Vlad G, Zhang L, Fang M, Kanagal-Shamanna R, Broach JR. Optical genome mapping in acute myeloid leukemia: a multicenter evaluation. Blood Adv. 2023 Apr 11;7(7):1297-1307. doi: 10.1182/bloodadvances.2022007583. PMID: 36417763; PMCID: PMC10119592.

8. Ulrich Broeckel, M. Anwar Iqbal, Brynn Levy, Nikhil Sahajpal, Peter L. Nagy, et.al Multisite Study of Optical Genome Mapping of Retrospective and Prospective Constitutional Disorder Cohorts. medRxiv 2022.12.26.22283900; doi: https://doi.org/10.1101/2022.12.26.22283900.

9. R.E. Stevenson, J. Liu, A. Iqbal, B. DuPont, et al. Multisite evaluation and validation of Optical Genome Mapping for prenatal genetic testing. https://doi.org/10.1101/2022.12.19.2228355**2**

10. Neveling K, Mantere T, Vermeulen S, Oorsprong M, van Beek R, Kater-Baats E, Pauper M, van der Zande G, Smeets D, Weghuis DO, Stevens-Kroef MJPL, Hoischen A. Next-generation cytogenetics: Comprehensive assessment of 52 hematological malignancy genomes by optical genome mapping. Am J Hum Genet. 2021 Aug 5;108(8):1423-1435. doi: 10.1016/j.ajhg.2021.06.001. Epub 2021 Jul 7. PMID: 34237281; PMCID: PMC8387283.

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.

TYPE	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
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.