Bionano Solve™ v3.8.2 Release Notes

DOCUMENT NUMBER:
RNOTE-00007

DOCUMENT REVISION:
D

Effective Date:
05/03/2024

bionano

Table of Contents

Revision History 4

Introduction 5

Compatibility 5

References 5

Improvements for Solve 3.8 5

Support for New Reference Genomes 5

Singularity 5

Improved Support for Standard Genomics File Formats 6

Control Database 6

Gene Annotations 6

Y-PAR Masked Human References 6

Stable Region Analysis 7

Known Issues and Limitations 7

Improvements for Solve 3.8.1 7

Guided Assembly 7

RedHat 9 and Slurm Support 7

Additional Issues 7

Improvements for Solve 3.8.2 8

Quality Score Standardization 8

Copy Number Variant Quality Assessment 8

Other Improvements 8

Technical Assistance 9

Legal Notice **10**

Patents 10

Trademarks 10

Revision History

REVISION	NOTES
A	Initial release of document.
B	Updated for Solve 3.8.1 release including information on Guided Assembly, RedHat9/Slurm support and issues addressed since Solve 3.8
C	Updated to add information under Known Issues and Limitations section.
D	Updated for Solve 3.8.2 release

Introduction

This document describes the release of Bionano Solve™ 3.8. This is an overview of the fixes and improvements of the Bionano Solve analysis tools and pipelines to provide a better understanding of the impact of moving to this version of the software. Should any questions arise, please contact support@bionano.com.

Bionano Tools and Bionano Solve are combined and branded as Bionano Solve. Bionano Solve is installed on Saphyr® Compute, Bionano Compute, and Bionano Access™ Servers before server shipment and installation.

Bionano Solve (folder “tools”) is located at the /home/bionano directory on the Compute server. The folder contains a collection of tools and scripts. Each individual tool is versioned independently. These tools together perform bioinformatics analyses on the Compute server.

Compatibility

Bionano Solve 3.8 is compatible with Bionano Access 1.8 only.

References

Visit https://bionano.com/software-and-data-analysis-support-materials// for file format specifications and Theory of Operation documents.

Improvements for Solve 3.8

Support for New Reference Genomes

TELOMERE-TO-TELOMERE CONSORTIUM REFERENCE

- Implemented system level support for the T2T-CHM13v2.0 reference genome
- Integrated T2T-CHM13v2.0 gene annotation
- Created Bionano Control Database versions for *de novo* Assembly and Rare Variant Analysis pipeline

MM39 REFERENCE

- Implemented system level support for the mouse mm39/GRCm39 reference genome
- Integrated mm39 gene annotation
- Created Bionano Control Database versions for *de novo* Assembly and Rare Variant Analysis pipeline

Singularity

- Adopted Singularity for dependency management of the Solve pipeline (replacing Docker)
- Provides method for portable, consolidated dependency management across supported Compute hardware

- Designed for ease-of-use on high-performance computing (HPC) systems
- Provides improved security for customers by removing requirements for running with elevated or root user privileges

Improved Support for Standard Genomics File Formats

- Introduced OGM BAM output (binary version of the Sequence Alignment/Map format) for molecule-to-reference alignments.
- Updated OGM Variant Call File (VCF) with improved representations of variants and quality filters.
- Updated OGM VCF to include Solve AOH calls for *de novo* and Guided Assembly – constitutional analyses.
- Updated OGM VCF to contain clustered variant calls only to remove redundant calls.
  - Translocations not included in clustering for initial release.
- Enables import of OGM data into VIA as well as any standards-compliant genomics software that supports BAM and VCF.

Control Database

- Added DLE-1 samples to human control database for a total of 285 individuals.
- Complete reanalysis of control databases for *de novo* Assembly and Rare Variant Analysis for hg19 and hg38.
- Introduction of control database for T2T-CHM13v2 reference genome.

Gene Annotations

- Updated gene annotation used by Variant Annotation Pipeline to harmonize gene annotations used by Access and VIA.
- Updated gene annotation to latest RefSeq builds for hg38 and hg19.
- Introduced gene annotation for T2T-CHM13v2.0 and mm39.
- Streamlined gene annotation to provide a single, high quality, current annotation for human analyses while allowing use of custom annotations if desired.

Y-PAR Masked Human References

- Added versions of human reference genomes that mask out the pseudoautosomal regions (PAR) on chromosome Y as do approaches used in many Next-Generation Sequencing (NGS) analysis workflows.
- This is done to address the sequence homology in these regions with the corresponding regions on chromosome X which can interfere with map and molecule alignments to the reference. Masking of these regions has been shown to improve structural and copy number variant calling for genes such as *CRLF2* that are in or near the region.
- Masked references are provided as options for hg19 and hg38 and by default for T2T-CHM13v2.0. For best results with SV detection performance and annotation with the Bionano control databases, the masked references are recommended for hg19 and hg38.

Stable Region Analysis

- Incorporated analysis of stable regions as a quality control step.
- Initially implemented in the EnFocus™ FSHD and EnFocus™ Fragile X applications, assessment of stable regions is now included in the *de novo* and Guided Assembly whole genome analysis pipelines.
- Output included in informatics report

Known Issues and Limitations

- Control databases were generated using the Y-PAR masked versions of hg19 and hg38. These databases can be used to annotate data analyzed using the unmasked versions of the reference genome, but SV calls in Y PAR1 and PAR2 will show as having 0% match in the control database.

Improvements for Solve 3.8.1

Guided Assembly

- Introduction of Guided Assembly pipeline as an alternative to *de novo* assembly and Rare Variant Analysis
- Mode of *de novo* Assembly that uses the reference genome as seed for refinement and extension
- Constitutional analysis targeted for 400 Gb and 800 Gb coverage tiers
- Low Allele Fraction (LAF) – targeted for 1.5 Tb coverage tier for cancer analysis
- Improves variant detection in low-allele fraction samples with improvements in insertions/deletions and duplication detection
- Dedicated control database for both modes for hg19, hg38 and T2T-CHM13v2.0 reference genomes

RedHat 9 and Slurm Support

- Established support for RedHat 9 Linux operating system. RedHat Linux is a commercially supported distribution and is offered as a replacement to CentOS Linux 7 which will reach end of life (EOL) on June 30, 2024.
- Implemented use of the Slurm Workload Manager for all analysis pipelines. Slurm replaces Sun Grid Engine in the RedHat 9 version of the Solve pipeline.
- CentOS and SGE installers are provided for Solve 3.8.1.

Additional Issues

- Corrected defect in dual and trio modes of Variant Annotation Pipeline where molecule support was not calculated for case or parental samples.
- Corrected issue where CNV calls for mouse reference mm10 reported with confidence scores of zero.
- Corrected issue so that translocation calls are clustered in the VCF as are other variant types.
- Updated representation of translocations in VCF to reflect directionality of breakpoints more accurately.
- Corrected confidence intervals for translocations in VCF so that each breakpoint is assessed separately.

- Remove redundant samples from human DLE-1 control databases. Solve 3.8 reported 384 individuals while number of unique individuals is 285.
- Corrected Variant Annotation Pipeline human gene annotation files to contain all transcripts rather than only primary transcript.

Improvements for Solve 3.8.2

Quality Score Standardization

- Implemented uniform quality scoring in VCF for structural variants. High confidence variants of any type are defined as variants with a QUAL score >= 20.
- Implemented confidence score for insertion and deletions 300bp – 1kbp.
- Implemented confidence score for duplication structural variants.

Copy Number Variant Quality Assessment

- Added coverage quality assessment to detect presence of telomeric gain artifacts. Assessment reported in the informatics report of all human and mouse whole genome pipelines as well as EnFocus™ FSHD and EnFocus™ Fragile X.

Other Improvements

- Fixed issue with HybridScaffold perl locale warnings.
- Corrected VCF representation of deletion calls produced by the RVA pipeline to show genotype as undetermined rather than defaulting to heterozygous.
- Fixed bug in de novo assembly pairwise alignment stage that in some cases increased fragmentation of assembly. While structural variant calls do not appear affected, total number of genome maps was elevated in some cases.
- Refreshed de novo assembly control databases for hg19, hg38 and T2T-CHM13v2.0 reference genomes.
- Added improvements to pipeline reproducibility to reduce variation between analyses.
- Added improvements to pipeline robustness to remove silent failures and improve reporting of root cause errors.
- Updated handling of molecule sizing errors and molecules with outlying label density to improve compute performance.

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 Monday through Friday, 9:00 a.m. to 5:00 p.m., CET UK: +44 115 654 8660 (UK) 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.

Patents

Products of Bionano Genomics® 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™, Bionano Genomics®, Bionano Solve™, Saphyr®, Saphyr Chip®, Bionano Access™, VIA™ software, and Bionano EnFocus™ 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.