Bionano Solve v3.5 Release Notes

Document Number: 30323

Document Revision: B

Table of Contents

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 Genomics logo and names of Bionano Genomics products or services are registered trademarks or trademarks owned by Bionano Genomics in the United States and certain other countries.

Bionano Genomics®, Irys®, IrysView®, IrysChip®, IrysPrep®, IrysSolve®, Saphyr®, Saphyr Chip®, Bionano Access®,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 Genomics is given or implied. Users are not permitted to use these trademarks without the prior written consent of Bionano Genomics. 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.

Revision History

Revision	Notes
A	Initial release of document.
B	Solve 3.5.1 additions.

Updates

The following updates have been made to the 3.5.1 release of Bionano Solve.

Update	Component	Change
3.5.1	Rare Variant Pipeline	Updated RVP run parameters file (SingleMoleculePipelineParameters.xml) to specify the amount of memory allowed for the “alignmolvrefsv” stage
		Updated RVP cluster parameters files (SingleMoleculePipelineClusterParameters*.xml) to optimize performance for the “alignmolvrefsv” stage

Introduction

This document describes the release of Bionano Solve v3.5. We provide an overview of the fixes and improvements as they impact RefAligner, the *de novo* assembly pipeline, and Rare Variant Pipeline. The latest addition is the Bionano EnFocus™ FSHD Analysis Pipeline.

References

Visit https://bionanogenomics.com/support-page/bionano-solve/ for file format specifications and Theory of Operation documents. A new FSHD-specific JSON output file format specification document has been added.

Bionano EnFocus™ FSHD Analysis Pipeline

- Added a new pipeline to analyze regions relevant to facioscapulohumeral muscular dystrophy (FSHD), such as the chr4 D4Z4 repeat region. The pipeline makes use of a new local assembly procedure to assemble regions of interest. The resulting genome maps are analyzed. By focusing on maps that align to the chr4 D4Z4 region, the pipeline sizes the D4Z4 repeat arrays and assigns haplotype to the alleles. The pipeline can differentiate between the permissive A alleles and the non-permissive B alleles. It can measure D4Z4 repeats to within 1 unit in most cases.
    - Note: when run in Bionano Access, the pipeline only supports DLE-1. When run on the command line, the pipeline could analyze BssSI data.
- Added reporting of SVs and/or CNVs proximal to the chr4 DZ4Z region, and CNVs proximal to the SMCHD1 gene, which may be relevant in FSHD Type 2 cases.

- Added evaluation of data quality based on the molecule alignment quality (based on molecule N50 > 150 kbp, map rate, and effective coverage), consensus map alignment quality (based on analysis of stable regions in the genome), and the inferred sex of the sample (based on copy number analysis). See Theory of Operation for details.
- Added a check to flag and reject input bnx files containing data from multiple flowcells or chip runs.
    - Note: If chip run information is absent in the input, the pipeline would output a warning message, but it would proceed normally otherwise.

RefAligner

- Improved handling of tandem repeats in the maps when making translocation breakpoint calls; previously, only tandem repeats on the reference were considered. These tandem repeats previously led to FP translocation breakpoint calls, where a map could incorrectly align to multiple reference regions.
- Updated assembly parameters xml such that the xml headers inside the files would match the filenames; previously, they were not always consistent.
- Updated default assembly parameters for non-human assemblies. In some cases, when the pre-assembly option was enabled, only a small fraction of the genome was assembled.

De novo assembly pipeline

- Updated the compression script to use zip instead of gzip for compression for faster import. The file structure in the compressed output has changed slightly as a result. For example, the molecule-to-map alignment files are no longer separated compressed. However, generally, the same files are being included. There is now a new _debug.zip output that contains additional files that may be useful for debugging purposes.
- Updated the alignmolvref stage (molecule-to-reference alignment) to run as a single job for improved runtime performance.
- Fixed a bug related to using the bypass option when the pre-assembly option is enabled. The pipeline could not correctly find existing files when both options were on.

Rare Variant Pipeline (RVP)

- Updated extraction of reference regions from the initial duplication clusters so that the SVs can be correctly assembled in the consensus check step.
- Updated SV count table in informatics report so that it is consistent with the *de novo* assembly informatics report.

- Improved runtime performance for the computation of the molecule alignment statistics. Previously, when input coverage was very high, the pipeline would appear to be stuck in this stage.
- Fixed a bug related to the pipeline incorrectly reporting that there were errors when there was none.
- Fixed a bug related to the effective coverage not correctly computed.

Other known issues and limitations

- Adaptive memory usage algorithm could be activated during assembly under different server environments. As a result, there could be slight differences in the SV output.
- Users with non-standard cluster or server configurations may experience suboptimal runtime performance.
- Hybrid Scaffold output FASTA/AGP files may contain header lines with whitespaces and would not pass NCBI AGP validation.
- Haplotype-aware refinement on non-human datasets is not a supported feature, and its use may have unintended consequences.
- Large heterozygous duplications may be called as homozygous when the allele with a single copy is not successfully assembled, if too few molecules span across the entire duplication region.
- We observed that with the Rare Variant Pipeline, PPV was slightly lower (~80%) for deletions under 25 kbp. This was not observed with the *de novo* assembly pipeline.

Technical Assistance

For technical assistance, contact Bionano Genomics 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@bionanogenomics.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.bionanogenomics.com/support**

Bionano Genomics, Inc.
9540 Towne Centre Drive, Suite 100
San Diego, CA 92121