How to Align a BNX to a Reference

Document Number: 30194

Document Revision: B

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®, and Bionano Access® 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.

Compatibility

These instructions were written for Bionano Solve® 3.2 and Bionano Access® 1.2.

Work Instructions

In Bionano Access® it is possible to align CMAPs to each other. The query CMAP in an alignment can be contigs or molecules. However, it is currently not possible to align a BNX file to a reference CMAP using Bionano Access. However, you can do this alignment on the command line and import the results. Follow the instructions below to do a molecule to reference CMAP alignment.

1.) SSH to your computing environment where Bionano Solve 3.2 is installed.

2.) Go to the Pipeline directory inside your Bionano Solve 3.2 installation. You should see a script named align_bnx_to_cmap.py. On our compute servers the path to this file would be /home/bionano/tools/pipeline/1.0/Pipeline/1.0. This is the script we will be using to generate the alignment. You can issue this command to see the command line help:

```
Python align_bnx_to_cmap.py --h
```

You will get the following response:

```
usage: align_bnx_to_cmap.py [-h] --prefix PREFIX --mol MOL --ref REF --ra RA
                            --nthreads NTHREADS [--output OUTPUT]
                            [--optArgs OPTARGS] [--snrFilter SNRFILTER]
                            [--snrThreshold SNRTHRESHOLD [SNRTHRESHOLD ...]]
                            [--dosnrThresh][--color COLOR]
                             [--pipeline PIPELINE]


Wrapper for running molecule-to-reference alignment using Pipeline scripts.
Final alignment results are in <output>/contigs/alignmolvref/merge/

optional arguments:
  -h, --help            show this help message and exit
  --prefix PREFIX       sample name prefix (required)
  --mol MOL             input molecule bnx (required)
  --ref REF             input reference cmap (required)
  --ra RA               RefAligner directory (required)
  --nthreads NTHREADS   Number of threads (required)
  --output OUTPUT       output dir (optional, default pwd)
  --optArgs OPTARGS     optArguments.xml file (optional, default
                        "optArguments_haplotype_saphyr.xml" in RefAligner dir)
  --snrFilter SNRFILTER
                        Label SNR filter method: 0 for histogram (default), 1
                        for fixed threshold (see below), 2 for NO filter
  --snrThreshold SNRTHRESHOLD [SNRTHRESHOLD ...]
                        Label SNR threshold: if 1 specified for previous
                        argument (or 0 specified, but insufficient data), use
                        argument as fixed threshold (ignored if histogram
                        method used or no SNR filter); if 2-color bnx, can
                        specify threshold for second color as optional second
                        argument [default 3.5, 3.5]
  --dosnrThresh         determines whether to create threshold labels based on SNR
  --color COLOR         Color channel for alignment: replace -usecolor X in
                        optArgs with this, must be either 1 or 2 [default OFF]
  --pipeline PIPELINE   Pipeline directory (optional, defaults to script dir)
```

3.) Issue the following command to align a BNX file to a reference. Replace the values in '<< >>' with the values for your circumstance.

```
python align_bnx_to_cmap.py --mol <<bnx file with path here>> --ref <<reference cmap with path here>> --prefix test --pipeline <<path to pipeline>> --ra <<path to refaligner>> --nthreads 64 --output <<output path>> --optArgs <<optArgs file to use>> --snrFilter 0 --snrThreshold 3.5 3.5 --dosnrThresh –-color 1
```

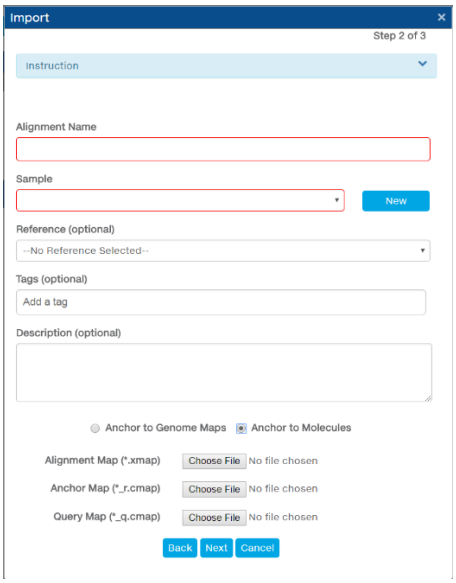
example:

```
python align_bnx_to_cmap.py --mol /home/users/jsmith/dev/GreenRed/RawMolecules.bnx --ref /home/users/jsmith/genomes/human/hg19/hg19_BSPQI_0Kb_0labels.cmap --prefix test --pipeline /home/users/jsmith/Solve3.2/Solve3.2.2_08222018/Pipeline/08222018/ --ra /home/users/jsmith/Solve3.2/Solve3.2.2_08222018/RefAligner/7782.7865rel/ --nthreads 64 --output /home/users/jsmith/myoutput --optArgs /home/users/jsmith/Solve3.2/Solve3.2.2_08222018/RefAligner/7782.7865rel/optArguments_haplotype_saphyr_human.xml --snrFilter 0 --snrThreshold 3.5 3.5 --dosnrThresh --color 1
```

4.) When the command has completed you will find the maps in the /contigs/alignmolvref/merge folder in the output directory you provided.

5.) An alignment consists of three files the reference CMAP (_r.cmap), the query map (_q.cmap), and the alignment map (.xmap). If your command was against a single genome map it should generate one alignment trio (the three files mentioned). If the reference map you provided has multiple contigs the command will generate multiple map files. Typically in this use case it will generate just one alignment trio. Localize these alignment files to your workstation so they are accessible from your browser. Note: all .xmap and .cmap files outputted are also put into the zip file, "alignments.tar.gz".

6.) To view the maps in Bionano Access, navigate to the desired project then click the Import button. Choose to import an alignment. Then give your alignment a name and select a sample. Then select the xmap, _r.cmap, and _q.cmap files to upload from your command line output. Be sure you select 'Anchor to Molecules'. Then click Next and your import will start. The system will notify you when the import has completed.



7.) To view the alignment after the import has completed select the alignment object in your project and click the View Molecule Alignment link in the options pane. This will open the three level viewer. When viewing molecules aligned directly to the reference there may be an excessive amount of label information to render. Be patient for the data to load. You may need to zoom in before the labels will appear.

8.) When the alignment is viewed in Access, only RefContigID=1 is available to view. To display a different RefContigID, the xmap file must be edited to move (not copy) the data line that contains the RedContigID whose data is to be displayed to the first data line.

```
# XMAP File Version:  0.2
# Label Channels: 1
# Reference Maps From:  /home/users/johndoe/data/test_BssSI_20170421_final/mol_to_ref_test/mol_to_ref_test_r.cmap
# Query Maps From:  /home/users/johndoe/data/test_spleen_BssSI_20170421_final/mol_to_ref_test/mol_to_ref_test_q.cmap
#h XmapEntryID  QryContigID RefContigID QryStartPos QryEndPos RefStartPos RefEndPos Orientation Confidence  HitEnum QryLen  RefLen  LabelChannel  Alignment
#f int          int         int         float       float     float       float     string      float       string  float   float   int           string
1 4665155 1 40827.8 295808.8  3276687.0 3529411.0 + 14.53 1M1I1D4M1I5M1D4M  301214.4  195471971.0 1 (16,1) (17,1) (19,3) (20,4) (21,5) (22,6) (23,8) (24,9) (25,10) (26,11) (27,12)
2 2596457 1 36705.5 356205.6  3276687.0 3594665.0 + 11.31 1M2I1D3M1I2M1D3M3I3D3M1I1M1D3M  381616.1  195471971.0 1 (16,1) (17,1) (19,4) (20,5) (21,6) (22,8) (23,9) (25,10) (26,11
3 4680895 1 389305.4  1192.3  3289471.0 3679400.0 - 13.89 1M2I2M1D1M2D1M2I2M1D4M1I2M1D1M1I1D3M1I1D4M  848489.1  195471971.0 1 (18,28) (19,25) (20,24) (22,23) (25,22) (26,19) (
4 1947302 1 135414.7  6590.2  3366990.0 3497654.0 - 11.31 6M1I3M  432897.4  195471971.0 1 (19,10) (20,9) (21,8) (22,7) (23,6) (24,5) (25,3) (26,2) (27,1)
5 2678251 1 192465.9  20.0  3366990.0 3560807.0 - 12.50 3M1D1M1D3M1D7M  281628.2  195471971.0 1 (19,14) (20,13) (21,12) (23,11) (25,10) (26,9) (27,8) (29,7) (30,6) (31,5) (32,4) (3
6 9014139 1 541372.6  751369.0  3366990.0 3573450.0 + 14.19 3M2I6M1D3M1D1M1I1M1D2M  795686.6  195471971.0 1 (19,84) (20,85) (21,86) (22,89) (23,90) (24,91) (25,92) (26,93) (27,9
7 6340157 1 286280.6  17516.6 3383653.0 3648740.0 - 13.64 7M1D3M1D1M1I1M1D1M1I1M2I1M2I2M2I1D2M  289111.4  195471971.0 1  (21,27) (22,26) (23,25) (24,24) (25,23) (26,22) (27,21)
8 805244  1 829.5 173417.7  3391288.0 3560807.0 + 14.83 6M1D5M1I2M  173437.7  195471971.0 1 (22,1) (23,2) (24,3) (25,4) (26,5) (27,6) (29,7) (30,8) (31,9) (32,10) (33,11) (34,13) (3
```

Above: Original XMAP file, where the first data line has a RefContigID=1.

Below: Modified XMAP file, where the first data line has a RefContidID=7. This line was moved to the top in a text editor. When this is now imported into Access, the viewer will display the alignments to RefContigID 7.

```
# XMAP File Version:  0.2
# Label Channels: 1
# Reference Maps From:  /home/users/johndoe/data/test_BssSI_20170421_final/mol_to_ref_test/mol_to_ref_test_r.cmap
# Query Maps From:  /home/users/johndoe/data/test_spleen_BssSI_20170421_final/mol_to_ref_test/mol_to_ref_test_q.cmap
#h XmapEntryID  QryContigID RefContigID QryStartPos QryEndPos RefStartPos RefEndPos Orientation Confidence  HitEnum QryLen  RefLen  LabelChannel  Alignment
#f int          int         int         float       float     float       float     string      float       string  float   float   int           string
8281  858255  7 1693.1  249235.8  37723635.0  37974014.5  + 14.94 1M3I1D5M1I1D7M2I2D3M1I2M2D1M1D2M1D1M  867817.3  182113224.0 1 (3693,1) (3695,5) (3696,6) (3697,7) (3698,8)
1 4665155 1 40827.8 295808.8  3276687.0 3529411.0 + 14.53 1M1I1D4M1I5M1D4M  301214.4  195471971.0 1 (16,1) (17,1) (19,3) (20,4) (21,5) (22,6) (23,8) (24,9) (25,10) (26,11) (27,12)
2 2596457 1 36705.5 356205.6  3276687.0 3594665.0 + 11.31 1M2I1D3M1I2M1D3M3I3D3M1I1M1D3M  381616.1  195471971.0 1  (16,1) (17,1) (19,4) (20,5) (21,6) (22,8) (23,9) (25,10) (26,1
3 4680895 1 389305.4  1192.3  3289471.0 3679400.0 - 13.89 1M2I2M1D1M2D1M2I2M1D4M1I2M1D1M1I1D3M1I1D4M  848489.1  195471971.0 1 (18,28) (19,25) (20,24) (22,23) (25,22) (26,19)
4 1947302 1 135414.7  6590.2  3366990.0 3497654.0 - 11.31 6M1I3M  432897.4  195471971.0 1 (19,10) (20,9) (21,8) (22,7) (23,6) (24,5) (25,3) (26,2) (27,1)
5 2678251 1 192465.9  20.0  3366990.0 3560807.0 - 12.50 3M1D1M1D3M1D7M  281628.2  195471971.0 1 (19,14) (20,13) (21,12) (23,11) (25,10) (26,9) (27,8) (29,7) (30,6) (31,5) (32,4) (3
6 9014139 1 541372.6  751369.0  3366990.0 3573450.0 + 14.19 3M2I6M1D3M1D1M1I1M1D2M  795686.6  195471971.0 1 (19,84) (20,85) (21,86) (22,89) (23,90) (24,91) (25,92) (26,93) (27,9
```

Technical Assistance

For technical assistance, contact Bionano Genomics Technical Support. You can retrieve documentation on Bionano products, Safety Data Sheets, certificates of analysis, frequently asked questions, and other related documents from the Support page 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) 888-7600
Website	www.bionanongenomics.com/support