Continuation of Processing From the eBIC AutoProcessing Pipeline

Revision:	1.0
CryoSPARC version:	4.5.3
Relion Version:	5.0
Revision date:	12 February 2025
Revision summary:	1) New guide

Comments to:	Peter Harrison Stephen Riggs Dan Hatton

Acknowledgements:	We would like to acknowledge all eBIC staff both past and present for their invaluable efforts and contributions towards the writing of this documentation.

Introduction

This document is a guide for eBIC users to enable them to use and continue processing their data using the eBIC AutoProcessing pipeline as a start point.

Getting prepared

First download your data on to you home system. If your session was fewer than 40 days ago, your data will still be on the Diamond file system. Use Globus Online to transfer datasets over 20GB, or FTP, SFTP or rsync for under 20GB. If your visit was more than 40 days ago it has probably been removed from the beamline storage systems. Use TopCAT to access your data from the archive. For Diamond Light Source data restore help, learn more here.

For more information please visit https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data.html

Contents

1 Continuation of Processing Using Relion or Doppio

1.1 Viewing projects in Relion or Doppio

Auto-processed results are directly compatible with Relion and Doppio, as all the necessary star files to continue from are included. The `relion_murfey` project can be opened and viewed in either program.

The full project will contain many repeats of Class2D and Selection jobs as these are run in batches, which you might want to exclude.

☐ The project for Relion or Doppio should be set to `/path/to/your/data/relion_murfey/`

☐ To exclude the batched jobs, first rename `default_pipeline.star` to `full_pipeline.star`, then rename `short_pipeline.star` to `default_pipeline.star`

If you use the `short_pipeline.star` , then the remaining jobs in the project will be preprocessing jobs (MotionCorr, CtfFind, AutoPick, IceBreaker, Extract, Select/split), 3D classification jobs (Select/Best_particles, InitialModel, Class3D) and refinement jobs (Select/onvalue, Extract, Refine3D, MaskCreate, PostProcess).

1.2 Stages from which to continue processing

Any of the stages in the project can be continued from in Relion or Doppio.
If you are happy with the picks we recommend using either:

`/AutoPick/job007/autopick.star` This is the complete list of crYOLO picks.

`/Select/Best_particles/particles_all.star` These are the particles selected from the best 2D classes in the pipeline.

Alternatively, if you want to pick your own particles use either:

`/MotionCorr/job002/corrected_micrographs.star` This is the complete list of micrographs.

`/CtfFind/job006/micrographs_ctf.star` This contains the complete list of micrographs with CTF values.

3D volumes from Class3D, Refine3D or PostProcess jobs can be used as reference models to continue your own classification.

2 Continuation of Processing with CryoSPARC

2.1 Importing Motion Corrected Micrographs and Particles into CryoSPARC

Motion corrected micrographs from the AutoProcessing pipeline can be imported into CryoSPARC. We recommend using the motion corrected micrographs and then running CTF correction in CryoSPARC.

☐ In CryoSPARC, select the Import Micrographs job.

☐ In the 'micrographs data path' box input the following:

`/path/to/your/data/relion_murfey/MotionCorr/job002/Movies/GridSquare_*/Data/*.mrc`

☐ Input values for the pixel size, accelerating voltage, spherical aberration and total exposure.

☐ Run the job.

☐ Run CTFFIND4 or Patch CTF as normal in CryoSPARC.

Particles from the AutoProcessing pipeline can be imported using the 'Import Particles' job type. In effect, any `particles.star` file can be imported into CryoSPARC. However, it is not possible to import particles from the AutoPick (crYOLO) job as the pipeline does not write out the appropriate star file.
We recommend either of the following:

`/Extract/job008/particles.star` This is the complete particle stack of crYOLO picks.

`/Select/Best_particles/particles_all.star` These are the best particles from 2D classification in the pipeline.

Alternatively, any `particle.star` file from any Class2D, Class3D or Refine3D can be imported into CryoSPARC.

☐ In CryoSPARC, select the 'Import Particles' job type.

☐ In the 'particle meta path' box, give the path to your data.

`/path/to/your/data/relion_murfey/Select/Best_particles/particles_all.star`

☐ Select 'Ignore raw data', 'Ignore pose data' and 'Remove leading UID in input expose file name'.

☐ Unselect 'Enable Strict Checking'.

☐ Link your CTF Estimation job from above to the source exposures slot.

☐ Run the job.

This error message will be produced if importing from the Select job:

```
Warning: Output will not contain particle orientations since the input star file does not contain: 'rlnRandomSubset'. Finally, parameter 'psize_A' was not set.
```

This error message will be produced if importing from the Extract job:

```
Warning: Output will not contain particle orientations since the input star file does not contain: rlnAngleRot, rlnAngleTilt, rlnAnglePsi, either {'rlnOriginX', 'rlnOriginY'} or {'rlnOriginXAngst', 'rlnOriginYAngst'}, 'rlnRandomSubset'. Finally, parameter 'psize_A' was not set.
```

However, the job will complete and the outputs can be used. All of these particle star files contain CTF information. So, it is not strictly necessary to re-run CTF refinement in CryoSPARC. However, we recommend CTF refinement is re-run as it can aid downstream processing.

2.2 Importing Raw Movies and Particles into CryoSPARC

Raw movies can be imported into CryoSPARC. Motion Correction and CTF Estimation can then be performed. Particle picks from the AutoProcessing pipeline can then be imported and used in CryoSPARC. However, several changes need to be made depending on the Motion Correction route taken.

2.2.1 Using MotionCorr2

☐ Using the 'Import Movies' job, import the raw movies, gain reference (`gain.mrc` in the processing directory).

☐ Set the pixel size, accelerating voltage, spherical aberration and total exposure.

☐ Run MotionCorr2 through CryoSPARC.

☐ Run Patch CTF or CTFFIND4 through CryoSPARC.

☐ Select the 'Import Particle Stack' job type and in 'Particle meta path' give the path to the `particle.star` file.

☐ Select 'ignore raw data' and 'Ignore pose data'

☐ For the following options, input the corresponding values:

Length of input exposure file name prefix to cut	22
Length of input exposure file name suffix to cut	36
Length of rlnMicrographName base name prefix to cut for query	0
Length of rlnMicrographName base name suffix to cut for query	31

☐ Untick 'enable strict checking'.

☐ Run the job.

2.2.2 Using Patch Motion Correction and Full-Frame Motion Correction

Patch Motion Correction and Full-Frame Motion Correction in CryoSPARC introduce a flip on the micrographs, which results in the imported crYOLO picks not aligning correctly. To overcome this, movies must be flipped in x using imod before import into CryoSPARC.

☐ Use imod to flipx the raw movies and the gain reference.

For mrc:

```
clip flipx [input] [output]
```

For tiff:

```
clip flipx -f TIFF [input] [output_flippedx]
```

☐ The following assumes that `_flippedx` has been added to the suffix of the movie files.

☐ The newly flipped files can be imported into CryoSPARC using the Import Movies job type.

☐ Perform Full Frame Motion Correction or Patch Motion Correction as desired.

☐ Select the 'Import Particle Stack' job type and in 'Particle meta path' give the path to the `particle.star` file.

☐ Select 'ignore raw data' and 'Ignore pose data'.

☐ If you have used Path Motion Correction, for the following options, input the corresponding values:

Length of input exposure file name prefix to cut	22

Length of input exposure file name suffix to cut	50
Length of rlnMicrographName base name prefix to cut for query	0
Length of rlnMicrographName base name suffix to cut for query	31

☐ If you have used Full Frame Motion Correction, for the following options, input the corresponding values:

Length of input exposure file name prefix to cut	22
Length of input exposure file name suffix to cut	37
Length of rlnMicrographName base name prefix to cut for query	0
Length of rlnMicrographName base name suffix to cut for query	31

☐ Untick ‘enable strict checking’.

☐ Run the job.

2.3 Importing 2D Class Averages into CryoSPARC

2D class averages can be imported using the ‘Import Templates’ option

☐ Select the ‘Import Templates’ job type

☐ Give the path to the template file (`run_it025_classes.mrcs` file). For example:

`/path/to/your/data/relion_murfey/Class2D/job010/run_it025_classes.mrcs`

☐ Give the pixel size. This is the pixel size of the `mrcs` file ***not*** of the original data collection.

☐ If you are unsure of the pixel size, it can be found be running the `header` command in imod.

☐ Run the job.

2.4 Importing 3D Class Averages into CryoSPARC

3D volumes and masks can be imported using the ‘Import 3D Volumes’ option.

☐ Select the ‘Import 3D Volumes’ job type

☐ Give the path your desired 3D volume. For example:

`/path/to/your/data/relion_murfey/Class3D/job012/run_it025_class003.mrc`

☐ Set the type of data to be imported (generally from the AutoProcessing pipeline this will either be map or mask).

☐ Give the pixel size. This is the pixel size of the `mrc` file ***not*** of the original data collection.

☐ If you are unsure of the pixel size, it can be found by running the `header` command in imod.

☐ Run the job.