Workflow @
CryoEM Shared Resource
Howard Hughes Medical Institute
Janelia Research Campus

Zhiheng Yu
# Two types of work

<table>
<thead>
<tr>
<th>Routine data collection for users</th>
<th>Collaboration projects with collaborators</th>
</tr>
</thead>
<tbody>
<tr>
<td>We focus mainly on instrument operation and data collection.</td>
<td>We cover the entire cryoEM workflow from grids preparation to map generation.</td>
</tr>
</tbody>
</table>

To serve labs with cryoEM experiences and easy access to grids preparation/screening equipment.

Receive: ready-to-image cryo grids  
Deliver: large quantity of raw data

To serve labs with limited experience in cryoEM or no easy access to equipment.

Receive: solution samples  
Deliver: 3D density map or feedback to improve sample
Routine data collection for users

- **Specimen Preparation**
  - Users send cryoEM images to qualify for data collection time
  - Users send frozen grids to the facility
  - Facility staff often provide suggestions for optimization if necessary

- **Data Acquisition**
  - Grids clipping and loading into the microscope.
  - Facility staff take care of aligning and operating microscope, optimizing imaging condition/parameters and setting up automated data collection.
  - Users pick acquisition targets. (can be done remotely if not present on site)
  - Facility staff monitor the data collection process and do quality check regularly.
  - Data management and transfer

- **Image Processing**
  - Users take full responsibility
  - Facility staff often provide assistance by request (motion correction, CTF estimation)

- **Analysis**
  - Users’ full responsibility
More on data collection

Grid clipping and loading 4-8 grids per session

Grid screening

Atlas acquisition (for full grid and selected squares)

Microscope Alignment, GIF tuning, camera gain reference

Set up imaging parameters (exposure time, mag, frame rate, defocus and etc)

Add targets into queue

Automated Data Collection

Quick motion correction and quality evaluation

On the fly data transfer

Monitor run progress and troubleshoot if necessary

Facility Staff

Users
Data Transfer to Users

• rsync via ssh directly to remote server.
  • Yay: allow users to process data on-the-fly.
  • Nay: remote firewall and security policy. A guest account with password is required.

• Portable Hard Drives
  • Yay: easy to set up.
  • Nay: takes time.

We currently don’t grant access to users to download from our server due to Janelia network security policy.
Data Management and Storage

Raw movie stacks with super-resolution mode are generated as compressed TIF files, typically about 600MB per movie or 1.5TB per 3-day session.

- RAID on K2 computer/Falcon capture PC
- Janelia network storage 20TB
  - Continuously update
  - Daily backup by IT

- Janelia RAID on cluster
- CryoEM Facility Local RAID ~300TB

Rapid read/write disks for data processing, i.e. motion correction.

Data are kept for ~2 months

10G network either on fiber optics or copper
Post data collection follow-up

• Book keeping: parameters such as total dose, exposure time, magnification, grid type, issues encountered and etc.

• Voluntary user feedbacks.

• Discussions and suggestions for next session if needed.

• When data is published, user should send a link to the paper as well as a representative figure to the facility.
Collaboration projects

Collaborators send solution samples to the facility
Facility staff screen samples by negative stain or cryo and provide suggestions to optimize protein preparation. Multiple iterations might be needed till suitable cryo grids are obtained.

Facility staff carry out the data collection. Collaborators are welcome to join the data collection and on-the-fly early processing (help pick particles).

Facility staff lead in data processing but interact closely with collaborators.

Facility staff send collaborators intermediate results and final maps. Collaborators analyze maps, build atomic models, and interpret results.
The key determinant: sample preparation

Protein expression, purification, and complex assembly

Buffer exchange or dilution if necessary

Negative Staining Analysis

CryoEM Analysis

High quality high throughput Data Collection

Titan Krios

Model Building, analysis and interpretation

Provide suggestions for optimization and future direction

If sample shows potential

2D Class Average

2D Class Average

Image Processing /map generation

Facility Staff

Collaborators
One typical image processing workflow

Data decompression and gain correction
IMOD

Motion correction, dose filter and CTF determination
* cisTEM* (Unblur, sum_movie, CTFFIND4)

Particle picking and extraction
EMAN2, Relion

Reference-free 2D Classification
Relion, cisTEM*

3D Classification
* cisTEM*/FrealignX, Relion

3D Refinement
* cisTEM*/FrealignX, Relion

Postprocessing
FrealignX, Relion

Initial model generation
EMAN2, cisTEM*/FrealignX

* cisTEM from Niko Grigorieff’s lab is a complete package from raw movie stack to 3D classification and refinement with a GUI interface for single particle cryoEM. Unblur, sum_movie, ctffind4 and FrealignX are all part of cisTEM which also includes 2D classification and more.
One example: T3SS Basal Body Complex

Acknowledgement

Many factors contributing to the success of a shared facility: equipment, user base, management, good policy/proctcol and more importantly, the team members who carry out the day to day work and interact with users.

- Reed George and HHMI/Janelia management
- Niko Grigorieff & Tamir Gonen and lab members
- Mark Philip and Building Facility team at Janelia
- Payton Lewis, Haifeng He from FEI and Fred Ulmer, Daniel Ray from Gatan
- All users