

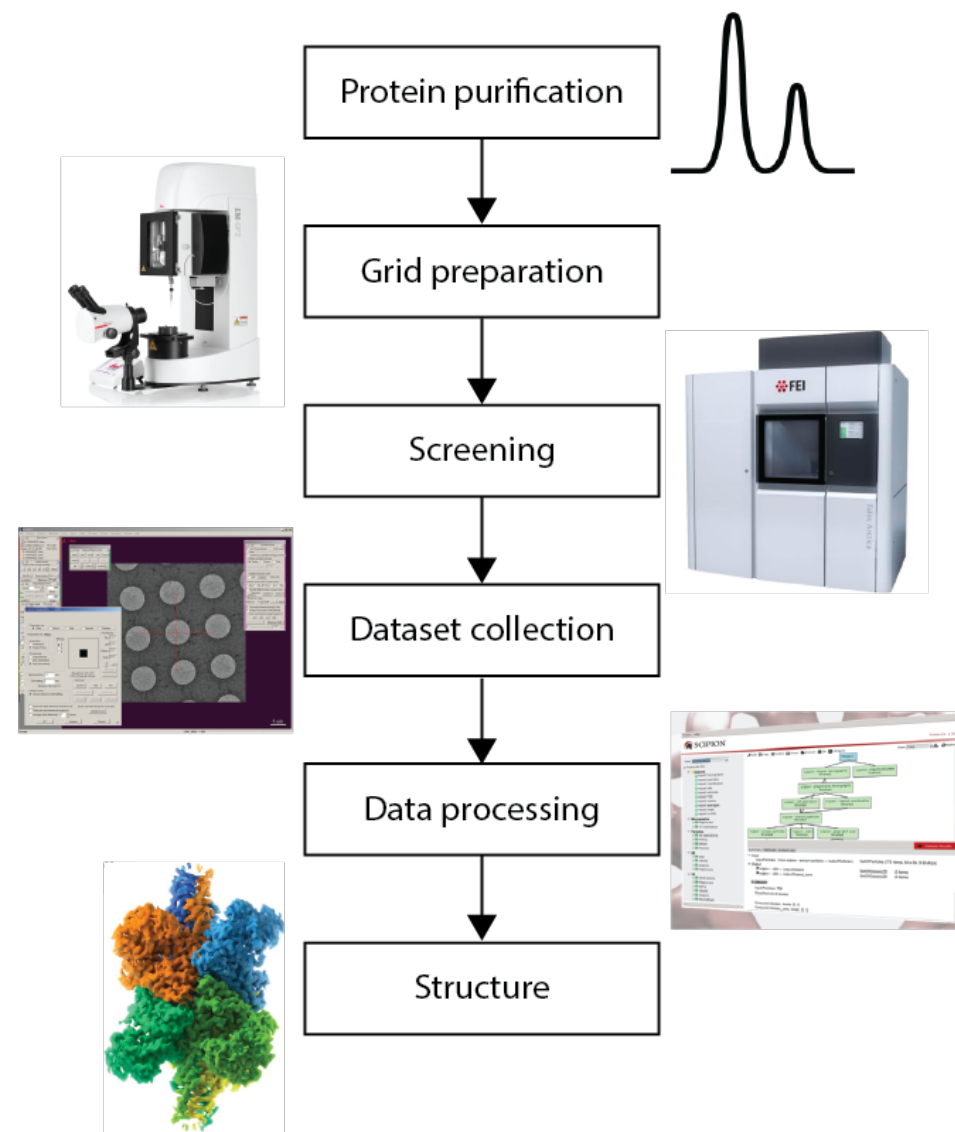


Automated systematic evaluation of cryo-EM specimens with SmartScope

Jonathan Bouvette
Ph.D, visiting fellow
National Institute of Environmental Health Sciences

April 6th 2022,
NYSBC

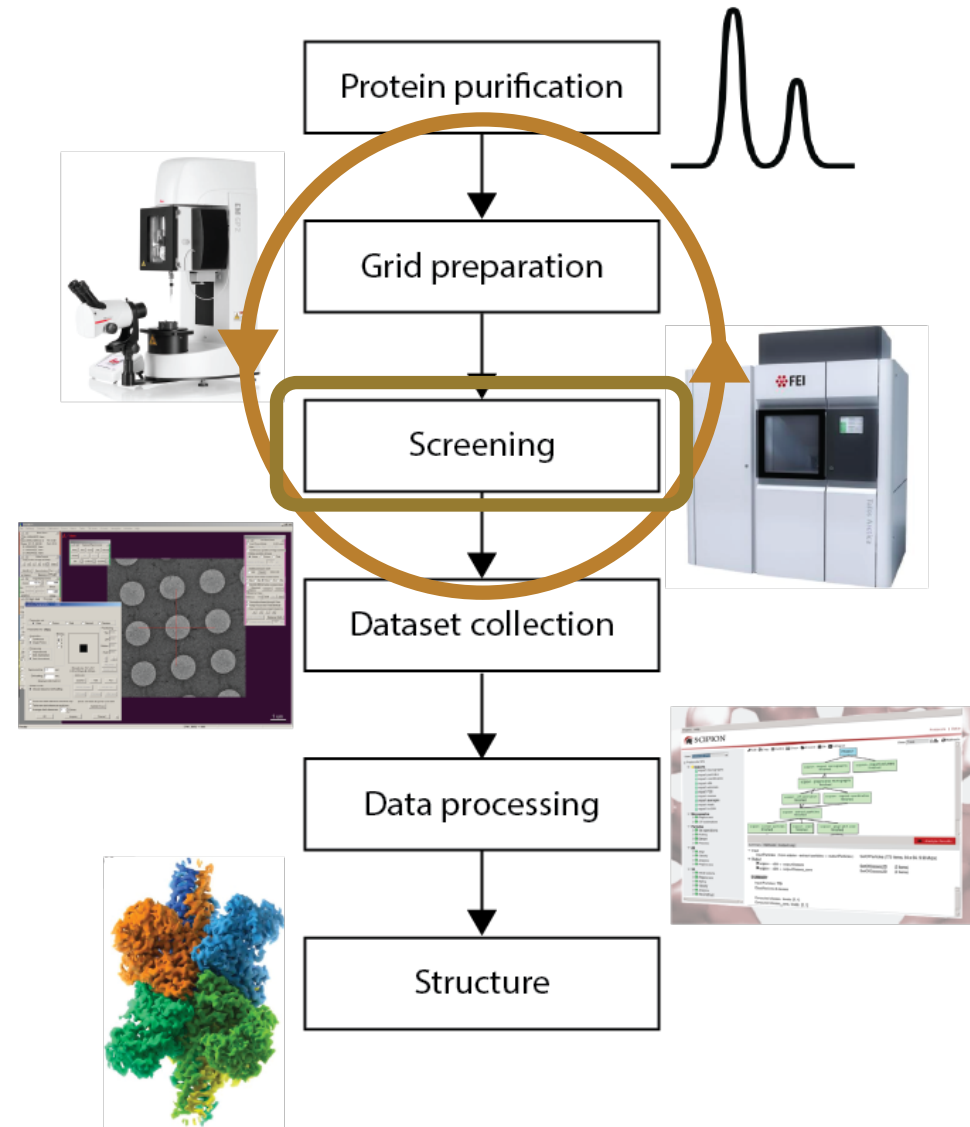
CryoEM workflow



CryoEM workflow

Sample optimization

- Multiple cycles are required to obtain a good sample
- Most projects require preparing and screening >100 grids
- Each grids take >30 min to screen



Grid Screening

Goal of a screening session

Learn as much as possible
about the specimen

Freezing conditions
Sample quality

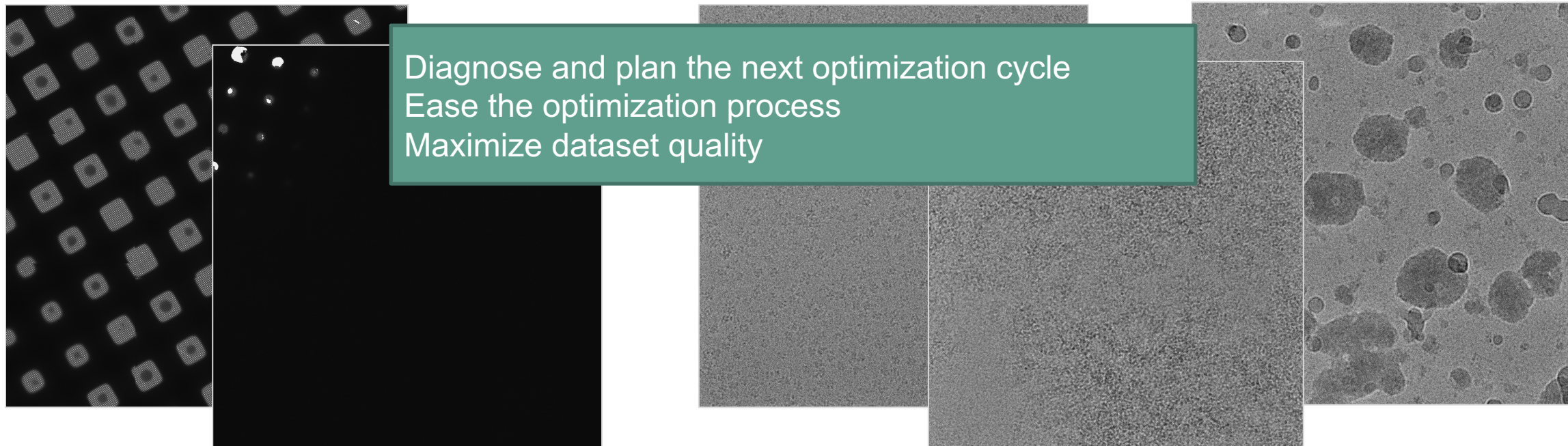
Thorough sampling

Different ice thickness
Find what is good
And what is bad

Good grid?

Where are the best areas?
Enough for a dataset?
Improvements?

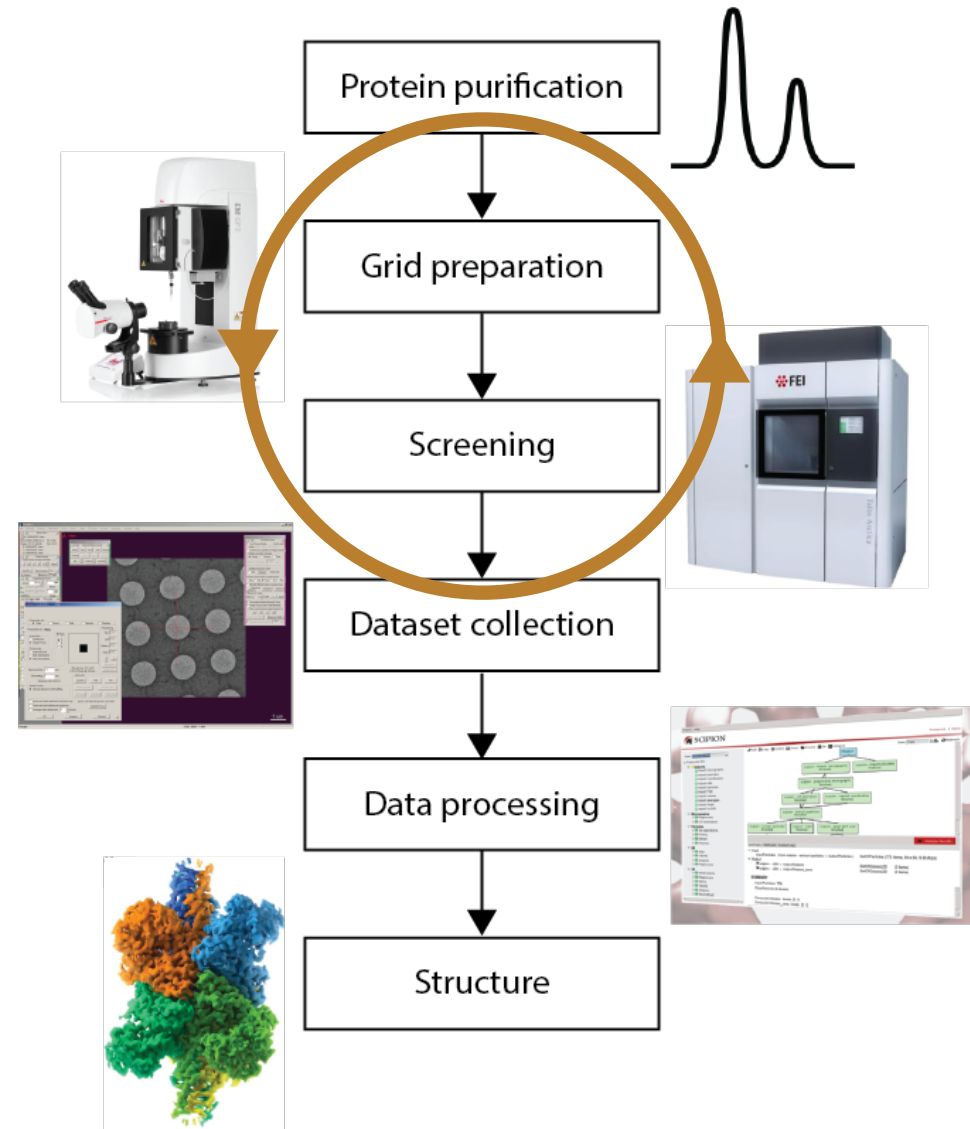
Diagnose and plan the next optimization cycle
Ease the optimization process
Maximize dataset quality



CryoEM workflow

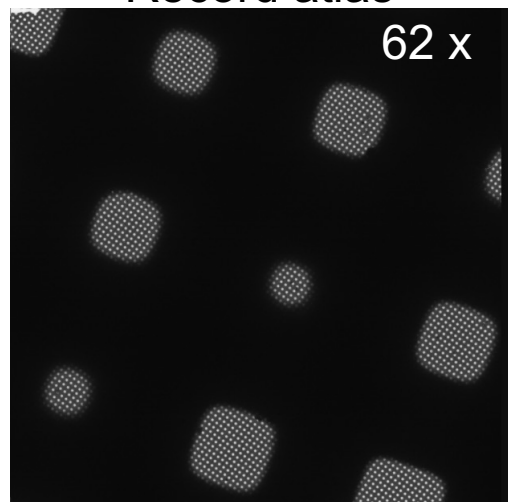
Weekly on the NIEHS Arctica

- 80-100 grids screened:
 - 30 hours of active screening
 - 10 hours of grid preparation
- ~4-7 grid collected:
 - 20 hours of active setup
 - 80 hours of collection

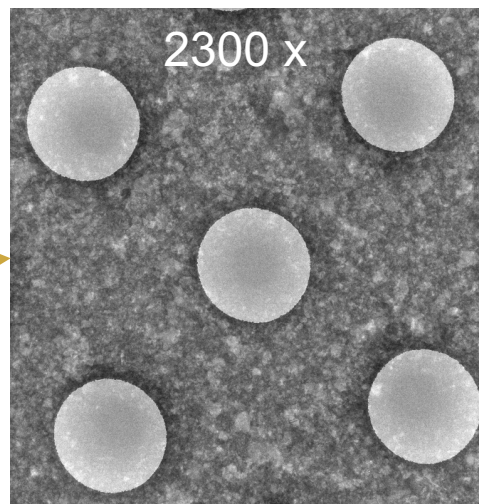


Grid Screening is repetitive

Record atlas

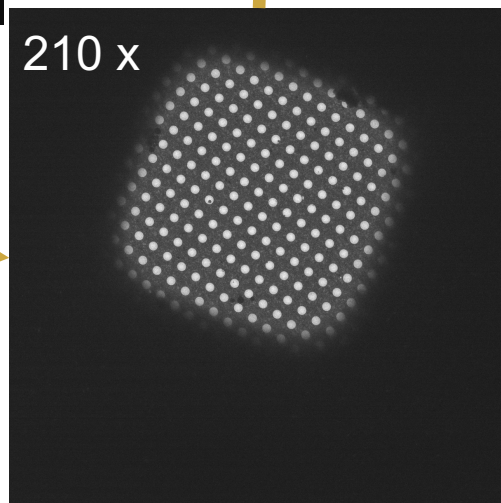


Save Image
Choose Area
Move stage

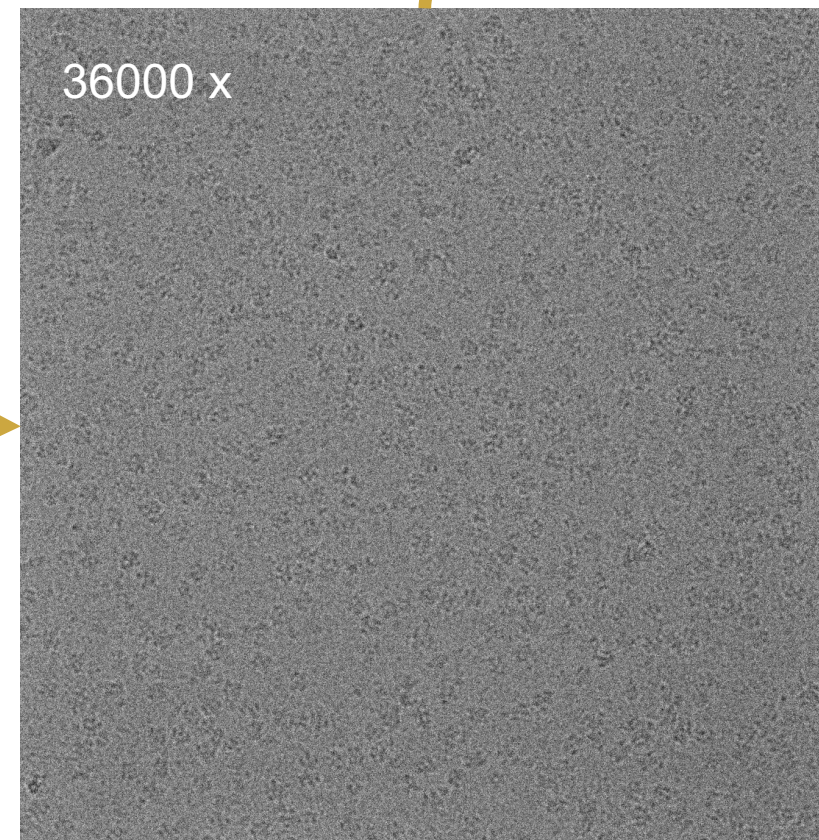


Save image
Star over

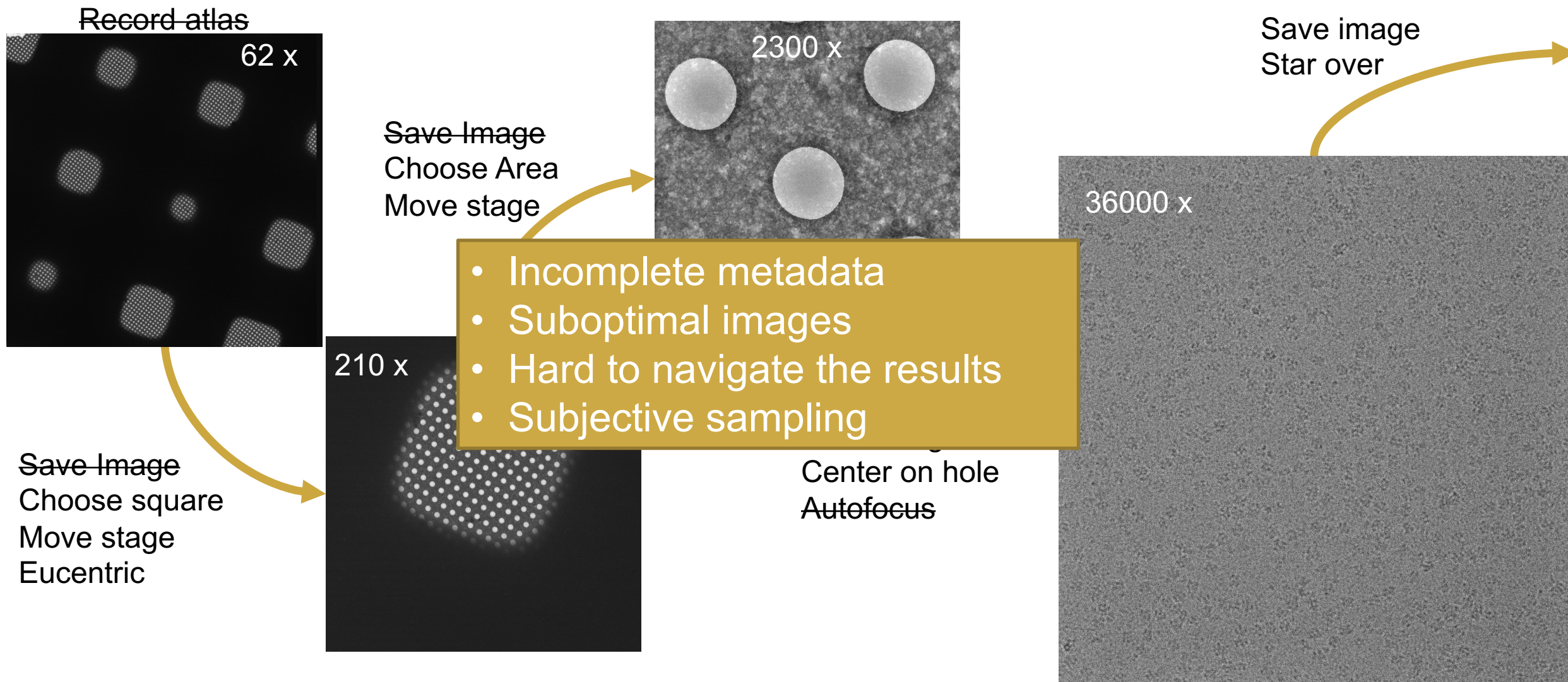
Save Image
Choose square
Move stage
Eucentric



Save image
Center on hole
Autofocus

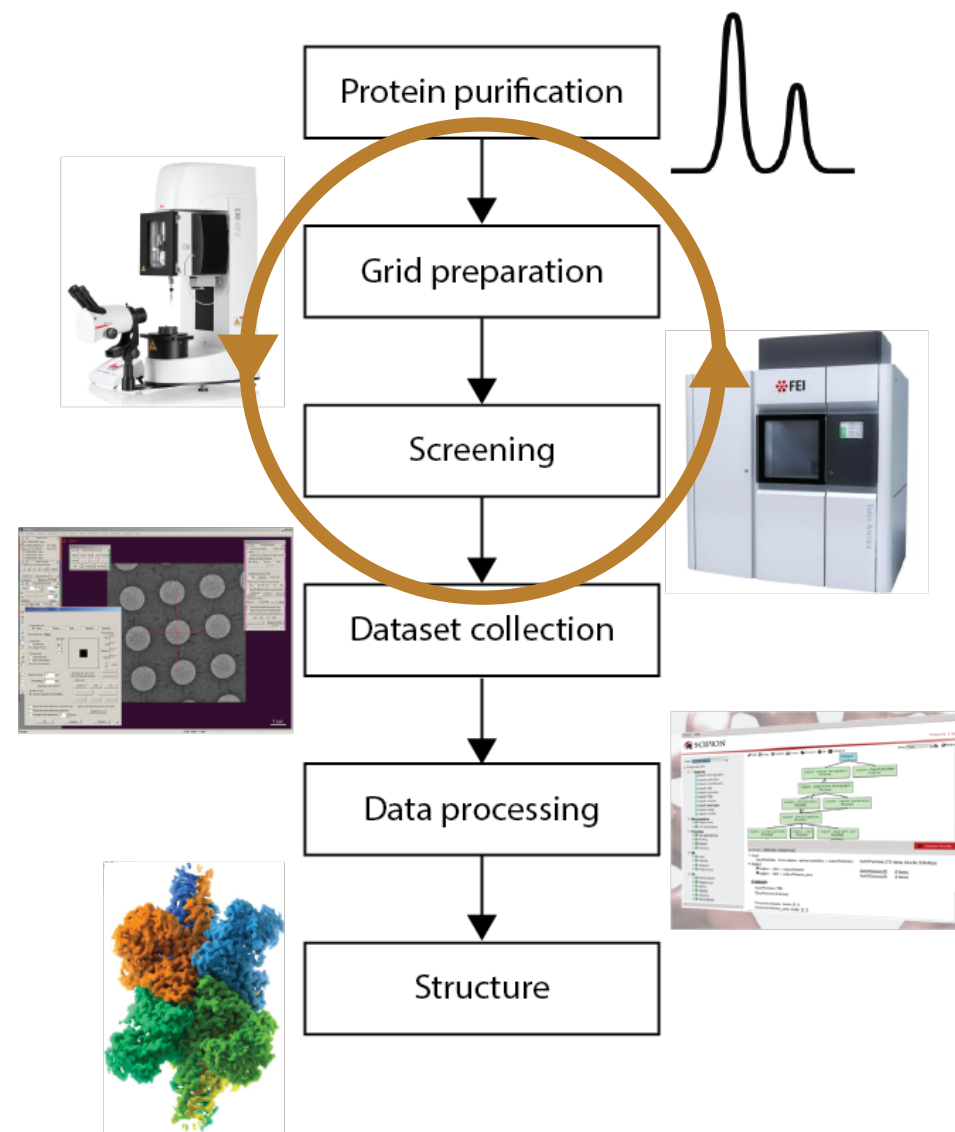


Manual grid screening – Cutting corners to speed up

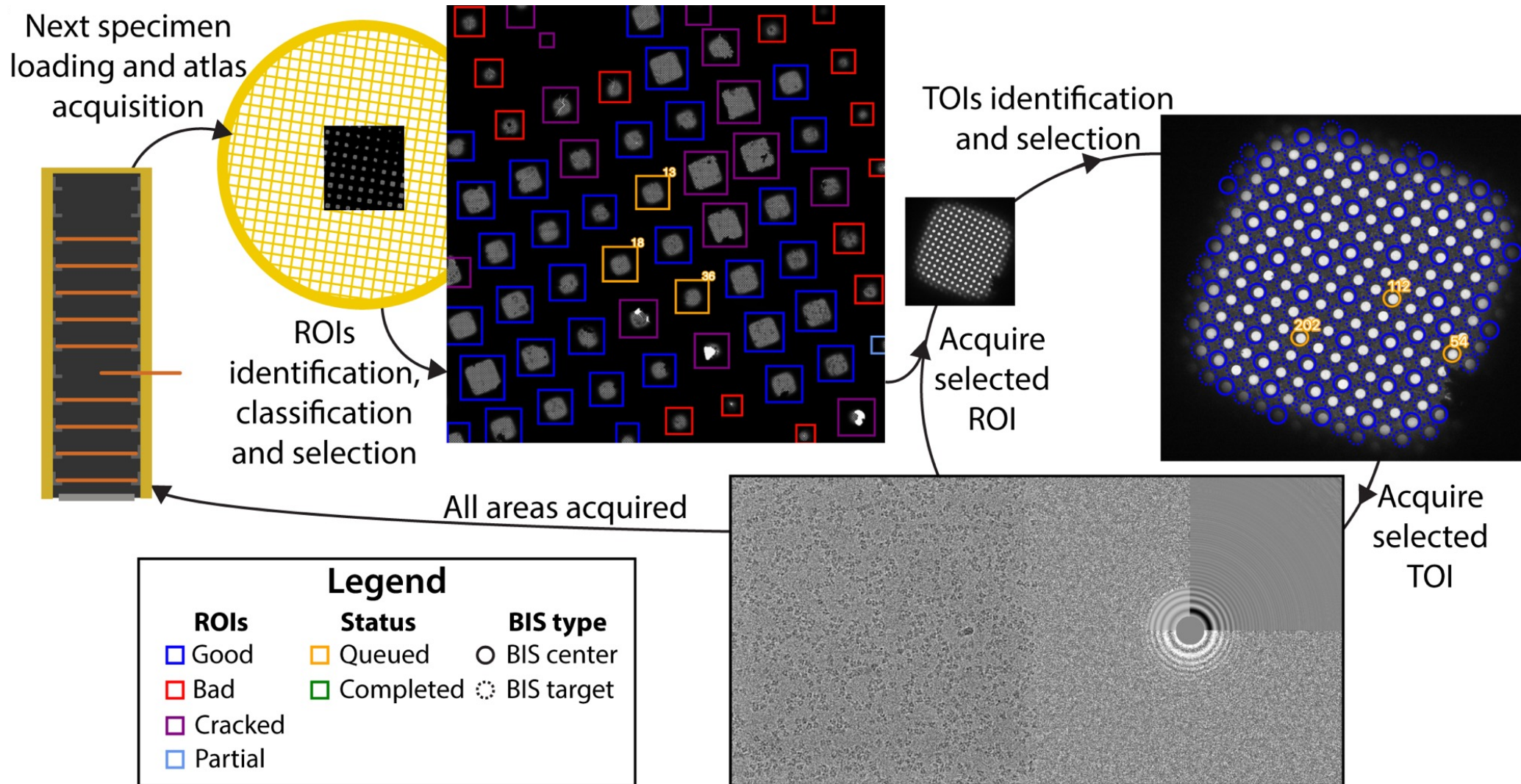


Goals

- Automate screening
- Provide good sampling
- Complete data
- Intuitive interface



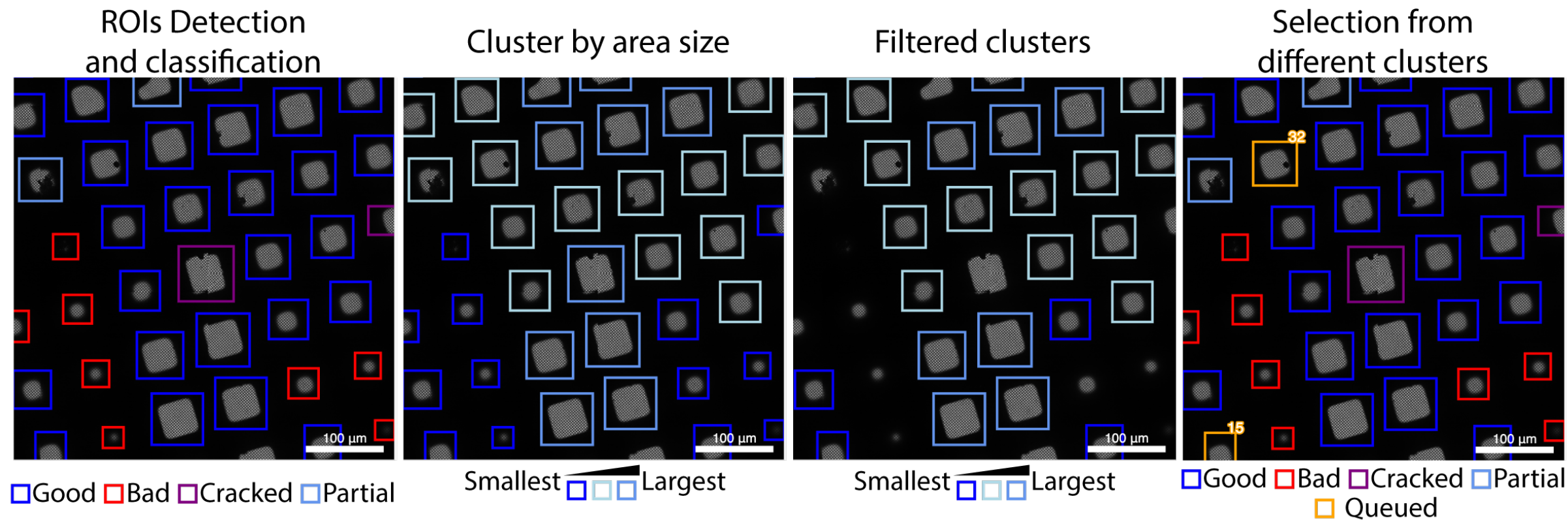
SmartScope – Automated workflow overview



SmartScope – Layered modular approach to area selection

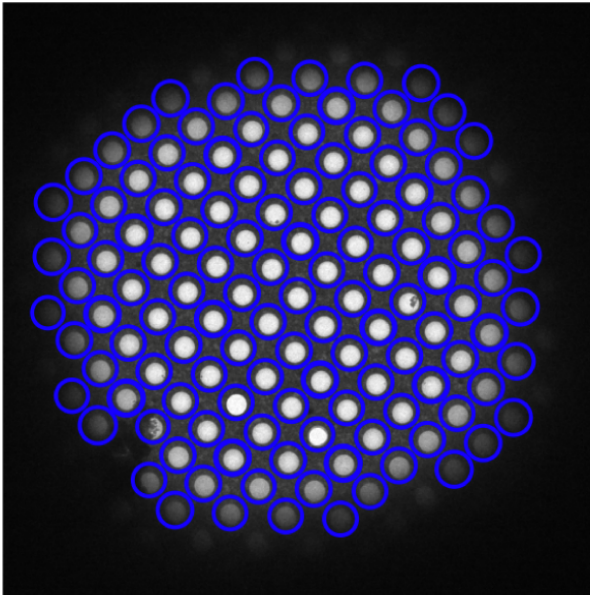
Finders	Classifiers	Selectors
Object detection	Named labels Finite number of categories	Clustering Tunable number of categories

SmartScope – Layered approach to area selection



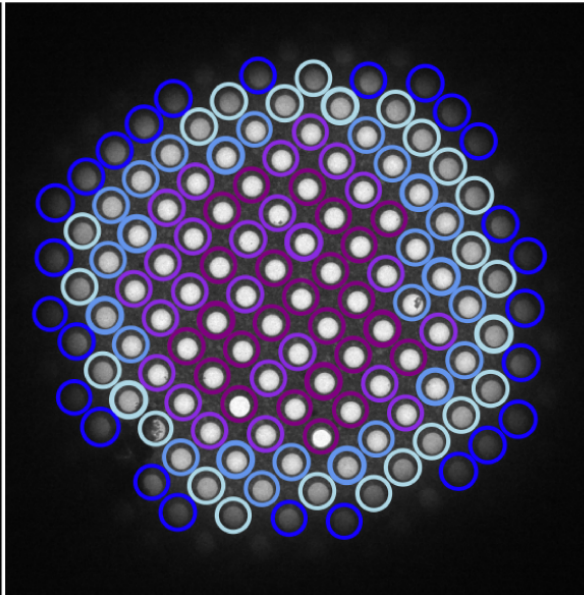
SmartScope – Layered approach to area selection

TOIs Detection



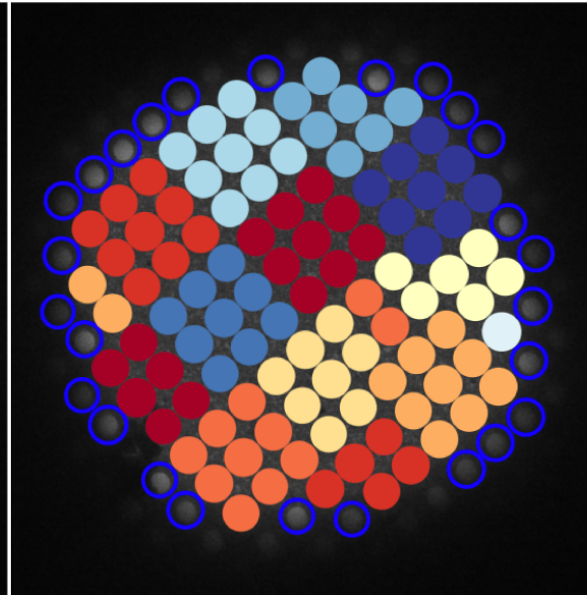
○ Target

Cluster by intensity

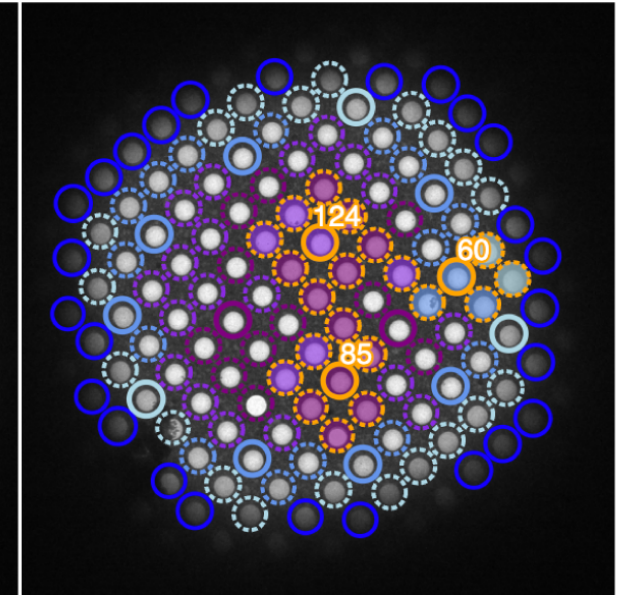


Darkest  Brightest

Group for BIS

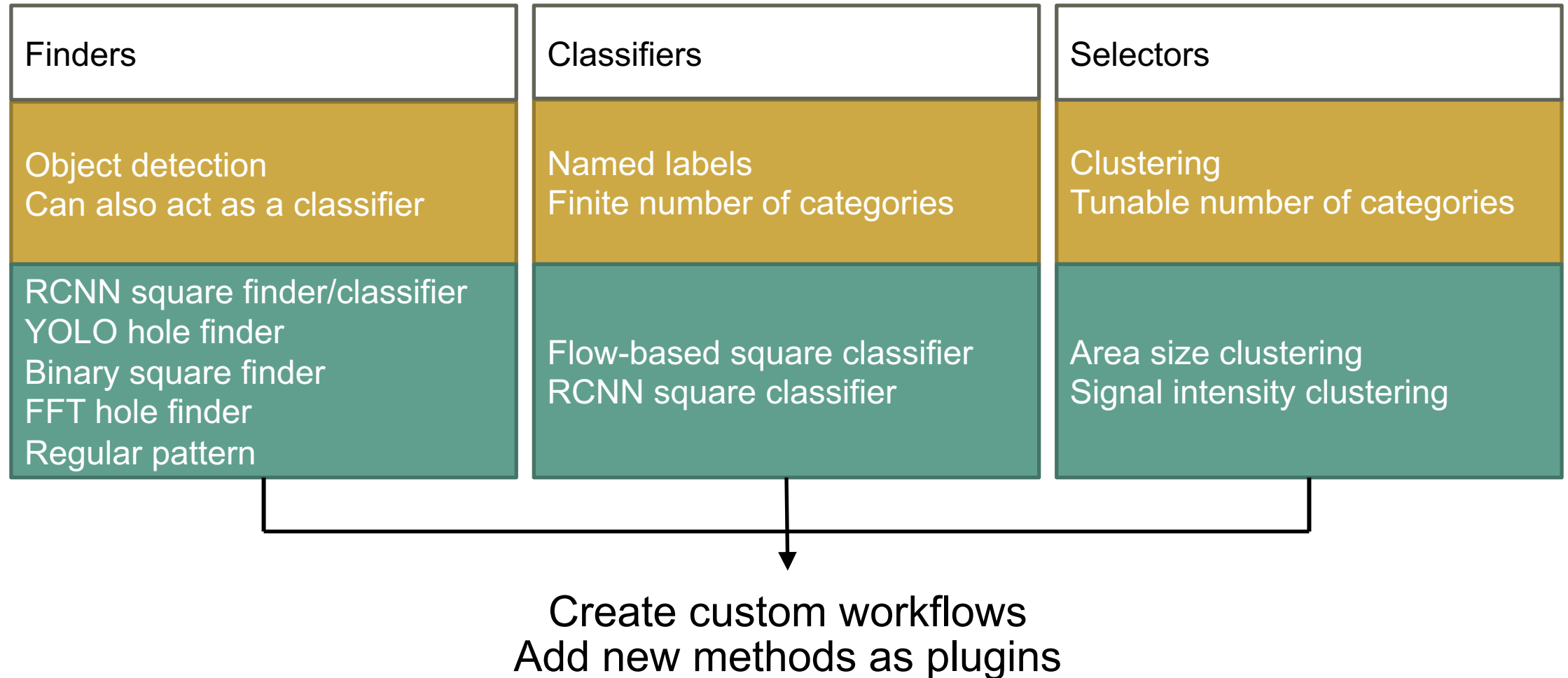


Selection from
different clusters



Darkest  Brightest
○ Queued

SmartScope – Layered modular approach to area selection



Web Interface

- Real-time tracking
- Microscope interaction

admin Session Browser Run Smartscope Change Log Logout

Screening: 1 grid1 [Go To](#)

Last update: 2022-03-29, 10:46:56
Status: complete [Stop](#) [Restart](#)

Quality: [Good](#) [Bad](#)

Notes: None [Save](#) [Show Stats](#) [+ Grid details](#) [Show legend](#)

GROUPS
testing-beta

SESSIONS
20220329_test2
20220328_testing

GRIDS
1_grid1
1_grid2

Atlas

Actions ▾

Legend
Good
Cracked
Small
Fractioned

Completion time: 2022-03-28, 15:01:11

Square 11

All Holes ▾ New targets ▾ Actions ▾

Legend
target

Completion time: 2022-03-28, 15:02:06

Hole

Rate ▾ Hole 121 [+](#) Rate ▾ Hole 158 [+](#) Rate ▾ Hole 147 [+](#)

Web Interface

- Real-time tracking
- Microscope interaction

admin Session Browser Run Smartscope Change Log Logout

Screening: 1 grid1 [Go To](#)

Last update: 2022-03-29, 10:46:56
Status: complete [Stop](#) [Restart](#)

Quality: [Good](#) [Bad](#)

Notes: None [Save](#) [Show Stats](#) [+ Grid details](#) [Show legend](#)

GROUPS
testing-beta

SESSIONS
20220329_test2
20220328_testing

GRIDS
1_grid1
1_grid2

Atlas

Legend
Cluster 2
Cluster 4
Cluster 0
Cluster 3
Cluster 1

Completion time: 2022-03-28, 15:01:11

Square 11

Legend
Cluster 2
Cluster 4
Cluster 0
Cluster 3
Cluster 1

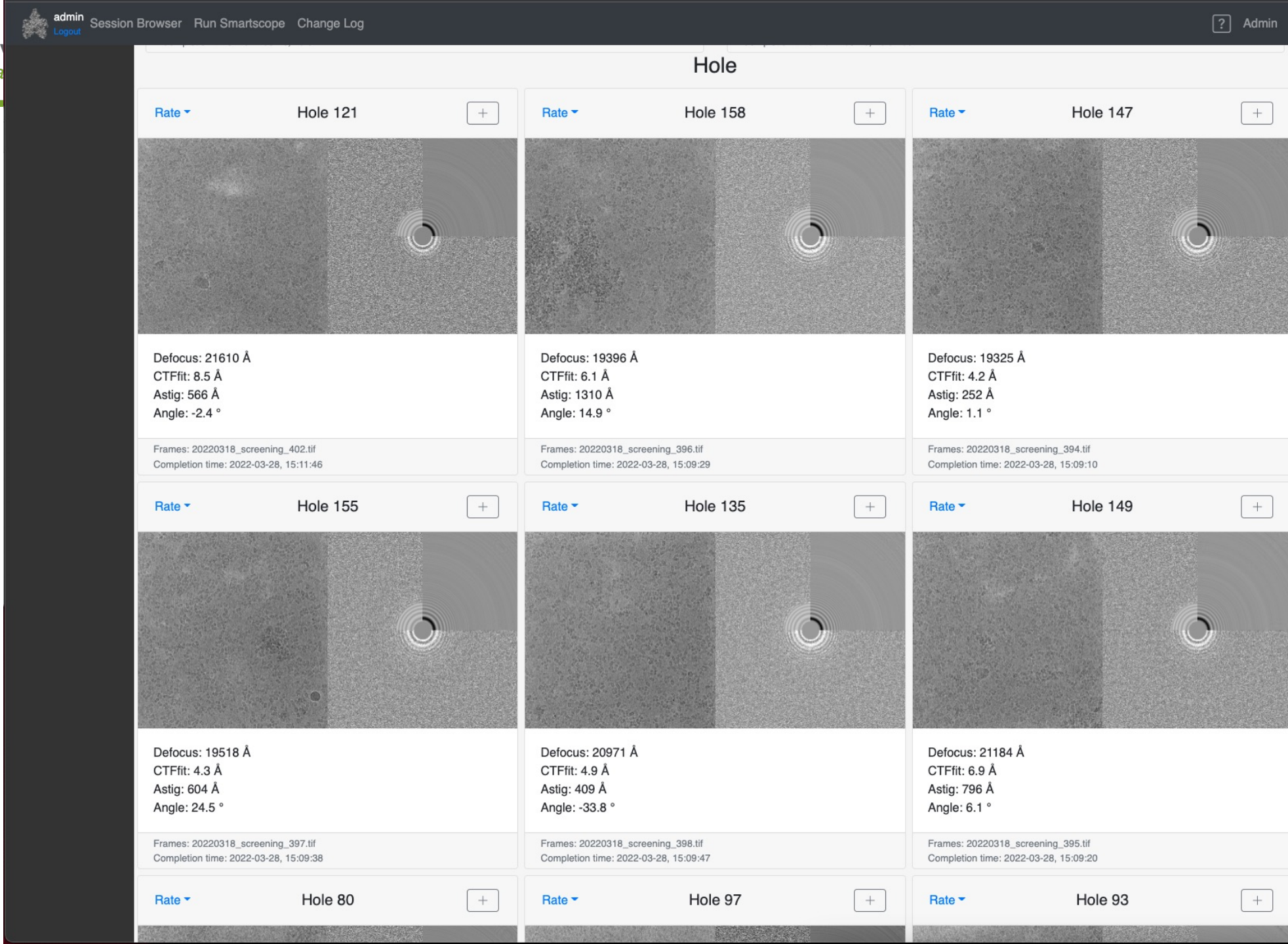
Completion time: 2022-03-28, 15:02:06

Hole

Rate ▾ Hole 121 [+](#) Rate ▾ Hole 158 [+](#) Rate ▾ Hole 147 [+](#)

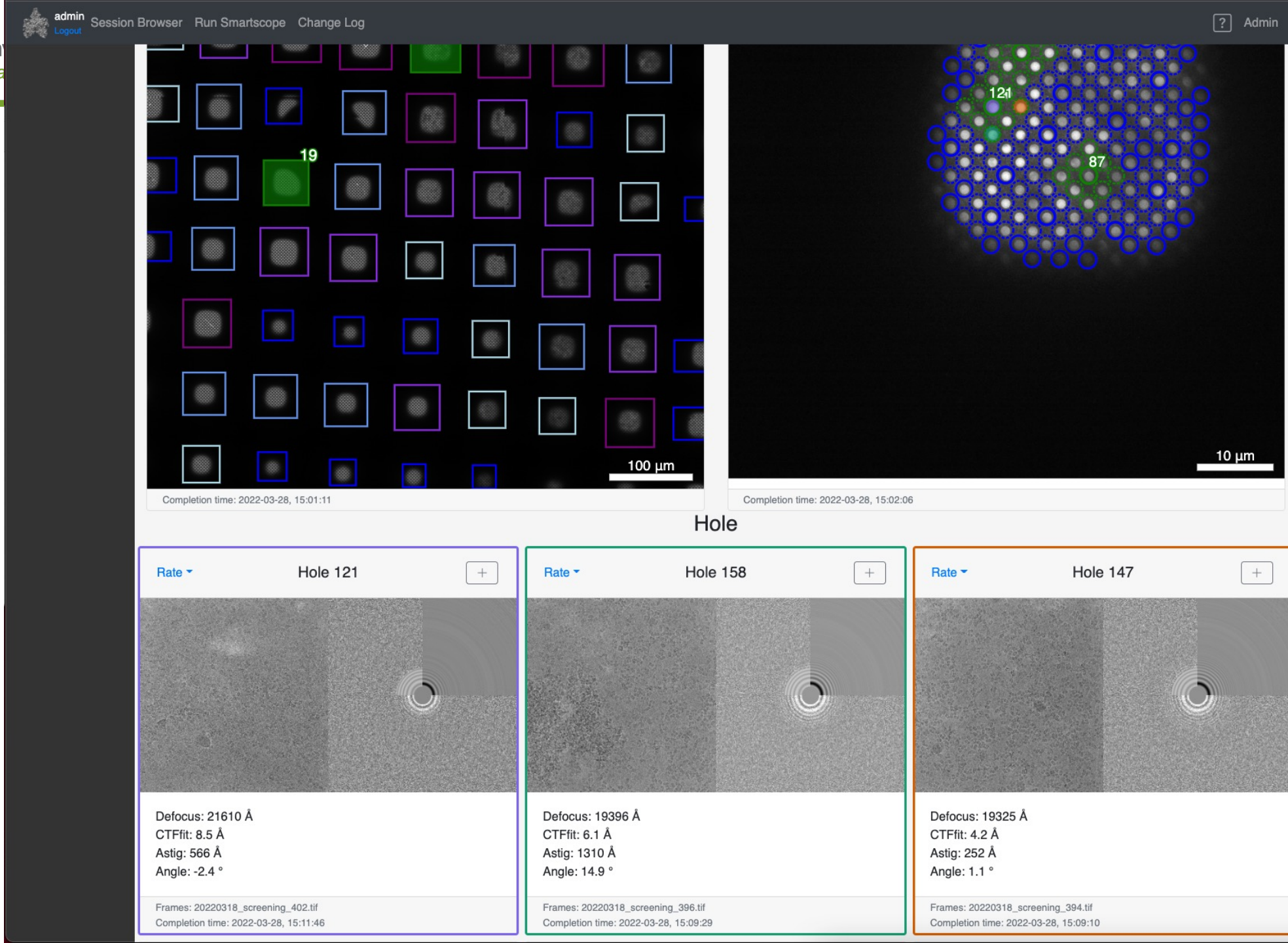
Web Interface

- Real-time tracking
- Microscope interaction
- Preprocessing



Web Interface

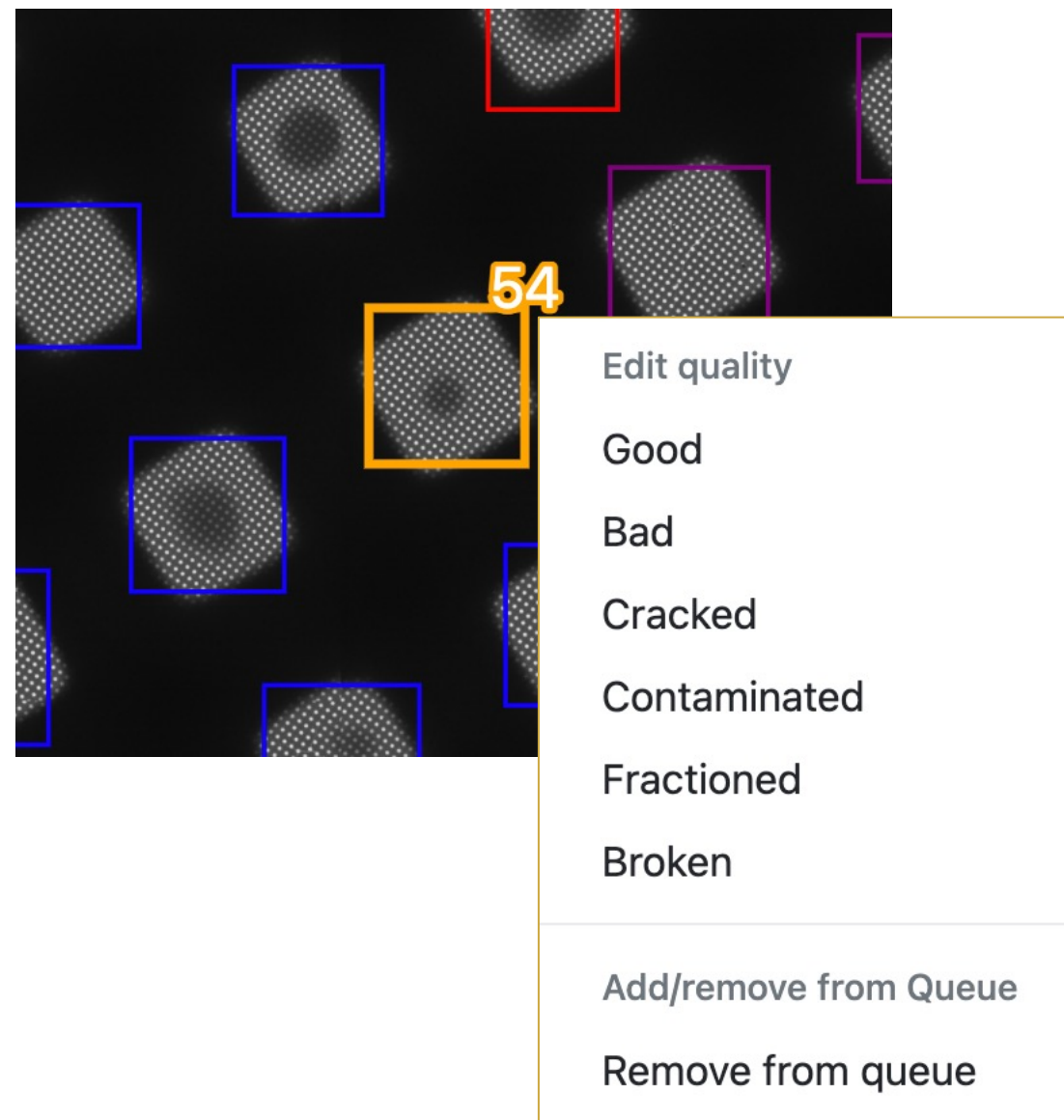
- Real-time tracking
- Microscope interaction
- Preprocessing
- Annotation



Supervised Automatic screening

Giving the users some freedom

- Change Label
- Modify selection
- Annotation
- Changing parameters
- Micrograph curation (still under work)



Automatic screening


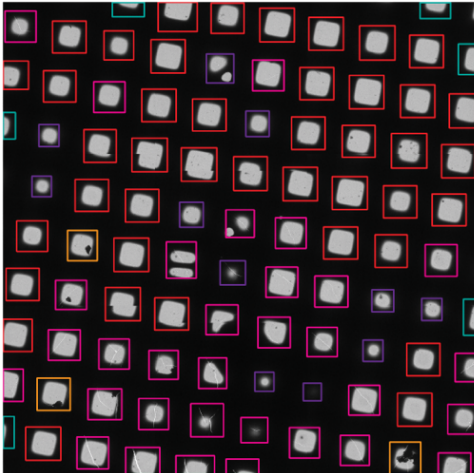
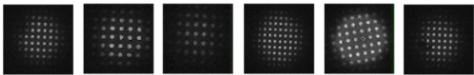
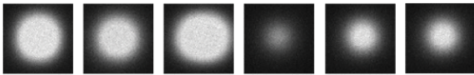

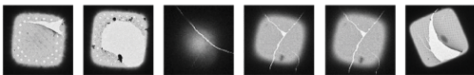
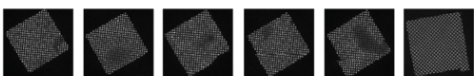
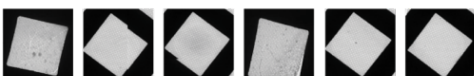
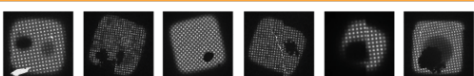
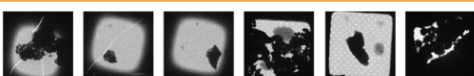


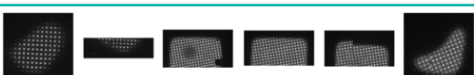

Leveraging early metadata

Faster R-CNN architecture

Identify and classify

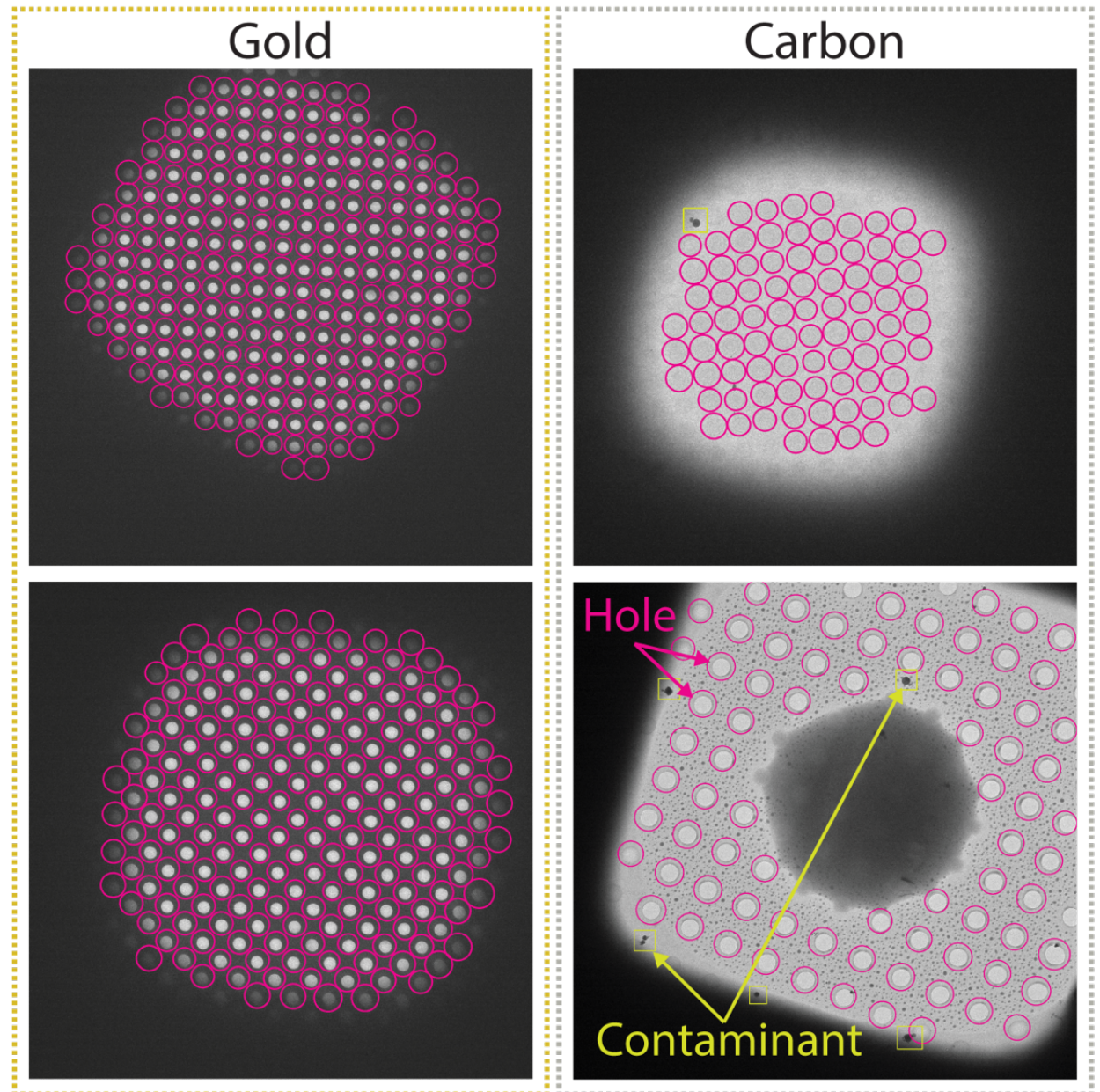
Training set:

~ 1500 labeled squares

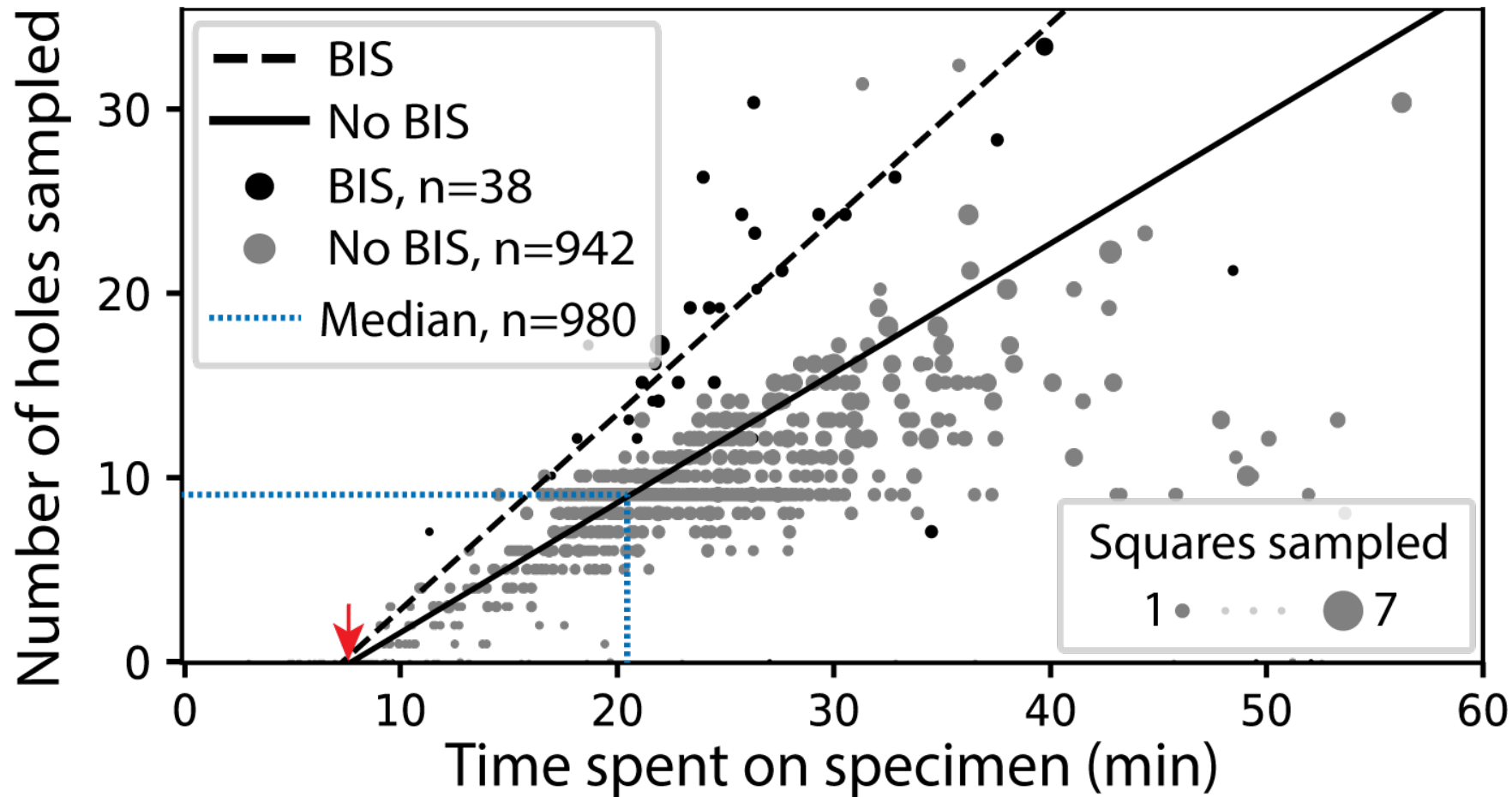
	Gold	Carbon	Detection Precision
Labels			
Small			73.4%
Cracked			76.5%
Dry			79.4%
Contaminated			50.4%
Good			81.2%
Partial			45.5%

Hole Finder

- YOLO-based architecture
- AI hole finder is being trained to find holes on multiple grid types.
- Currently 10 000 holes in the training set.
- Precision of 98%, 89% recall
 - Mean-average precision 87%



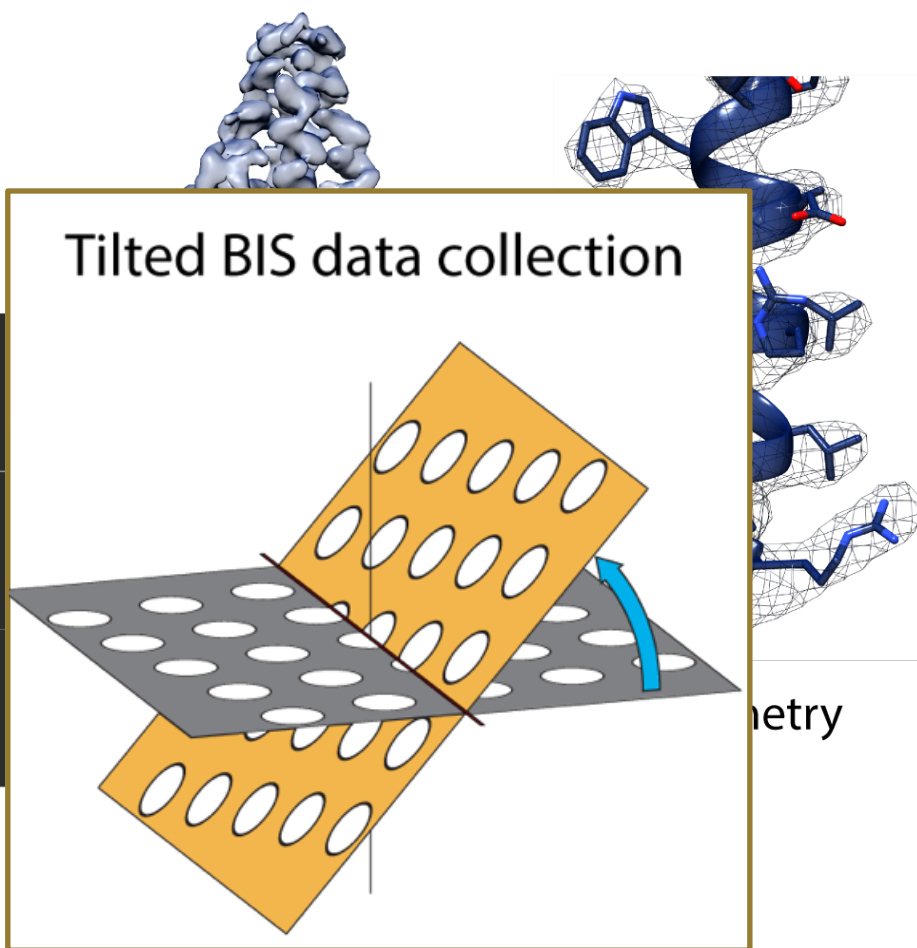
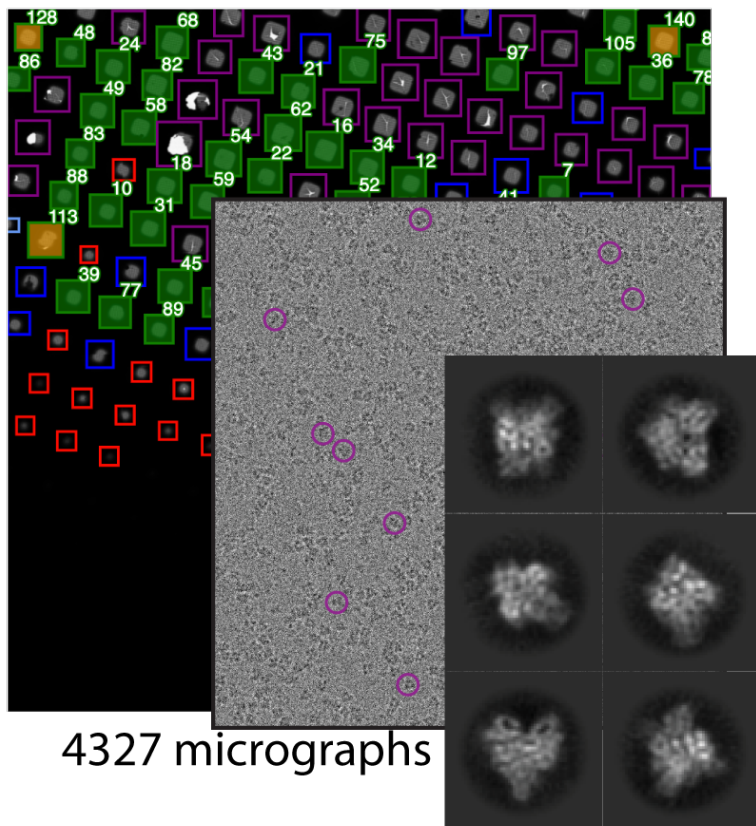
Screening statistics



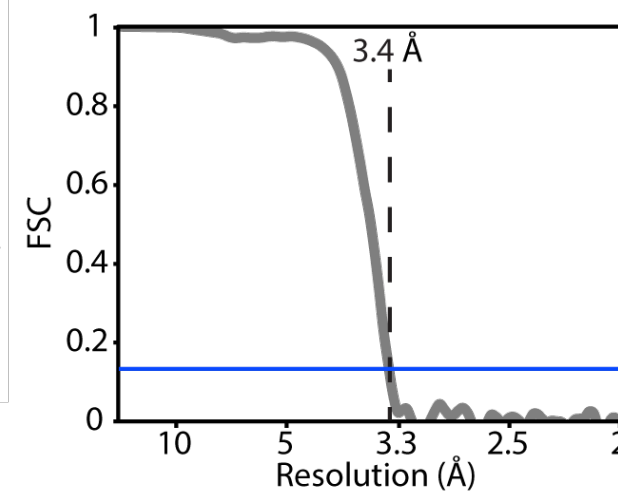
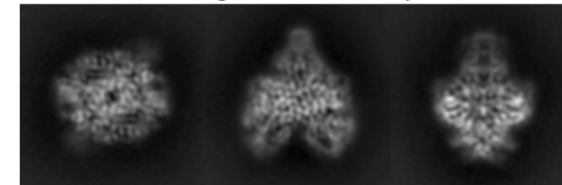
Talos Arctica
K2 detector

Automatic data collection

Quick setup and high-resolution capabilities



3D orthogonal projections



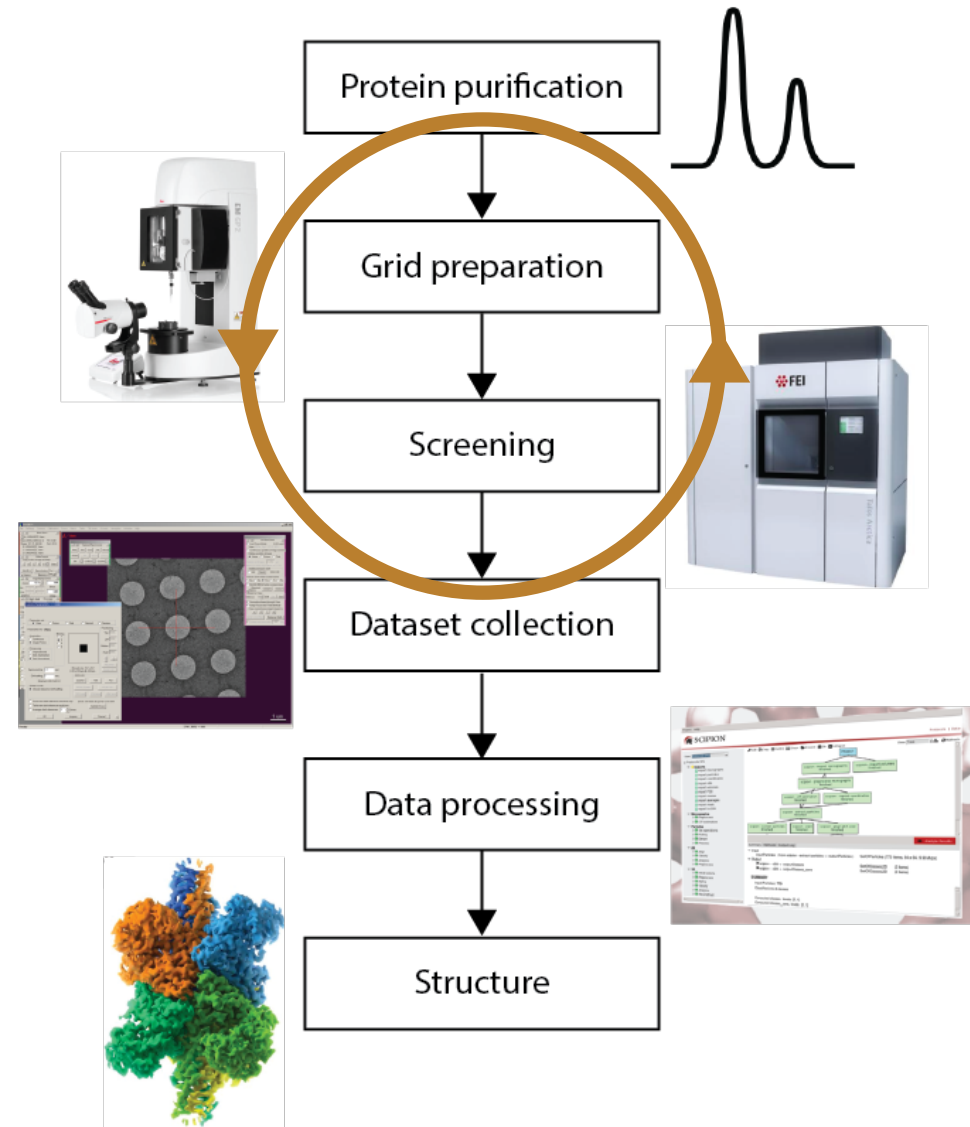
Conclusions

- Automated screening procedure
 - Square finder and classifier
 - Hole finder
 - Clustering methods
- Interactive interface
 - Ability to choose and modify area selection
 - Easy result access and complete bookkeeping
- Data persistence and organization
- Fast data collection setup
- Overnight screening sessions

CryoEM workflow

Weekly at the NIEHS Arctica

- >120 ~~80-100~~ grids screened:
 - ~~30 hours of active screening~~
 - Lightly supervised automatic screening
 - 10 hours of grid preparation
- ~4-7 grid collected:
 - <10 ~~20~~ hours of active setup
 - >90 ~~80~~ hours of collection



Short term goals – More Flexibility with modular protocols

Protocol recipe

Magnification level	Acquisition method
	Finder (1)
	Classifier (0 or more)
	Selectors (1 or more)



High magnification	Acquisition method
	Frames
	Preprocessing

- Allow easy addition of Finders, Classifiers, Selectors as external plugins.
- Add acquisition methods to the microscope interface also as plugins.
- Create protocols by mixing existing methods.

Ease the integration of new workflows
Sample variety: virions, filaments, cells
Tomography

Sample-specific navigation roadmap

1. Sample specific state selection
2. User annotation to drive the selection on-the-fly
3. Using preprocessing information as feedback to drive the selection
4. Train AI models to drive the selection and “learn” about the samples

Acknowledgements

Bartesaghi Lab:



Dr. Alberto Bartesaghi

Wendy Qinwen Huang



Funding: NIEHS/NIH

Copeland Lab:

Dr. William Copeland

Dr. Amanda Riccio

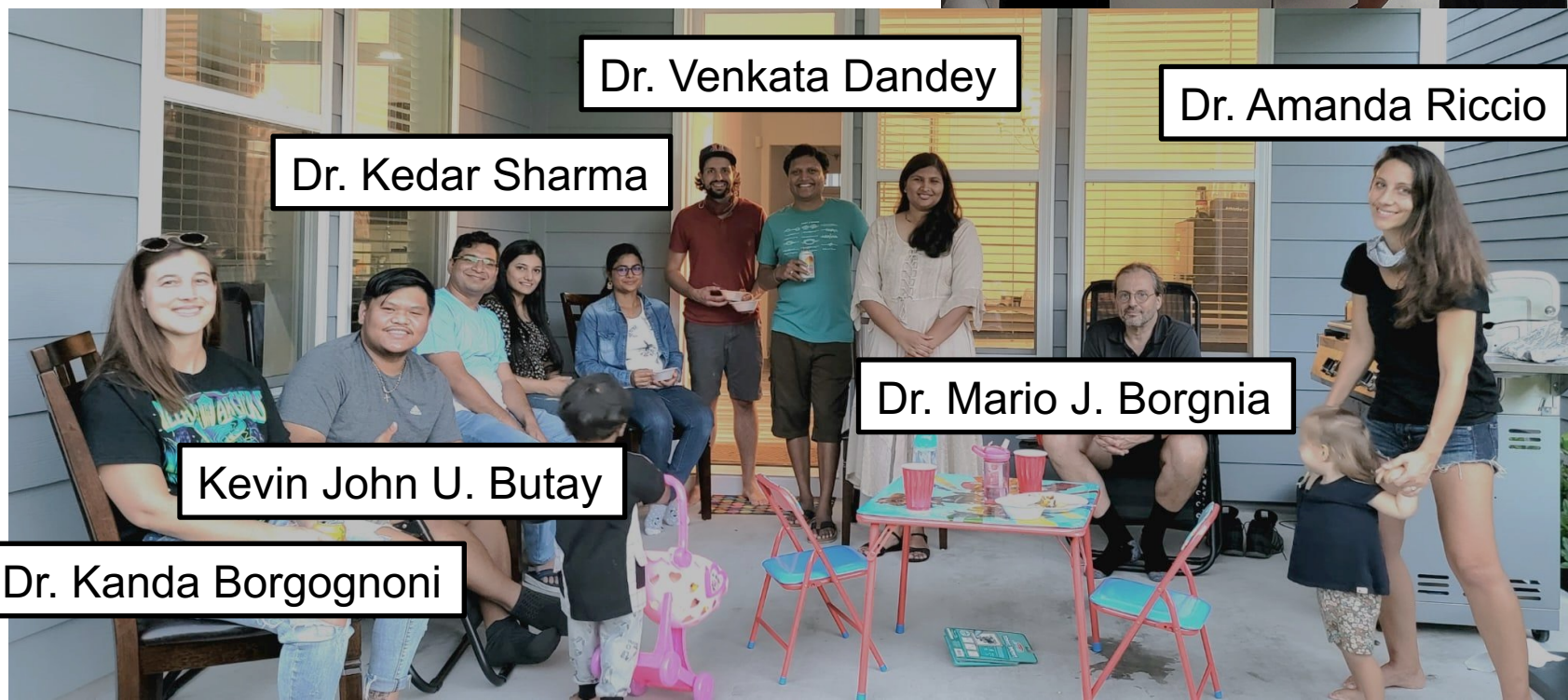
UNC CryoEM:

Dr. Joshua Strauss

Jared Peck



Elizabeth Viverette



Dr. Venkata Dandey

Dr. Amanda Riccio

Dr. Kedar Sharma

Dr. Mario J. Borgnia

Kevin John U. Butay

Dr. Kanda Borgognoni