

SuperResNET: Learning in-cell macromolecular architecture from SMLM data

Ismail M. Khater, Y. Lydia Li, Timothy H. Wong, Ghassan Hamarneh*, *Ivan R. Nabi**

*equal contribution

Protein structures are now being resolved at the atomic level, but deciphering their molecular organization in the cell remains a challenge. SuperResNET is an integrated machine learning-based analysis software for visualizing and quantifying 3D point cloud data acquired by single molecule localization microscopy (SMLM). The computational modules of SuperResNET include correction for multiple blinking of a single fluorophore, denoising, segmentation (clustering), and feature extraction, which are then used for cluster group identification, modularity analysis, blob retrieval and visualization in 2D and 3D. More recent updates to SuperResNET allow two-channel interaction distance analysis to determine how two proteins interact within macromolecular assemblies. SuperResNET can be effectively and easily applied to any SMLM event list from which it rapidly learns macromolecular architecture in the intact cell. SuperResNET makes super-resolution microscopy accessible to biologists, with a GUI version for analysis of individual data sets and a batch analysis version for statistical analysis of multiple replicates and conditions. SuperResNET represents an analysis tool to discover protein architecture and structural diversity in the cell.

SuperResNET: Learning in-cell macromolecular architecture from SMLM data



Ismail M. Khater¹, Y. Lydia Li², Timothy H. Wong², Shane Eastwood³, Randy Panopio³, Ghassan Hamarneh^{3*}, Ivan R. Nabi^{2*}

*equal contribution; ¹Computer Engineering, Birzeit University, Ramallah, Palestine; ²Life Sciences Institute, University of British Columbia, Vancouver, BC, Canada; ³School of Computing Science, Simon Fraser University, Vancouver, BC, Canada

Protein structures are now being resolved at the atomic level, but deciphering their molecular organization in the cell remains a challenge. SuperResNET is an integrated machine learning-based analysis software for visualizing and quantifying 3D point cloud data acquired by single molecule localization microscopy (SMLM). The computational modules of SuperResNET include correction for multiple blinking of a single fluorophore, denoising, segmentation (clustering), and feature extraction, which are then used for cluster group identification, modularity analysis, blob retrieval and visualization in 2D and 3D. More recent updates to SuperResNET allow two-channel interaction distance analysis to determine how two proteins interact within macromolecular assemblies. SuperResNET can be effectively and easily applied to any SMLM event list from which it rapidly learns macromolecular architecture in the intact cell. SuperResNET makes super-resolution microscopy accessible to biologists, with a GUI version for analysis of individual data sets and a batch analysis version for statistical analysis of multiple replicates and conditions.

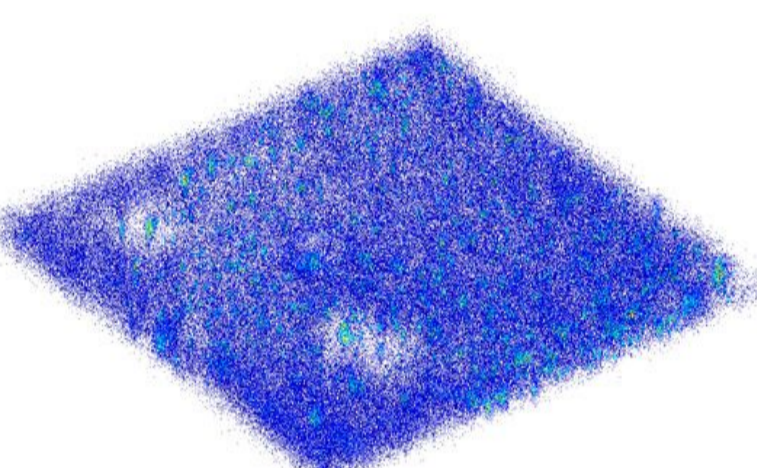
SuperResNET is an analysis tool to discover protein architecture and structural diversity in the cell

SuperResNET Modules

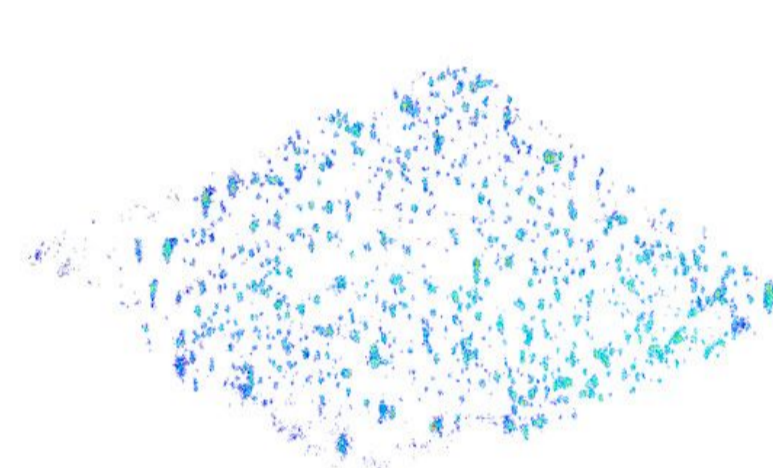
Data Analysis

X	Y	Z
...
...
...
...
...
...
...
...
...
...

Load data
loads point cloud data and provides histogram visualization of localization and the associated metadata



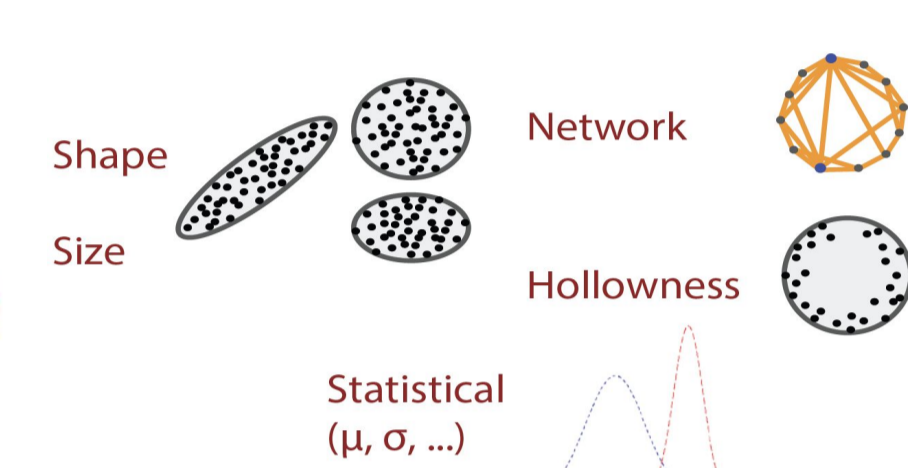
Merge & network analysis
removes artifacts from multiple blinks by iteratively merging localizations within a merge threshold, and assesses scale of clustering by Ripley's H-function



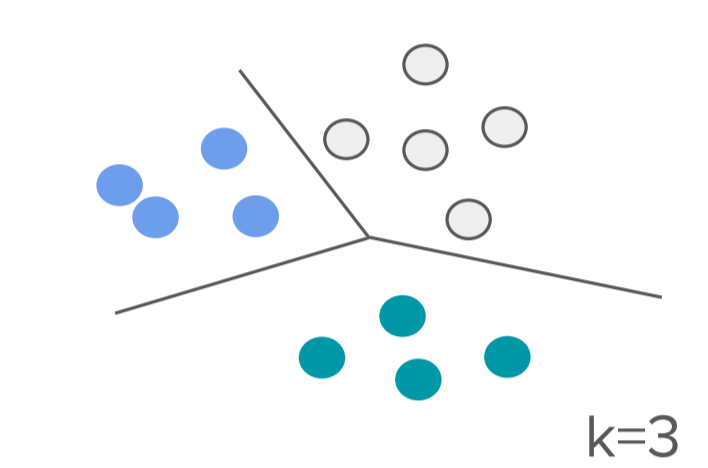
Filter
denoises data based on comparison with degree distribution of a random network



Segment
provides Mean-shift segmentation for blob-like structures and DBSCAN to segment structures with other shapes

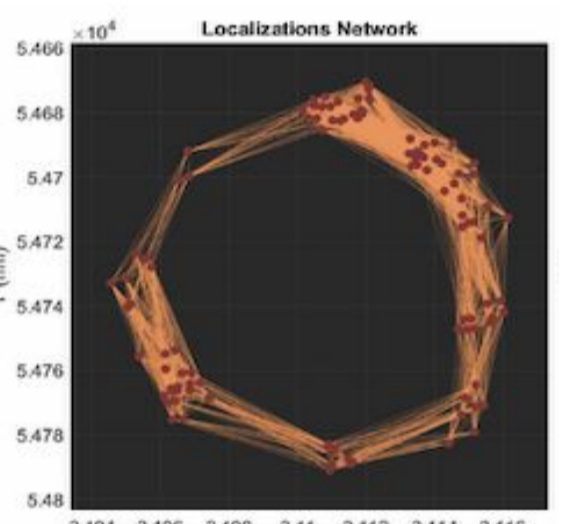


Blob features
extraction, quantification and visualization of 30 features comprising size, shape, topology/hollowness, statistical and network features of segmented blobs and reports histograms of the features

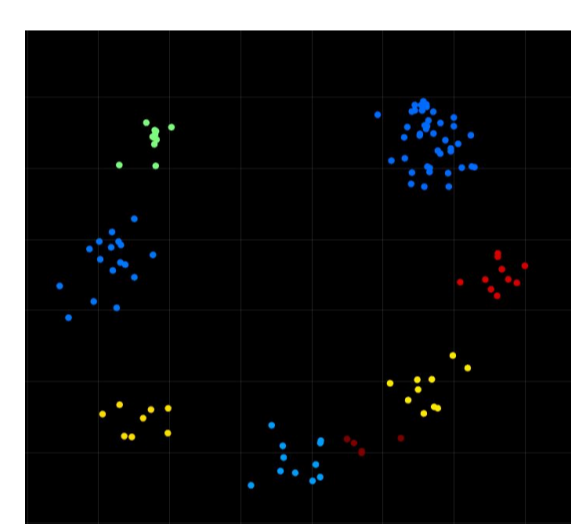


Group
assigns group to segmented blobs based on selected features through K-means clustering, reports color-coded histograms of grouped features, assesses grouping using t-SNE or silhouette methods

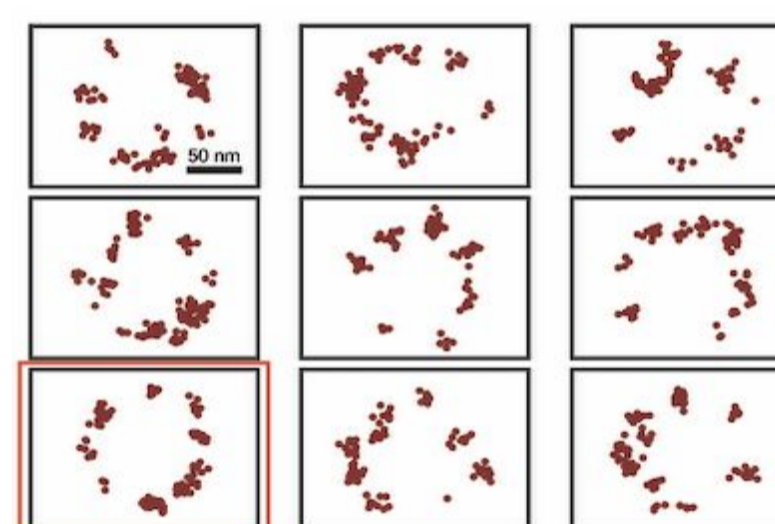
Data Output & Visualization



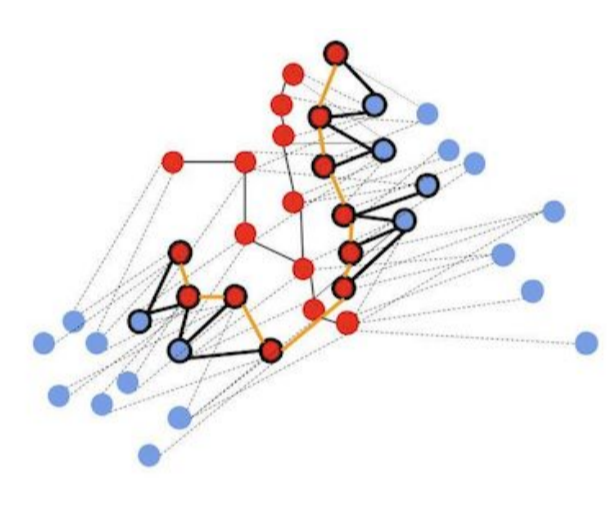
Individual blobs
visualizes individual blobs with network connection at user-defined proximity thresholds



Blob modules
extracts and visualizes modules of interacting molecules at user-defined proximity thresholds for individual blobs



Blob retrieval
retrieves representative blobs from each group with visualization of localization, network, and boundaries in 2D and 3D

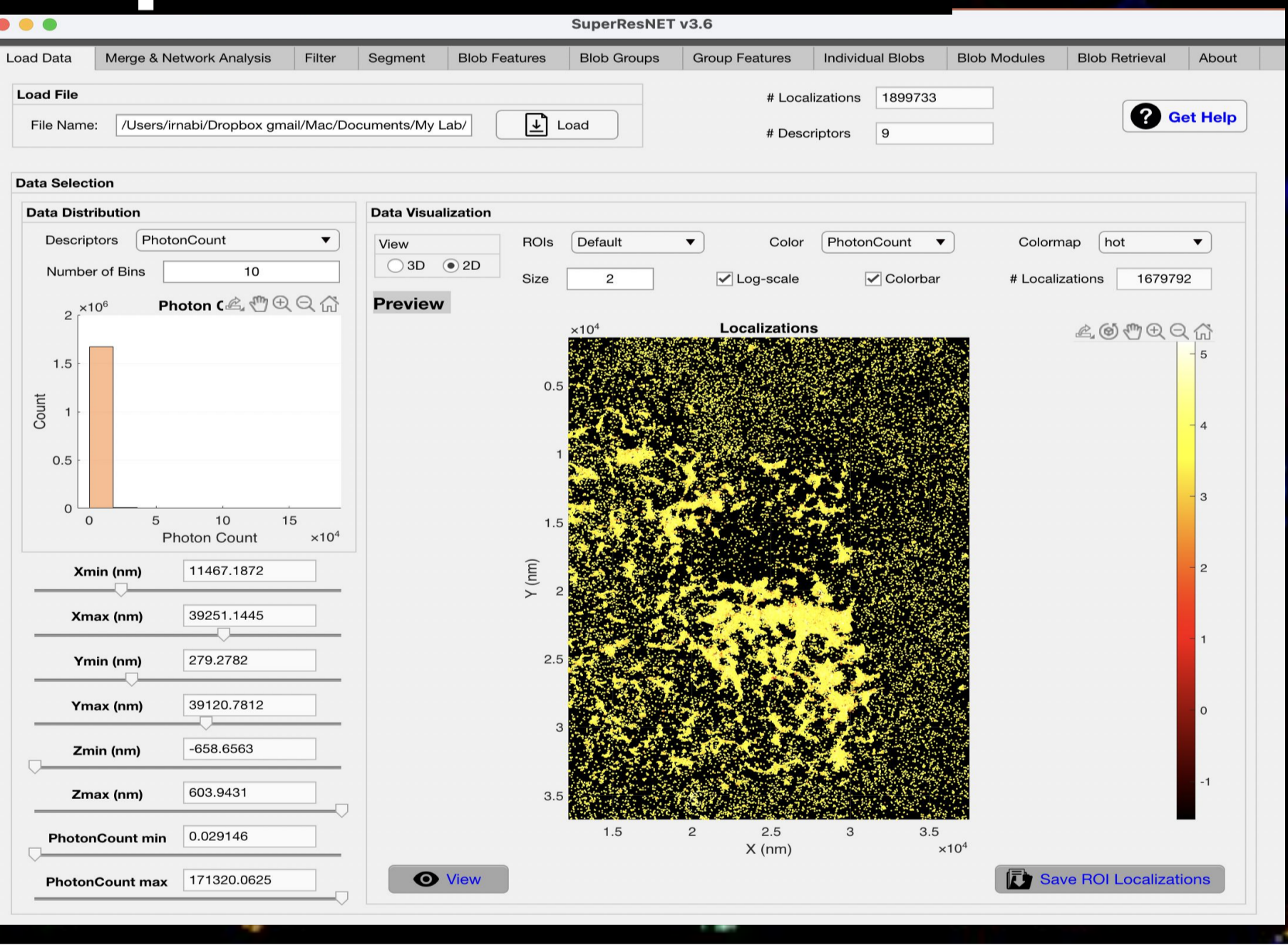


Blob interaction
loads multiple data sets and identifies localizations in close proximity within interacting blobs

Group ID	Num. Localization	Avg. Dist. to Centroid	Char. path	Avg. Clust. Coef.	PA	CL	CP	CS	Volume	Avg. Dist.
1.0000	12.0000	7.3475	7.3921	7.8922	0.9051	0.9141	0.9350	0	0	0
2.0000	12.0000	98.8726								

Quantitative output for statistical analysis
quantification of cluster features, grouping based on cluster features, network and modularity analysis of individual clusters

SuperResNET GUI



SuperResNET Features

- GUI and non-GUI batch analysis versions
- Imports file formats: .bin, .ascii, .xyz, .txt, .mat, .csv
- Exports high-resolution figures (tiff, png, pdf, eps...) and quantitative data (for Excel, R, Matlab, Python...)
- Load and easily switch between multiple datasets

SuperResNET-specific analysis methods

- Alpha filtering of random network-like blinks
- Merge analysis to correct for multiple blinking
- 30 features (Size, Shape, Topology, Network)
- Feature selection and normalization
- Convex hull analysis
- Modularity analysis
- Dual-channel interaction distance analysis

Established Analysis Methods

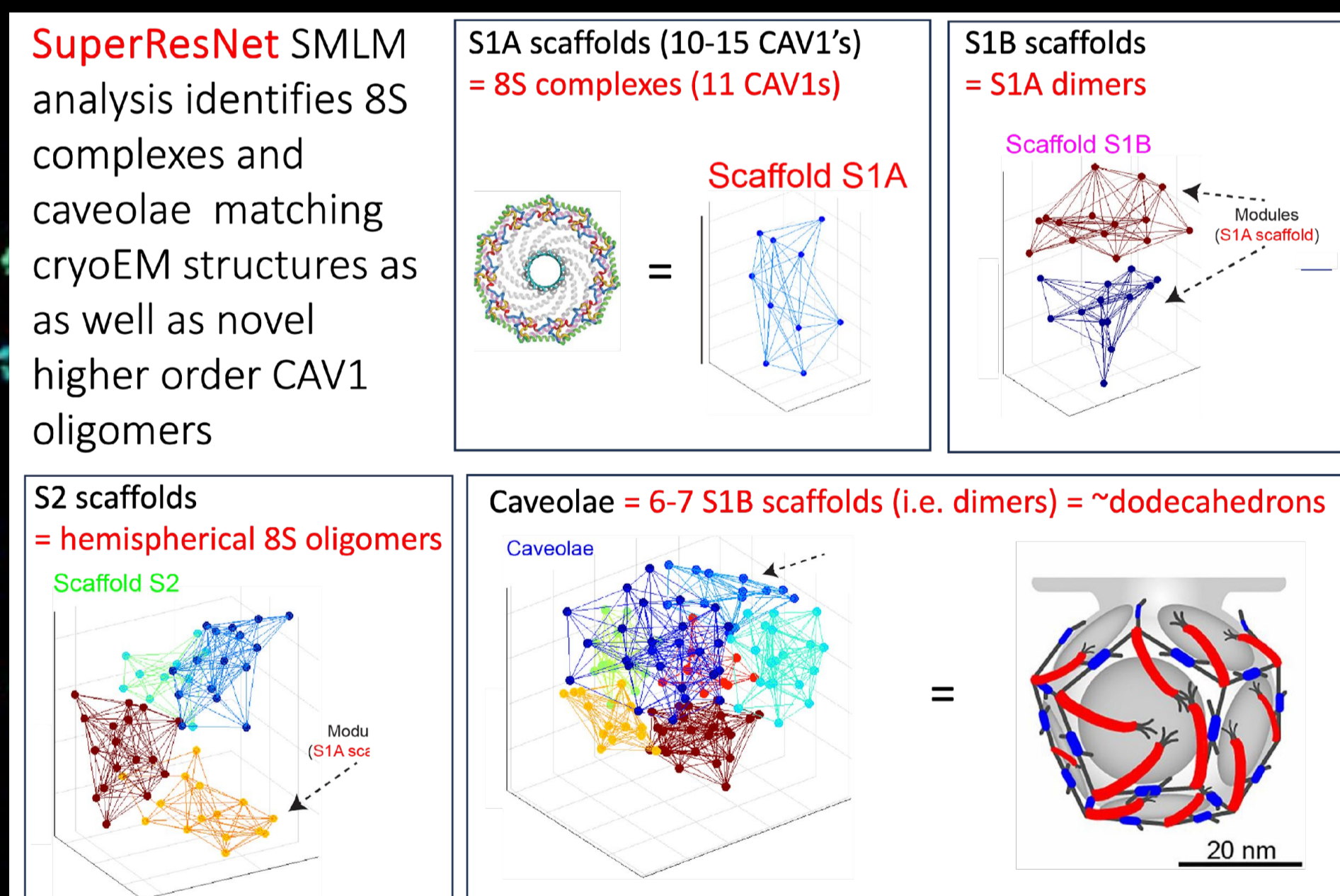
- Ripley's H function
- K-means and DBSCAN clustering
- Mean shift segmentation
- Network analysis

Visualization

- 2D and 3D point clouds
- Networks
- Pairwise feature visualization
- Convex hull
- Retrieval of most representative blobs
- Identification of blob communities (modularity)

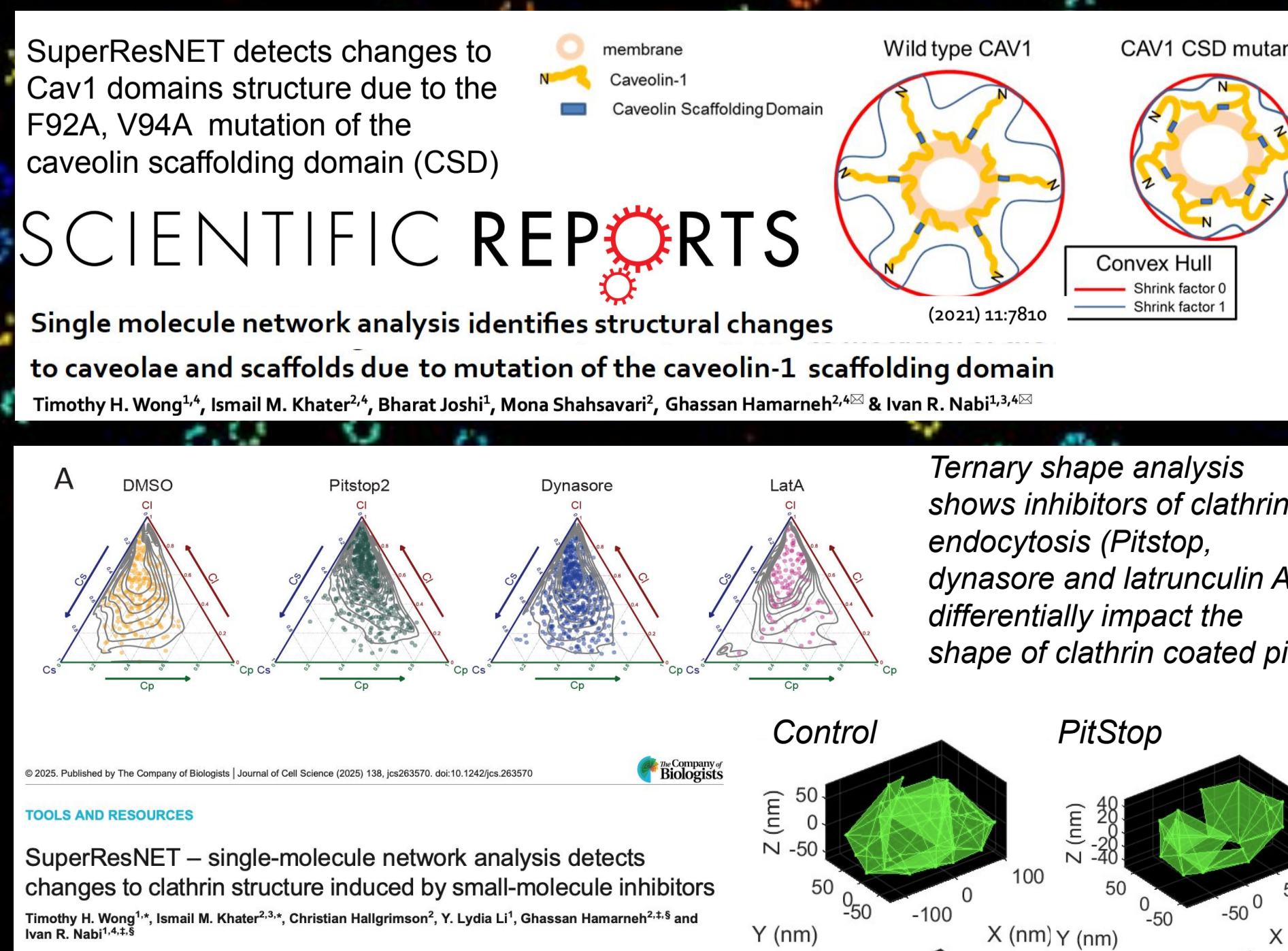
SuperResNET Examples

Identification of novel subcellular structures



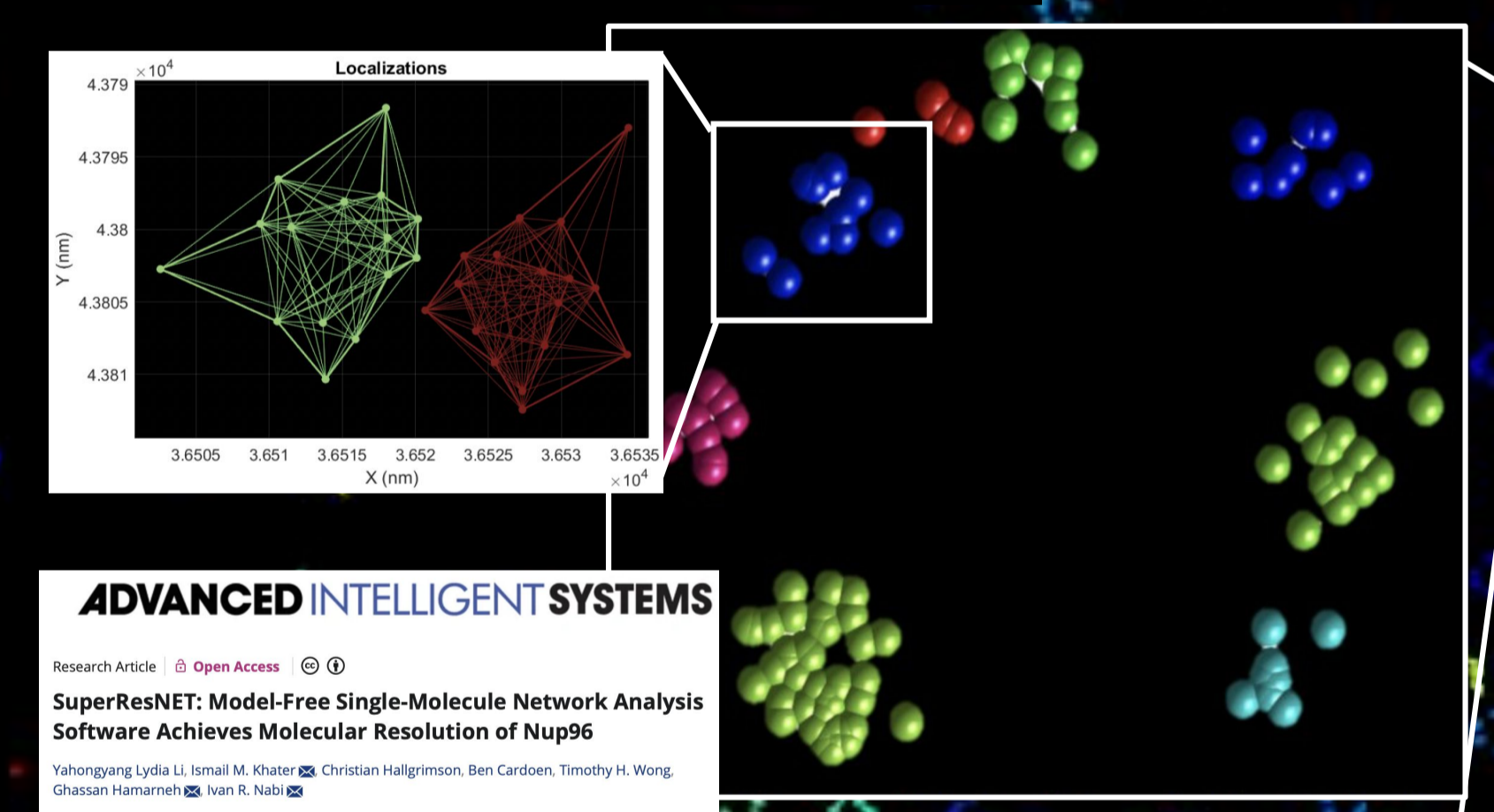
SCIENTIFIC REPORTS
Super-resolution modularity analysis shows polyhedral caveolin-1 oligomers combine to form scaffolds and caveolae

Detection of changes to molecular structure due to small molecule inhibitors and point mutations



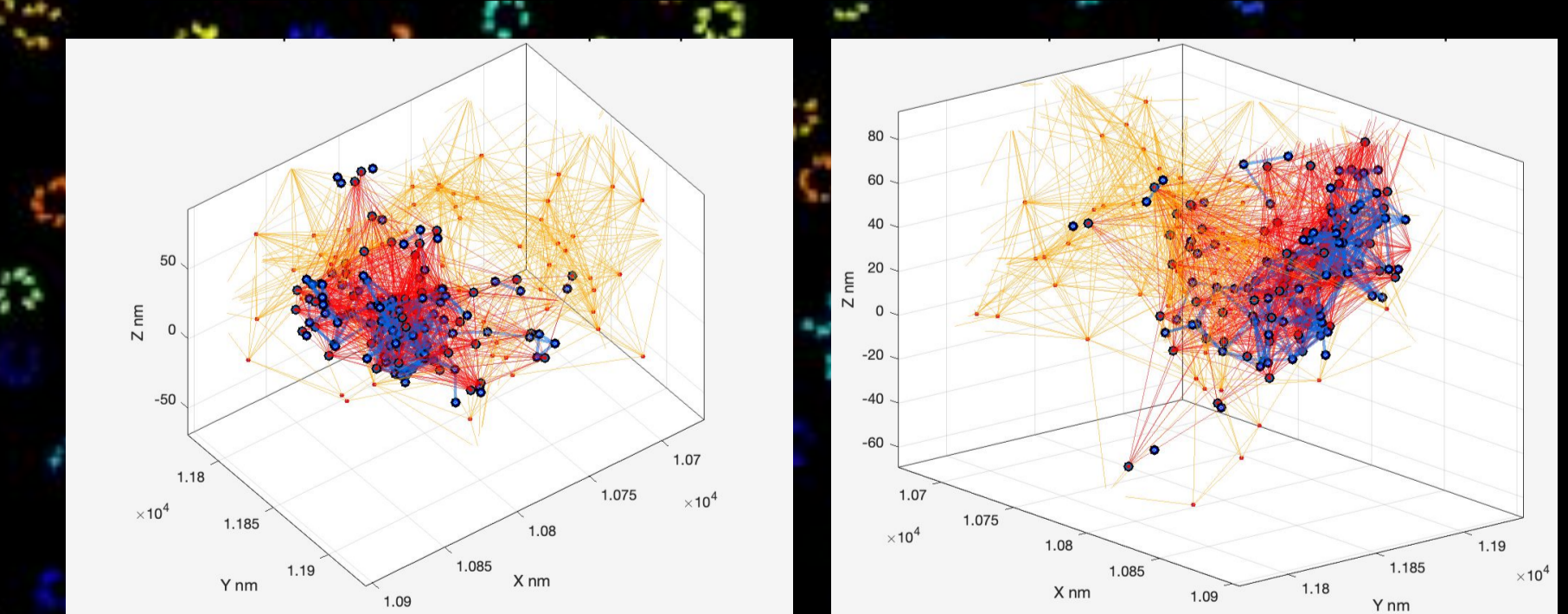
SuperResNET

Modularity analysis resolves Nup96 corners from dSTORM data



SuperResNET interaction analysis

- Processes two (or more) SMLM data sets and identifies localizations from the two data sets in close proximity
- Example shows cavin-1 (blue) interacting with a Cav1-labeled caveolae (yellow and red)
- Dense Cav1 points in proximity to cavin-1 are in red



Structural analysis of MinFlux data

