https://github.com/Mattheyses-Lab/Nawara_et_al._NatCommun_2022.git contains all the automatic data processing and analysis software used in the manuscript.

1. **System requirements** must be compatible with:

   - MATLAB R2018a and R2020b
     (https://www.mathworks.com/support/requirements/previous-releases.html)

   - FiJi (ImageJ,
     https://imagej.net/downloads#:~:text=ImageJ%20will%20run%20on%20any,Java%20installed%20from%20java.com)

   - For CMEanalysis UAB Cheaha Supercomputer was used
     (https://www.uab.edu/it/home/research-computing/cheaha)

2. **Installation** – no installation is required. For MATLAB programs – double click on the code of choice, for FiJi drag the dz_channel_generator.ijm into to FiJi window to open it.

The three main programs that are used to analyze data are in the following folders:

A. **CMEanalysis_STAR** (MATLAB R2018a) – clathrin accumulation detection and tracing

   - Contains optimized for STAR microscopy CMEanalysis software (source:
     https://github.com/DanuserLab/cmeAnalysis). Data was organized as suggested by
     Aguet et al., Dev. Cell 26(3), pp. 279-291, 2013. and runed using the following
     executable codes:

Example code run for control group:

```
>> data = loadConditionData('/data/user/tnawara/Data/Data analysis/AA_First
paper/Revisions/siRNA/Exp3/CLCa_CTRL_exp3', {'Ch1', 'Ch2', 'Ch3'}, {'EGFP',
'iRFP713', 'dZ'}, 'Parameters', [1.49 60 6.45e-6]);
>>[resCTRL, dataCTRL] = cmeAnalysis(data, 'ControlData', resCTRL, 'Overwrite',
false, 'TrackingRadius', [1 3], 'TrackingGapLength', 13);
```

Example code run for experimental group:

```
>> data = loadConditionData('/data/user/tnawara/Data/Data analysis/AA_First
paper/Revisions/siRNA/Exp3/CLCa_siRNA_exp3', {'Ch1', 'Ch2', 'Ch3'}, {'EGFP',
'iRFP713', 'dZ'}, 'Parameters', [1.49 60 6.45e-6]);
>>[resKD, dataKD] = cmeAnalysis(data, 'ControlData', resCTRL, 'Overwrite',
false, 'TrackingRadius', [1 3], 'TrackingGapLength', 13);
```
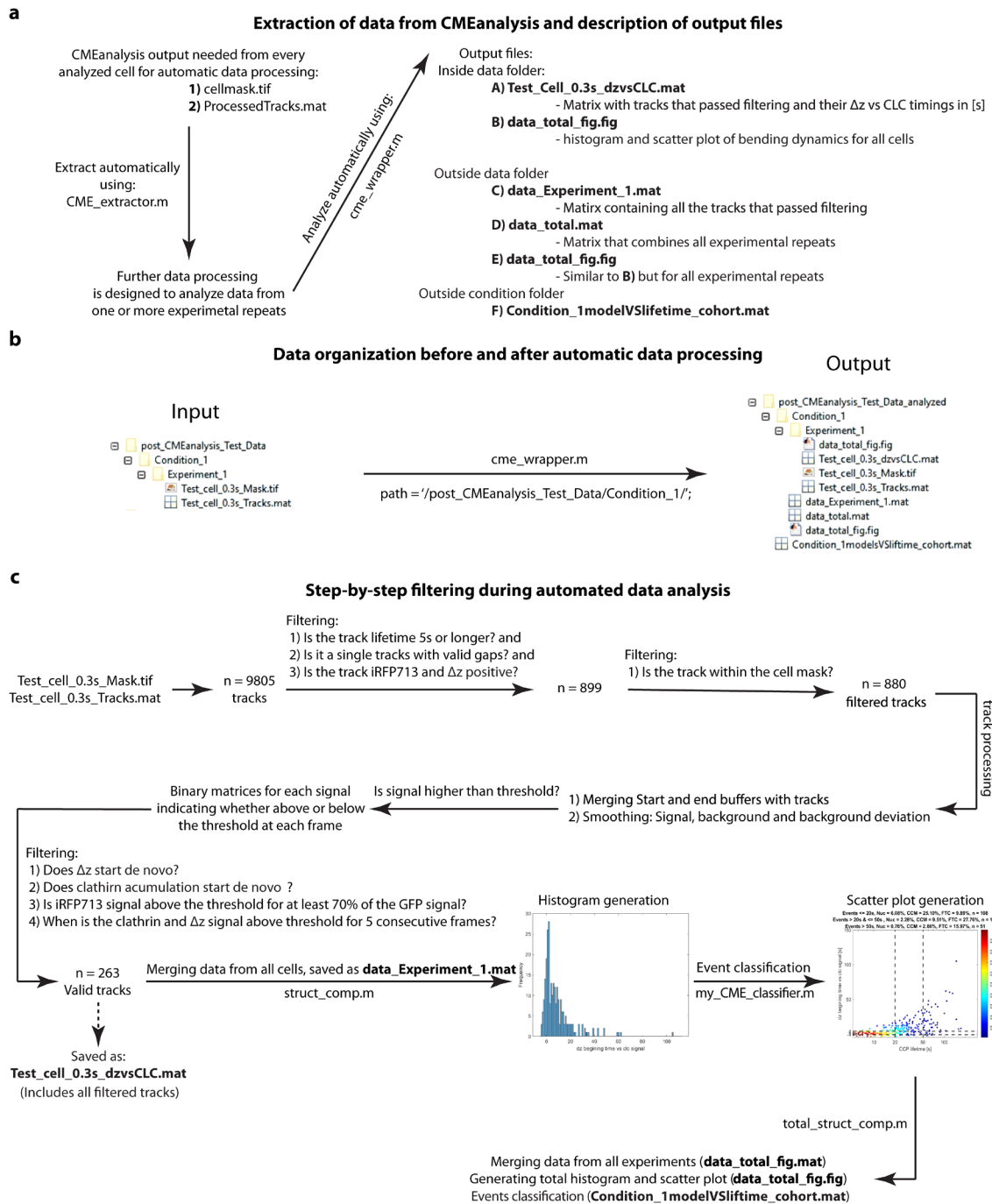
- Key parameters used for data processing are shown in Supplementary Table 1 (Nawara T. et al. 2022)
- Test data set:
  https://drive.google.com/drive/folders/1fd5narw6eAPgUsKROjOHWgRpWcxmboRb?usp=sharing

B. **post_CMEanalysis** (MATLAB R2020b) – processing of tracks detected by CMEanalysis
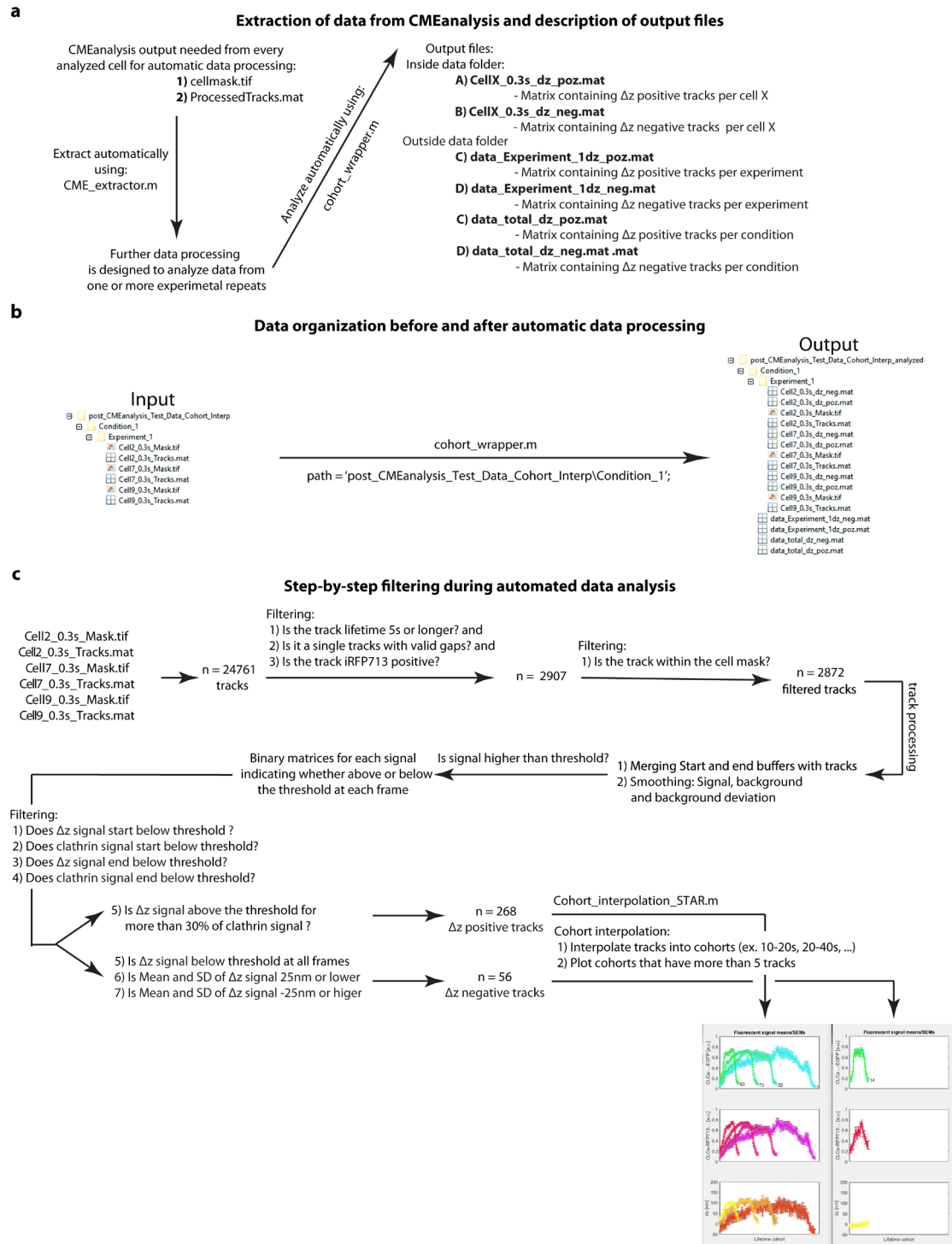
- **delta z Beginning** – measures when curvature develops in relation to clathrin accumulation. Data was organized and processed as follows (Sup Fig.9)

a

Extraction of data from CMEanalysis and description of output files

CMEanalysis output needed from every analyzed cell for automatic data processing:
1) cellmask.tif
2) ProcessedTracks.mat

Extract automatically using:
CME_extractor.m

Analyze automatically using: cme_wrapper.m

Further data processing is designed to analyze data from one or more experimetal repeats

Output files:
Inside data folder:
A) Test_Cell_0.3s_dzvsCLC.mat
 - Matrix with tracks that passed filtering and their Δz vs CLC timings in [s]
B) data_total_fig.fig
 - histogram and scatter plot of bending dynamics for all cells

Outside data folder
C) data_Experiment_1.mat
 - Matirx containing all the tracks that passed filtering
D) data_total.mat
 - Matrix that combines all experimental repeats
E) data_total_fig.fig
 - Similar to **B)** but for all experimental repeats
Outside condition folder
F) Condition_1modelVSlifetime_cohort.mat

b

Data organization before and after automatic data processing

Input

post_CMEanalysis_Test_Data
 Condition_1
  Experiment_1
   Test_cell_0.3s_Mask.tif
   Test_cell_0.3s_Tracks.mat

cme_wrapper.m
path = '/post_CMEanalysis_Test_Data/Condition_1/';

Output

post_CMEanalysis_Test_Data_analyzed
 Condition_1
  Experiment_1
   data_total_fig.fig
   Test_cell_0.3s_dzvsCLC.mat
   Test_cell_0.3s_Mask.tif
   Test_cell_0.3s_Tracks.mat
  data_Experiment_1.mat
  data_total.mat
  data_total_fig.fig
 Condition_1modelsVSliftime_cohort.mat

c

Step-by-step filtering during automated data analysis

Test_cell_0.3s_Mask.tif
Test_cell_0.3s_Tracks.mat

n = 9805 tracks

Filtering:
1) Is the track lifetime 5s or longer? and
2) Is it a single tracks with valid gaps? and
3) Is the track iRFP713 and Δz positive?

n = 899

Filtering:
1) Is the track within the cell mask?

n = 880 filtered tracks

track processing

Binary matrices for each signal indicating whether above or below the threshold at each frame

Is signal higher than threshold?

1) Merging Start and end buffers with tracks
2) Smoothing: Signal, background and background deviation

Filtering:
1) Does Δz start de novo?
2) Does clathirn acumulation start de novo ?
3) Is iRFP713 signal above the threshold for at least 70% of the GFP signal?
4) When is the clathrin and Δz signal above threshold for 5 consecutive frames?

n = 263 Valid tracks

Merging data from all cells, saved as **data_Experiment_1.mat**
struct_comp.m

Saved as:
Test_cell_0.3s_dzvsCLC.mat
(Includes all filtered tracks)

Histogram generation

Event classification
my_CME_classifier.m

Scatter plot generation

total_struct_comp.m

Merging data from all experiments (**data_total_fig.mat**)
Generating total histogram and scatter plot (**data_total_fig.fig**)
Events classification (**Condition_1modelVSliftime_cohort.mat**)

   i. Test data set:

https://drive.google.com/drive/folders/1v3B125oLSkMTZihWJZuGWquWwCAEv5XF?usp=sharing

- **flat curved sorting and interpolation** – separates and groups detected tracks on whether the curvature was induced (Flat vs. Curved CCSs - Sup Fig.10)

a

Extraction of data from CMEanalysis and description of output files

CMEanalysis output needed from every analyzed cell for automatic data processing:
1) cellmask.tif
2) ProcessedTracks.mat

Extract automatically using:
CME_extractor.m

Further data processing is designed to analyze data from one or more experimetal repeats

Analyze automatically using:
cohort_wrapper.m

Output files:
Inside data folder:
A) CellX_0.3s_dz_poz.mat
  - Matrix containing Δz positive tracks per cell X
B) CellX_0.3s_dz_neg.mat
  - Matrix containing Δz negative tracks per cell X
Outside data folder
C) data_Experiment_1dz_poz.mat
  - Matrix containing Δz positive tracks per experiment
D) data_Experiment_1dz_neg.mat
  - Matrix containing Δz negative tracks per experiment
C) data_total_dz_poz.mat
  - Matrix containing Δz positive tracks per condition
D) data_total_dz_neg.mat .mat
  - Matrix containing Δz negative tracks per condition

b

Data organization before and after automatic data processing

Input

post_CMEanalysis_Test_Data_Cohort_Interp
  Condition_1
    Experiment_1
      Cell2_0.3s_Mask.tif
      Cell2_0.3s_Tracks.mat
      Cell7_0.3s_Mask.tif
      Cell7_0.3s_Tracks.mat
      Cell9_0.3s_Mask.tif
      Cell9_0.3s_Tracks.mat

cohort_wrapper.m

path = 'post_CMEanalysis_Test_Data_Cohort_Interp\Condition_1';

Output

post_CMEanalysis_Test_Data_Cohort_Interp_analyzed
  Condition_1
    Experiment_1
      Cell2_0.3s_dz_neg.mat
      Cell2_0.3s_dz_poz.mat
      Cell2_0.3s_Mask.tif
      Cell2_0.3s_Tracks.mat
      Cell7_0.3s_dz_neg.mat
      Cell7_0.3s_dz_poz.mat
      Cell7_0.3s_Mask.tif
      Cell7_0.3s_Tracks.mat
      Cell9_0.3s_dz_neg.mat
      Cell9_0.3s_dz_poz.mat
      Cell9_0.3s_Mask.tif
      Cell9_0.3s_Tracks.mat
    data_Experiment_1dz_neg.mat
    data_Experiment_1dz_poz.mat
    data_total_dz_neg.mat
    data_total_dz_poz.mat

c

Step-by-step filtering during automated data analysis

Cell2_0.3s_Mask.tif
Cell2_0.3s_Tracks.mat
Cell7_0.3s_Mask.tif
Cell7_0.3s_Tracks.mat
Cell9_0.3s_Mask.tif
Cell9_0.3s_Tracks.mat

n = 24761 tracks

Filtering:
1) Is the track lifetime 5s or longer? and
2) Is it a single tracks with valid gaps? and
3) Is the track iRFP713 positive?

n = 2907

Filtering:
1) Is the track within the cell mask?

n = 2872 filtered tracks

track processing

1) Merging Start and end buffers with tracks
2) Smoothing: Signal, background and background deviation

Is signal higher than threshold?

Binary matrices for each signal indicating whether above or below the threshold at each frame

Filtering:
1) Does Δz signal start below threshold ?
2) Does clathrin signal start below threshold?
3) Does Δz signal end below threshold?
4) Does clathrin signal end below threshold?

5) Is Δz signal above the threshold for more than 30% of clathrin signal ?

n = 268 Δz positive tracks

5) Is Δz signal below threshold at all frames
6) Is Mean and SD of Δz signal 25nm or lower
7) Is Mean and SD of Δz signal -25nm or higer

n = 56 Δz negative tracks

Cohort_interpolation_STAR.m

Cohort interpolation:
1) Interpolate tracks into cohorts (ex. 10-20s, 20-40s, ...)
2) Plot cohorts that have more than 5 tracks

Other programs used I the manuscript are listed below:

C. **FiJi** (ImageJ) – contains a FiJi plug-in that is used for generating the Δz/curvature channel

- Requires:

    i. Split, background subtracted, flat field and bleach corrected 488 channel time-lapse (end extension "_488_Cor.tif")

    ii. Split, background subtracted, flat field and bleach corrected and registered 647 channel time-lapse (end extension "_647_Cor_Reg.tif")

    iii. Sigma (interpolation correction adjust at line 21)

- Change path to folder @ line 10

- Change folder name @ line 9

- Change folder cell number in the experiment to go through all the cells one at the time @ line 7

- Change the 1/ γ value @ line 47

- Test data set: https://drive.google.com/drive/folders/1FKiySKlWkElbX4NqiqW8CdV_1CEqmel1?usp=sharing

D. **Beads** (MATLAB R2020b) – Measures theoretical and STAR measured Δz of microspheres

- Based on Stabley et al. 2015 (https://www.nature.com/articles/ncomms9307)

- Beads needs to be split, background subtracted and flat field corrected for TRIF modes pictures, required files and extensions:

    i. TIRF 488 – "_TIRF_488_Cor_Blur.tif"

    ii. TIRF 647 – "_TIRF_647_Cor_Reg.tif"

    iii. EPI 488 – "_EPI_488.tif"

- Define path to data @ line 15

- Define pixel size in nm @ line 16

- EPI and TIRF images will overlay, pick beads that TIRF signal is nicely in the middle of EPI ring. Right click when picking the last bead.

- Select beads with uniform fluorophores distribution by accepting it 'y' or rejecting 'n'

- It will scan through folders automatically

- Test data set:

  https://drive.google.com/drive/folders/11KzyFfA1jUVBUlwiGU2KhLohTEa7B5qo?usp=sharing

E. **EPI.STAR** (MATLAB R2020b) – similar to **delta z Beginning** but supports 4 channels and measures and displays the cumulative difference between the disappearance of clathrin-coated structures form TIRF and EPI, data organization for CMEanalysis:

- Ch1 – TIRF 488
- Ch2 – TIRF 647
- Ch3 – Δz channel
- Ch4 – EPI 488

F. **Video annotation** (MATLAB R2020b):

- **image_annotation.m** - annotates time scale on live-cell videos
- **image_overlay.m** - overlays a definable inset with the full size video

G. **Monte Carlo of stochastic variation of tagged and untagged molecules** (MATLAB R2020b):

- **Delta_Z_Modeling_Figures.m** – prints the representative stages of vesicles formation
    i. Adjust the amount of snapshots @ line 4
    ii. Adjust percentage of vesicle coverage @ line 5
- **Delta_Z_vs_CM_Spaghetti_Plot.m** – Performs Monte Carlo simulation of random distribution of tagged molecules at the modeled vesicle
    i. Adjust the amount of tagged molecules by decreasing the number @ line 14

3. **Run time**:

   a. CMEanalysis – hours to days
   b. Post_CMEanalysis and EPI.STAR – several minutes
   c. Monte Carlo Simulation – few minutes