**Introduction**

This data set is for sharing all of the raw images analyzed in the paper, “***DASC, a sensitive classifier for measuring discrete early stages in clathrin-mediated endocytosis***”, <https://doi.org/10.7554/eLife.53686>. In the paper, we proposed a new method that can unbiasedly classify two populations of clathrin structures during endocytosis, clathrin coated pits and abortive coats. Based on this new classification, we measured phenotypes of siRNA knockdown of 11 endocytic accessory protein (EAP), analyzed clathrin-AP2 dual channel images and measured curvature formation during endocytosis.

Images in all four experiments are included in this data set, ‘EAP knockdown’, ‘alphaPIP2-’ as a positive control experiment, ‘CLC\_AP2\_dual’ as another control experiment, and ‘EpiTIRF’ for curvature measurement. Each experiment has various conditions, see the metadata file for details.

Readers can follow this README to obtain the entire data and then apply DASC, one experiment by one experiment, to reproduce the results in the associated paper. The software of DASC is available on GitHub: <https://github.com/DanuserLab/cmeAnalysis>. See the instruction on the GitHub page for how to use DASC to analyze the images.

Files in this data are compressed. Each ‘.tif.bz’ file is a compressed tif image sequence of fluorescently labeled cells. The names of these compressed files include descriptive information of ‘experiment’, ‘condition’, ‘date’, ‘cell name’ and ‘channel name’ (optional). These descriptors are separated by ‘#’.

To decompress the files, run the included Matlab script (‘decompression.m’) on Linux system.

1. Find and run the script in Matlab (Linux system only).
2. After the decompression is finished, the files are automatically reorganized into designated folders sorted by experiment, condition, date, cell name and channel name.
3. The file structure is for readers to conveniently apply the software ‘DASC’ to reproduce the results in the paper.

Author correspondence: Gaudenz Danuser (gaudenz.danuser@utsouthwestern.edu), Lyda Hill Department of Bioinformatics, UT Southwestern, Dallas, Texas, U.S.

**Methods and Materials**

***Cell culture and cell engineering***

ARPE19 and ARPE-19/HPV-16 (ATCC® CRL-2502™) cells were obtained from ATCC and cultured in DMEM/F12 medium with 10% (v/v) FBS at 37°C under 5% CO2. ARPE-19/HPV-16 cells were infected with recombinant lentiviruses encoding eGFP-CLCa in a pMIEG3 vector, and sorted by FACS after 72 hours (*Aguet F. et al, Dev. Cell 2013*). AP2 reconstitution was achieved by infecting the eGFP CLCa-expressed ARPE-19/HPV-16 cells (ARPE\_HPV16 eGFP\_CLCa) with retroviruses encoding siRNA resistant WT or PIP2- (K57E/Y58E) AP2 alpha subunit in a pMIEG3-mTagBFP vector and FACS sorted based on BFP intensity (*Kadlecova Z. et al, JCB 2016*). Western blotting was used to confirm reconstituted-protein expression and knockdown efficiency of the generated cell lines using anti-alpha-adaptin (Thermo Fisher Scientific, #AC1-M11) and anti-CALM (Abcam, #ab172962) antibodies. APRE19 cells with stable expression of mRuby2-CLCa and -eGFP-AP2 were also generated via lenti- and retroviral transduction, respectively.

***siRNA transfection***

200,000 ARPE-19/HPV-16 cells were plated on each well of a 6-well plate for ≥ 3 hours before transfection. Transfections for siRNA knockdown were assisted with Lipofectamin RNAiMAX (Life Technologies, Carlsbad, CA). Briefly, 6.5 µl of Lipofectamin RNAiMax and 5.5 µl of 20 µM siRNA were added separately into 100 µl OptiMEM and incubated separately for 5 min at room temperature. siRNA were next mixed with lipofectamin RNAiMAX and incubated at room temperature for another 10 min before being added dropwise to the cells with fresh medium. Measurements were performed at day 5 after plating cells following two rounds of siRNA transfection (time gap = 24-48 hrs between transfections). Western blotting confirmed that the knockdown efficiency for all proteins was over 80%. Control cells were transfected in parallel with control siRNA (siControl) purchased from QIAGEN (Germantown, MD).

***Microscopy imaging and quantification***

Total Internal Reflection Fluorescence (TIRF) Microscopy imaging was conducted as previously described (*Loerke D. et al, PLoS Bio. 2009*). Cells were grown on a gelatin-coated 22x22mm glass (Corning, #2850-22) overnight and then mounted to a 25x75mm cover slide (Thermo Scientific, #3050). Imaging was conducted with a 60X, 1.49-NA Apo TIRF objective (Nikon) mounted on a Ti-Eclipse inverted microscope equipped with an additional 1.8X tube lens, yielding a final magnification of 108X. Perfect focus was applied during time-lapsed imaging. For epi/TIRF imaging, nearly simultaneous two channel (488 epifluorescence/TIRF) movies were acquired with multi-dimension acquisition (MDA). Movies were acquired at the rate of 1 frame/s. cmeAnalysis was applied for CCP detection and tracking (*Aguet F. et al, Dev. Cell 2013, Jaqaman K. et al, Nat. Method 2008, Loerke D. et al, Traffic 2011*). Variation could arise from the heterogeneity of cover glass by itself and the gelatin-coating. After decompression, each tif file is an image sequence of 451 frames. For each file, the fluorescent marker and imaging method can be found in the metadata sheet by matching all the descriptors.