Data from: Inferring single cell behavior from large-scale epithelial sheet migration patterns

Journal of the Royal Society Interface

Rachel M. Lee, Haicen Yue, Wouter-Jan Rappel, and Wolfgang Losert

Contact Information

rmlee@umd.edu, wlosert@umd.edu

Abstract

Cell migration plays an important role in a wide variety of biological processes and can incorporate both individual cell motion and collective behavior. The emergent properties of collective migration are receiving increasing attention as collective motion's role in diseases such as metastatic cancer becomes clear. Yet, how individual cell behavior influences large-scale, multi-cell collective motion remains unclear. In our study, we provided insight into the mechanisms behind collective migration by studying cell migration in a spreading monolayer of epithelial MCF10A cells. We quantify migration using particle image velocimetry and find that cell groups have features of motion that span multiple length scales. Comparing our experimental results to a model of collective cell migration, we find that cell migration within the monolayer can be affected in qualitatively different ways by cell motion at the boundary, yet it is not necessary to introduce leader cells at the boundary or specify other large-scale features to recapitulate this large-scale phenotype in simulations. Instead, in our model, collective motion can be enhanced by increasing the overall activity of the cells or by giving the cells a stronger coupling between their motion and polarity. This suggests that investigating the activity and polarity persistence of individual cells will add insight into the collective migration phenotypes observed during development and disease. This dataset provides microscopy images and analysis to support the article in the Journal of the Royal Society Interface describing these migration behaviors.

Funding

This work was carried out with financial support to RML and WL from a NSF-Physics of Living Systems grant (PHY1205965). RML was additionally supported by the JCM Foundation through an ARCS/MWC Scholar Award. WJR and HY were supported by NIH Grant No. P01 GM078586 and NSF Grant No. DMS 1309542. The raw images analyzed in this work have been described previously [1] and were collected with funding and support from the Intramural Research Program of the Center for Cancer Research, NCI, National Institutes of Health.

[1] Lee RM, Kelley DH, Nordstrom KN, Ouellette NT, Losert W. *Quantifying stretching and rearrangement in epithelial sheet migration*. New J Phys. 2013;15(2):25036.

Dataset Overview

Raw Experimental Images: Example sets of images are provided as .zip files containing tif images

ExperimentalData.zip: Contains the subfolders CorrelationCurves, EdgeCoordinates, MonolayerRadius, PIV_FlowFields, and VelocityCurves

MatPIV161.zip: MATLAB code used for PIV analysis

SimulationData.zip: Contains the subfolders CorrelationCurves, FlowFields, MonolayerRadius, SimulationCoordinates, and VelocityCurves

Raw Experimental Images

The raw images analyzed in this work have been described previously [1] and were collected with funding support from the Intramural Research Program of the Center for Cancer Research, NCI, National Institutes of Health. Example images are provided .zip files containing as multipage tiff files. The frames were taken every two minutes and a pixel corresponds to 0.65 μm. The four files from 2012-03-15 consist of the four fields of view corresponding to one monolayer in 1:1 media conditions, while the four files from 2012-06-02 correspond to one monolayer in 1:5 media conditions.

ExperimentalData.zip

All files begin with the date the data was taken and an identifier for the monolayer (Dot1, Dot2, etc.). The imaging field of view (L = left, R = right, etc.) is indicated where appropriate. Files dated 2012-03-15 correspond to monolayers migrating in 1:1 media conditions and the remaining files correspond to experiments conduced at 1:5 media conditions.

CorrelationCurves

For each monolayer, the file *correlationData.mat is a MATLAB data file containing:

- bin_mids: The location in μm of the bin center for the averaged correlation values
- Cr_mean: The averaged correlation values at the distance specified by bin_mids
- A, B, lc1, and lc2: The fit values obtained by fitting Cr_mean to a double exponential of the form

$$C(\Delta r) = Ae^{-\frac{r}{L_{c1}}} + Be^{-\frac{r}{L_{c2}}} + (1 - A - B)$$

EdgeCoordinates

For each imaging field of view which contains the monolayer edge, the file *edgeData.mat is a MATLAB data file containing a data structure variable, ‘edge_data.’ This structure contains the following fields:

- edge_data.points: The coordinates of the monolayer leading edge specified as a three dimensional matrix. The third dimension is the frame of imaging, while the second dimension is the x and y coordinates of the edge in units of pixels. The first dimension is along the length of the monolayer edge and is padded with zeros which can be removed in further analysis.
- edge_data.imname: The image file name used to generate this structure
- edge_data.imsize: The pixel size of the images used to generate the data structure
- edge_data.side: The side of the image which contains the cell monolayer

MonolayerRadius

For each monolayer, the file *radiusData.mat is a MATLAB data file containing:

- X: The center x-axis location of the monolayer for each frame
- Y: The center y-axis location of the monolayer for each frame
- R: The monolayer radius for each frame in μm
- x_pos: The x-axis location of each field of view taken of this monolayer
- y_pos: The y-axis location of each field of view taken of this monolayer

PIV_FlowFields

For each imaging field of view, the file *pivData.mat is a MATLAB data file containing the structure piv_data. This structure contains the following fields:

- piv_data.x: The x-coordinates for the PIV flow field (in pixels)
- piv_data.y: The y-coordinates for the PIV flow field (in pixels)
- piv_data.u: The x-axis velocity for the PIV flow field (in pixels/frame)
- piv_data.v: The y-axis velocity for the PIV flow field (in pixels/frame)
- piv_data.imname: The file name of the images analyzed for this structure
- piv_data.firstframe: The first frame of the image sequence analyzed
- piv_data.lastframe: The last frame of the image sequence analyzed

VelocityCurves

For each monolayer, the file *velocityCurves.mat is a MATLAB data file containing:

- bin_mid: The middle of the bins used to take velocity averages in units of *r/R,* where *R* is the monolayer radius and *r* is the location within the monolayer
- speed_mean: The mean speed at each bin location in μm/min
- speed_std: The standard deviation of speed at each bin location μm/min
- v_r_mean: The mean radial velocity at each bin location in μm/min
- v_r_std: The standard deviation of radial velocity in μm/min
- N: The number of velocity vectors averaged in each bin
- firstframe: The first frame analyzed
- lastframe: The last frame analyzed

MatPIV161.zip

This folder contains the version of the MatPIV toolbox (J. Kristian Sveen, GNU general public license) used to create the files in the PIV_FlowFields folder.

SimulationData.zip

All files begin with an identifier which indicates the simulation parameters used in that set of data. The key to these file names is provided in ReadMe_SimulationFileNames.pdf.

CorrelationCurves

For each set of simulation parameters, the file *correlationData.mat is a MATLAB data file containing:

- bin_mids: The location of the bin center for the averaged correlation values
- Cr_mean: The averaged correlation values at the distance specified by bin_mids
- A, B, lc1, and lc2: The fit values obtained by fitting Cr_mean to a double exponential of the form

$$C(\Delta r) = Ae^{-\frac{r}{L_{c1}}} + Be^{-\frac{r}{L_{c2}}} + (1 - A - B)$$

FlowFields

For each set of simulation parameters, the file *flowFields.mat is a MATLAB data file containing the structure flow_data. This is derived by interpolating the simulation coordinates data to a grid and is saved as a structure which contains the following fields:

- flow_data.x: The x-coordinates for the PIV flow field
- flow_data.y: The y-coordinates for the PIV flow field
- flow_data.u: The x-axis velocity for the PIV flow field
- flow_data.v: The y-axis velocity for the PIV flow field

Monolayer Radius

For each set of simulation parameters, the file *edgeData.mat is a MATLAB data file containing:

- points: A cell array where each cell indicates a frame of the simulation and within each cell array are the coordinates for the monolayer boundary
- X: The center x-axis location of the monolayer for each frame
- Y: The center y-axis location of the monolayer for each frame
- R: The monolayer radius for each frame

SimulationCoordinates

For each set of simulation parameters, the corresponding .xyz file is a space delimitated text-file with the following columns describing each particle in the simulation:

1. x-position
2. y-position
3. z-position (zero for all simulations presented here)
4. instantaneous x-axis velocity
5. instantaneous y-axis velocity
6. instantaneous z-axis velocity
7. particle index
8. mean x-axis velocity with a relaxation time (this is the velocity used in all analysis)
9. mean y-axis velocity with a relaxation time (this is the velocity used in all analysis)
10. x-axis motility force (when the total motility force is zero, the cell is 'asleep')
11. y-axis motility force (when the total motility force is zero, the cell is 'asleep')
12. cell index (The cell to which this particle belongs; each cell is made of two particles)
13. average relative position of the neighbor cells along the x-axis
14. average relative position of the neighbor cells along the y-axis

VelocityCurves

For each set of simulation parameters, the file *velocityCurves.mat is a MATLAB data file containing:

- bin_mid: The middle of the bins used to take velocity averages in units of *r/R,* where *R* is the monolayer radius and *r* is the location within the monolayer.
- speed_mean: The mean speed at each bin location in μm/min
- speed_std: The standard deviation of speed at each bin location μm/min
- v_r_mean: The mean radial velocity at each bin location in μm/min
- v_r_std: The standard deviation of radial velocity in μm/min
- N: The number of velocity vectors averaged in each bin
- firstframe: The first frame analyzed