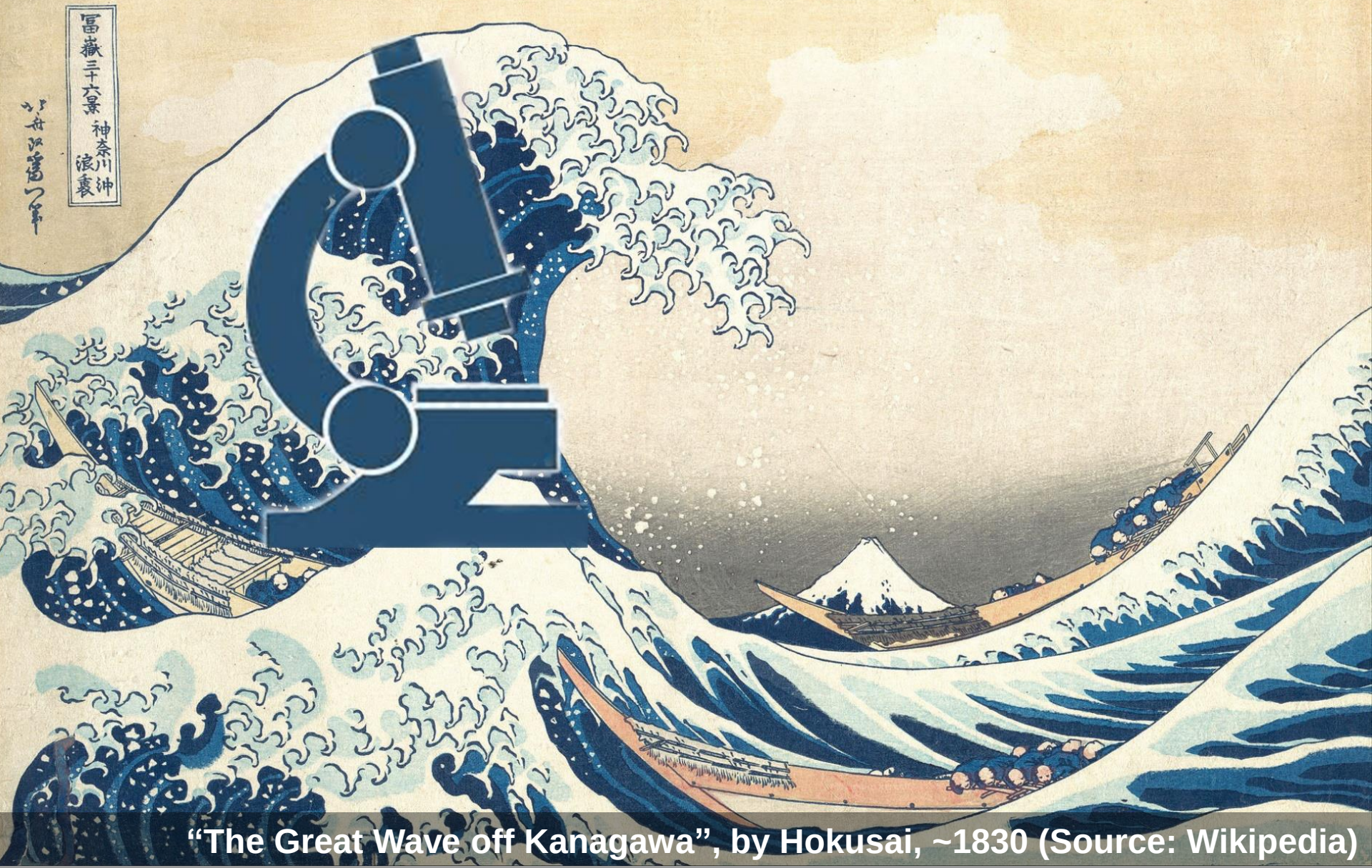
Data science in cell imaging

Lecture 13: course summary, public data repositories and reuse



“The Great Wave off Kanagawa”, by Hokusai, ~1830 (Source: Wikipedia)

All slides are open under the cc-by license

PPTX slides available here

Course’s theme

This course will review the state-of-the-art in visualizing, processing, integrating and mining cell image data sets, **deciphering complex patterns and turning them into new biological knowledge**. It will include a variety of computational approaches applied to diverse bio-imaging data.

Before we start: quick Q&A

Q	A
Why are you speaking English?	Science is communicated in English (also useful for industry)
Do I need any background in biology?	I assume no prior knowledge in biology. I do expect curiosity and interest in the domain.
Do I need computational background?	Yes. I am encouraging students from diverse background to join, I will try to keep it as simple and as intuitive as possible. But this course main target audience are computational students.
Is it a good fit for undergraduate students?	Feedback from last year indicates that the course was quite challenging for (excellent) 3rd year students

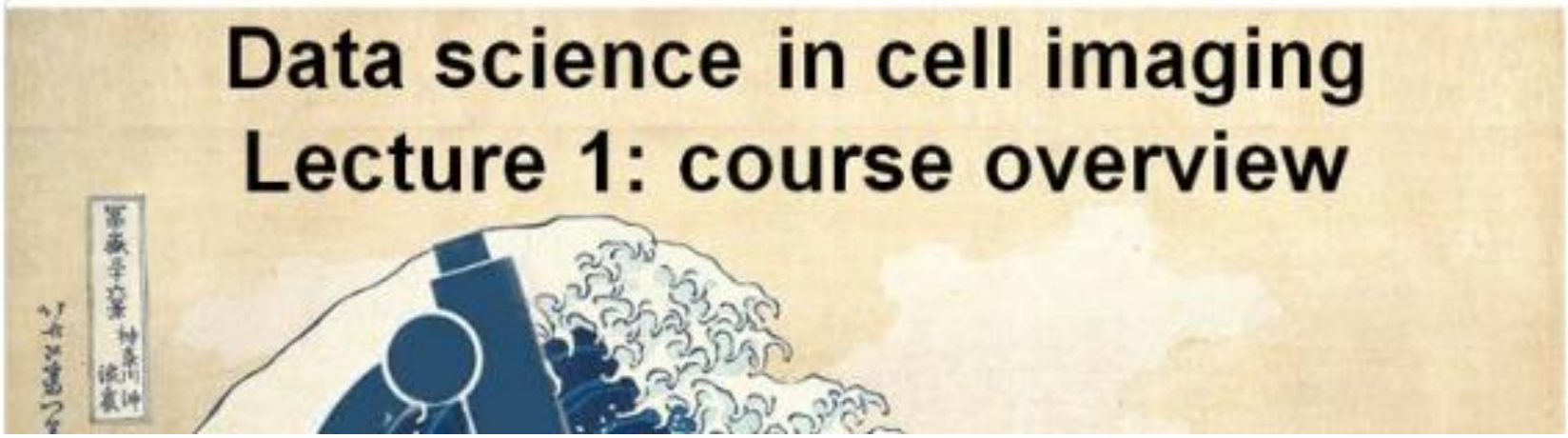
Before we start: quick Q&A

Q	A
I am here to learn about deep learning!	The focus is not of DL (or any specific technique). Also, the course does not involves lab/hands on experience (beyond the final project).
Obligatory attendance?	No! But you'll find it hard to follow if you do not attend
I am looking for an easy course	Look elsewhere
High grades?	Last year average was 90, and it was an elite group of students

Assaf Zaritsky
@AssafZaritsky

Preparing slides for the first lecture in my brand new course on Data Science in Cell Imaging starting next week @bengurionu!

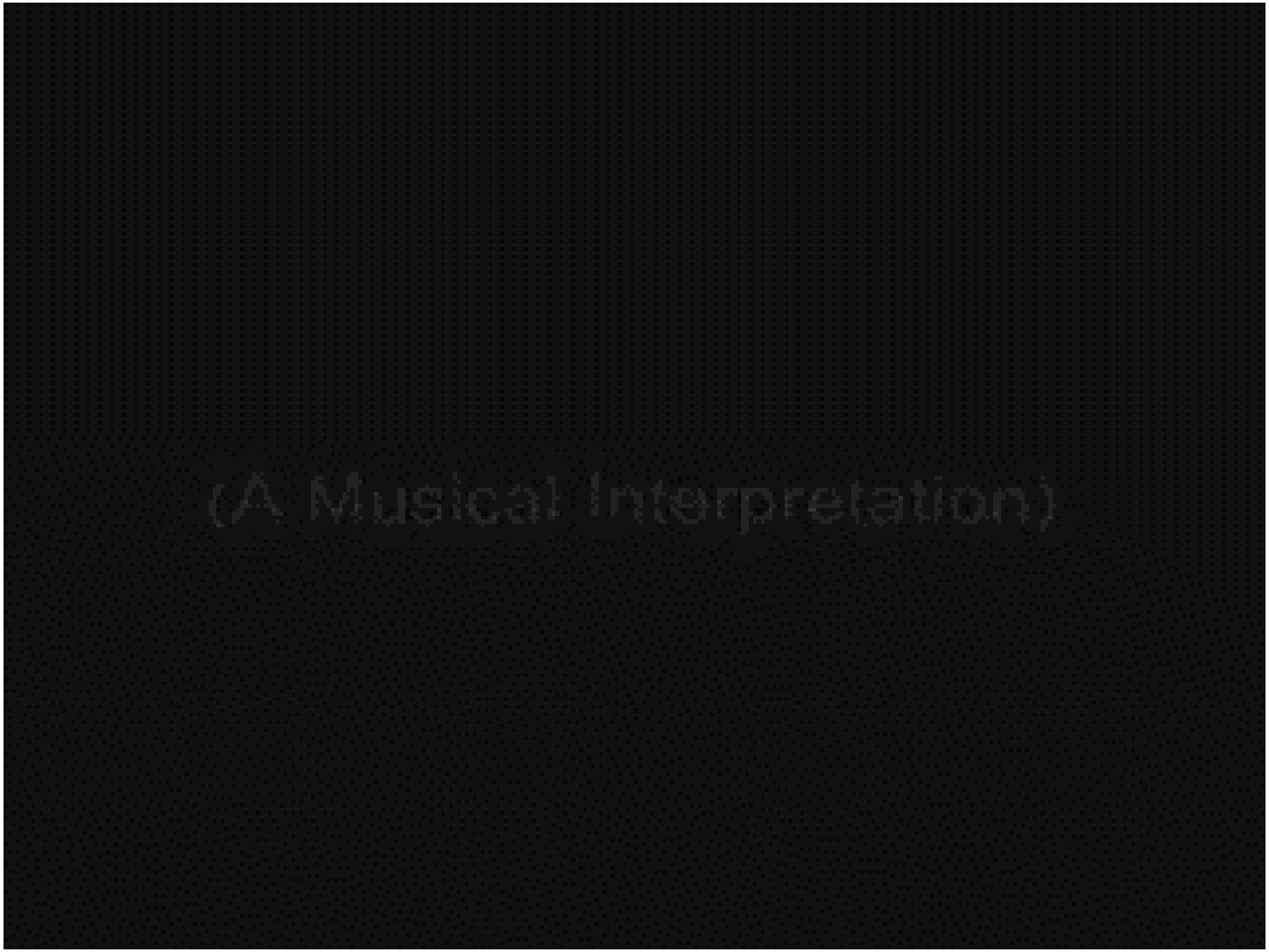
Very curious to see if I can get computational grad students (with no BIO background) interested in the interface of computation and cell biology :-)



Broad introduction

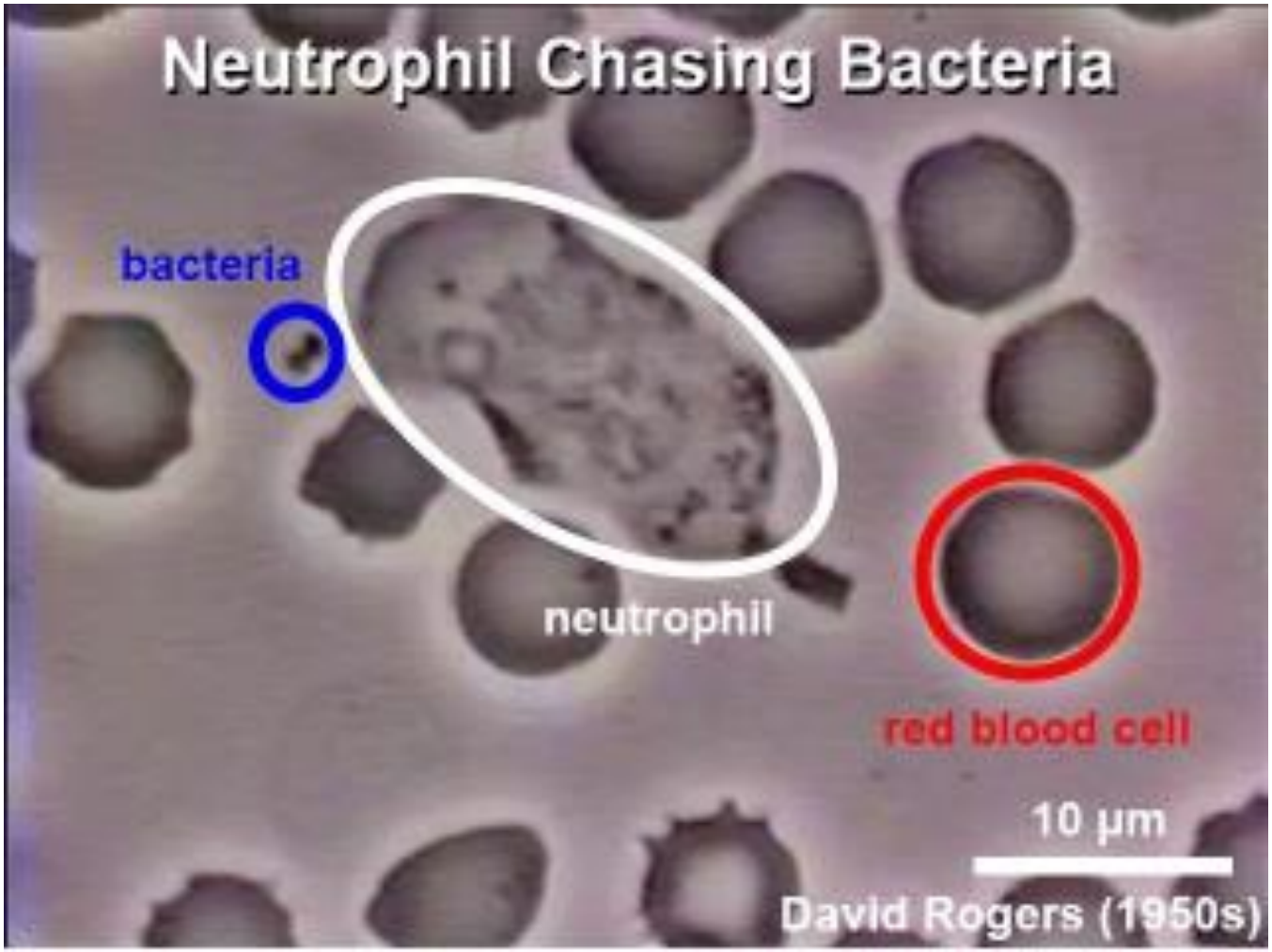
Neutrophil (white blood cell) chasing bacteria

sensing, information processing, decision making



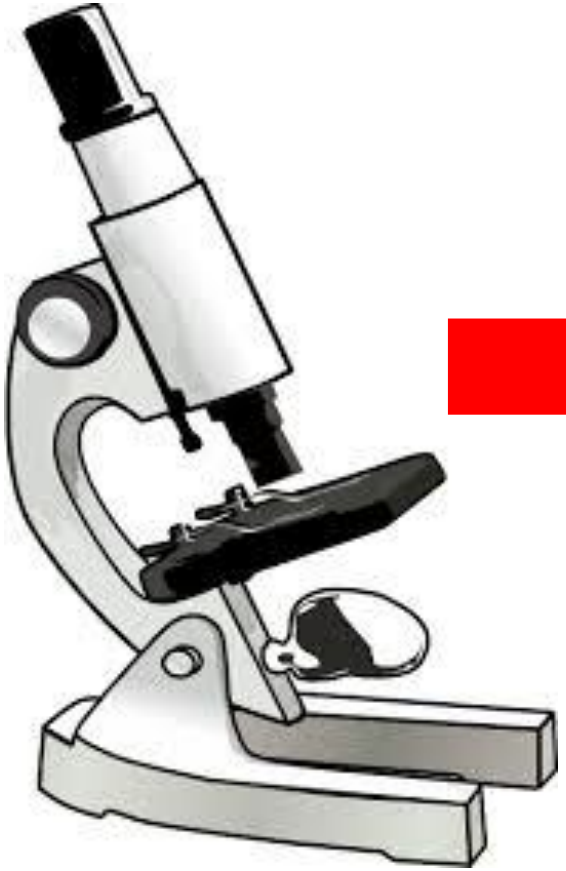
Original movie made in the 1950s by the late David Rogers at Vanderbilt University

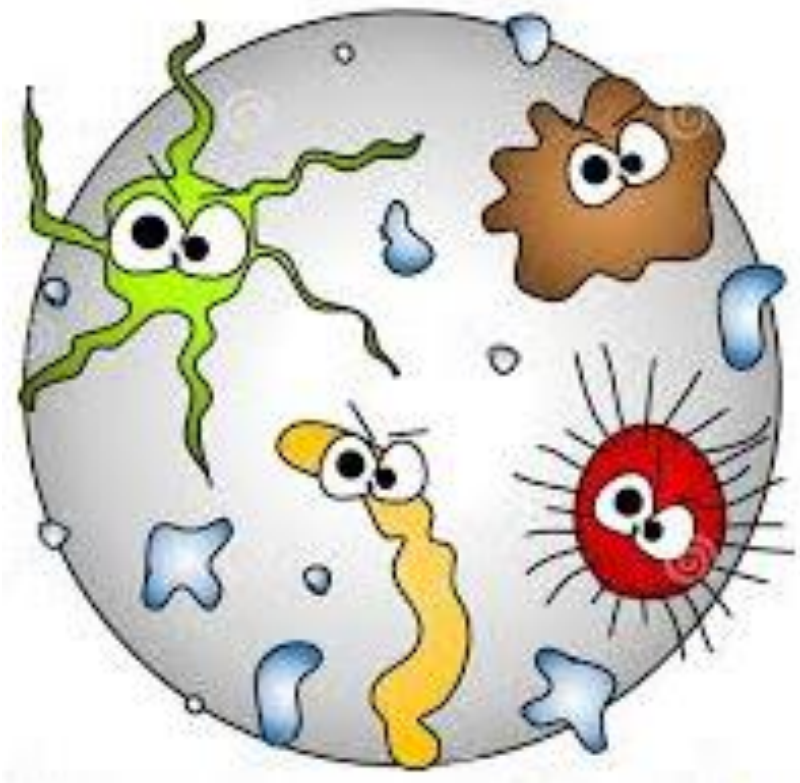
Sensing, information processing, decision making



Original movie made in the 1950s by the late David Rogers at Vanderbilt University

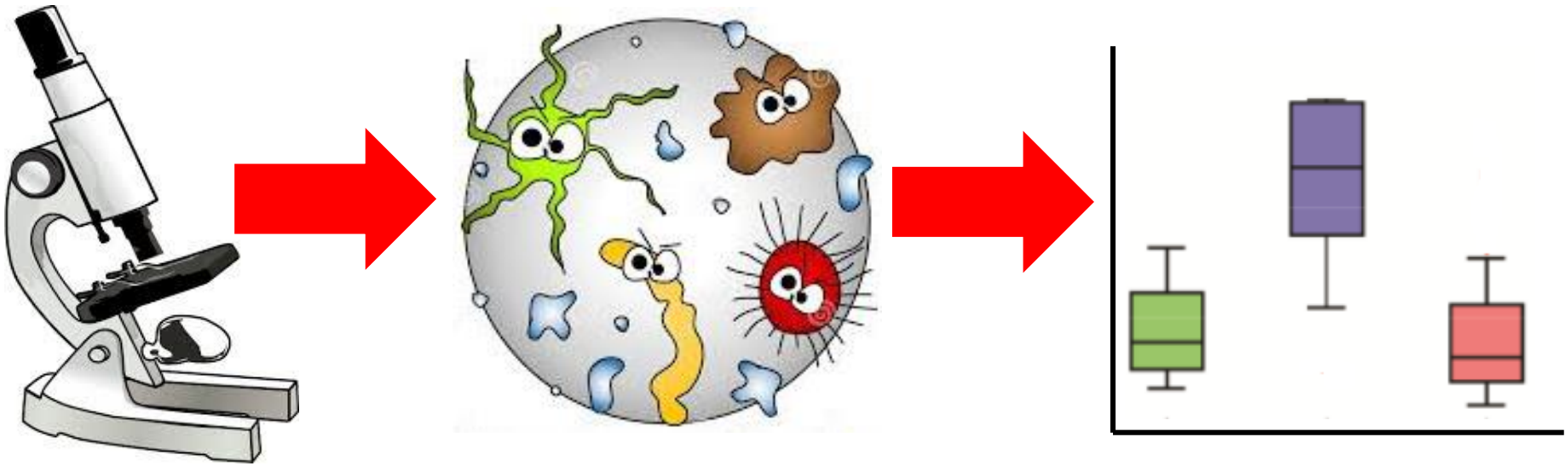
Seeing is believing





John Clarke (1639)

“Seeing is believing, quantifying is convincing”



quantitatively test hypotheses based on observations

Sbalzarini (2016)

Bioimage analysis tools

ImageJ

Image Processing and Analysis in Java

CellProfiler

cell image analysis software

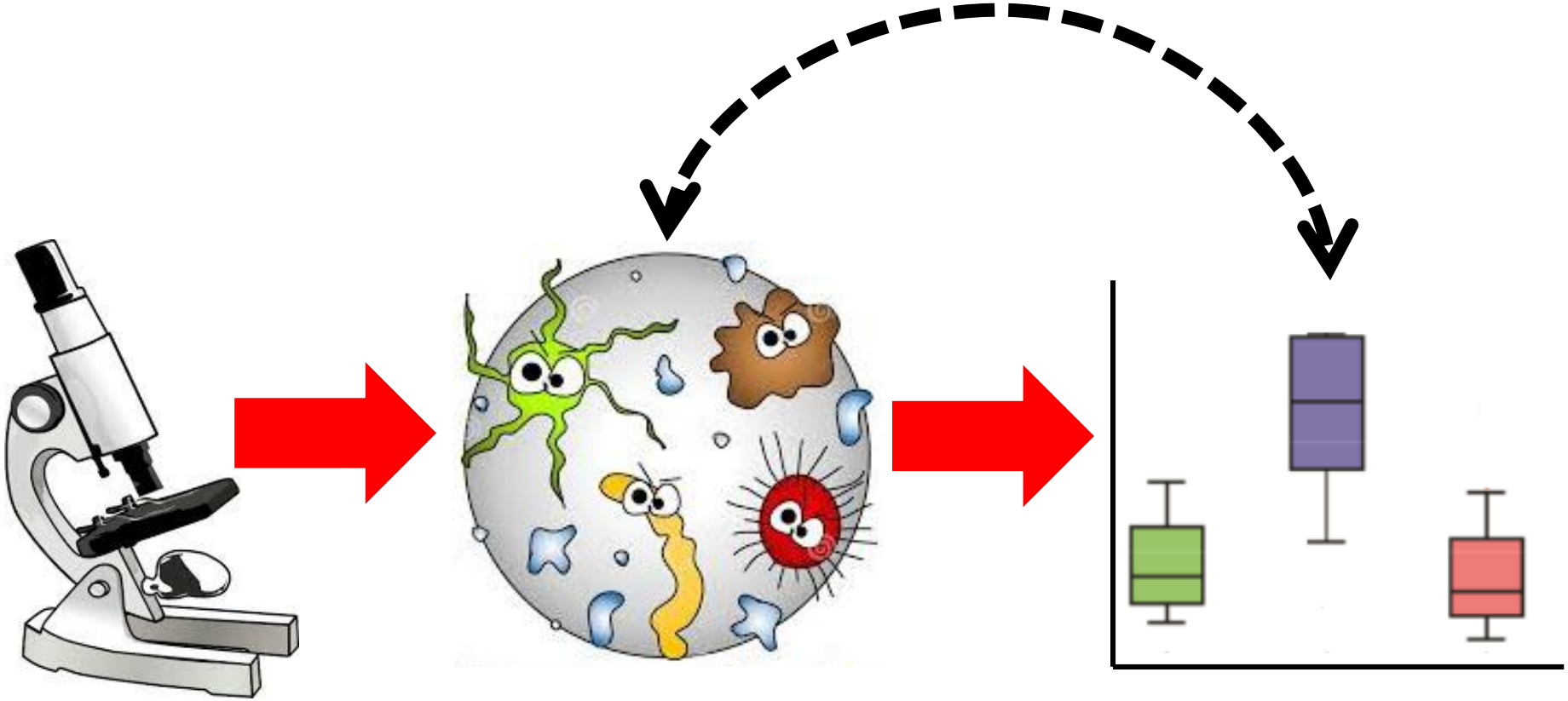
FEI

part of Thermo Fisher Scientific

BITPLANE

SCIENTIFIC SOLUTIONS

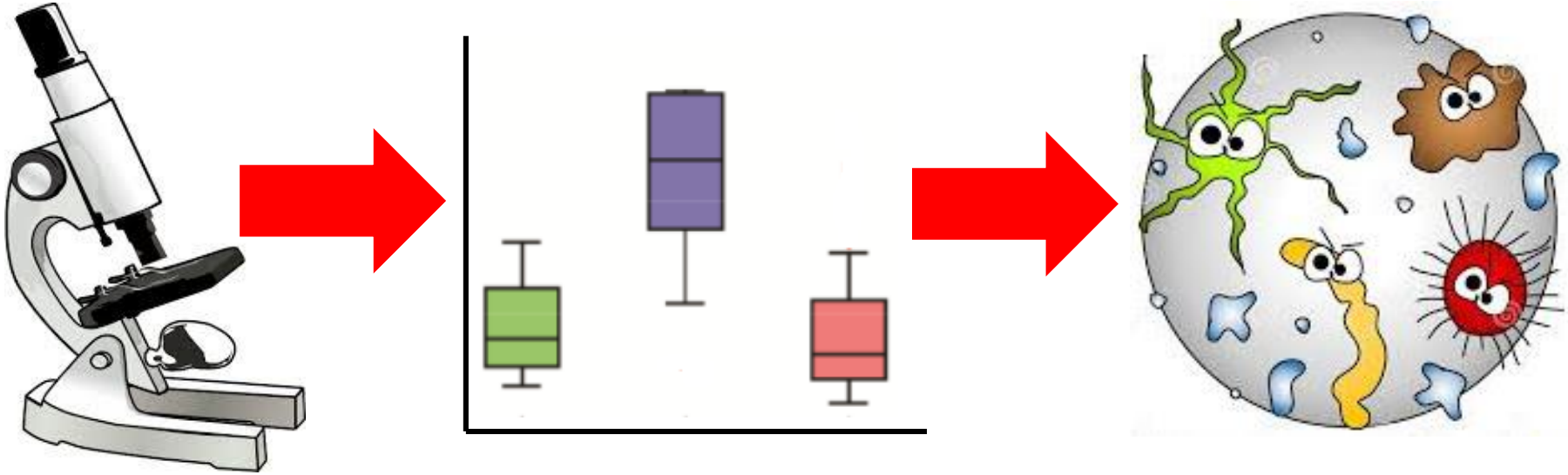
Quantifying the invisible
(and then, sometimes, seeing it)



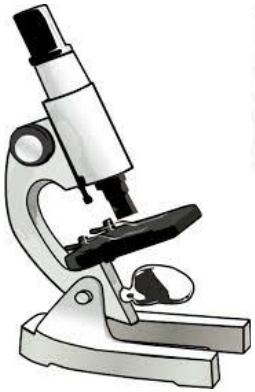
Sbalzarini (2016)

Quantifying the invisible
(and then, sometimes, seeing it)

Domain knowledge



Automation



Completeness

Invisible patterns

Danuser (2011)

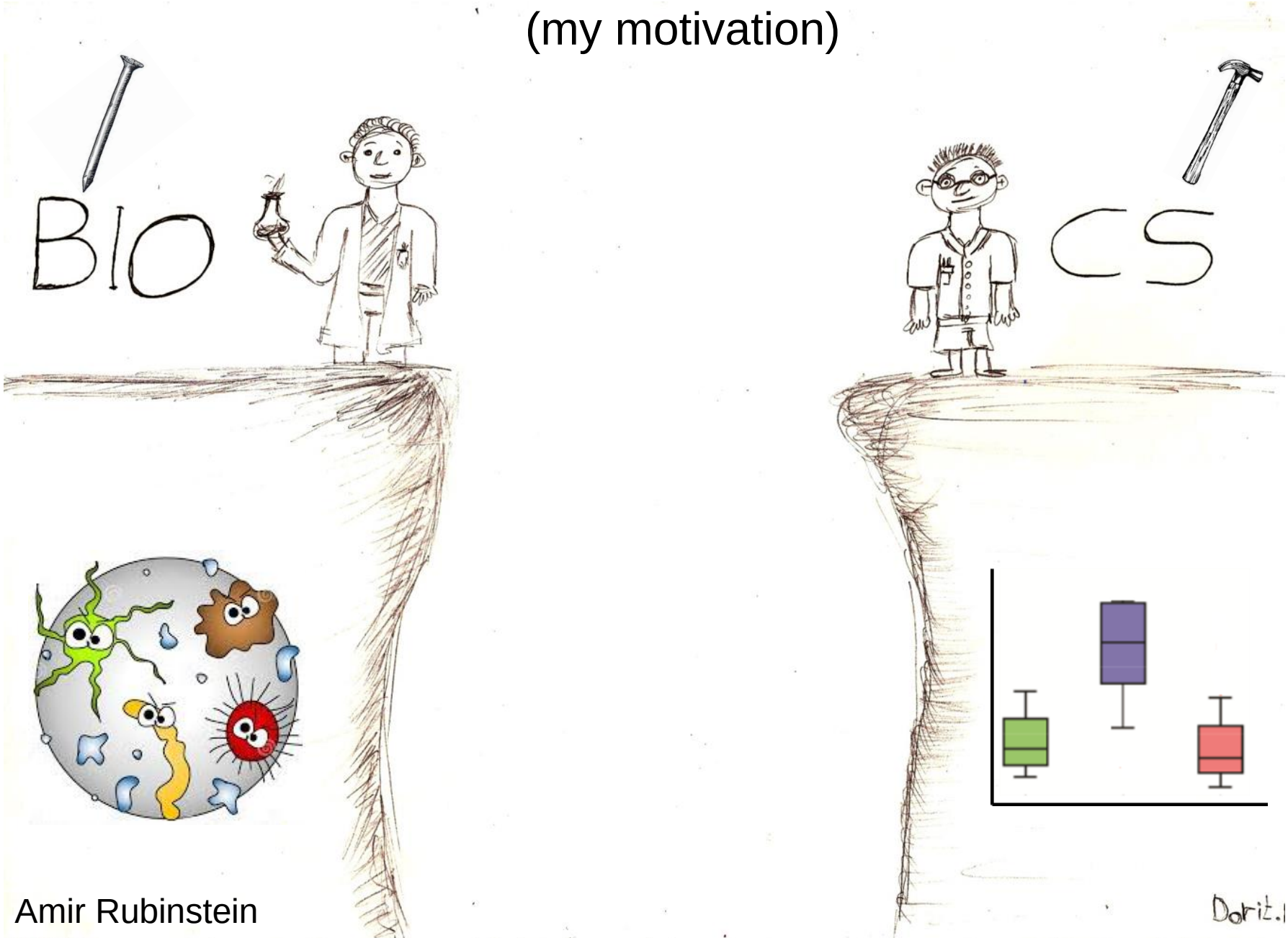
Lab of computational cell dynamics (and the focus of this course)

Motivated by fundamental questions in cell biology our lab produces biological insights along with specialized analytic tools that reveal hidden patterns in dynamic cell imaging data

Course objectives and admin.

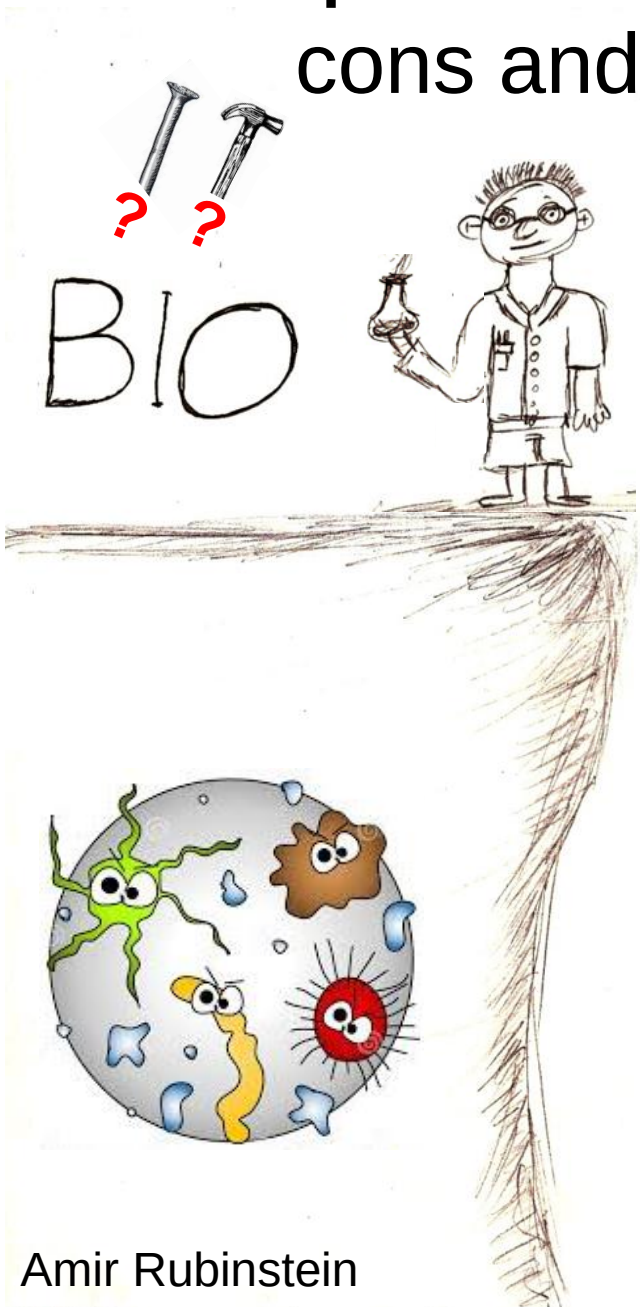
Course objective

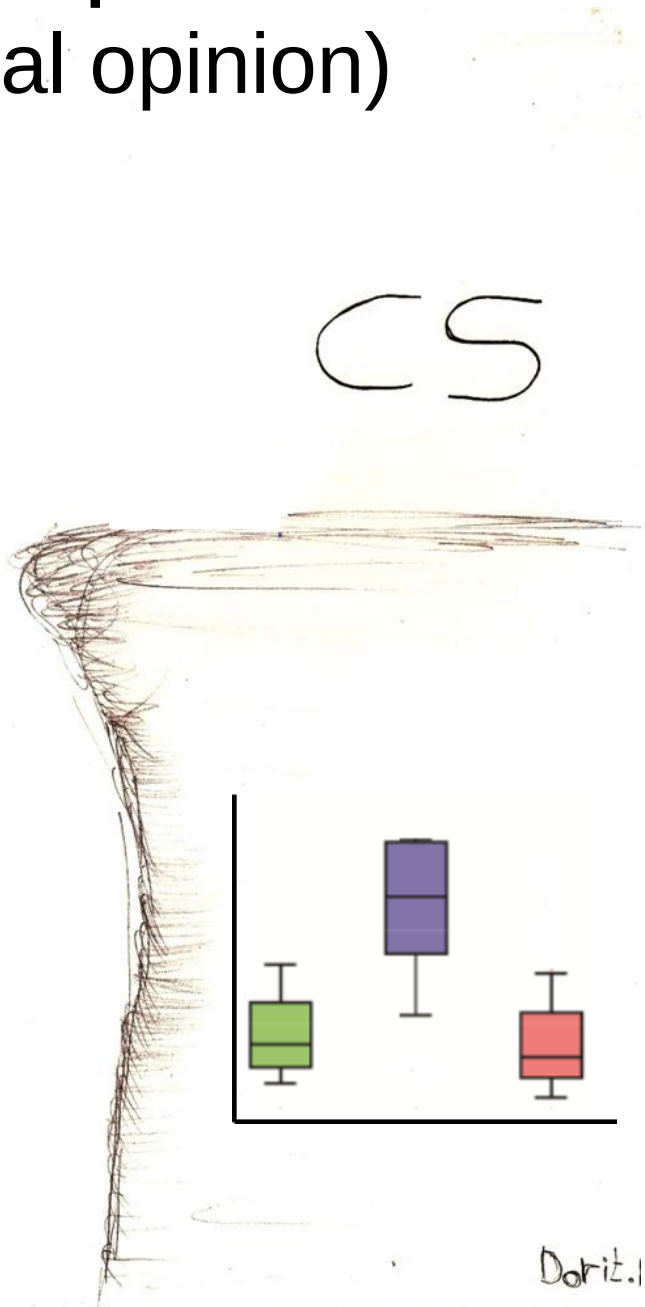
(my motivation)



Amir Rubinstein

Experiments and computation:
cons and pros (personal opinion)

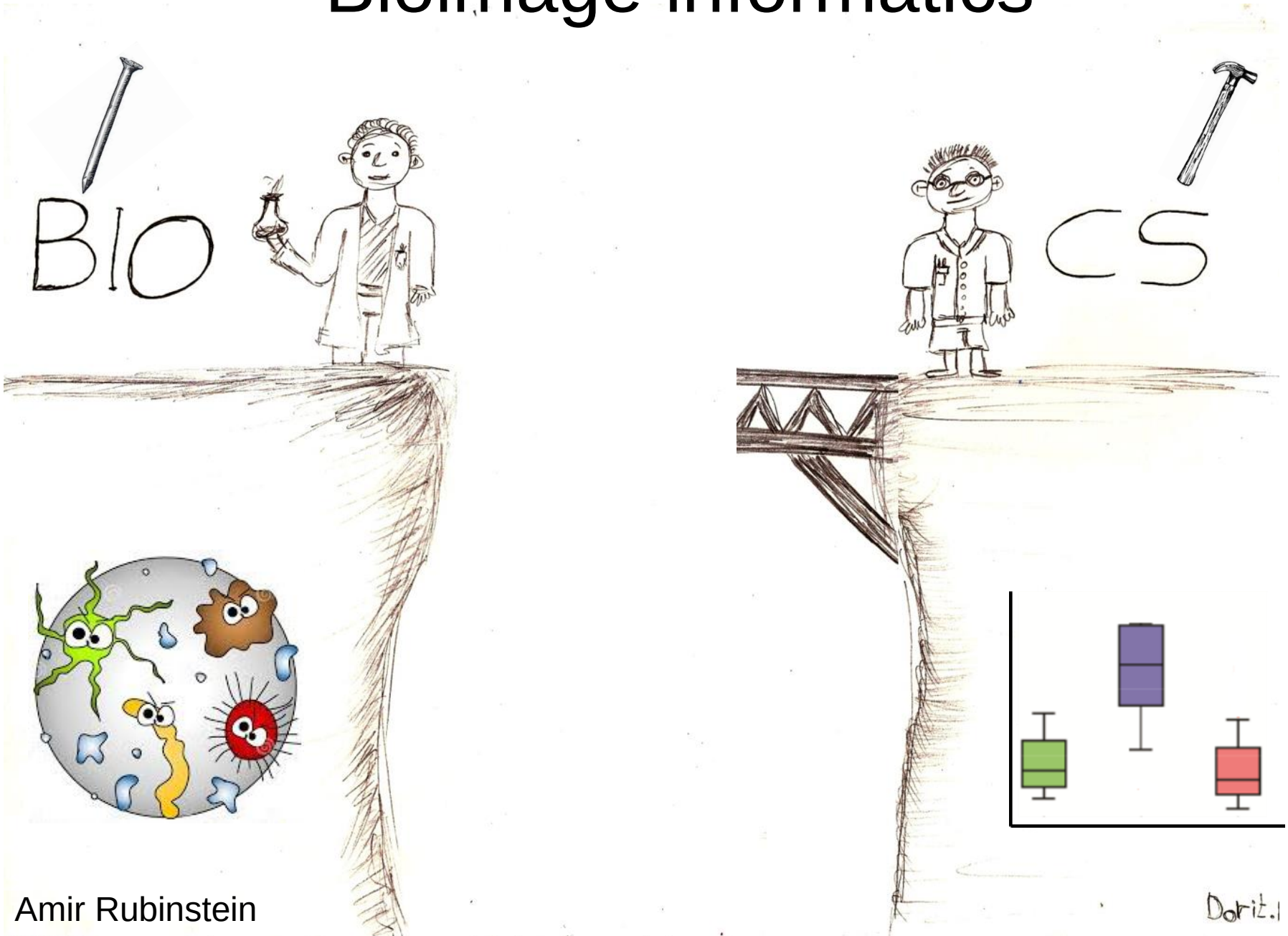




Amir Rubinstein

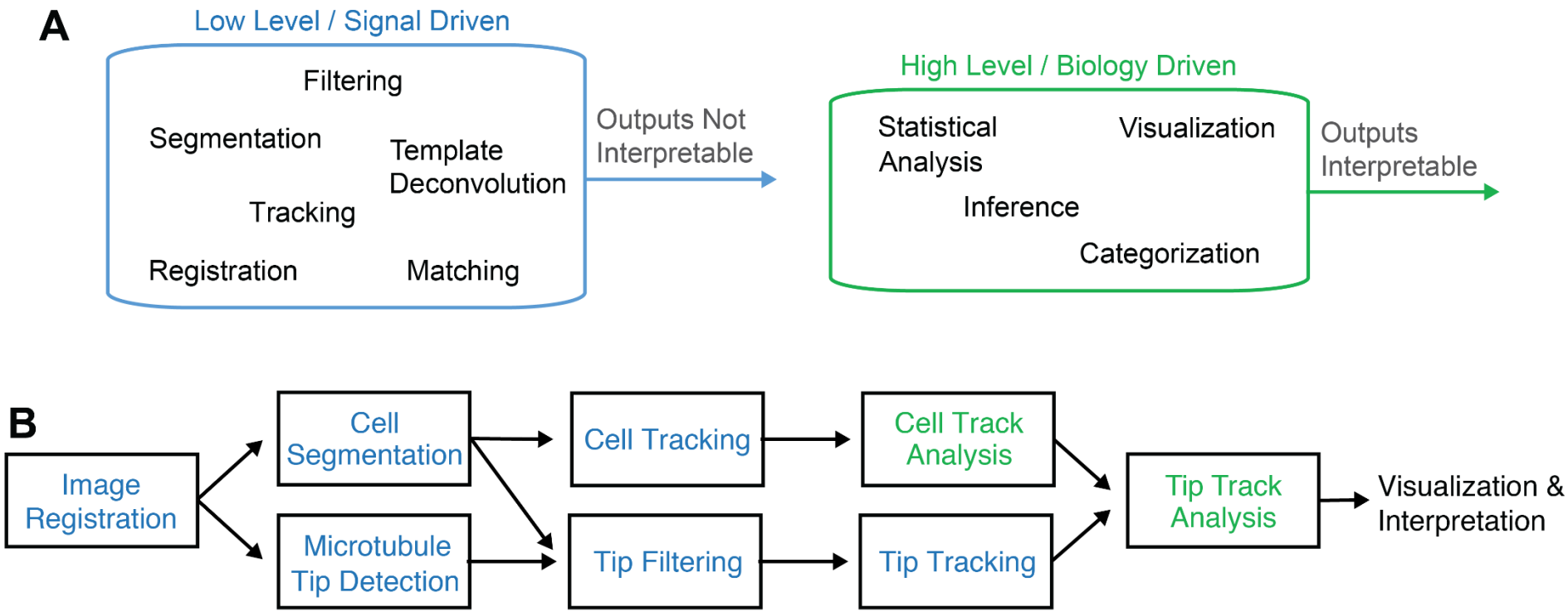
“The ultimate goal is to **replace human labor as much as possible with computer calculations**, so that **biologists can focus fully on formulating high-level hypotheses and designing increasingly sophisticated experiments**, while improving the objectivity and reproducibility of these experiments.” (Meijering et al., 2016)

Bioimage informatics



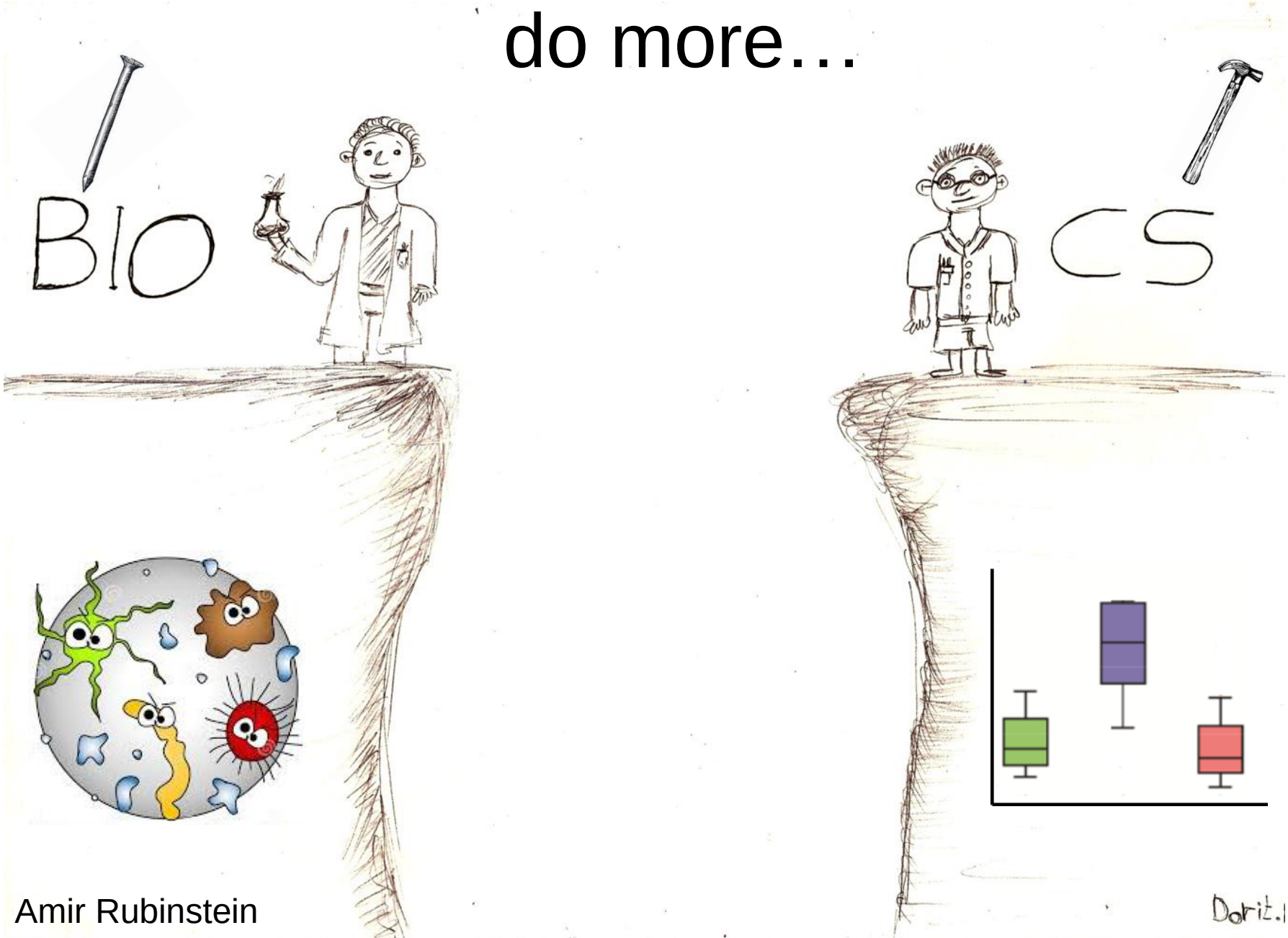
Amir Rubinstein

Hierarchies of data processing in microscopy



Driscoll and Zaritsky (2021)

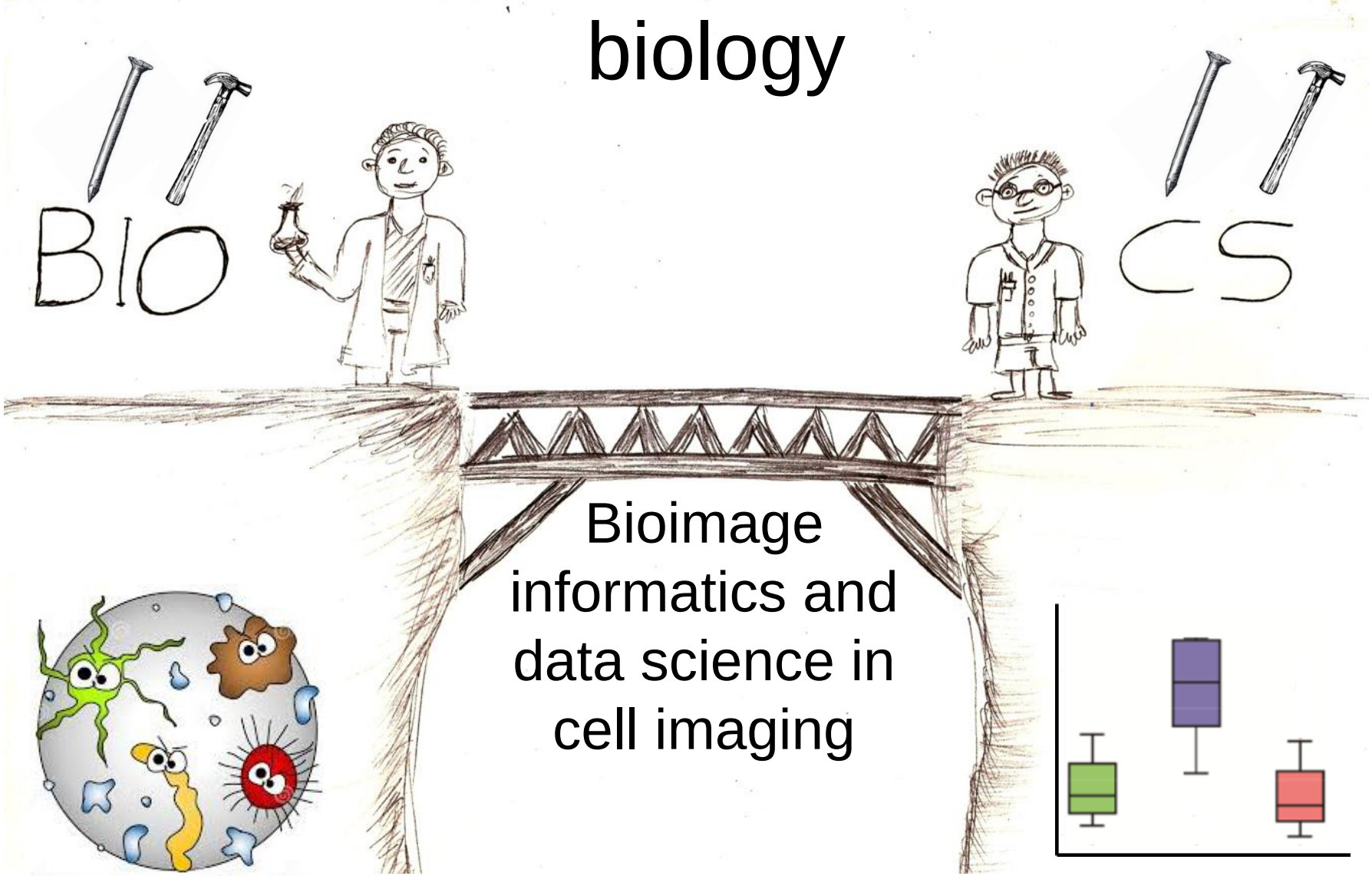
We (computational scientists) can do more…



Amir Rubinstein

The interplay between development of quantitative tools ("hammers“) and identifying open important questions in cell biology ("nails“)

Bridging computer science and cell biology



Read: https://www.nature.com/news/cell-biologists-should-specialize-not-hybridize-1.20277

Amir Rubinstein

Course description

- Presenting a broad set of (microscopy-driven) **biological questions** and the diverse set of **computational tools** to solve them
- Not going deep into the biology nor into the computational aspects (unless students push for it)
- Computational aspects include techniques from computer vision, machine learning, time series analysis, network algorithms, etc.
- Full disclosure: topics selection is biased toward my personal interests (and research))

Grades

- Paper presentation – 20%
- Final research project – 80%
  - In groups (1-3 students)
  - Open data exploration / collaboration / tool building/validation/comparison
  - “Freestyle”
  - Project presentation day
  - Written report
  - Each participant will clearly articulate their contribution
- Grades

Research project types

- Data mining (and/or integration) in public repositories
- Data mining of datasets explored in the lab
- Collaborative bioimage analysis projects
- Tool building projects

Course (tentative) schedule and overview

Tentative topics list
(we will not cover all!)

- Cell biology and microscopy (Natalie Elia, BGU)
- Bioimage informatics (Jean Yves Tinevez, Pasteur)
- Deep learning in microscopy
- High content single cell phenotypic profiling
- Advanced representations of cell shapes, intracellular organization and trajectories data
- Atlases and public data repositories
- Information processing in multicellular systems
- Computer vision in cell imaging (Tammy Riklin Raviv, BGU)
- Importing ideas from systems biology (Tal Shay, BGU)
- Integrating microscopy and omics (Paul Villoutreix, CENTURI)
- Misc. topics: reusing cell image data, automated microscopy, high content simulations, medical imaging, 3D image analysis, visualization

Tentative schedule

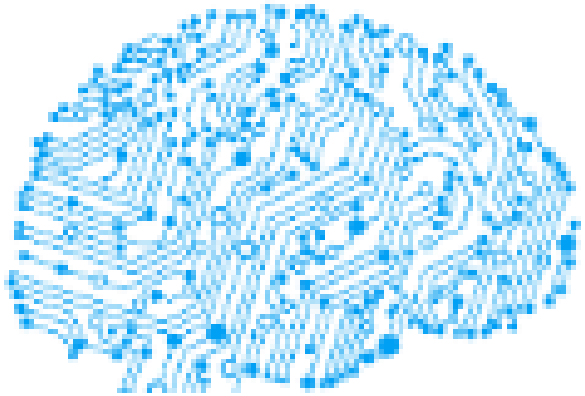
Class #	Topic
1	Introduction to data science in cell imaging
2	Introduction to cell biology & microscopy
3	Deep learning in microscopy
4	Deep learning in microscopy
5	Deep learning in microscopy
6	Image-based phenotypic profiling
7	Image-based phenotypic profiling
8	Advanced representations of cell shapes, intracellular organization and trajectories data
9	Atlases and public data repositories
10	Information processing in multicellular systems
11	Bioimage informatics, spatiotemporal analysis
12	Importing ideas from systems biology
13	Misc. topics

Misc. topics may include: reusing cell image data, computer vision in cell imaging, data harmonization, integration and fusion, automated microscopy, high content simulations, medical imaging

Cell biology (microscopy) is behind, in terms of the application of modern computer vision – an opportunity!

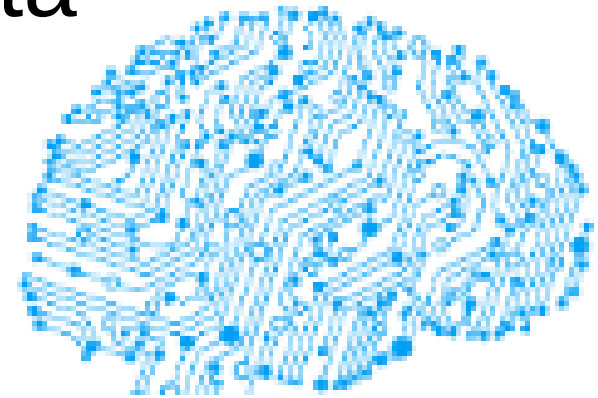
Most computer vision tasks

>



Identifying patterns in complex cell image data

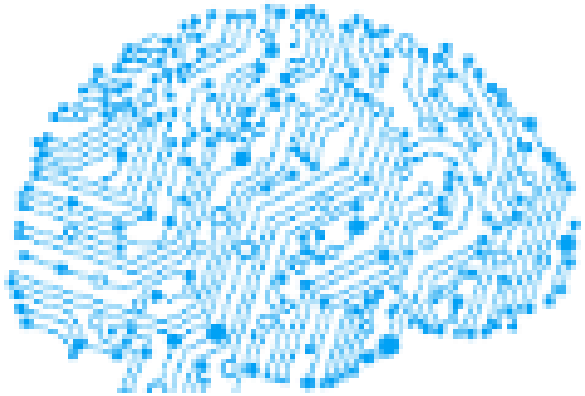
<



Cell biology (microscopy) is behind, in terms of the application of modern computer vision – an opportunity!

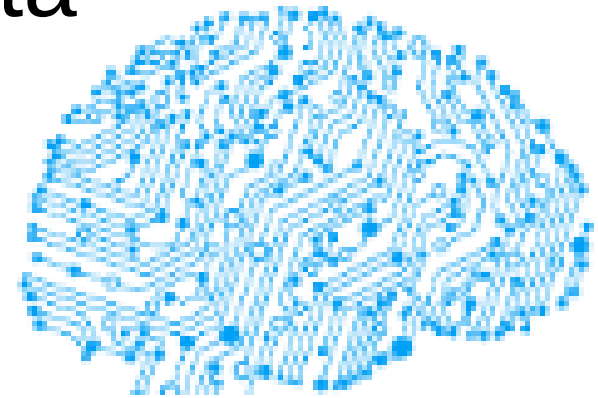
Most computer vision tasks

>



Identifying patterns in complex cell image data

<



Research outcome

Algorithmic elegance and performance

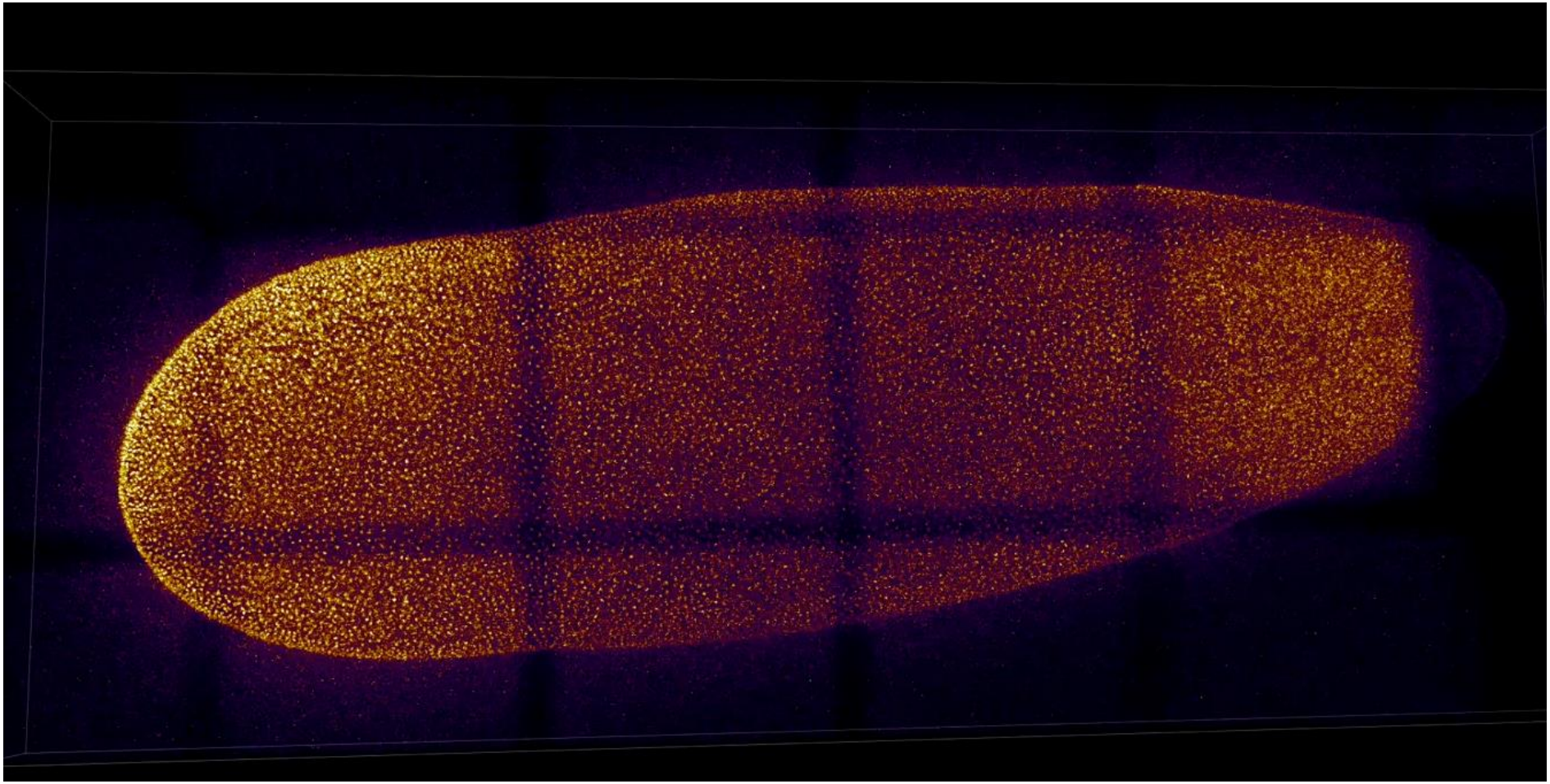
versus

discovery!

What is special about cell image data? - discussion

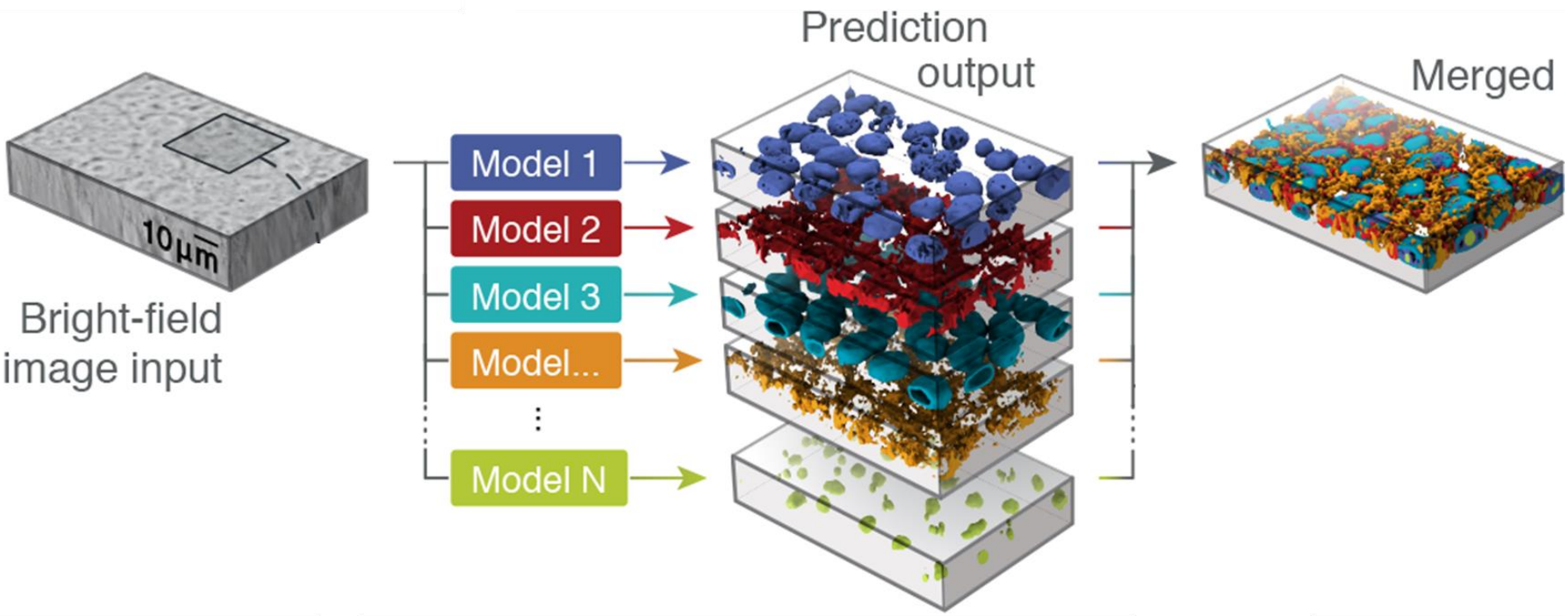
Brief browsing through <u>some</u> of the course highlights

Content aware image restoration



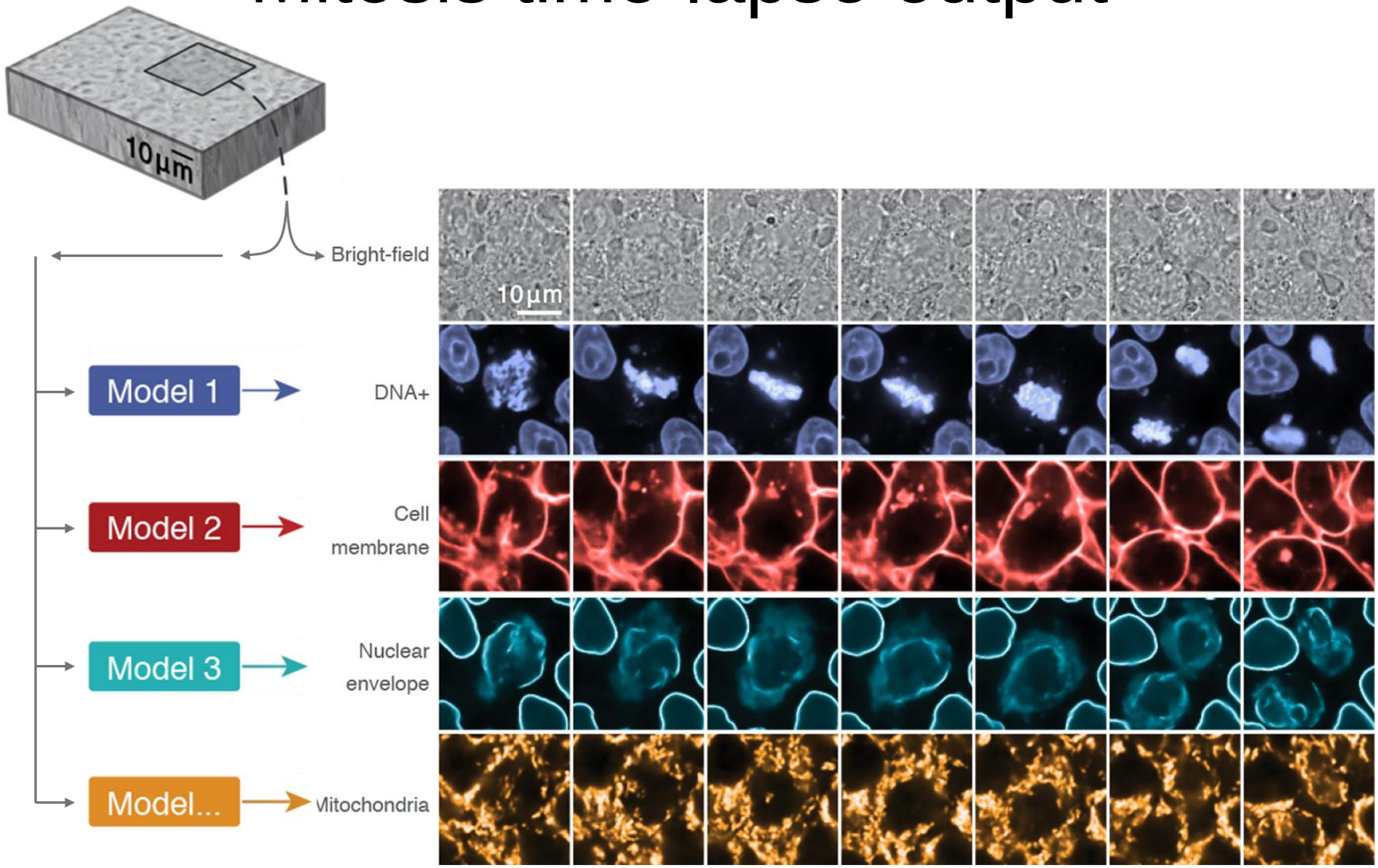
Weigert et al. (2018)

Label-free images contain information on the molecular organization of the cell



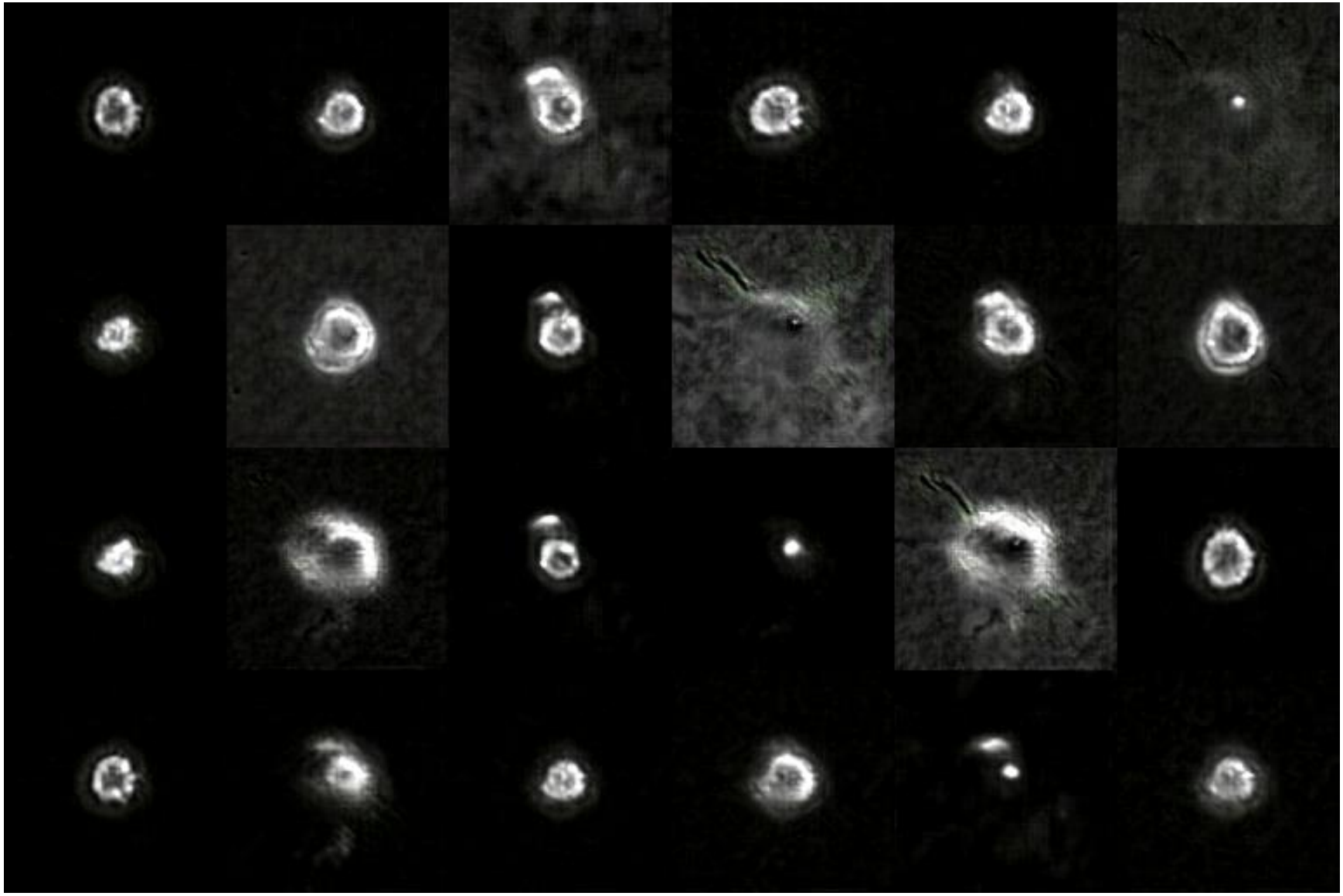
Greg Johnson

Mitosis time-lapse output



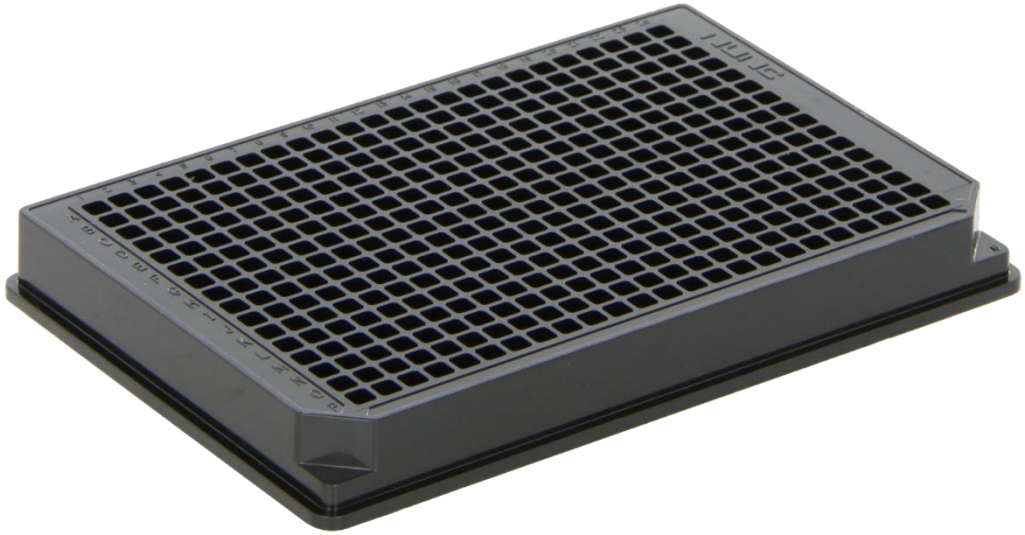
Greg Johnson

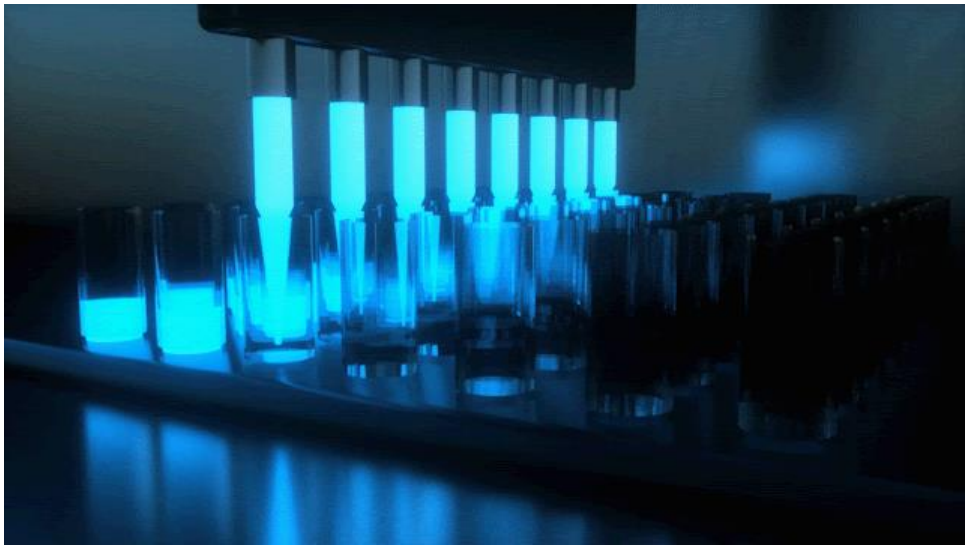
Interpretable machine learning for melanoma classification



Discovering drugs in high throughput

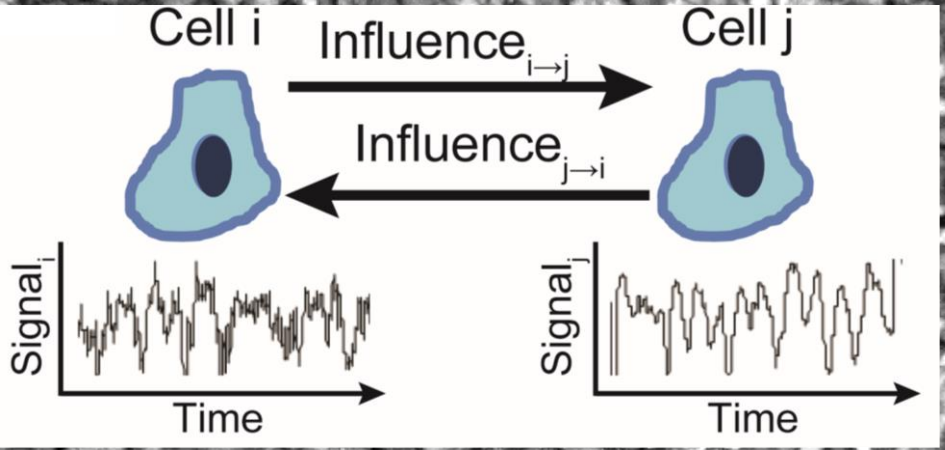
X 384 =

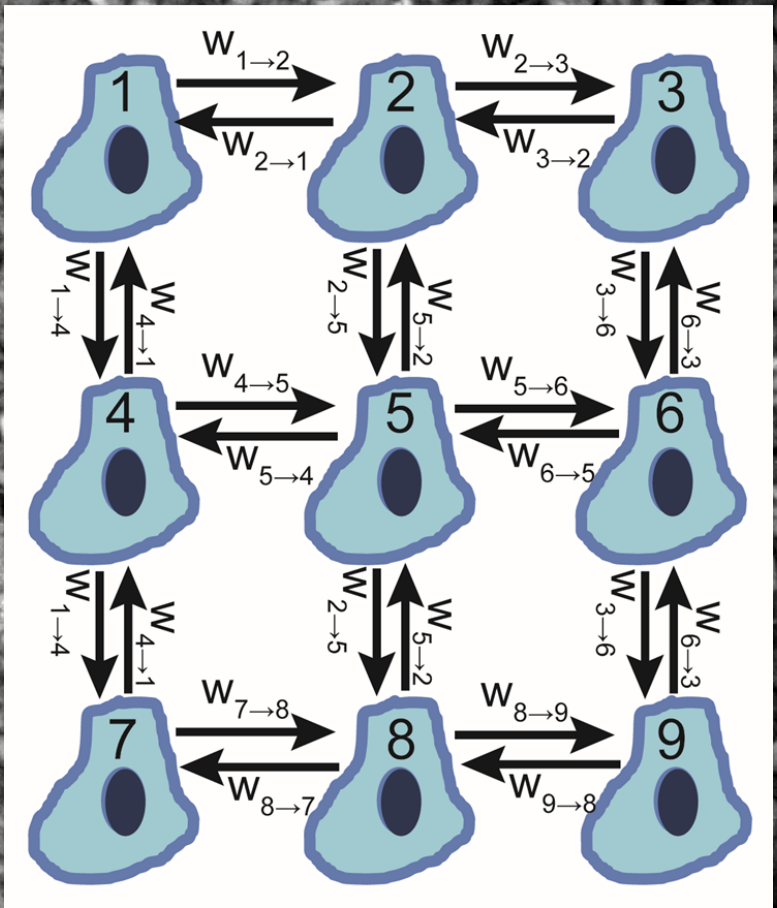




Anne Carpenter, Image: Nalgene; video: Chemistry World

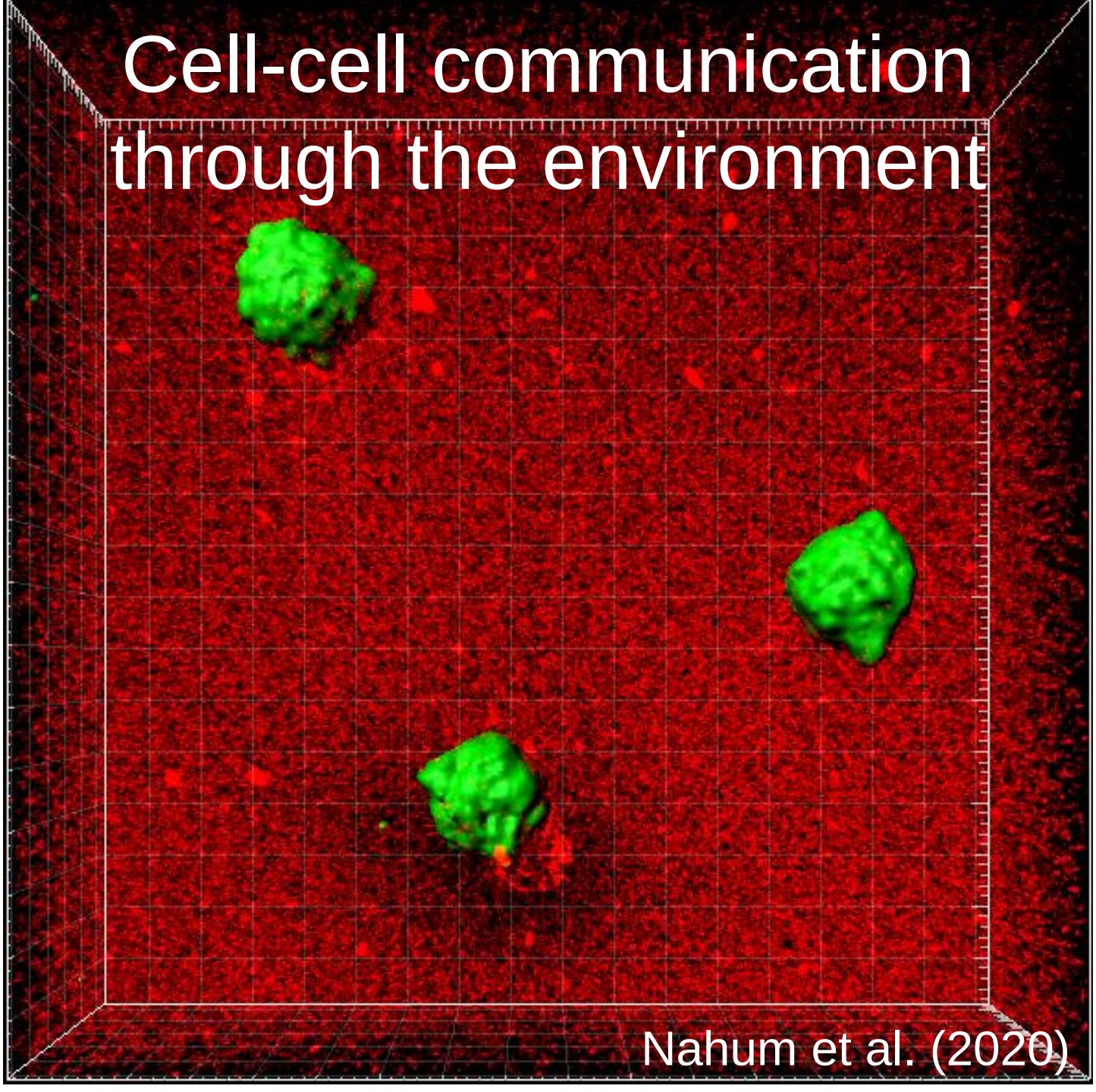
Modeling networks of spatial influence at the single cell resolution





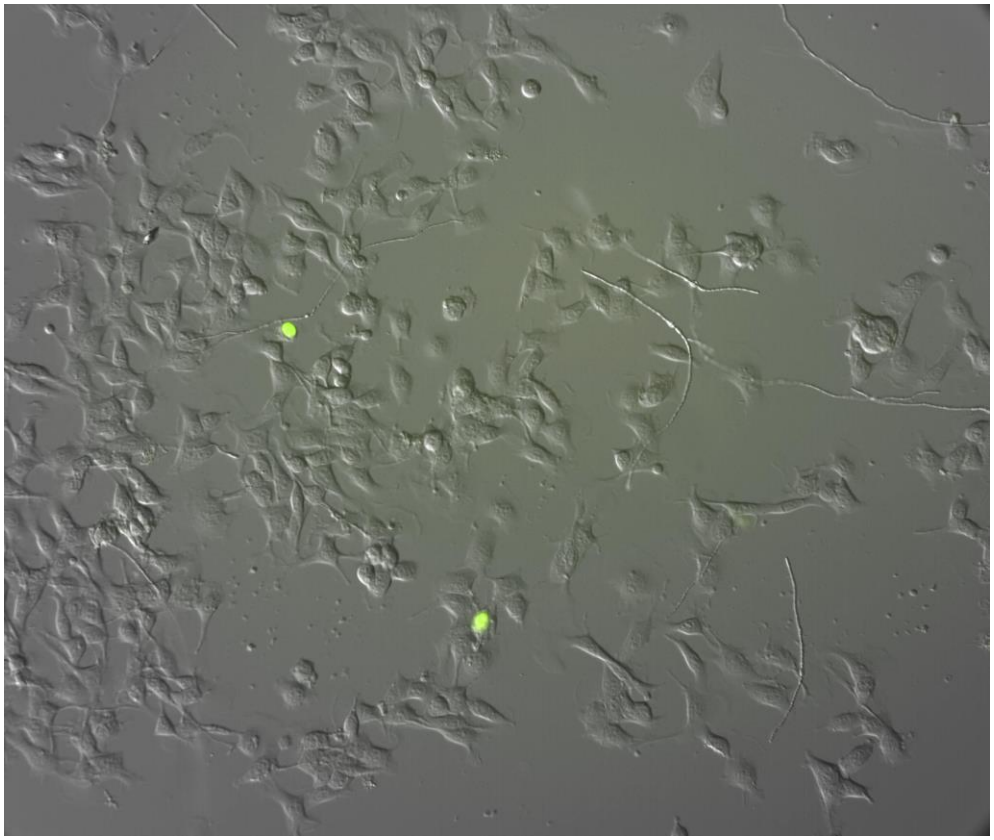
Zamir et al. (2020)

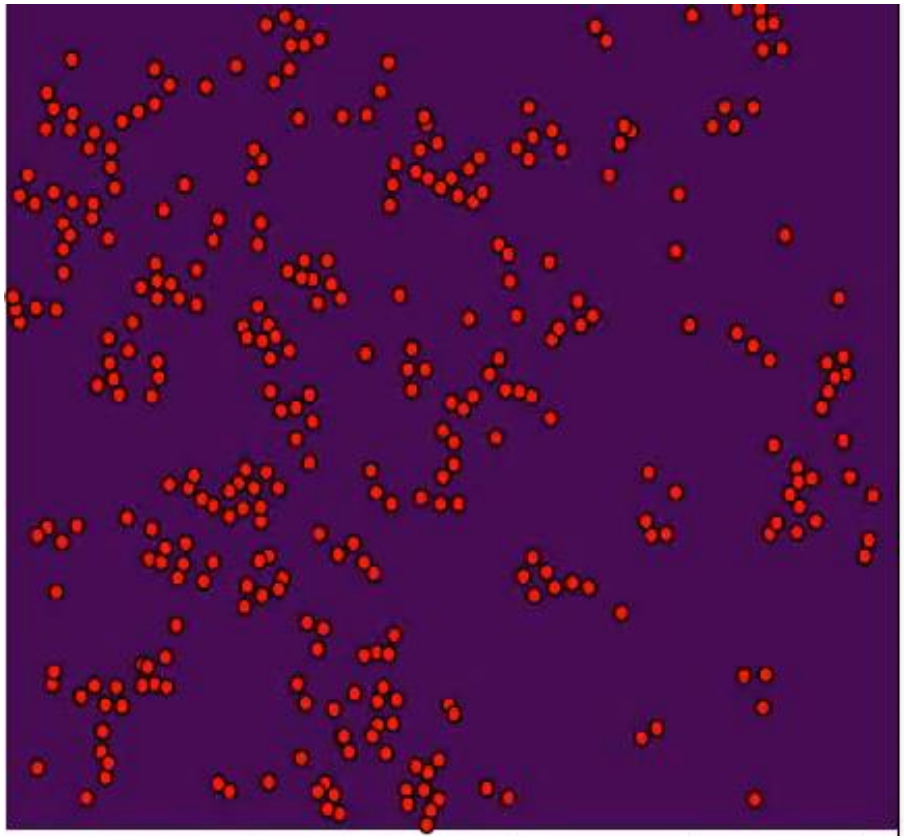
Cell-cell communication through the environment



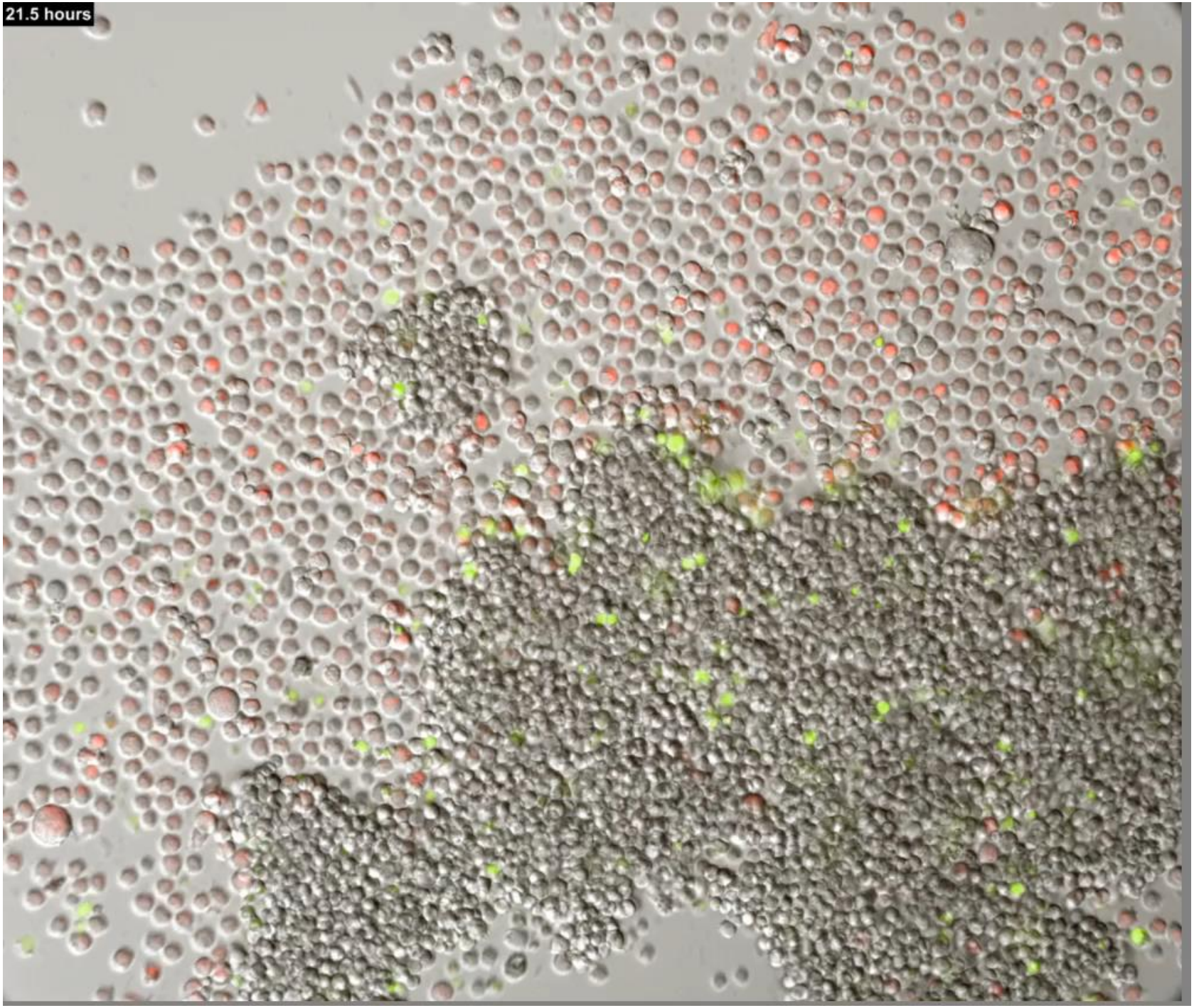
Nahum et al. (2020)

Collective cell death



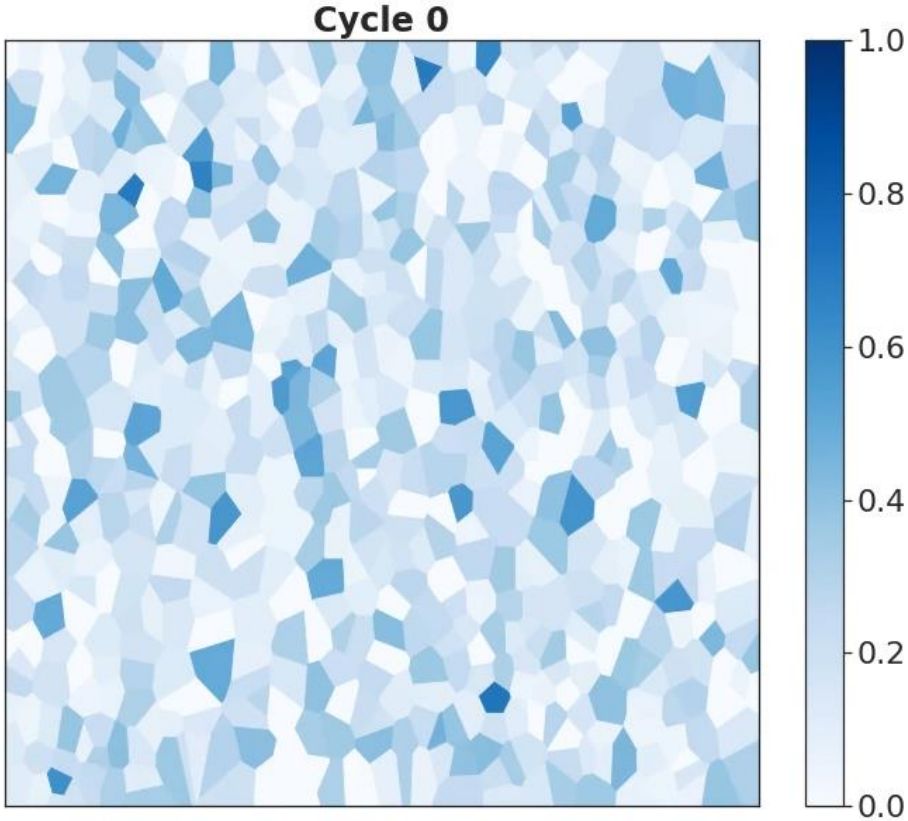


Yishaia Zabary

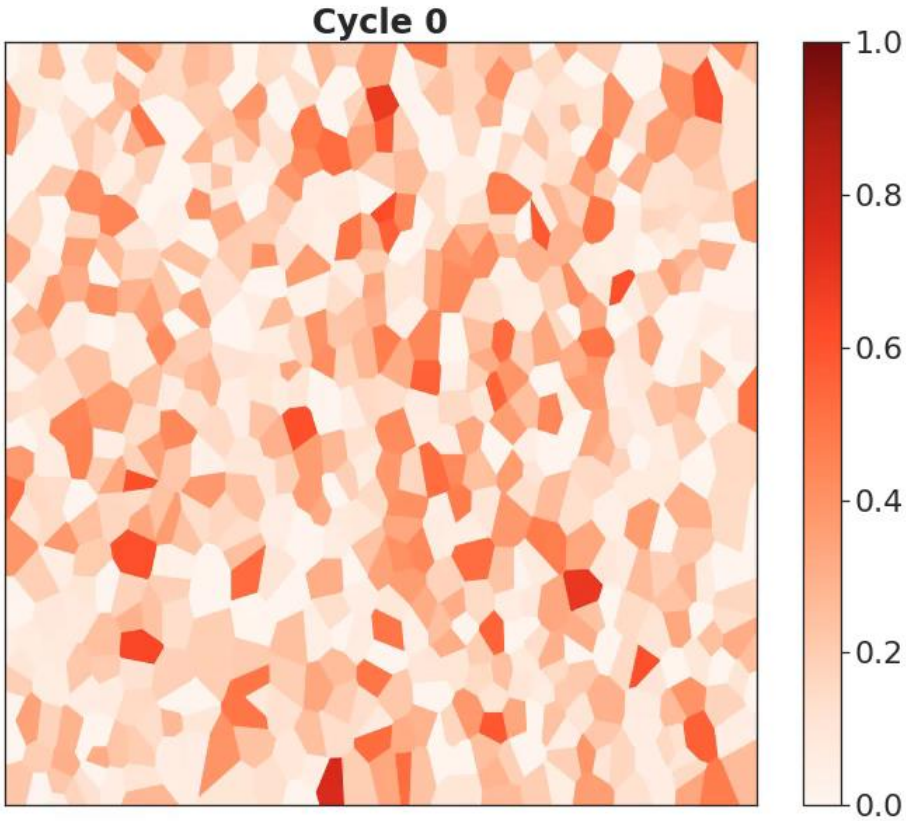

21.5 hours

Synchronized multicellular signaling

Transmission score



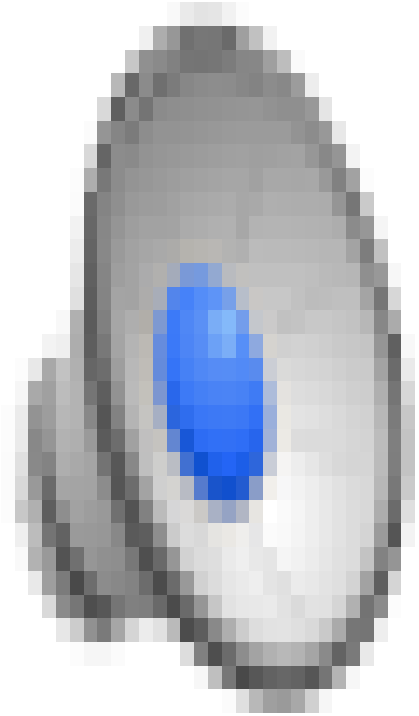
Receiver score



Amos Zamir

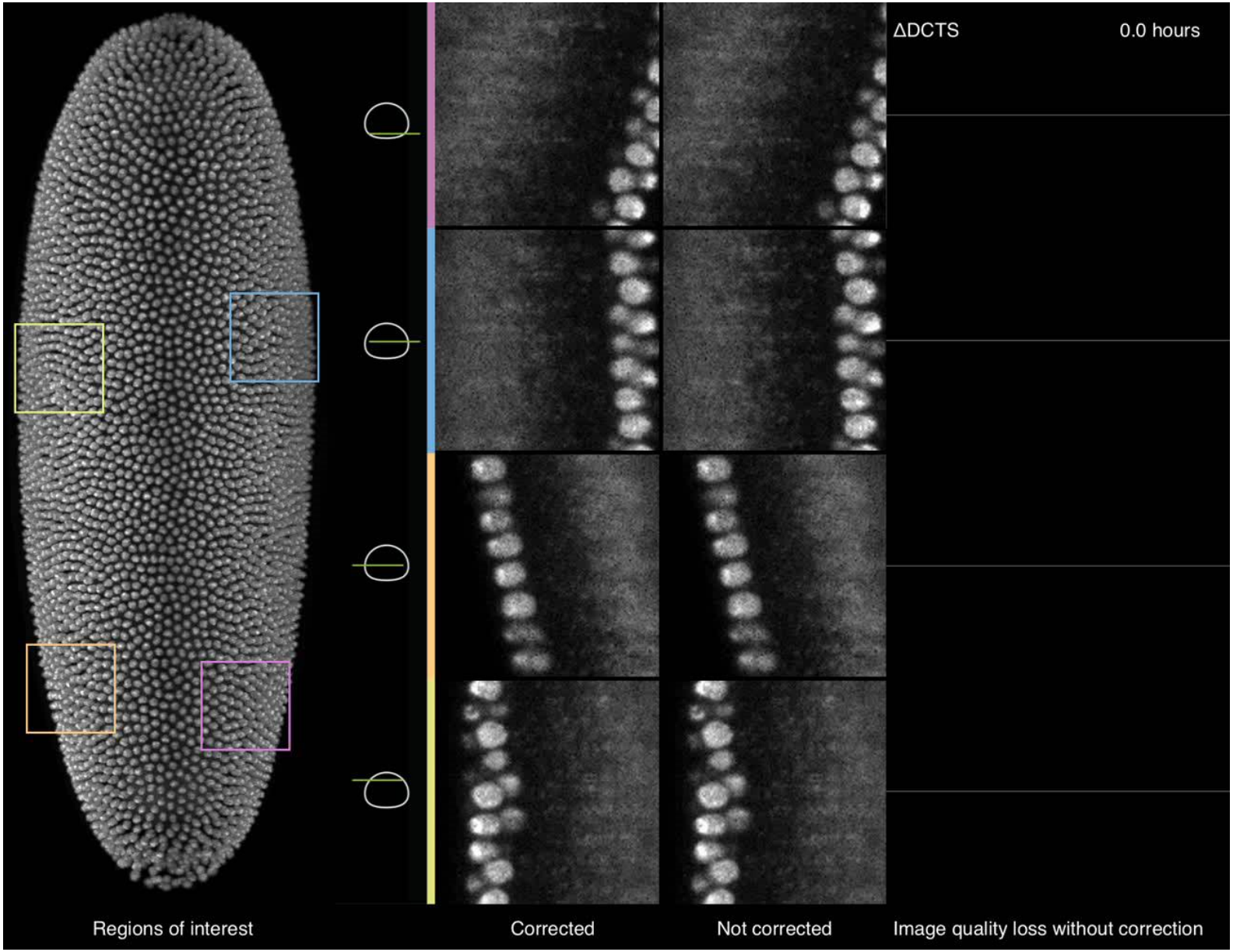
Cell shape in 3D

Immune cell in 3D



Renkawitz et al. (2019)

Adaptive (event-driven) microscopy

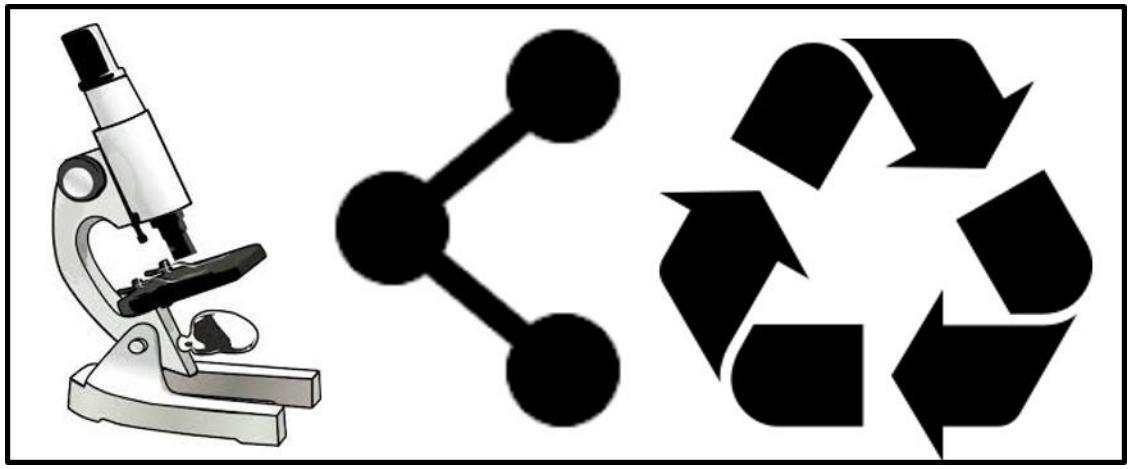


Royer et al. (2016)

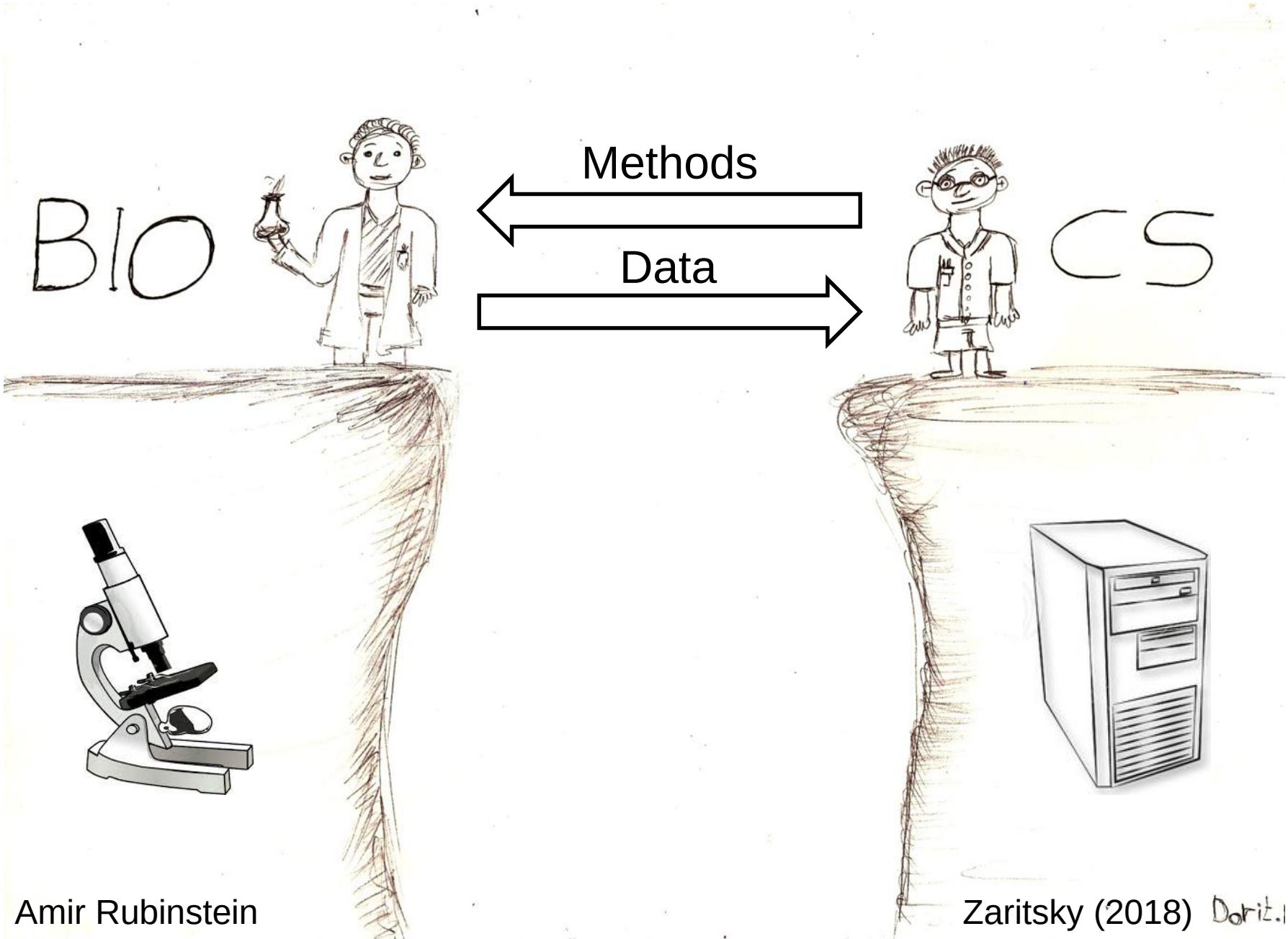
Image data resources and reuse

“The Great Wave off Kanagawa”, by Hokusai, ~1830 (Source: Wikipedia)

Reusing cell image data for new biological insight (and tool development, and reproducibility)



Zaritsky (2018)



Amir Rubinstein

Zaritsky (2018)

How?

Culture

- Open data
- Appreciation of secondary analysis

Infrastructure

- Repositories
- Data formats
- APIs
- Visualization

Data science

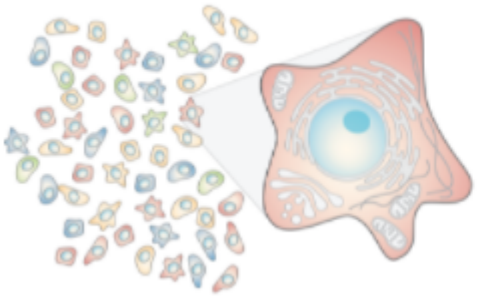
- Tool dev.
- Integration
- Data mining

Zaritsky (2018)

The time is now!

Massive data collection efforts for cell phenotyping

Several dataset discussed/mentioned in the course



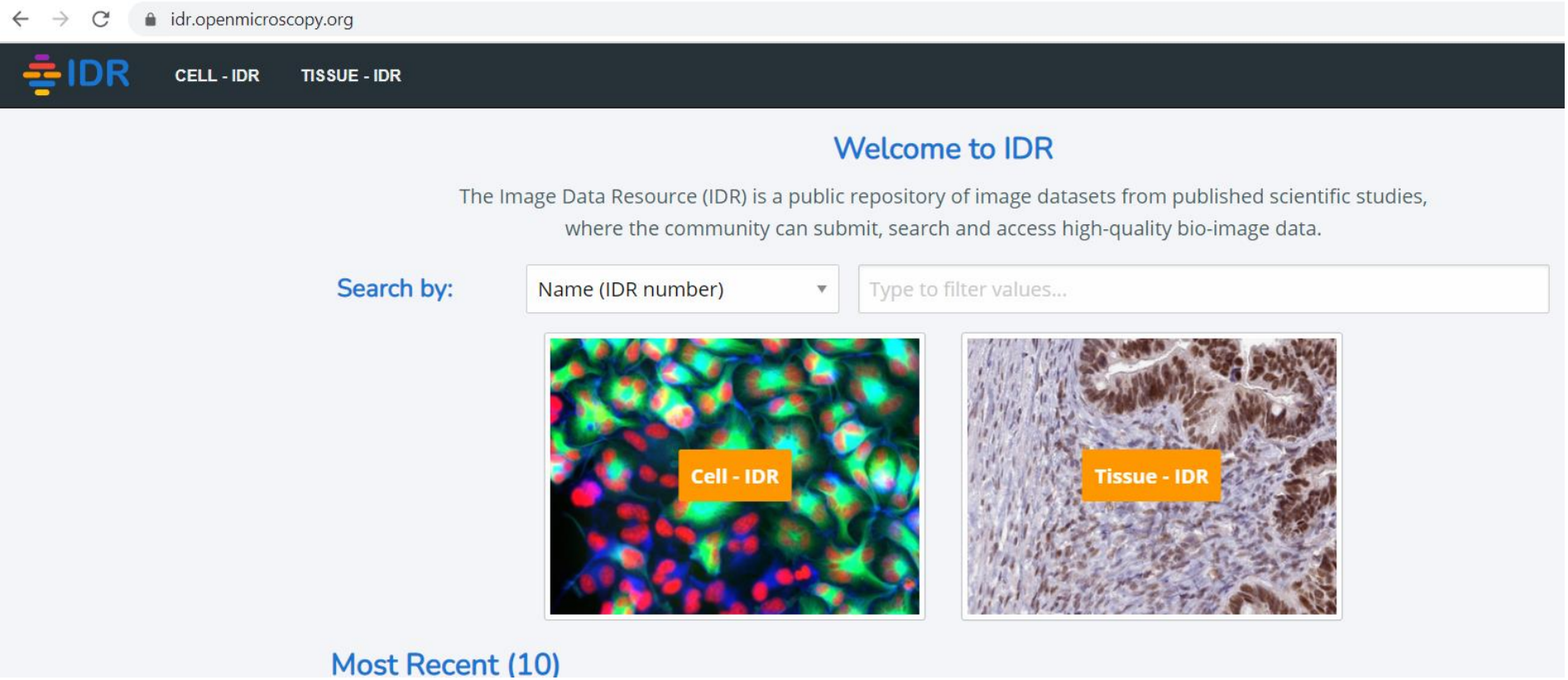
JUMP-Cell Painting Consortium

Joint Undertaking in Morphological Profiling

https://jump-cellpainting.broadinstitute.org/

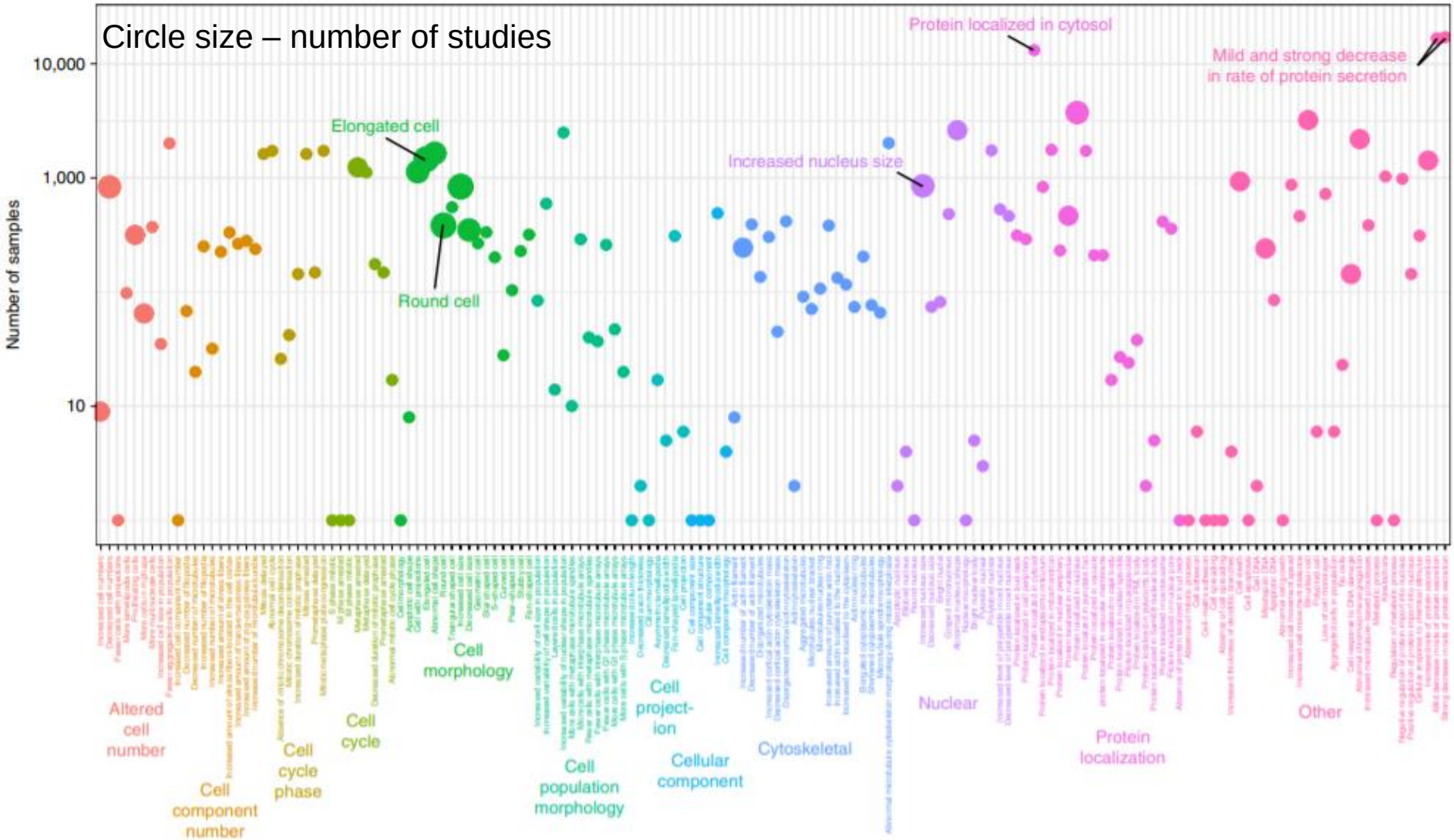
CPJUMP1: new resource of matched chemical and genetic perturbations, Chandrasekaran et al. (2021)

The image data resource (IDR)



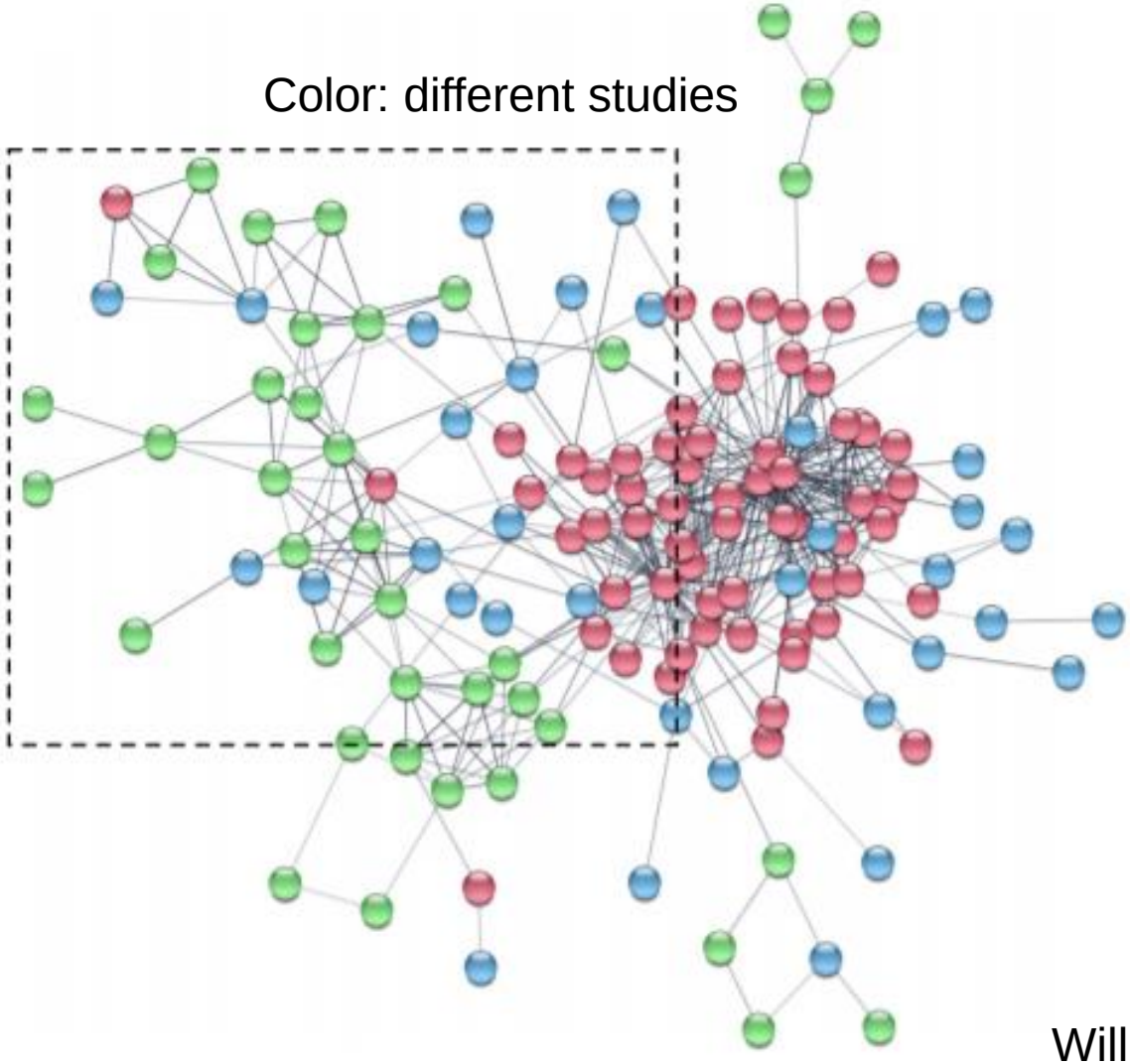
https://idr.openmicroscopy.org/

Sampling of phenotypes in the IDR



Williams et al. (2017)

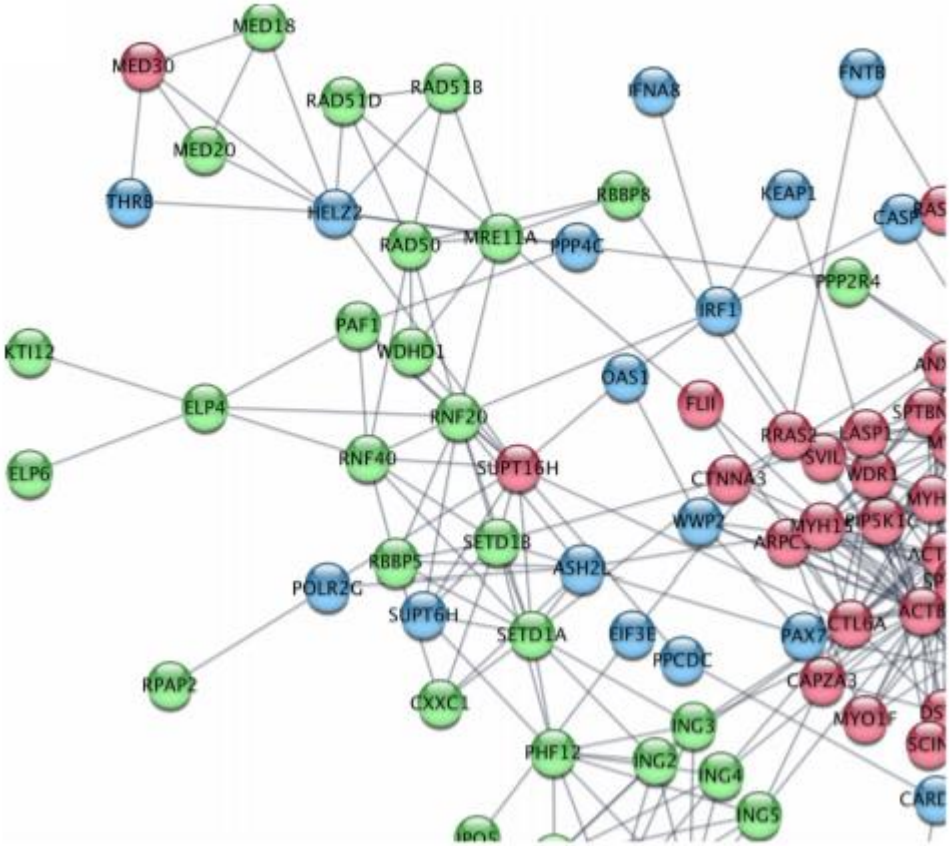
Network analysis of genes linked to the elongated cell phenotype in the IDR



Williams et al. (2017)

Network analysis of genes linked to the elongated cell phenotype in the IDR

Color: different studies



Williams et al. (2017)

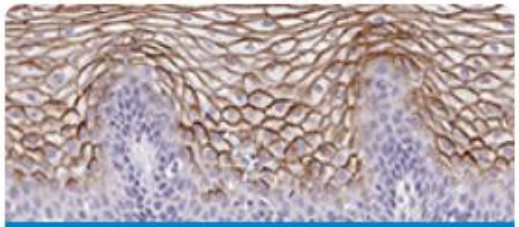
THE HUMAN PROTEIN ATLAS

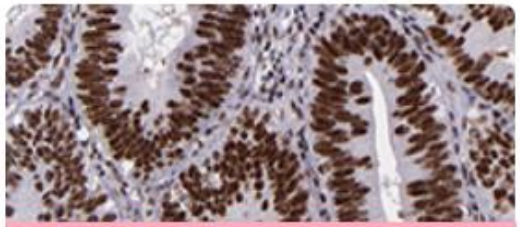
MENU HELP NEWS

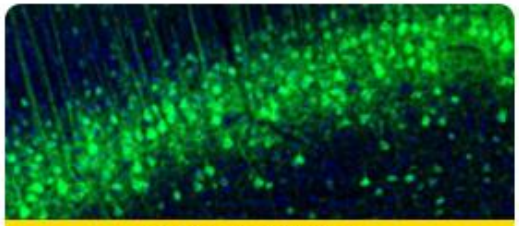
SEARCH

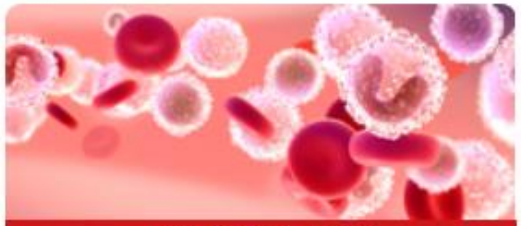
Search Fields »

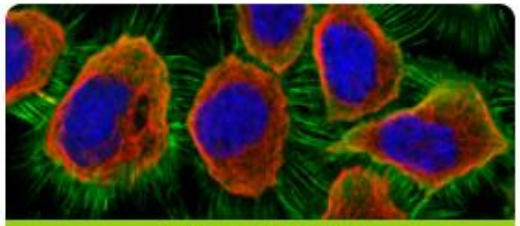
e.g. ACE2, GFAP, EGFR











https://www.proteinatlas.org/

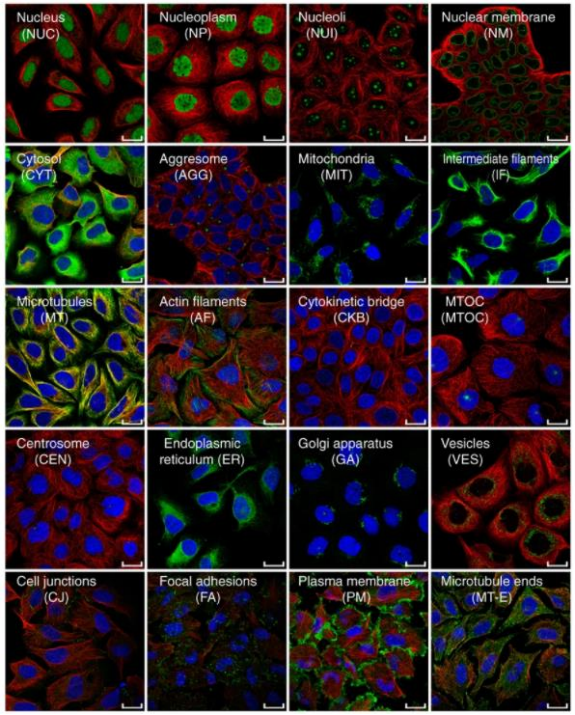
Citizen science in cell biology

322,006 players

Training/tutorial

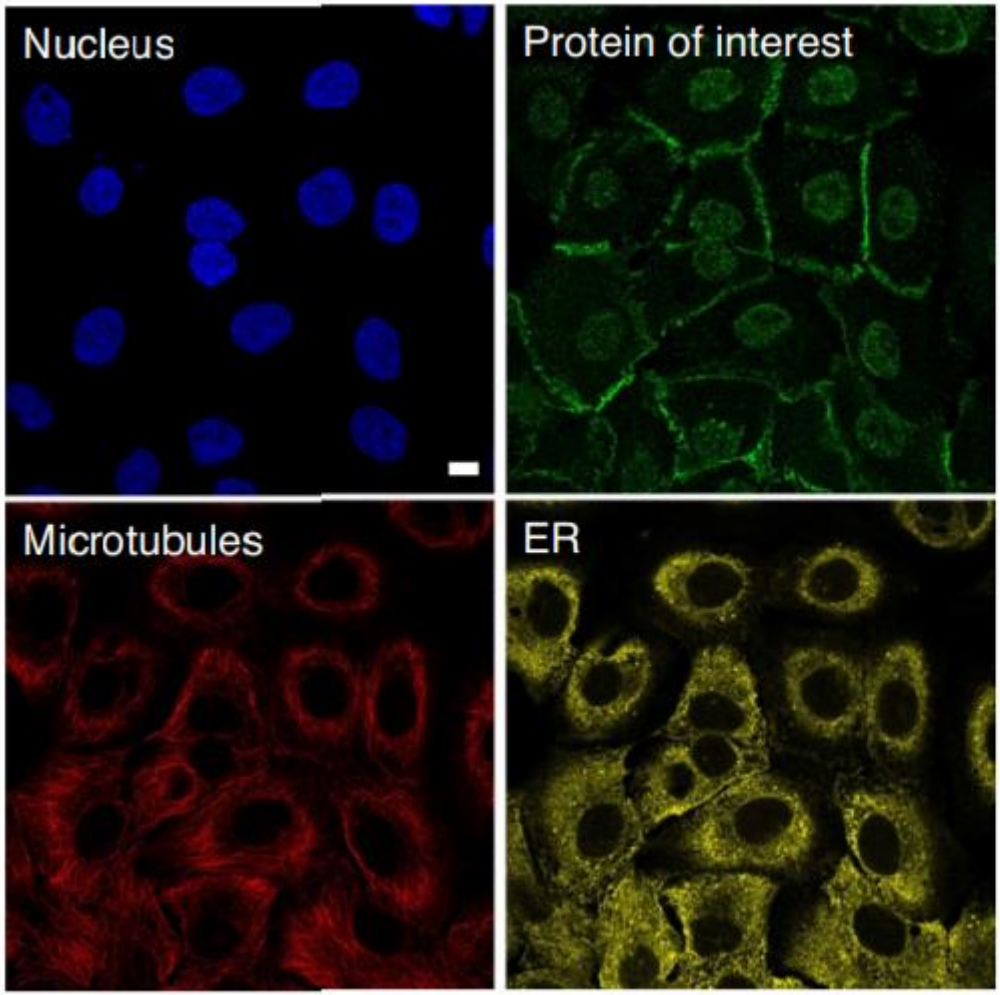
59,901 players

23.7 million image classifications



Sullivan and Winsnes (2018)

Protein localization classification competition



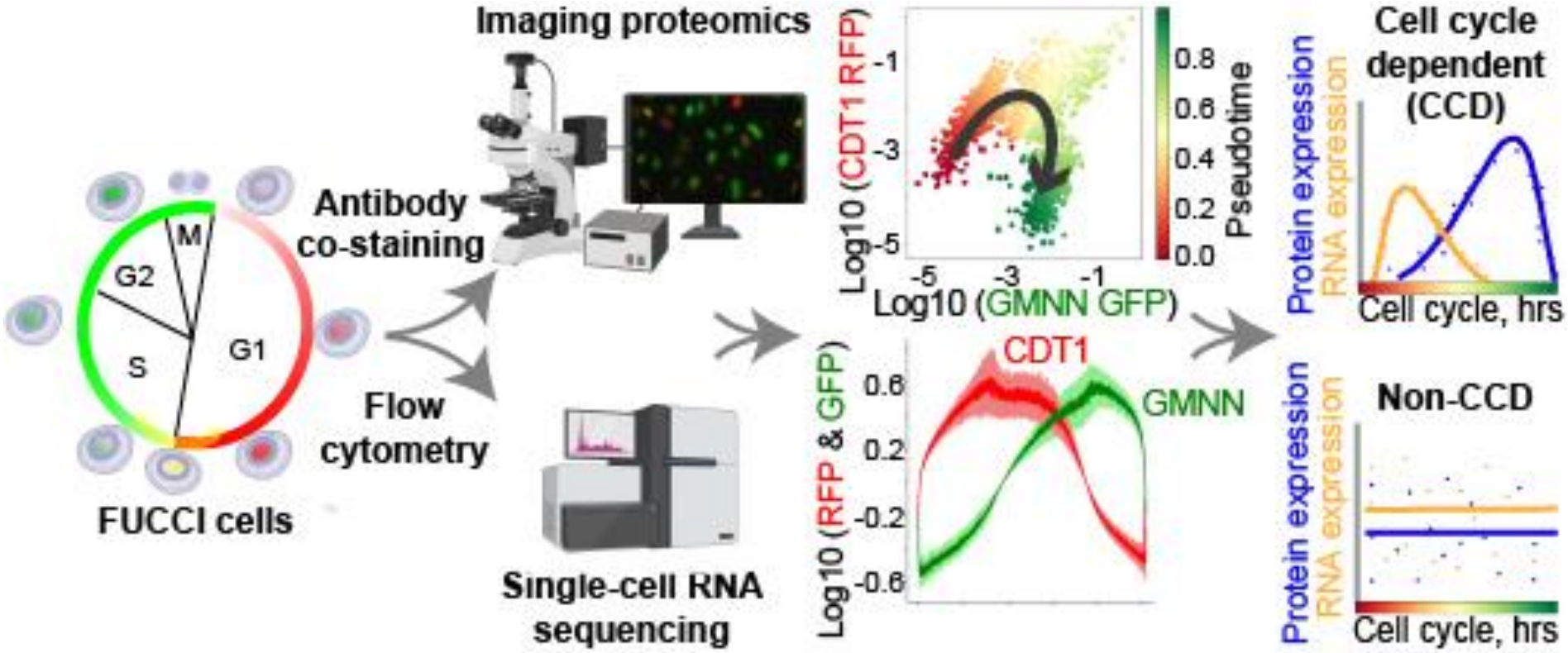
Classifier →

Multi-label prediction

Nucleoplasm
Cytosol
Plasma membrane
Nucleoli
Mitochondria
Golgi apparatus
Nuclear bodies
Nuclear speckles
Nucleoli fibrillar c.
Centrosome
Cell junctions
Actin filaments
...

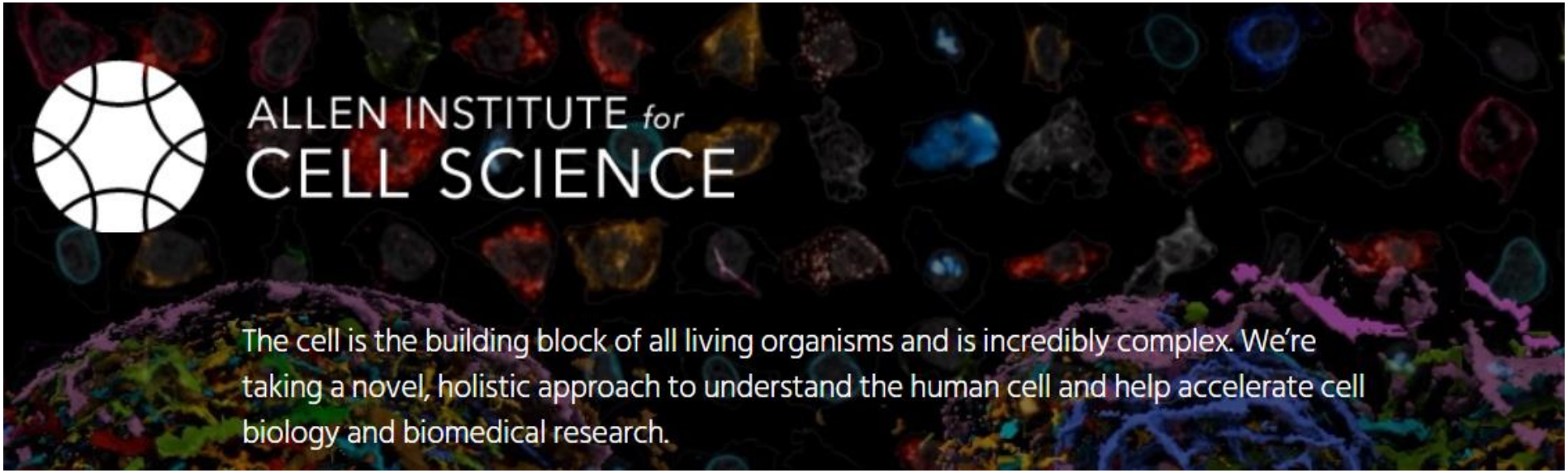
Ouyang et al. (2019)

Integrating proteomics at subcellular resolution with single-cell transcriptomics to construct a spatiotemporal map of human proteomic heterogeneity by



Mahdessian, Cesnik et al. (2021)

Allen Institute of Cell Science

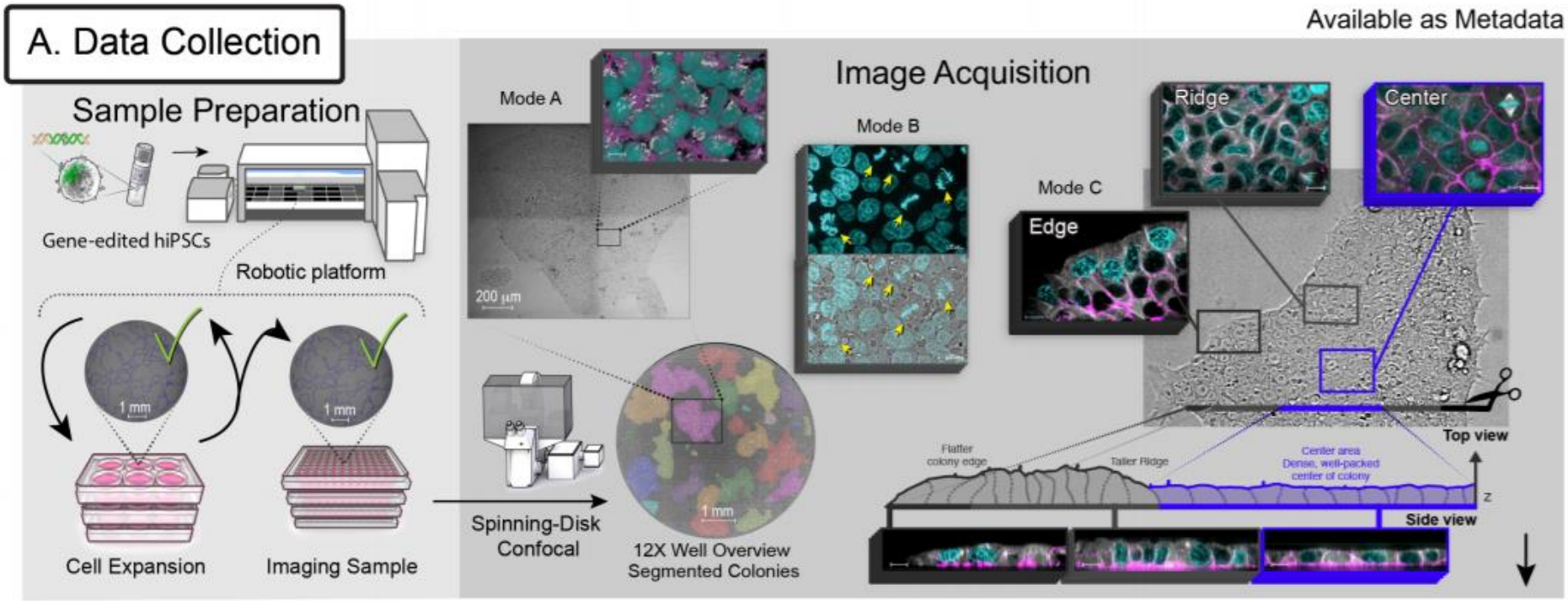


Resources:

- Cell catalog https://www.allencell.org/cell-catalog.html
- Analysis and visualization tools, education resources https://www.allencell.org/
- Data https://www.allencell.org/data-downloading.html
- Next: characterizing integrated intracellular organization

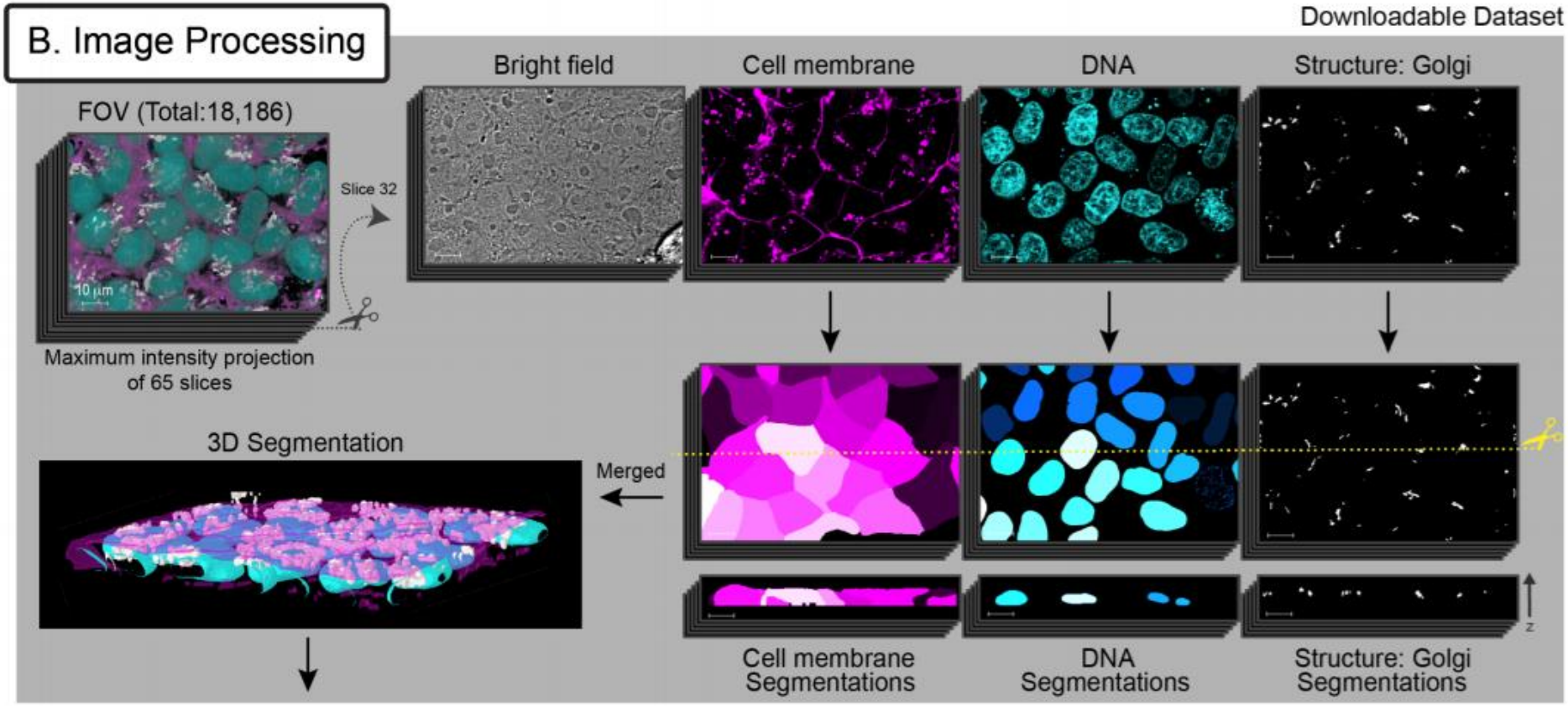
https://alleninstitute.org/what-we-do/cell-science/

Pipeline 1



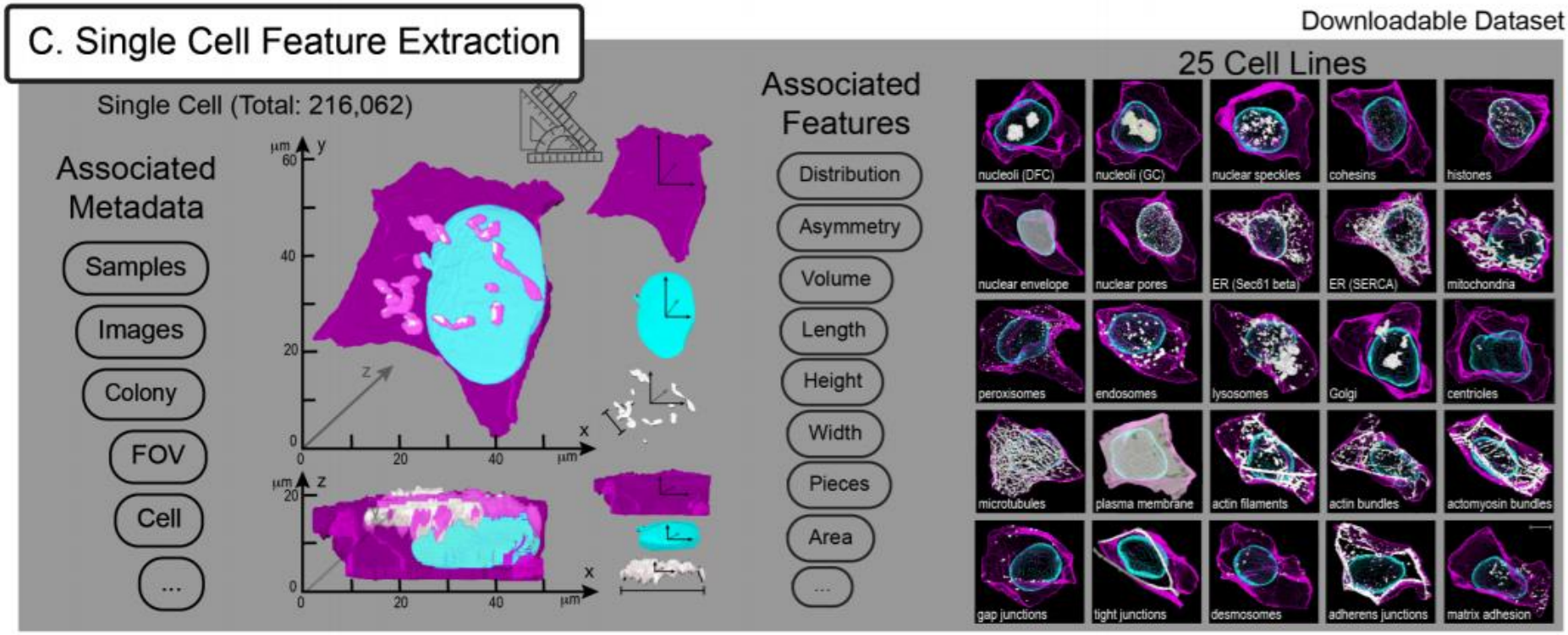
Viana et al. (2021)

Pipeline 2



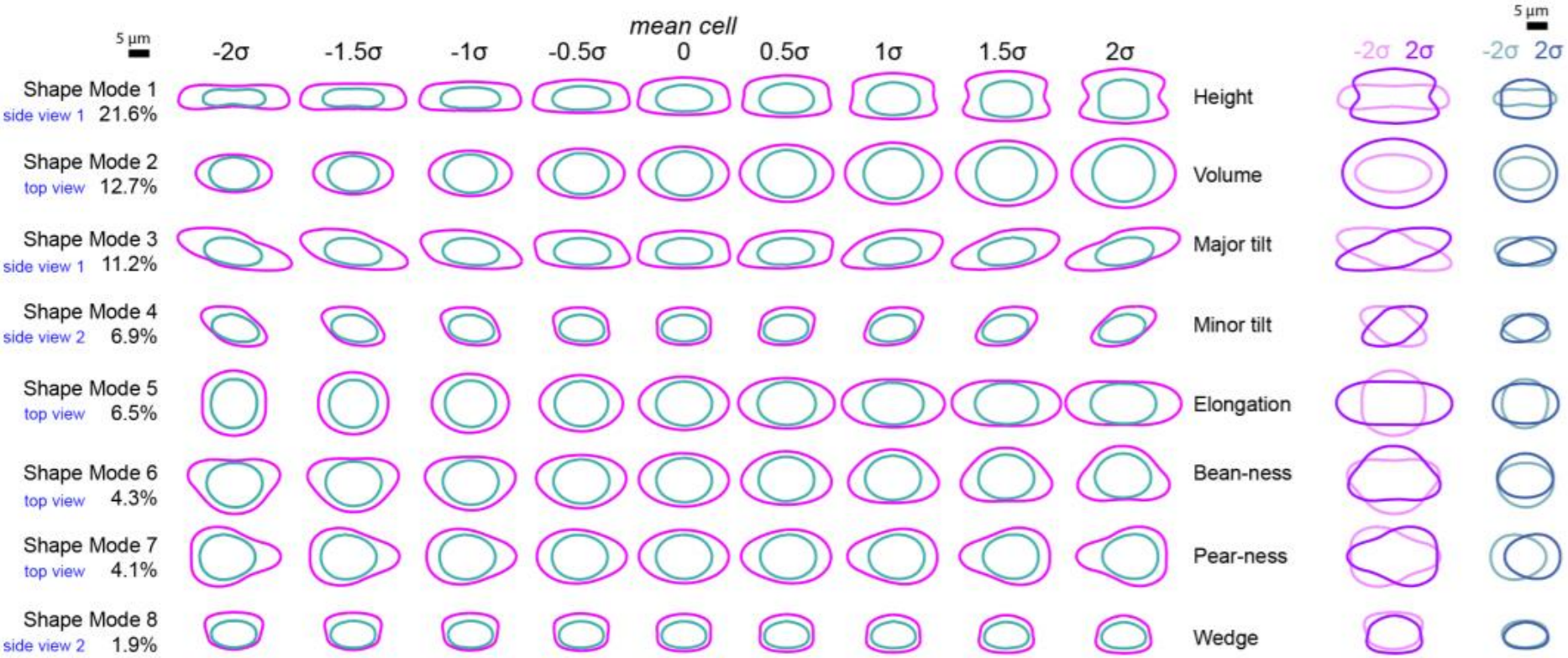
Viana et al. (2021)

Pipeline 3



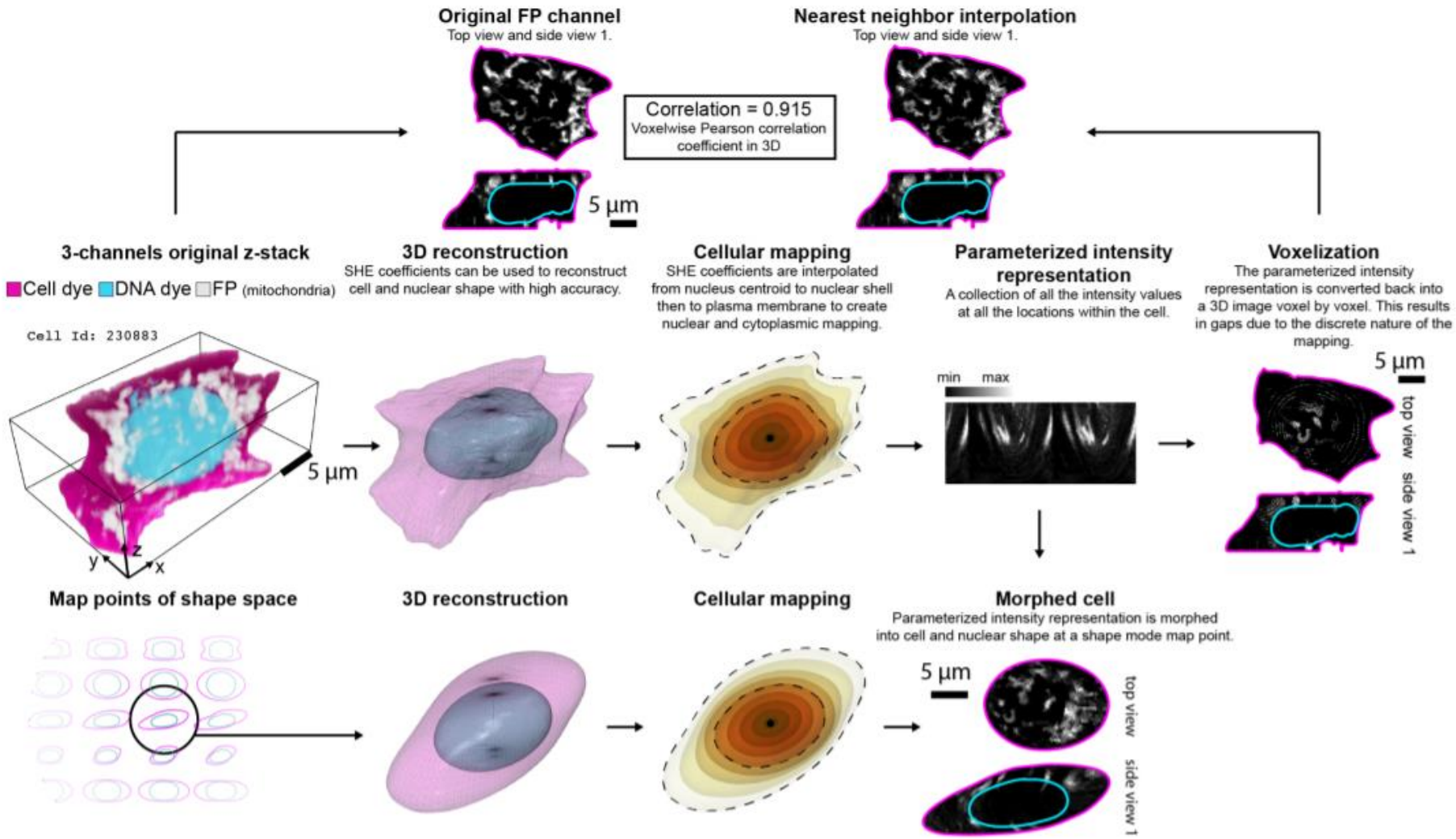
Viana et al. (2021)

Modes of shape variations



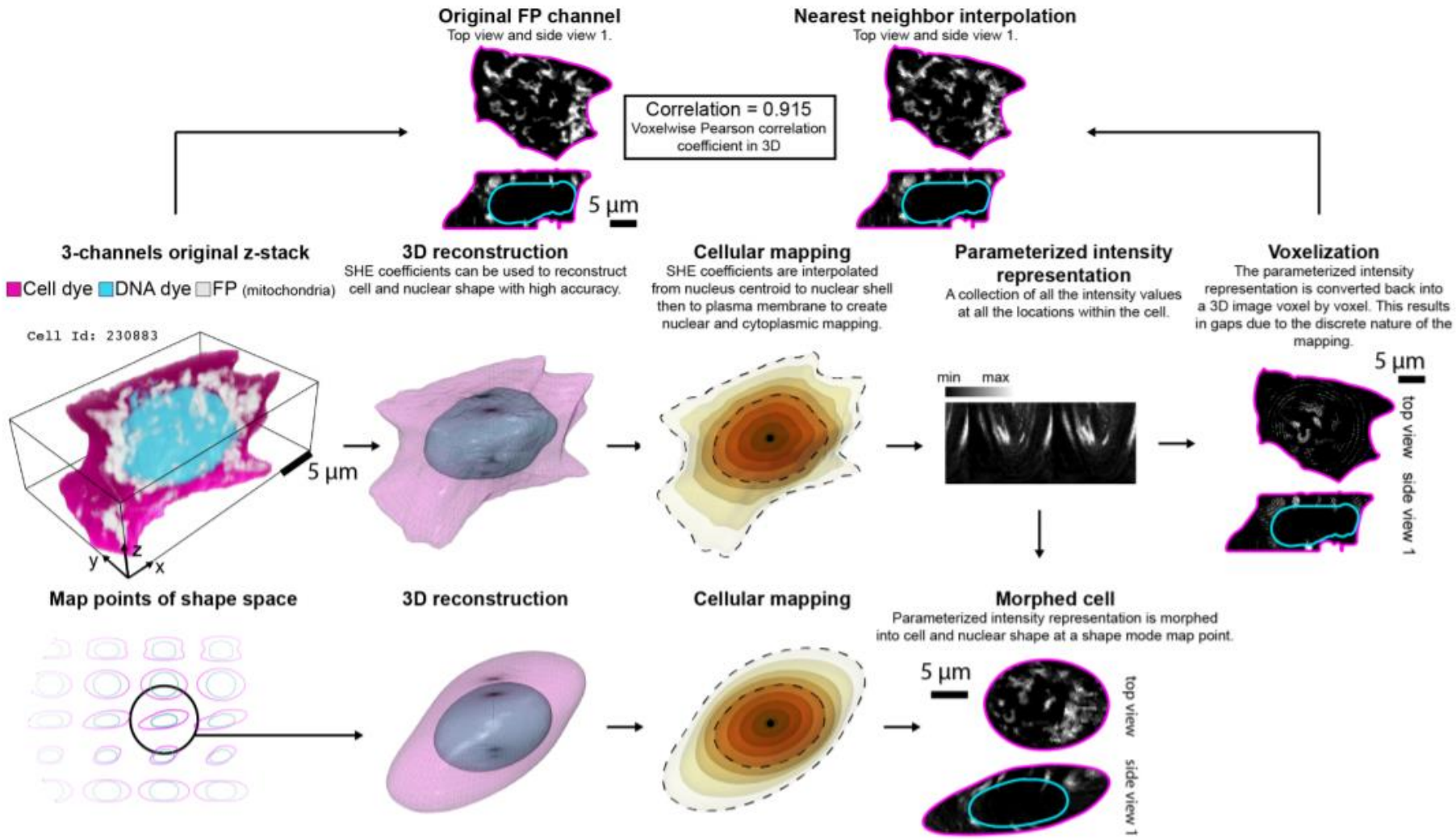
Viana et al. (2021)

Building integrated average cells throughout the shape space



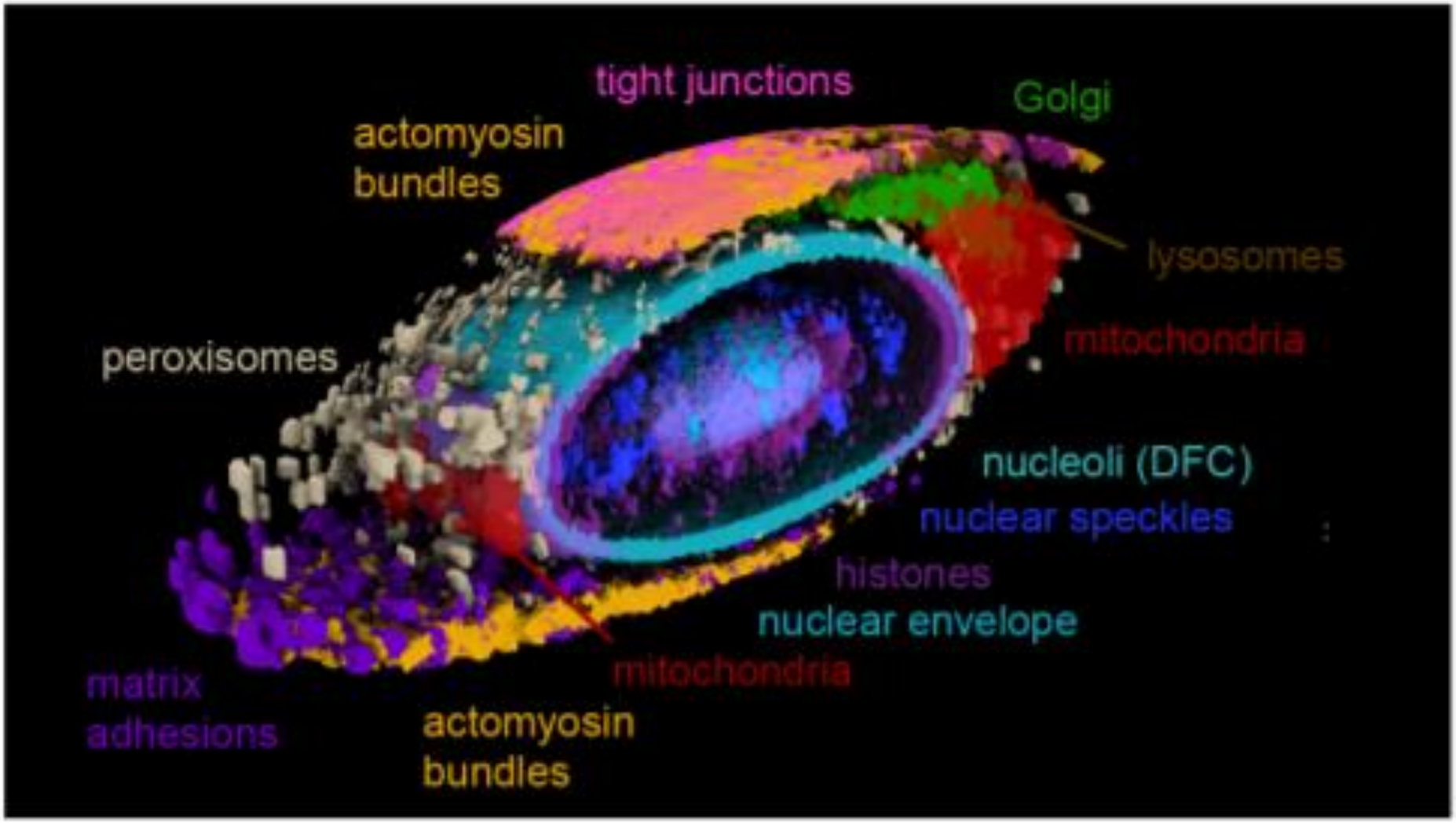
Viana et al. (2021)

Building integrated average cells throughout the shape space



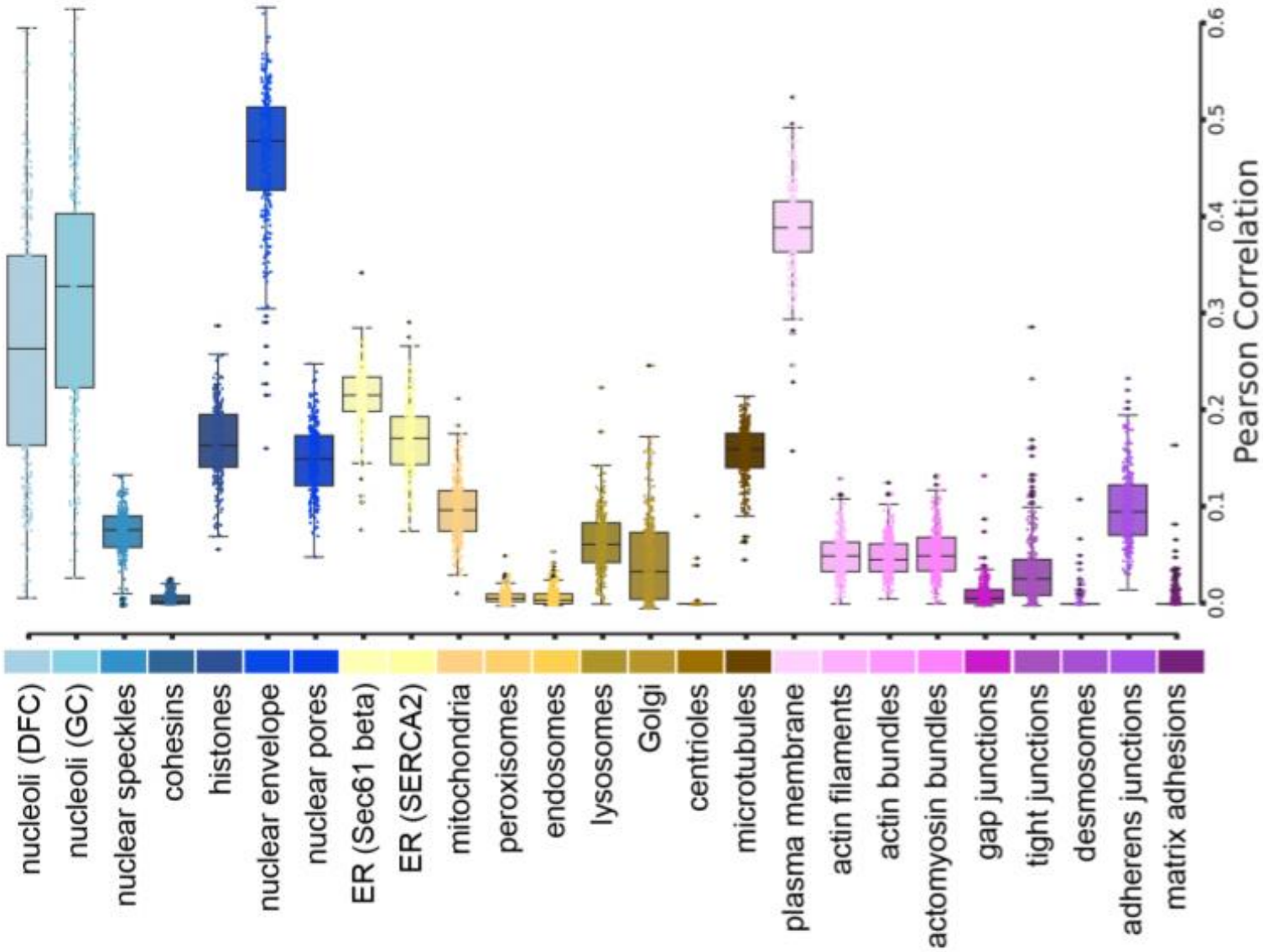
Viana et al. (2021)

Building integrated average cells throughout the shape space



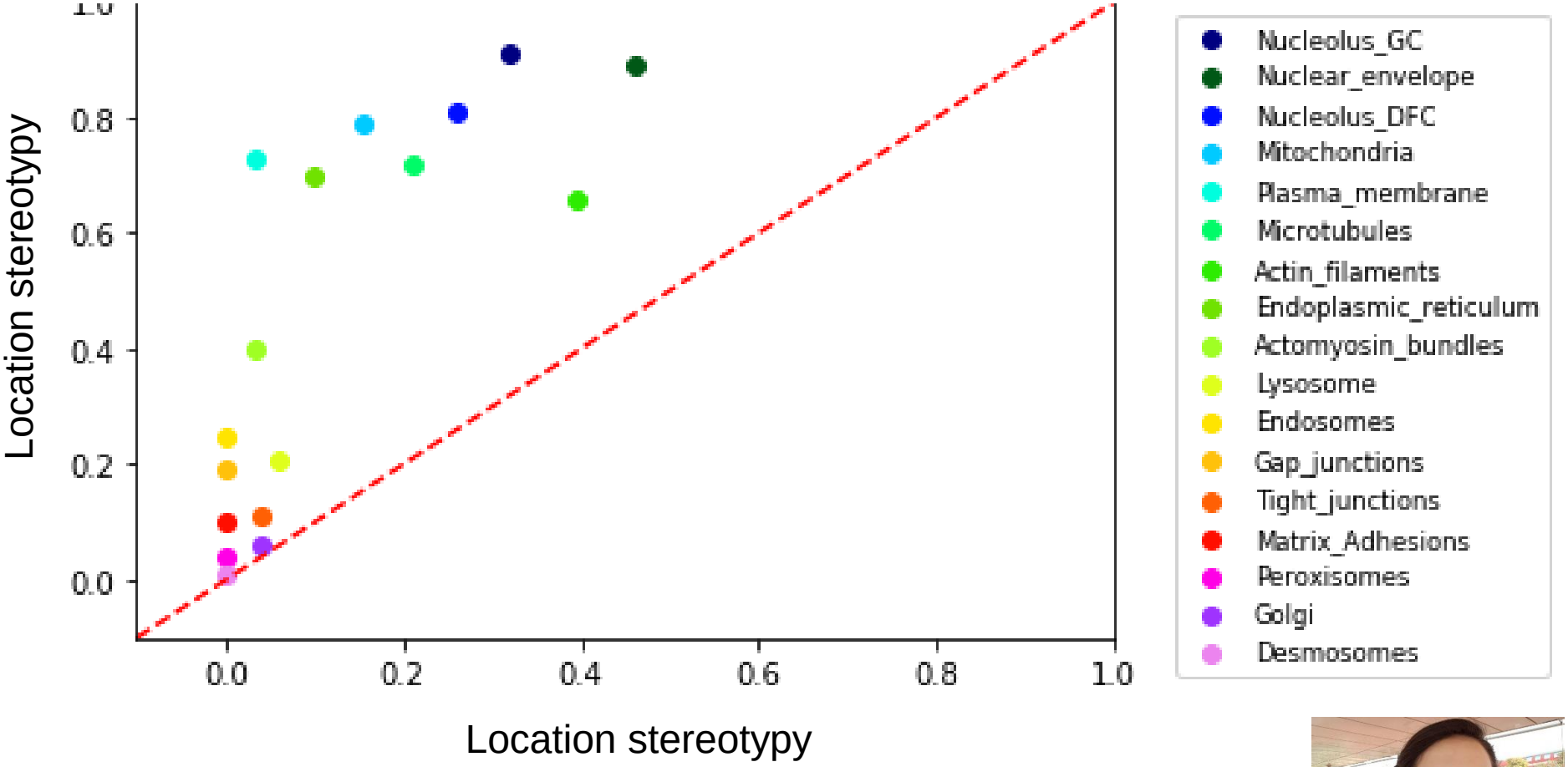
Viana et al. (2021)

Location stereotypy of cell structures



Viana et al. (2021)

Location stereotypy and in silico labeling

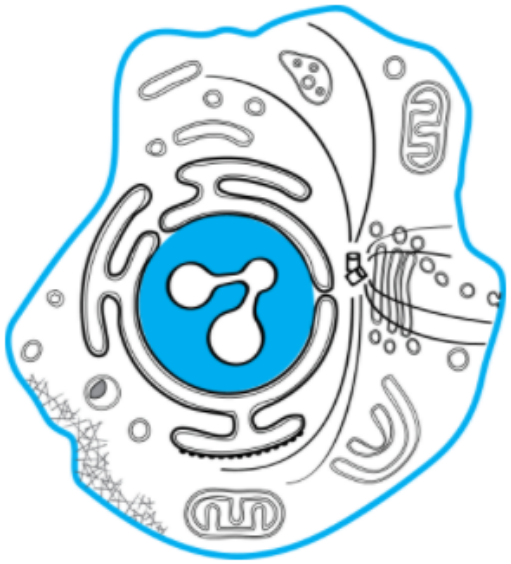


Open Cell (CZI)

OpenCell

Proteome-scale measurements of human protein localization and interactions

Targets | Gallery | How to | About



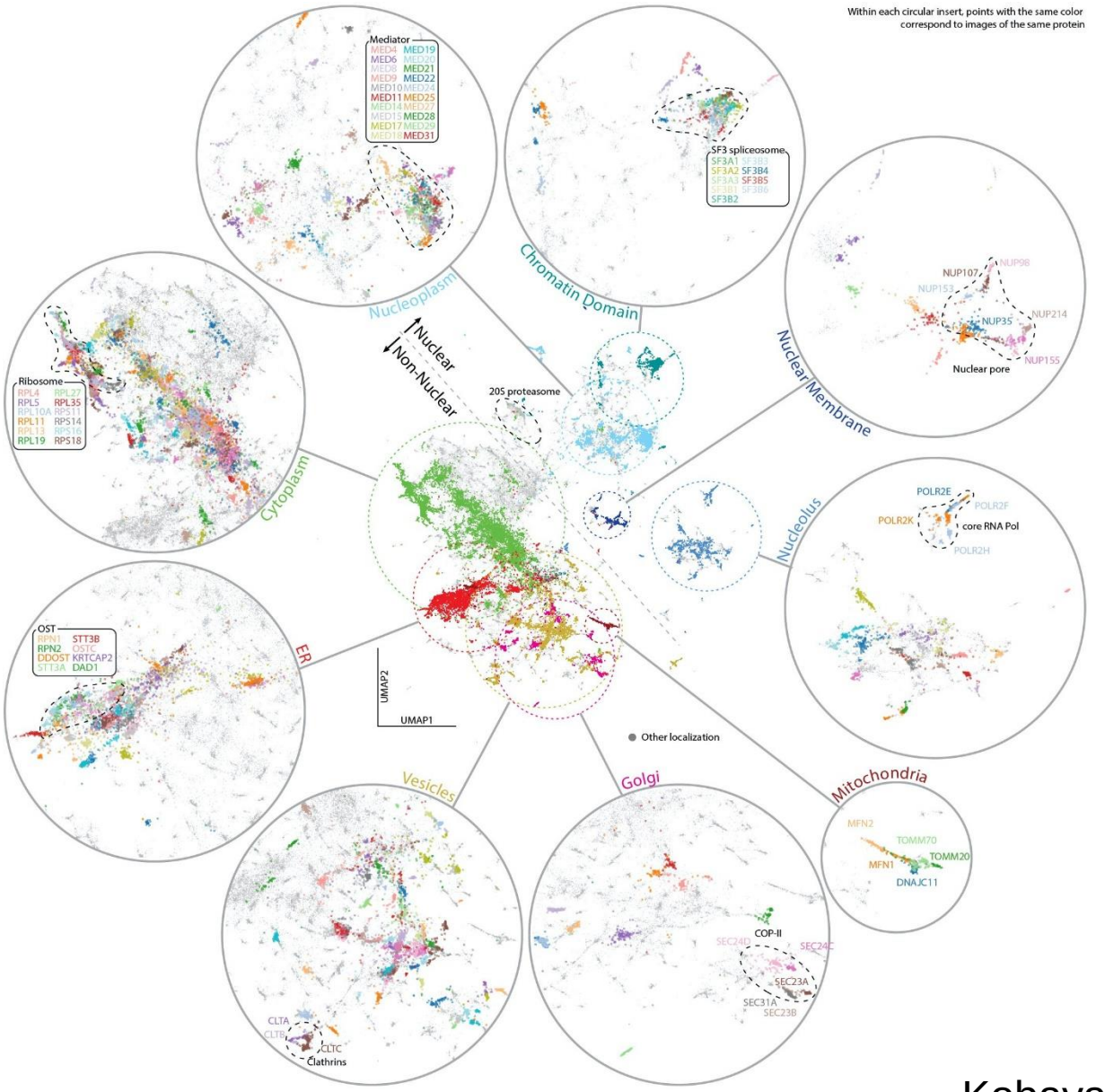
Search for a protein

For example: MAP4, POLR2F, Golgi, mediator complex

1,311	30,293	5,912
Tagged proteins	Protein interactions	3D images

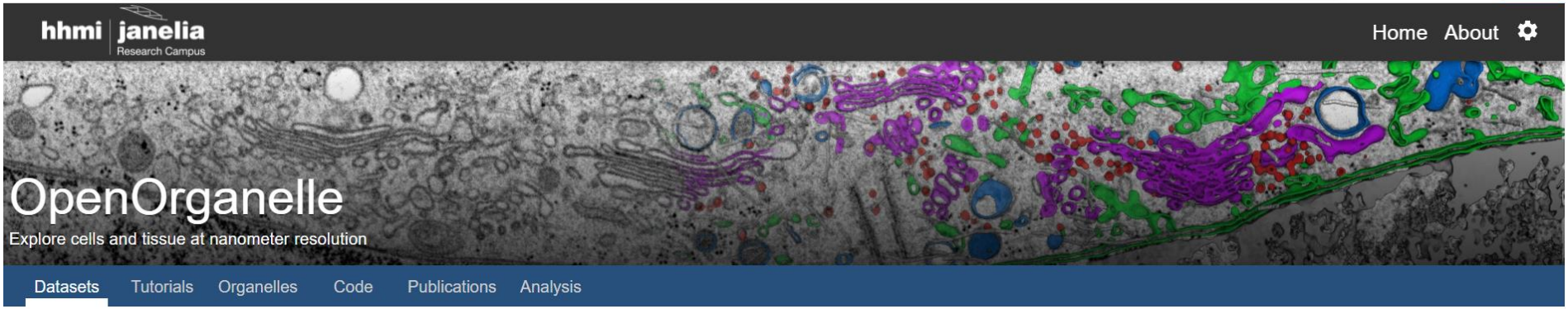
https://opencell.czbiohub.org/

Protein localization atlas (via autoencoding)

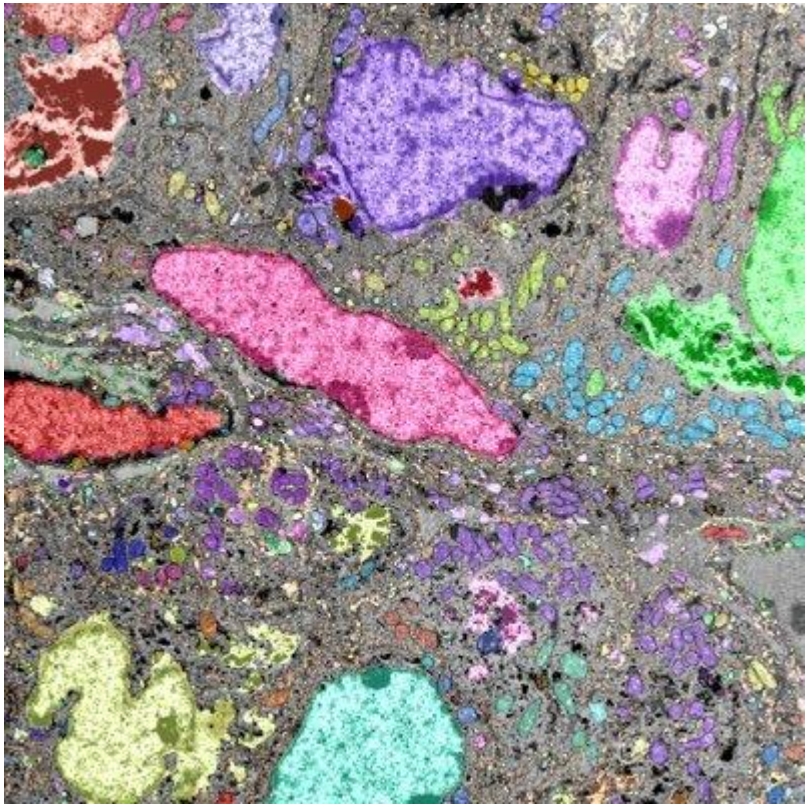


Kobayashi el al. (2021)

OpenOrganelle (HHMI Janelia)

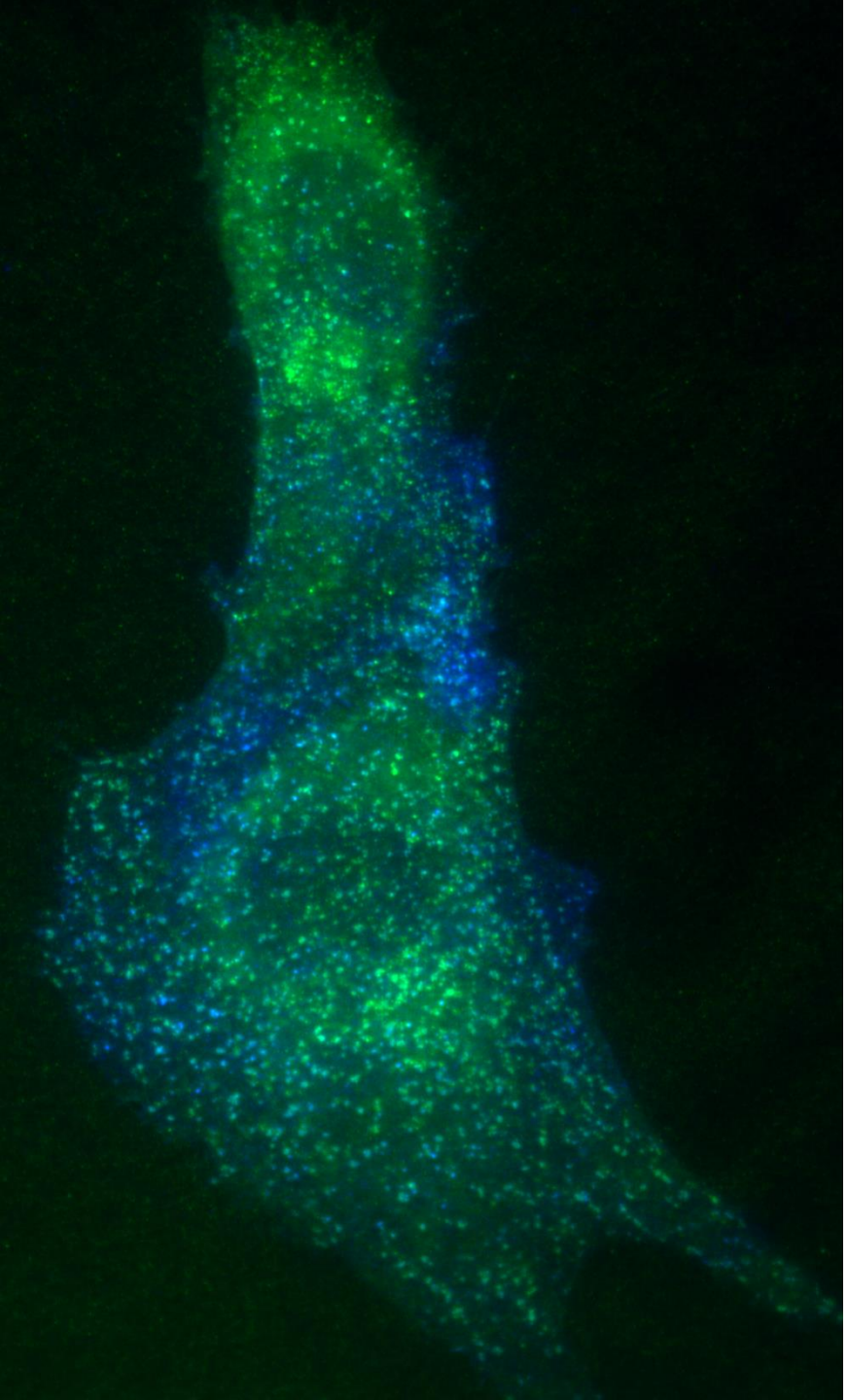


https://openorganelle.janelia.org/



Cho, Cheveralls, Brunner, Kim, Michaelis, Raghavan, et al. (2021)

Tool development, generating new insight from old data – my personal experiences

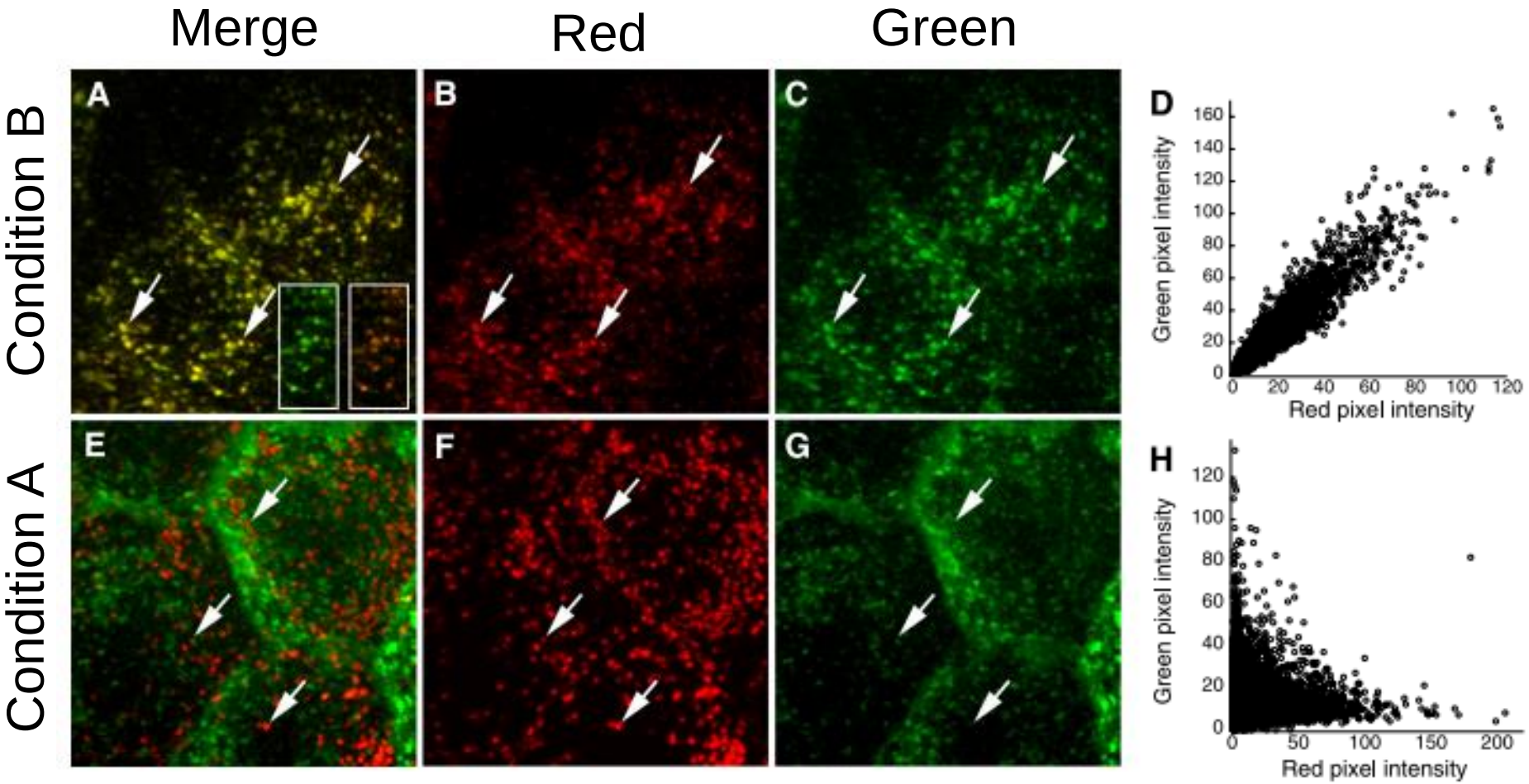


Decoupling global biases and local interactions between cell biological variables

LI GI

DeBias

Quantifying protein-protein co-localization



Dunn et al. (2011)

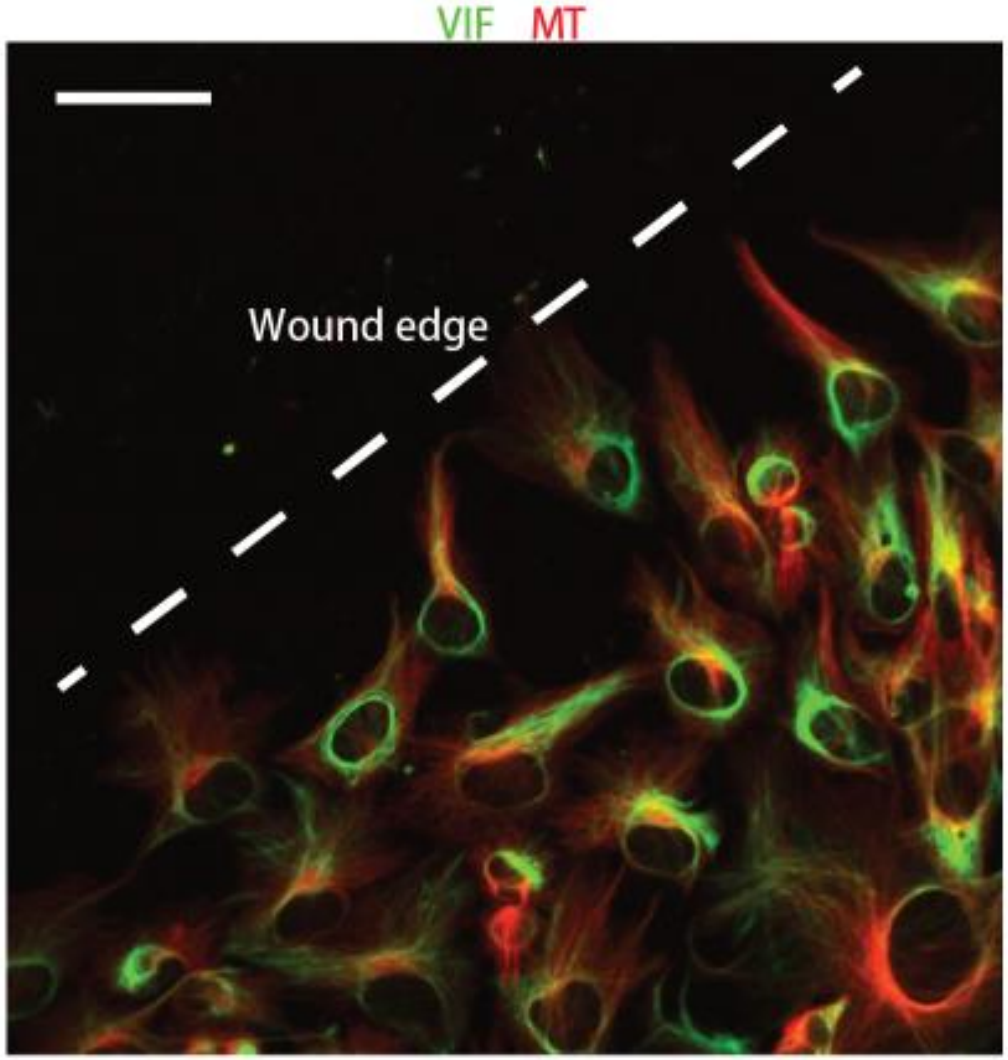
What additional information is hidden in co-localization data?

LI GI

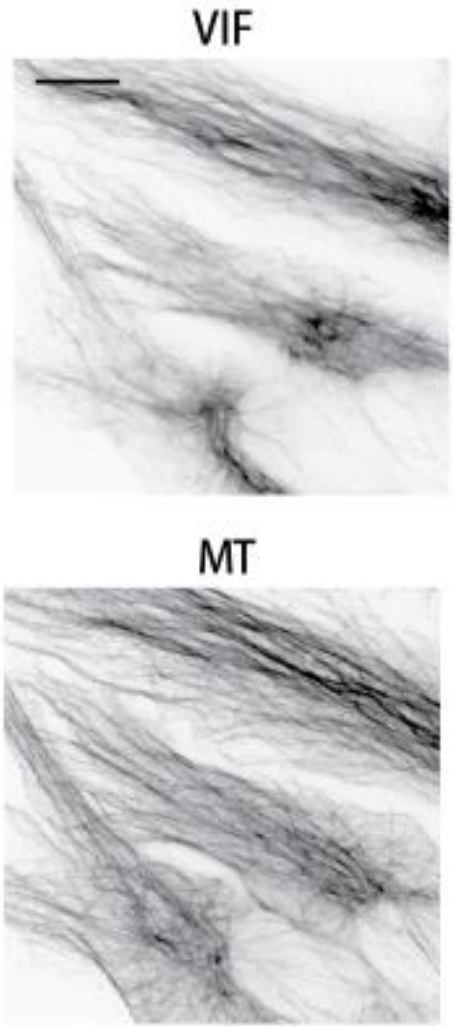
DeBias

Co-orientation of intracellular cytoskeletal networks in migrating cells

Vimentin provides a structural template for microtubule growth

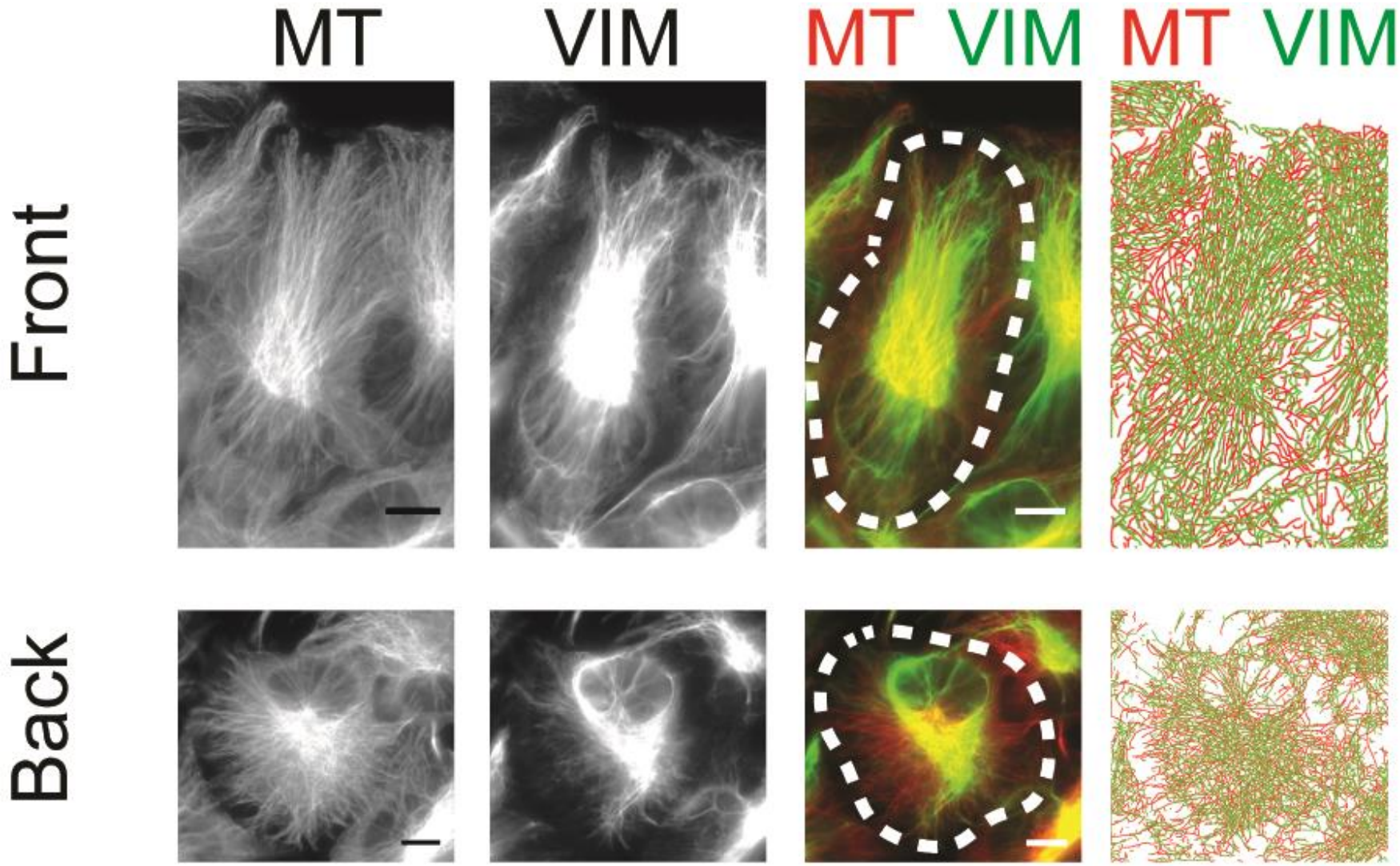


Genome-edited Retinal Pigment Epithelial (RPE) cells



Gan, Ding and Burckhardt et al. (2016)

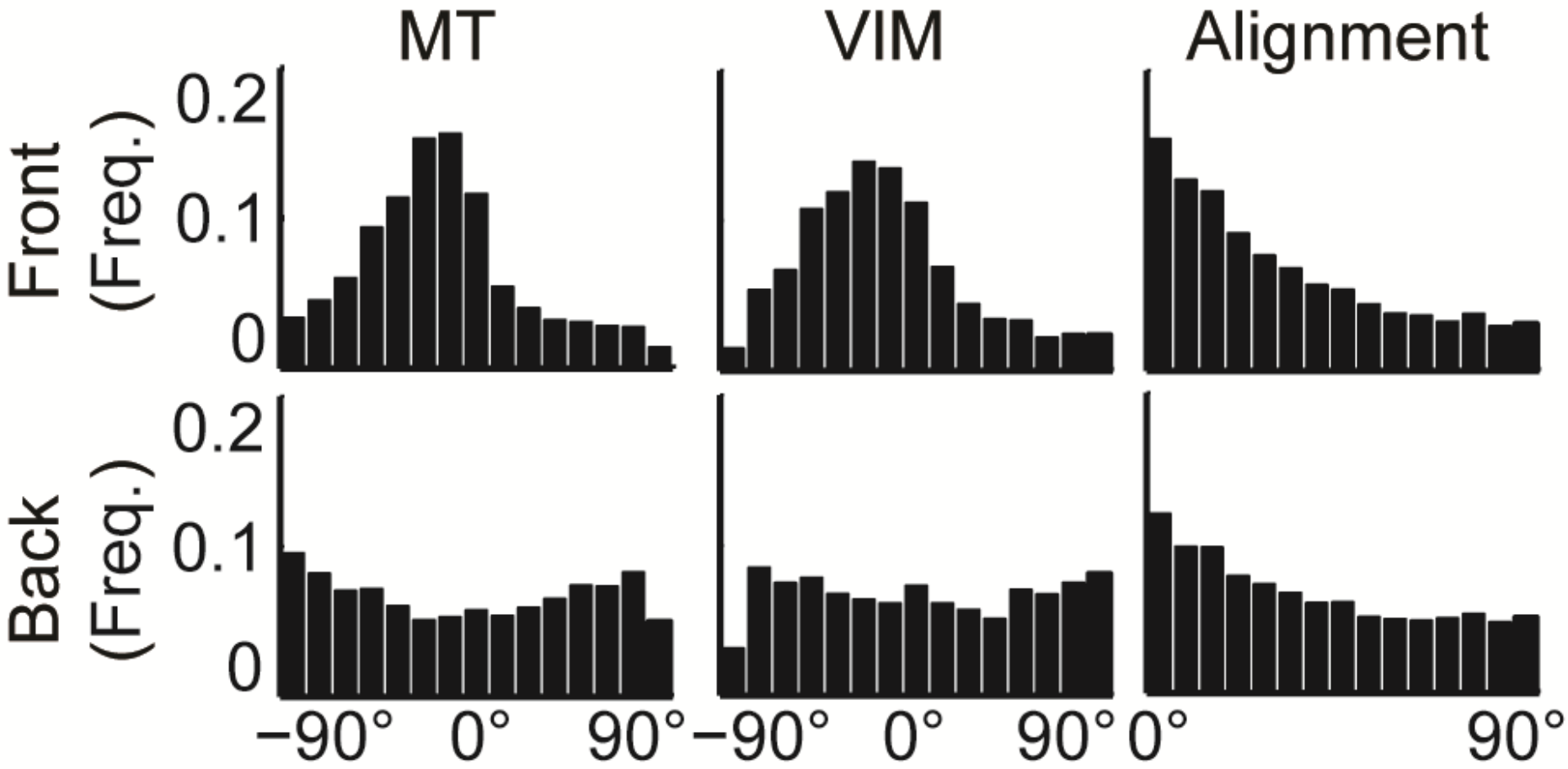
A relation between cell polarity and vimentin-microtubule interaction?



Zhuo Gan

Polarity-independent interaction of vimentin and microtubules

Polarity-independent interaction of vimentin and microtubules



What do we want to achieve?

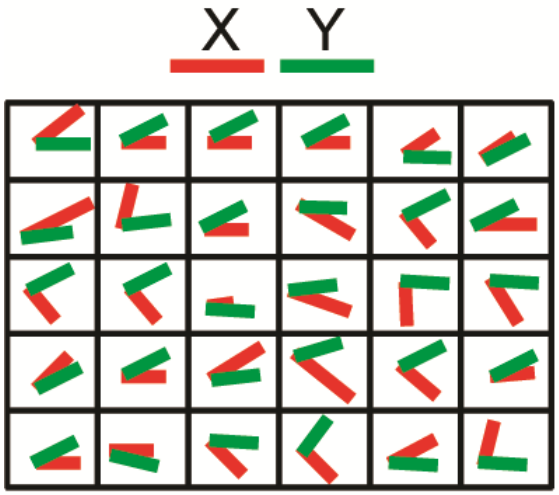
- Simultaneous investigation of mechanisms that drive global bias and local interactions

How?

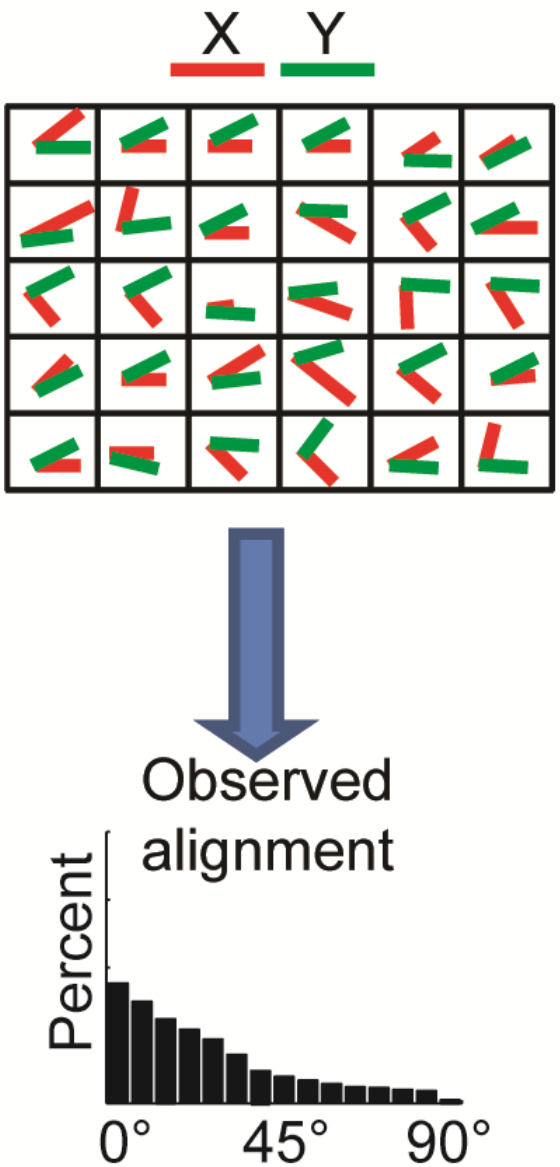
- By modeling the observed agreement between matched variables as the cumulative global and local components

Observed colocalization = Global bias + Local interaction

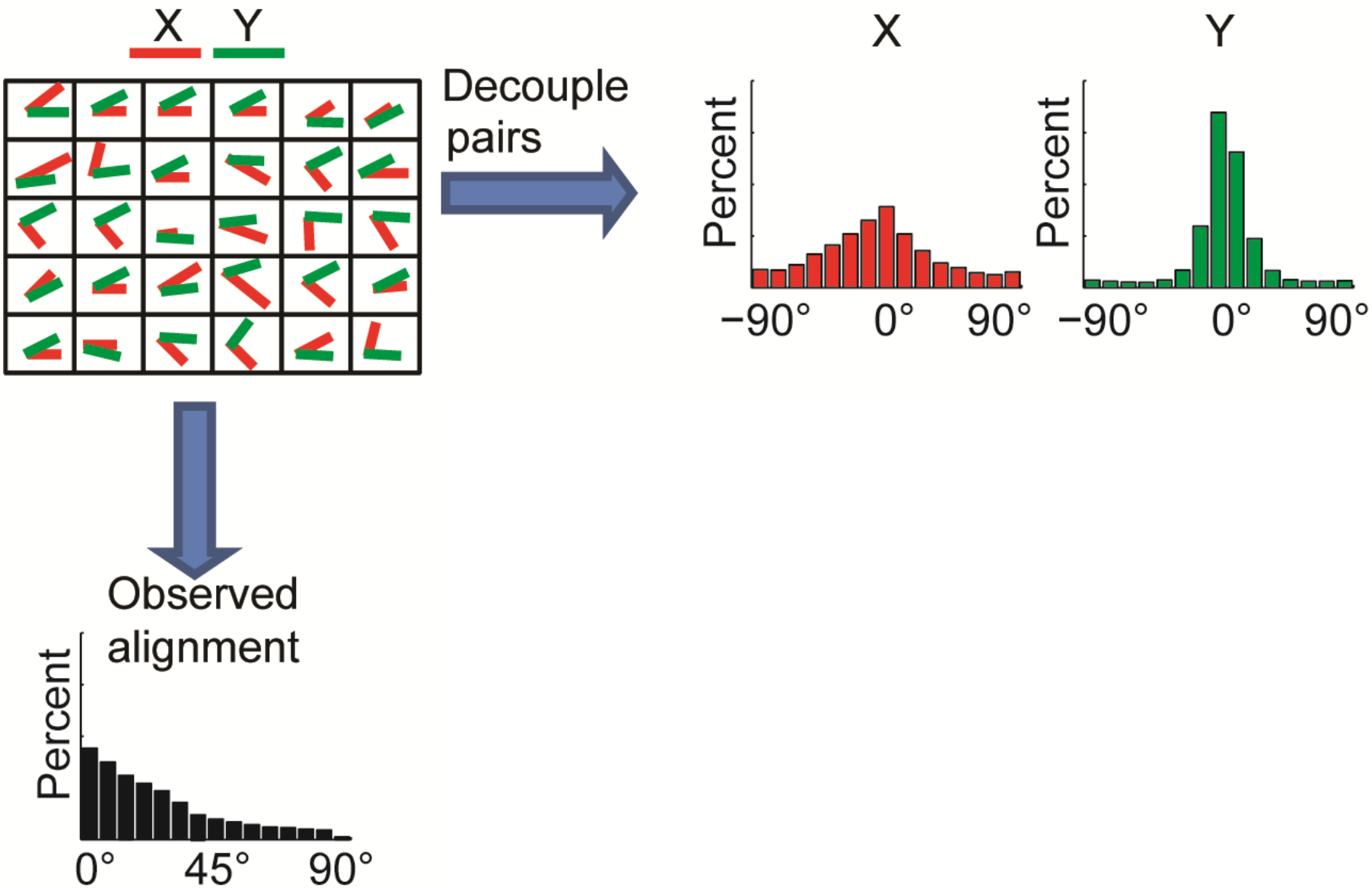
DeBias



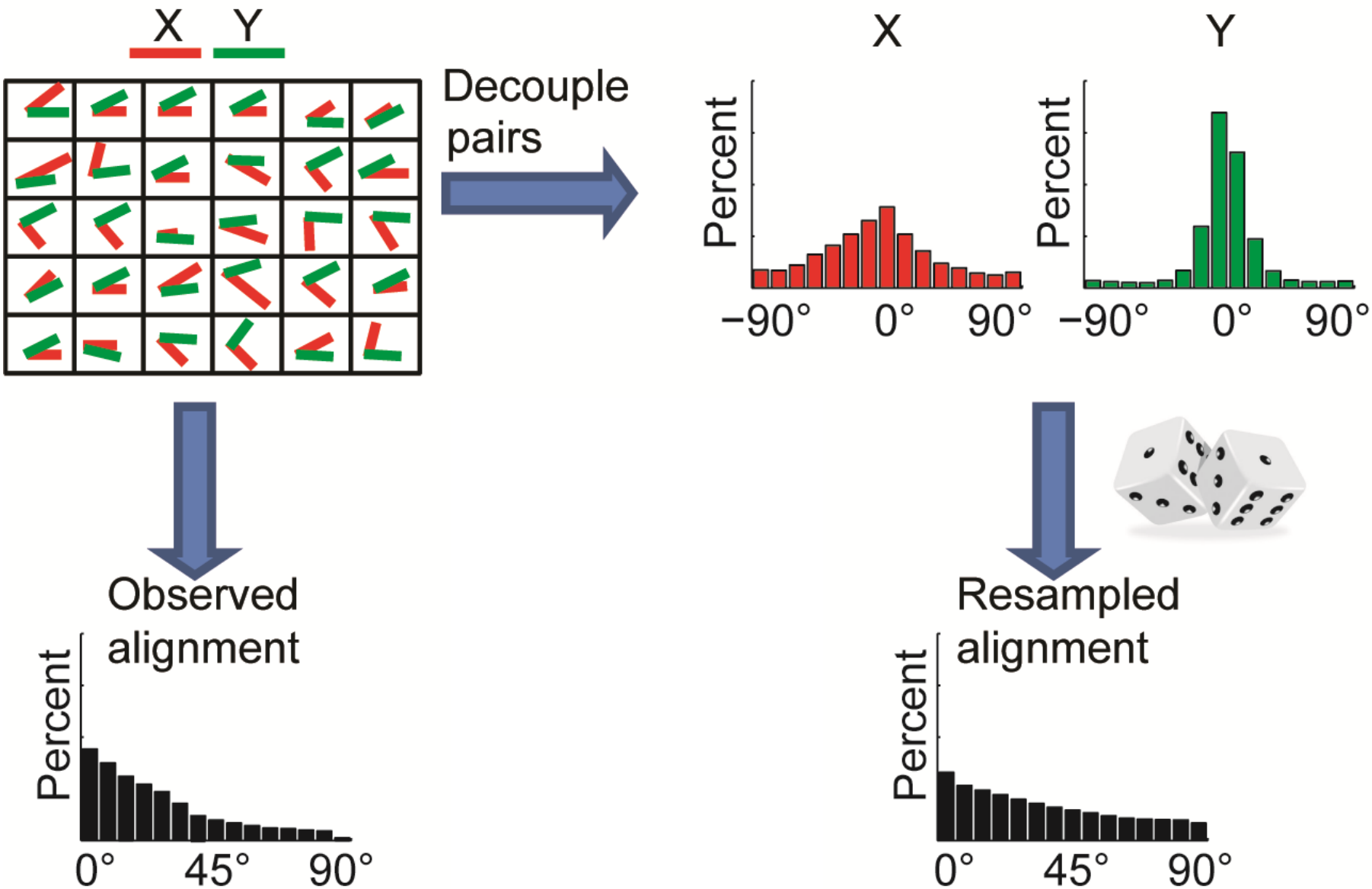
DeBias



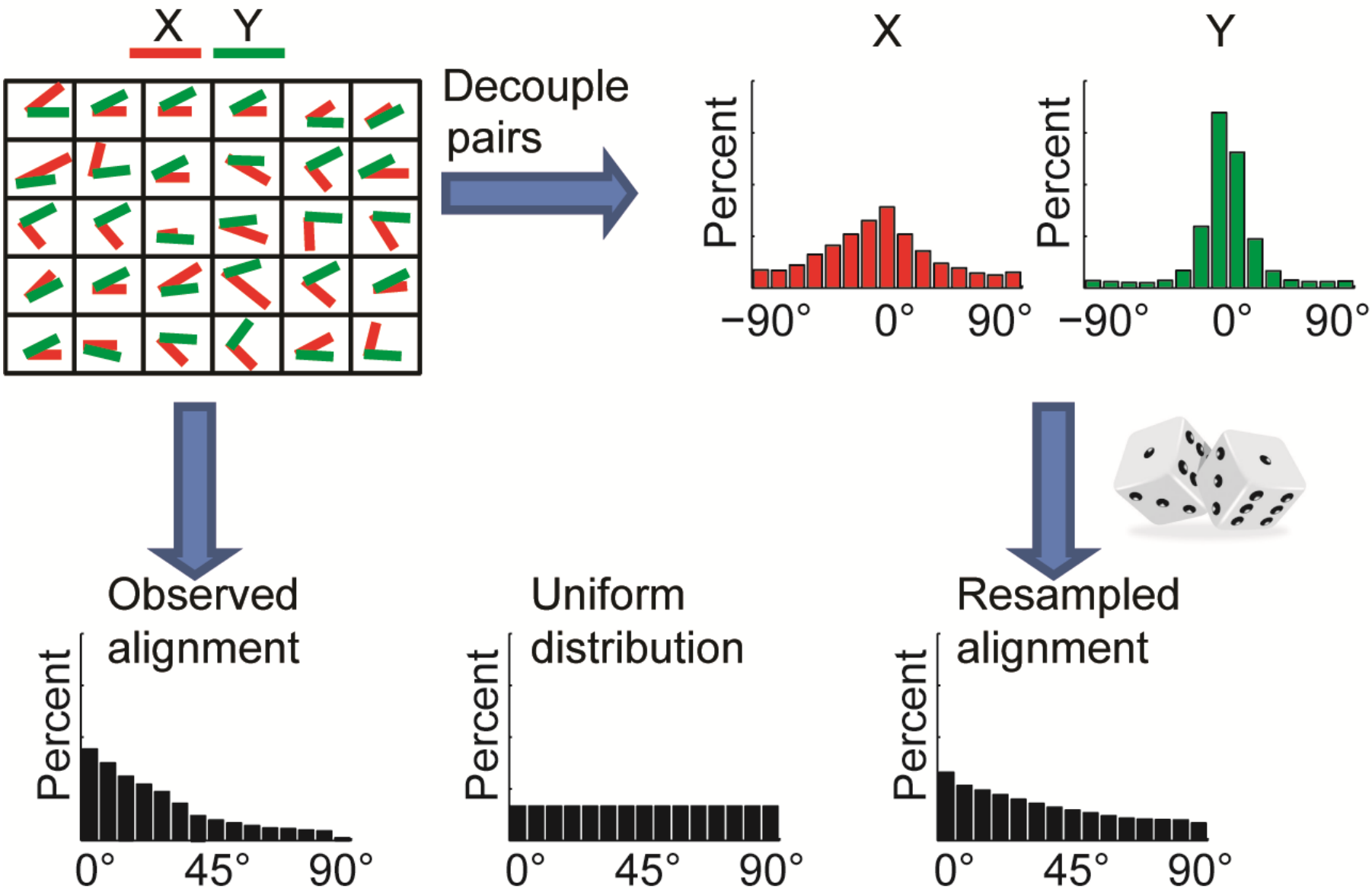
DeBias


X
Y
Decouple pairs
X
Y
Percent
−90°
0°
90°
Percent
−90°
0°
90°
Observed alignment
Percent
0°
45°
90°

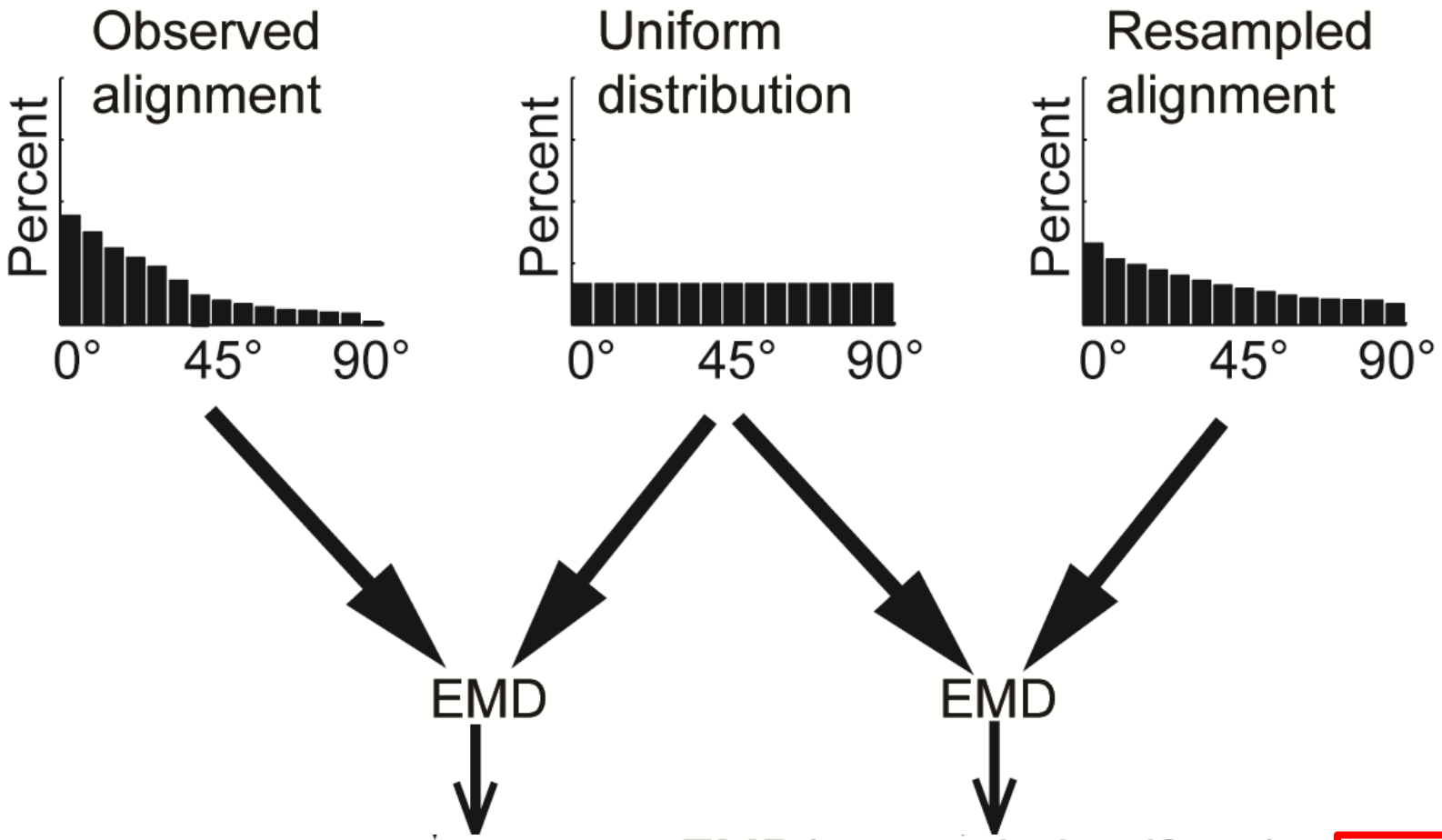
DeBias


X
Y
Decouple pairs
X
Y
Percent
−90°
0°
90°
Percent
−90°
0°
90°
Observed alignment
Percent
0°
45°
90°
Resampled alignment
Percent
0°
45°
90°

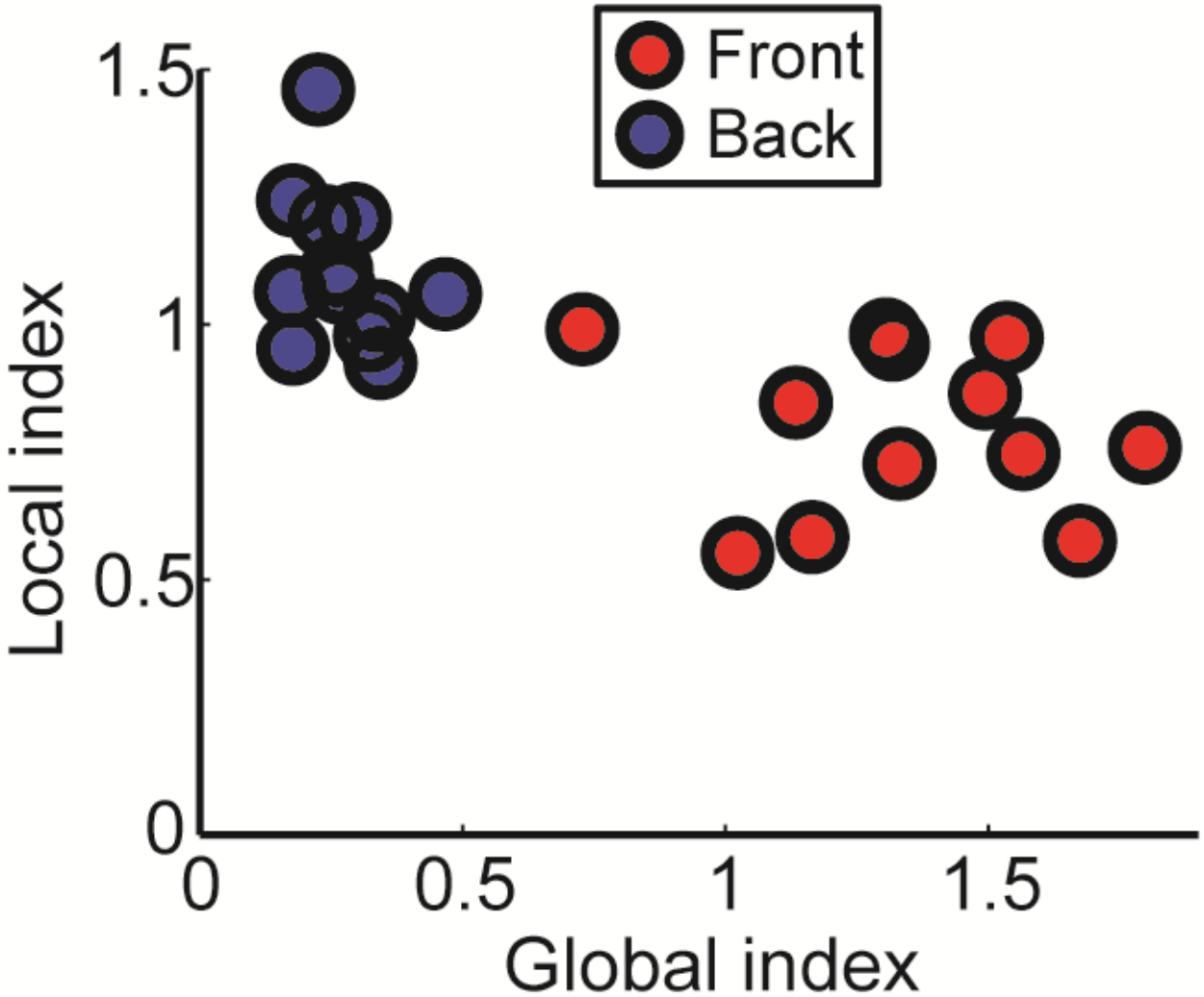
DeBias


X
Y
Decouple pairs
X
Y
Percent
−90°
0°
90°
Percent
−90°
0°
90°
Observed alignment
Percent
0°
45°
90°
Uniform distribution
Percent
0°
45°
90°
Resampled alignment
Percent
0°
45°
90°

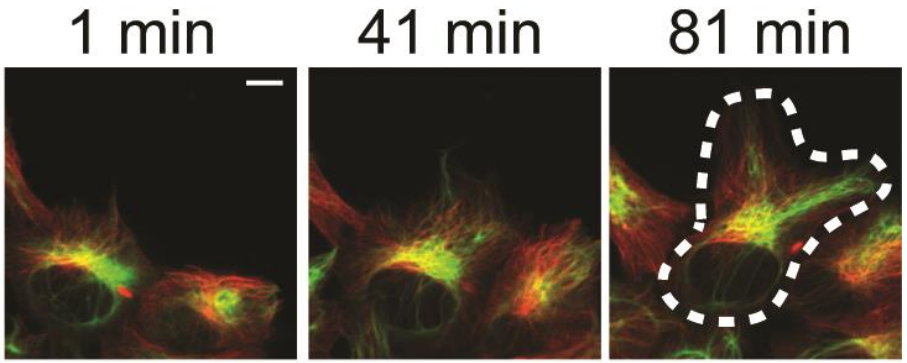
DeBias

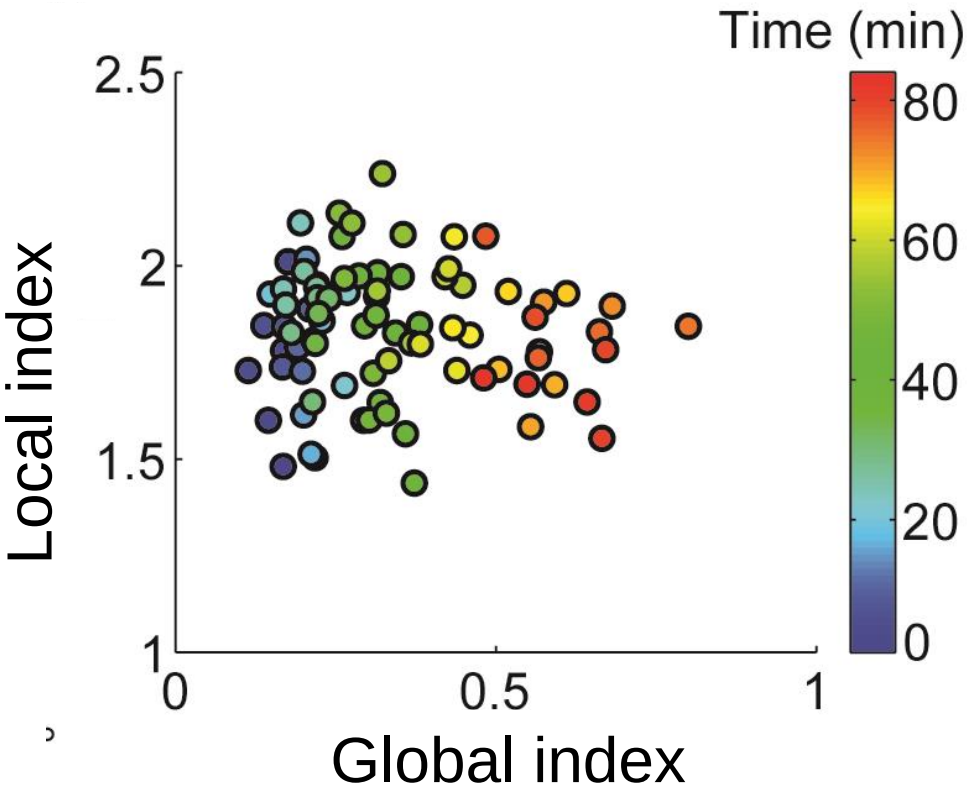

Observed alignment
Percent
0°
45°
90°
Uniform distribution
Percent
0°
45°
90°
Resampled alignment
Percent
0°
45°
90°
EMD
EMD
EMD(observed,uniform) - EMD(resampled,uniform) = local index
global index

Polarity-independent interaction of vimentin and microtubules



Polarity-independent interaction of vimentin and microtubules

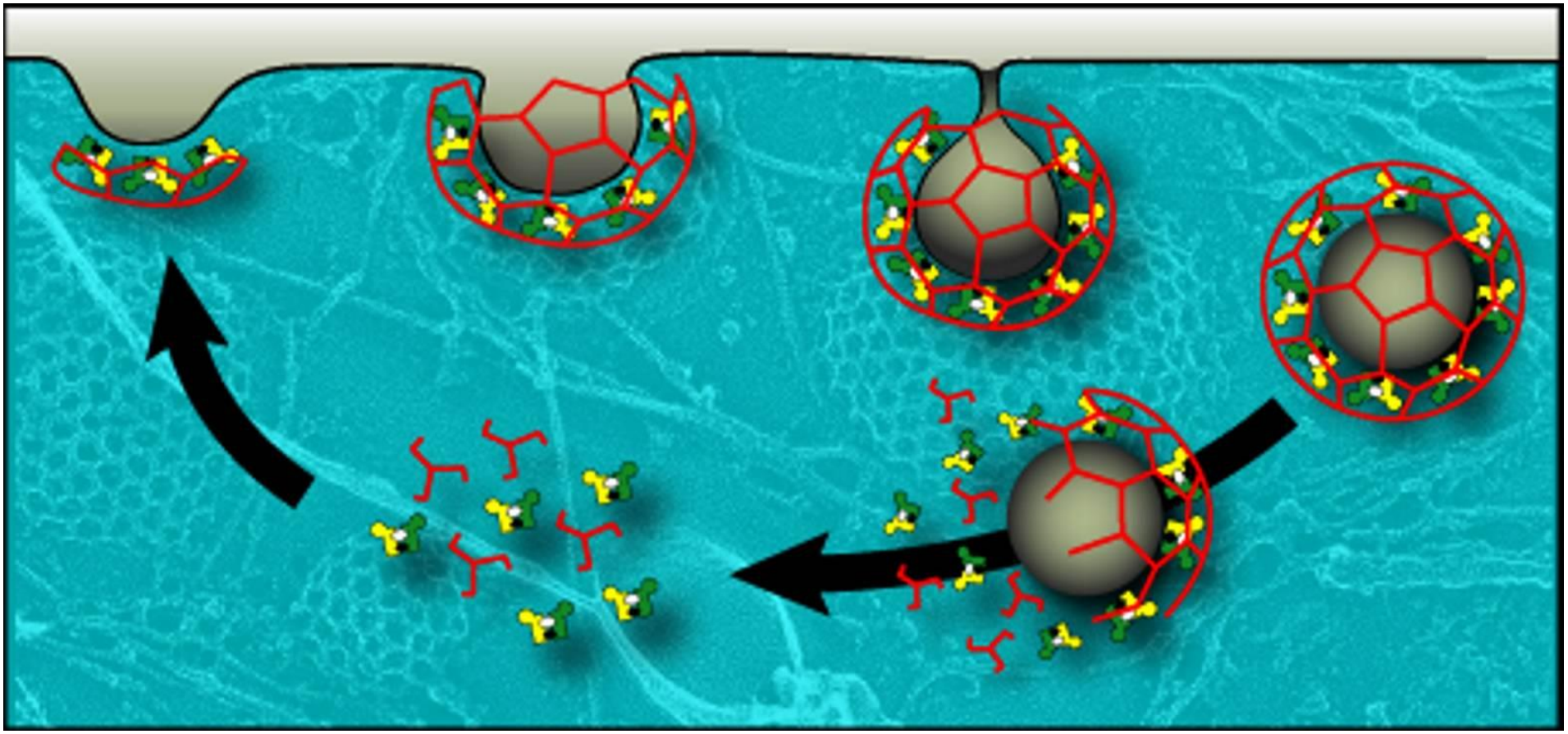




Global bias: cell polarity

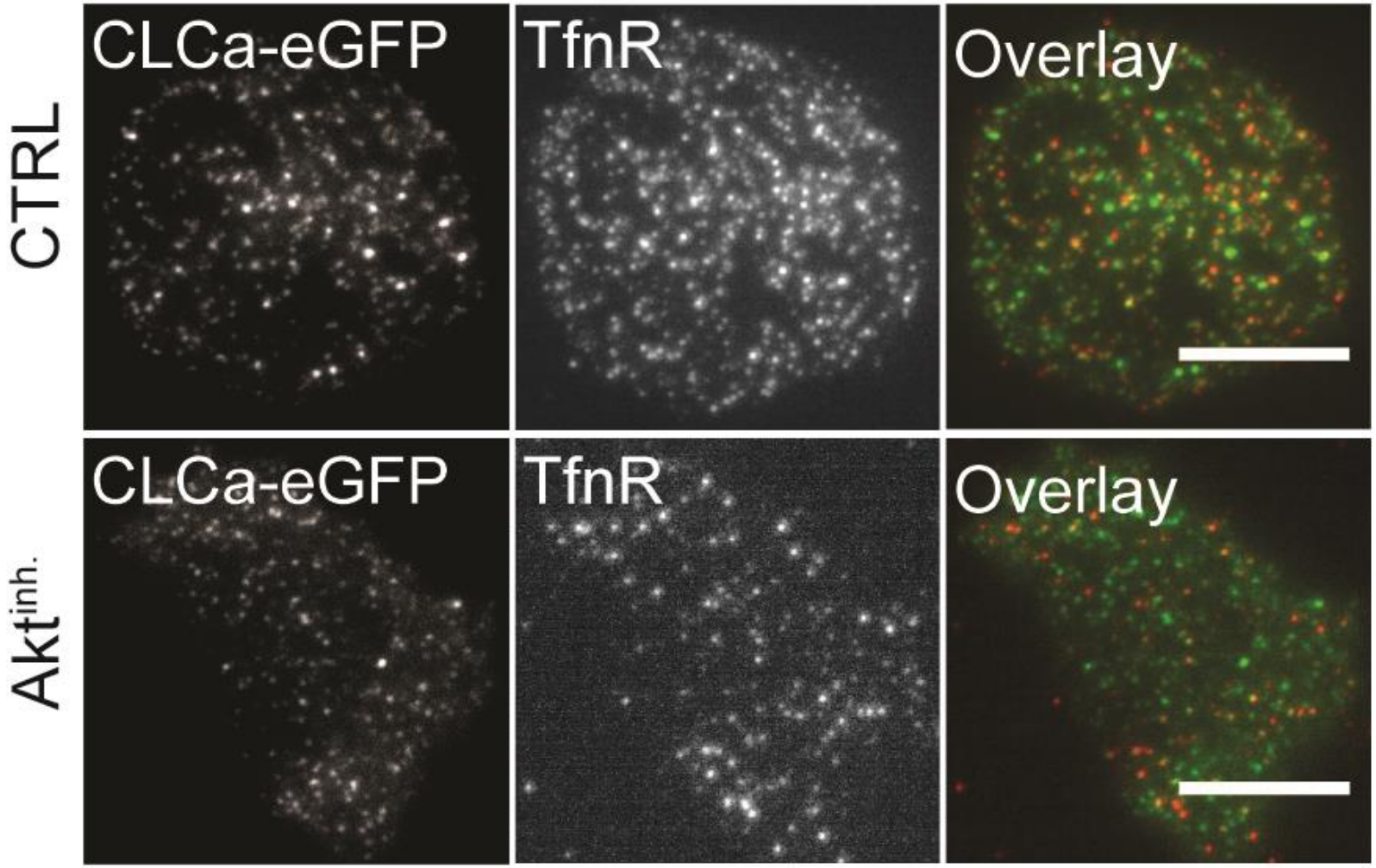
Inferring co-localization and predicting dynamics from fixed cells during clathrin-mediated endocytosis (CME)

Maturation of a clathrin coated pit



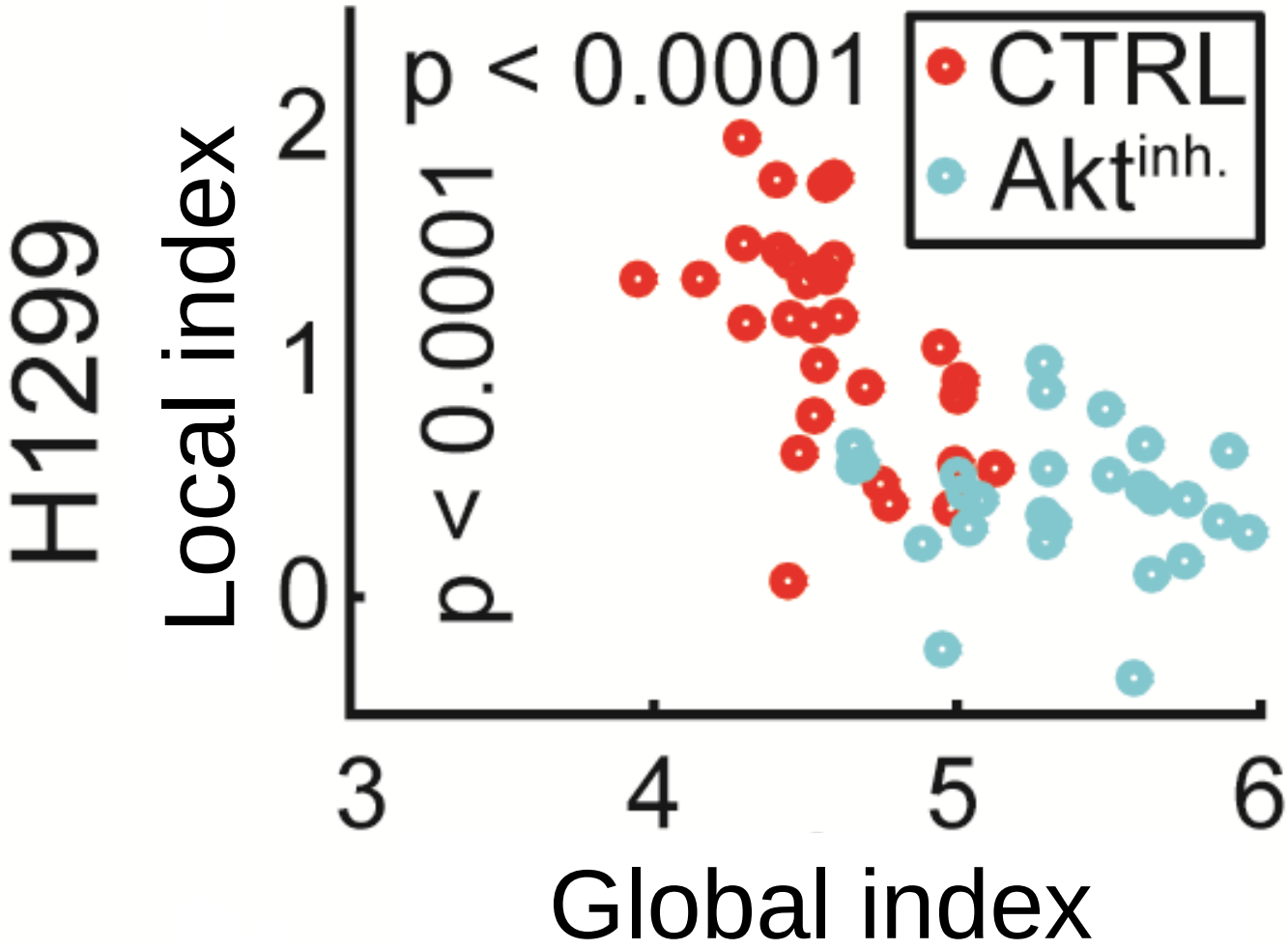
Marcel Mettlen

Cross-talk between signaling receptors (AKT) and components of the endocytic machinery

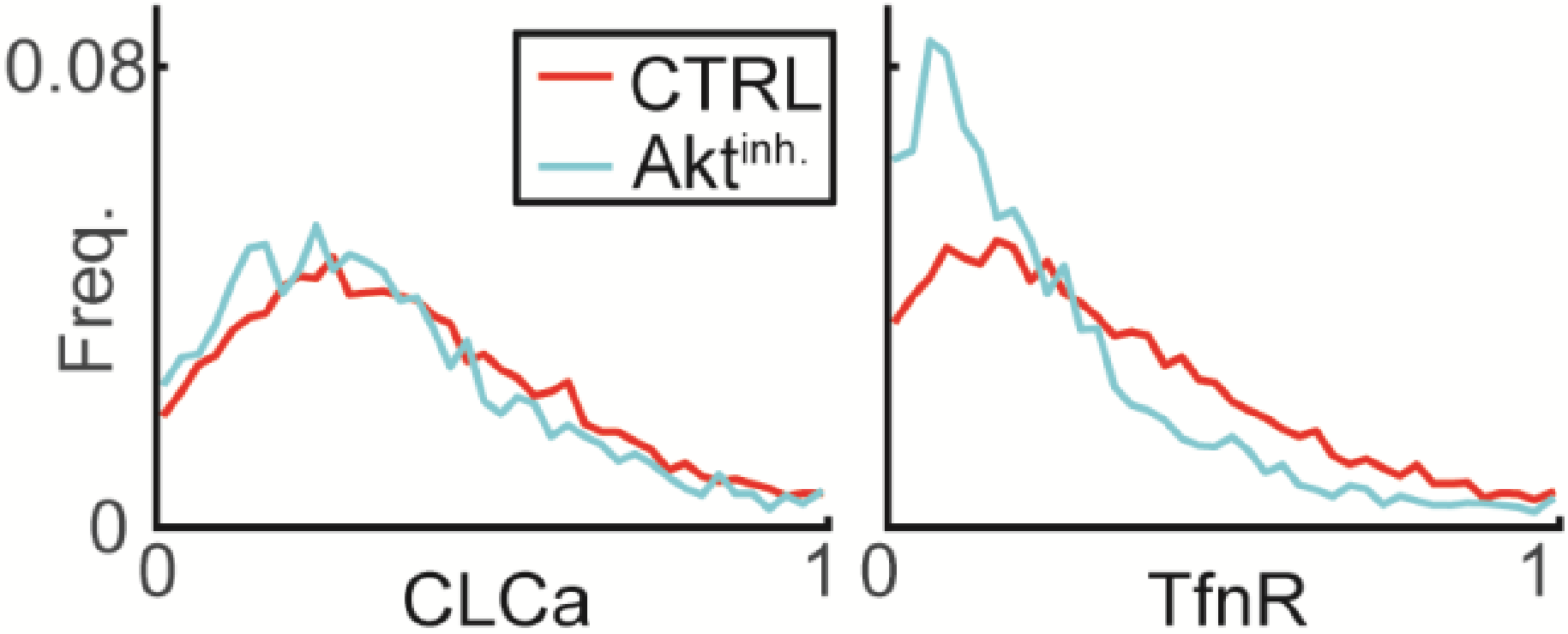


Carlos Reis, H1299 cells

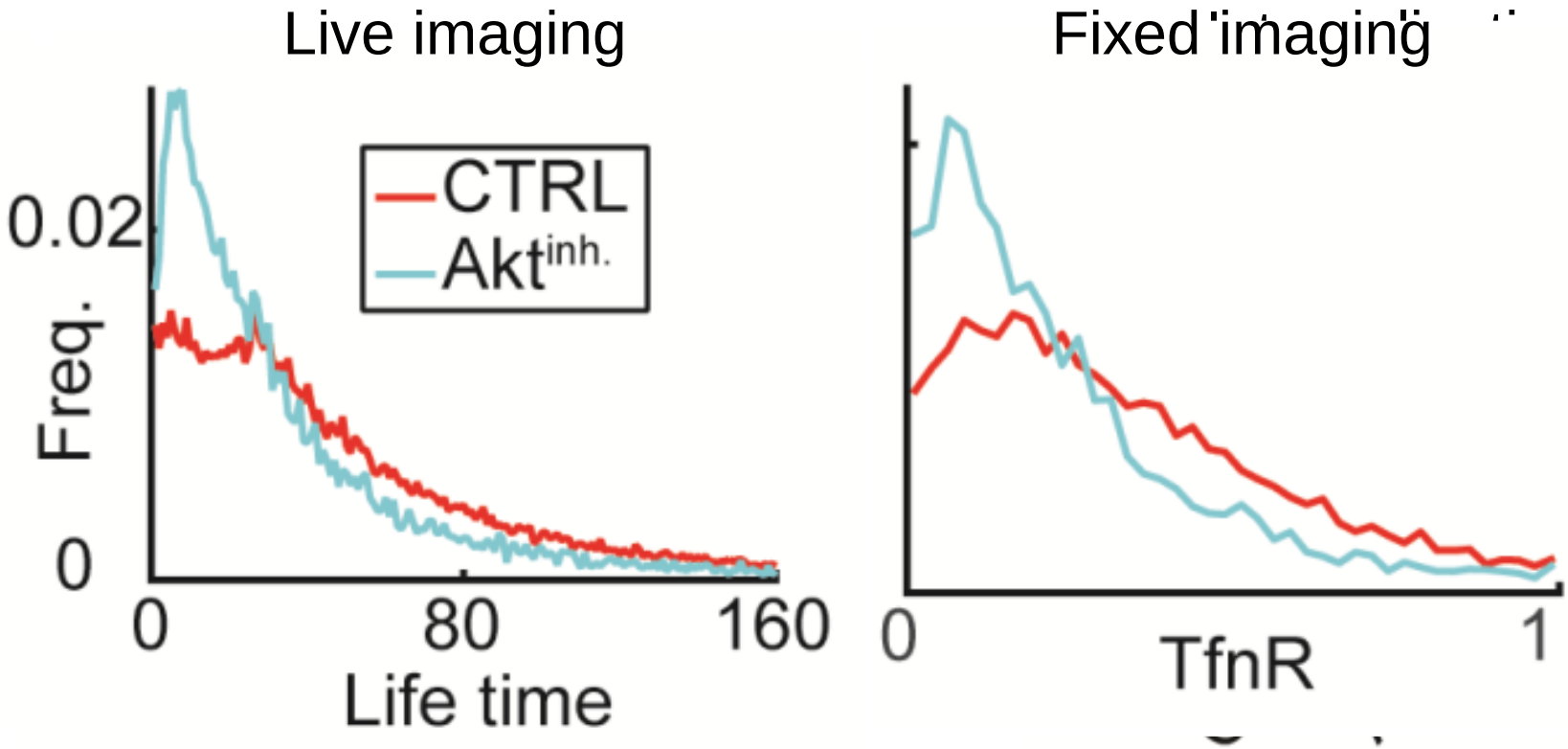
Inferring co-localization and predicting dynamics from fixed cells during clathrin-mediated endocytosis (CME)



Global bias: reduced TfnR in CCPs upon Akt inhibition

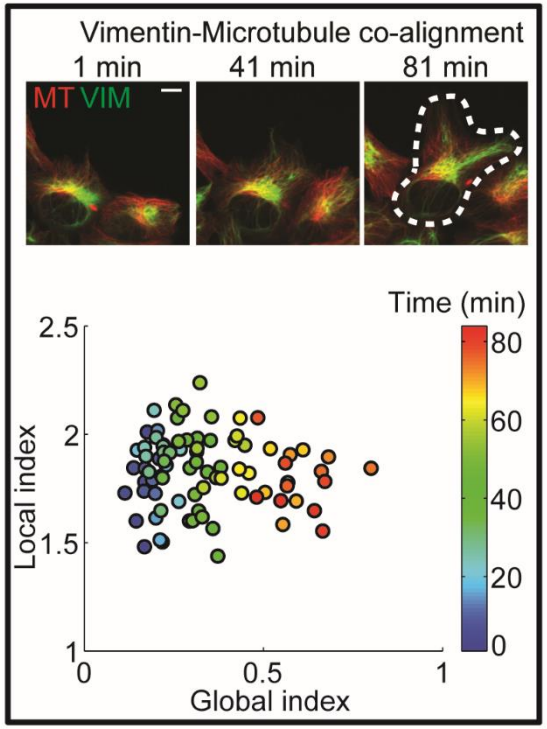


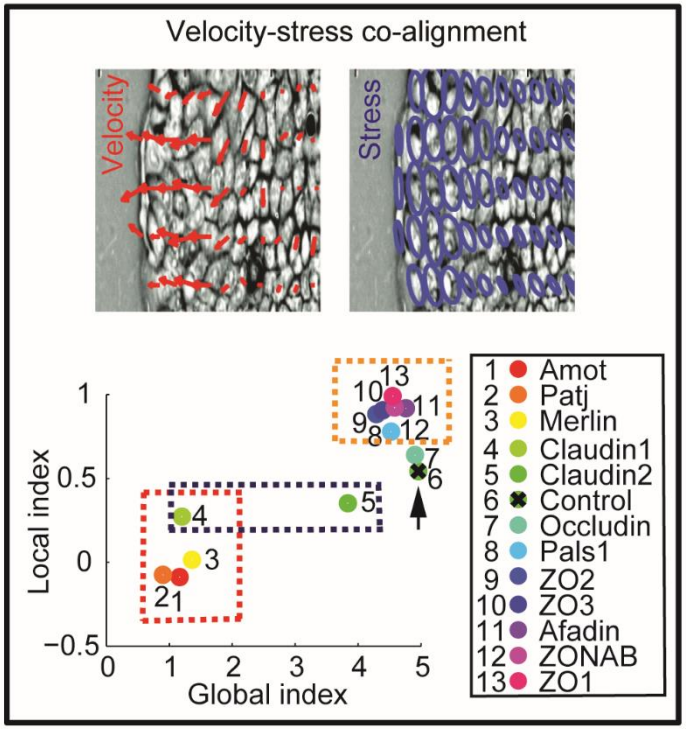
More CCPs containing less TfnR alter CCPs dynamics upon AKT inhibition

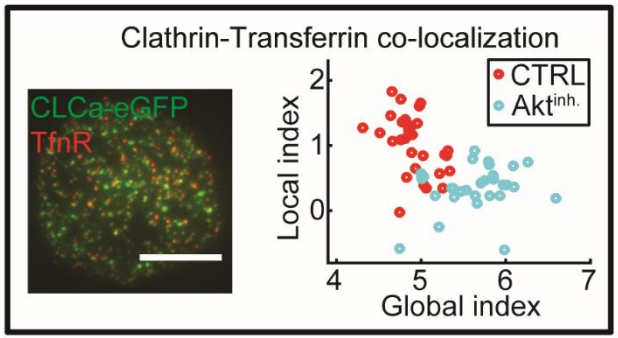


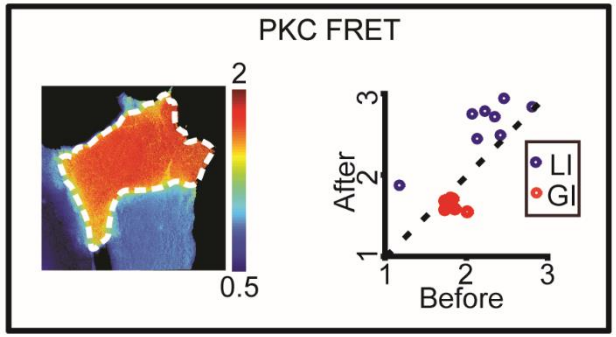
Reduced TfnR in CCPs upon Akt inhibition increased short-lived, (most likely) abortive events → decrease in CME efficiency

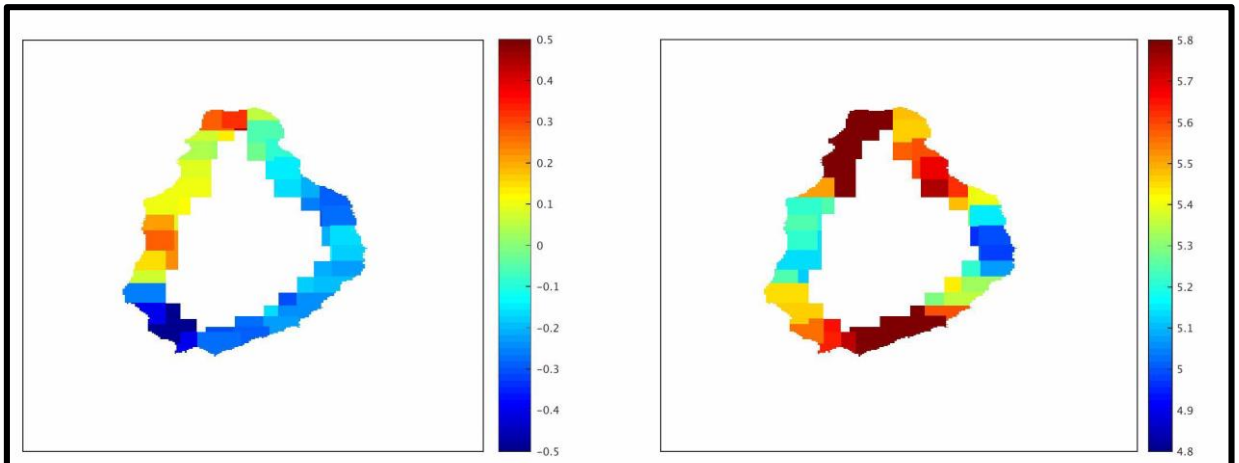
Global bias: more CCPs with less TfnR upon Akt inhibition











Spatio-temporal co-localization of Rac1 and Vav1 activity in a migrating cell (With Dan Marston, UNC)

Resources

Webserver: https://debias.biohpc.swmed.edu/

Source code: https://github.com/DanuserLab/DeBias

https://elifesciences.org/content/6/e22323

Take home message
DeBias enables identifying the gorilla

Simons and Chabris (1999)

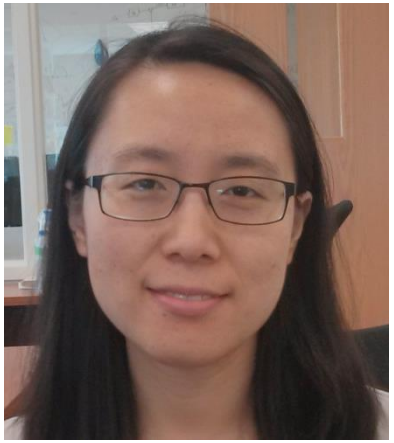
Acknowledgments

UT SOUTHWESTERN MEDICAL CENTER

Uri Obolski, (Theory)

Carlos Reis (Endocytosis)

Zhuo Gan, (Vimentin, PKC)



Yi Du (Webserver)

Tamal Das

Joachim Spatz

Liqiang Wang

Liya Ding

Christoph Burckhardt

Sandy Schmid

Gaudenz Danuser

https://elifesciences.org/content/6/e22323

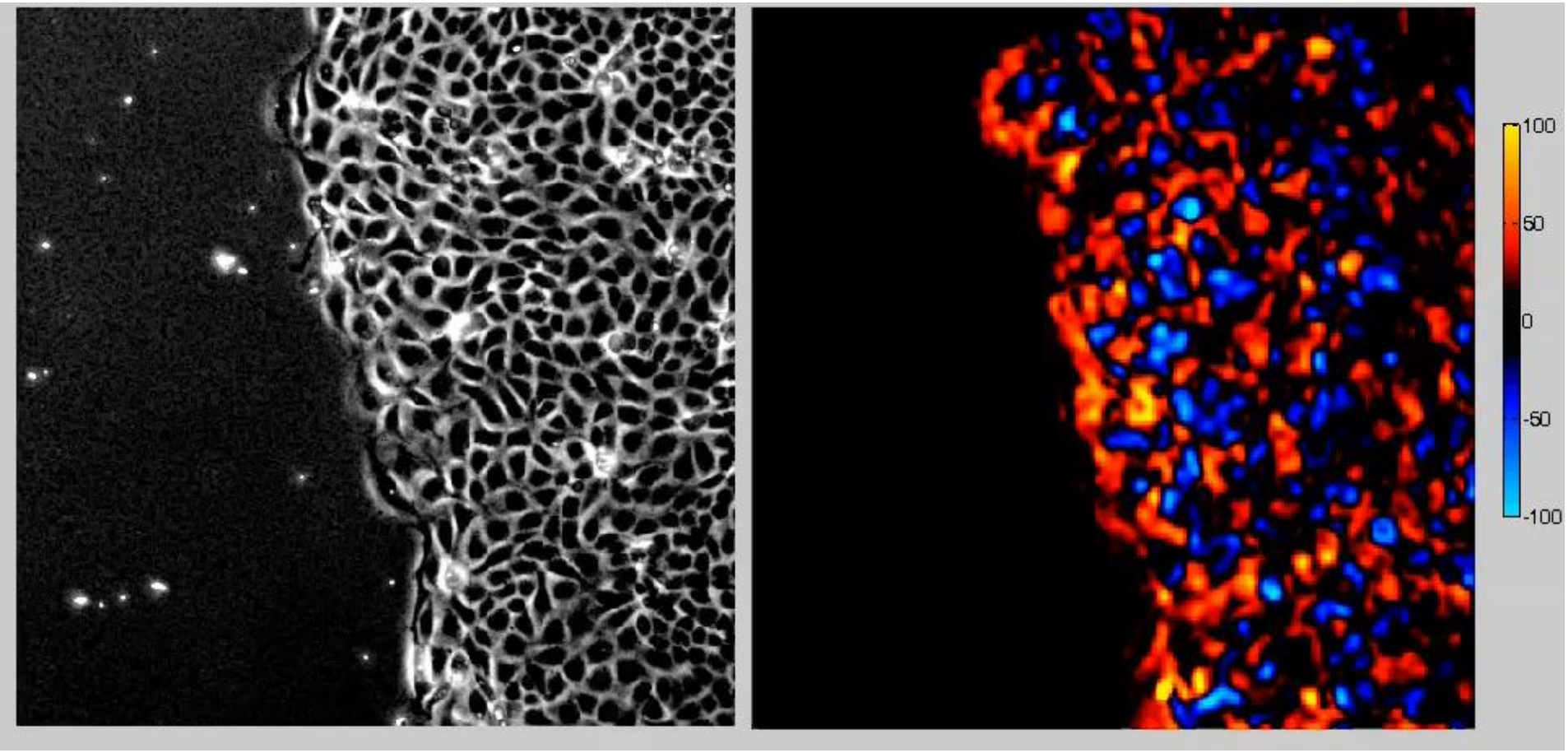
Origins of DeBias:
Reusing cell image data for new biological insight

The interplay between development of quantitative tools ("hammers“) and identifying open important questions in cell biology ("nails“)

How (global) coordination emerges from (local) heterogeneous traction forces?

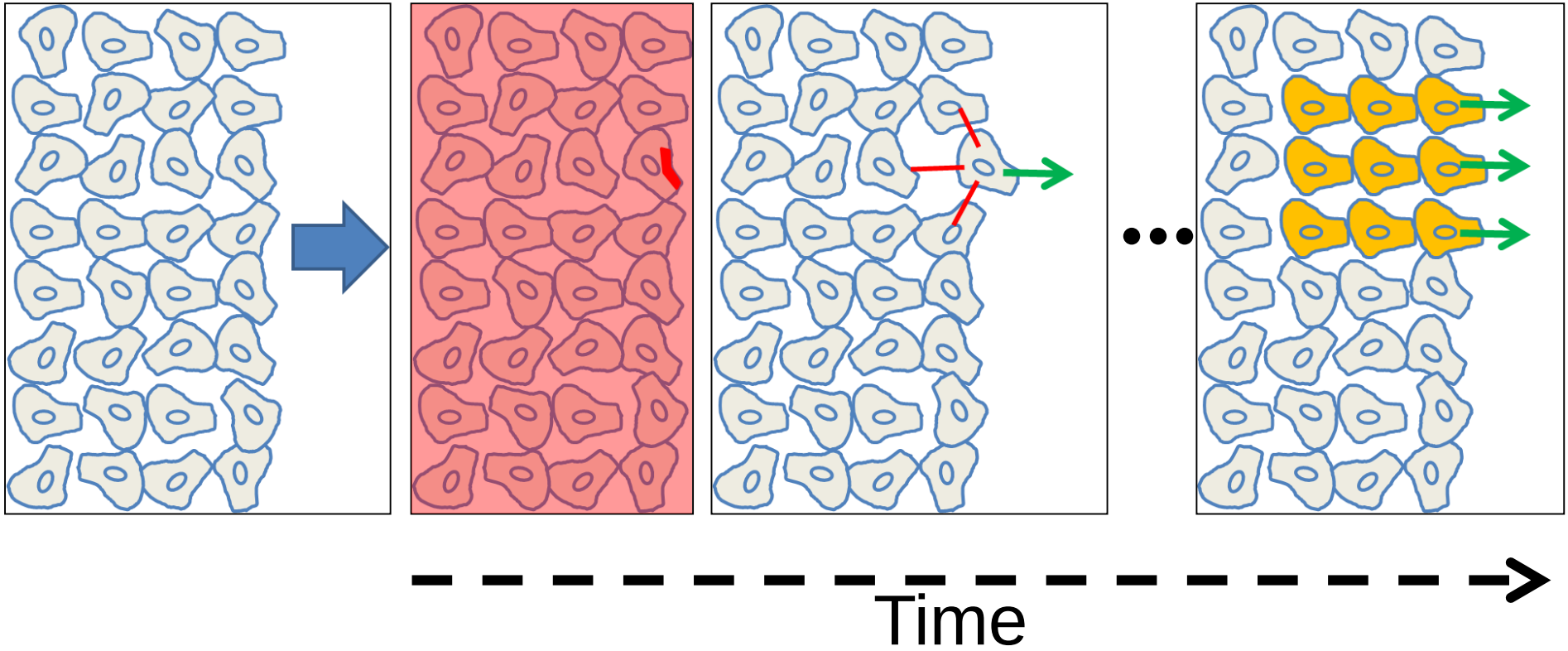
Phase Contrast

Traction T_x (P_a)



Trepat et al. (2009)

Suggested model

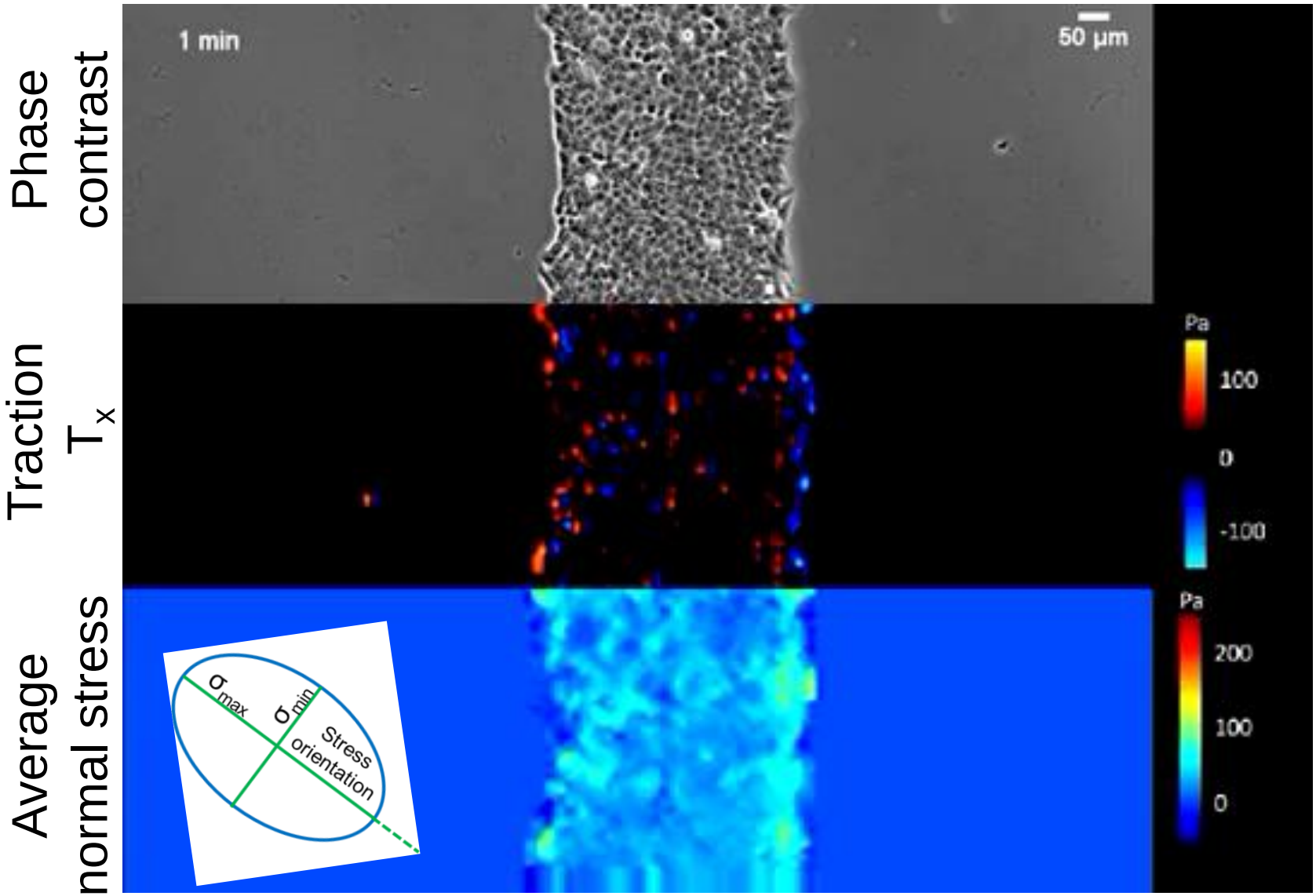


Stochastic force exertion transform to directional migration

Strain on neighbors coordinate their movement

Propagation in time and space to guide groups of cells

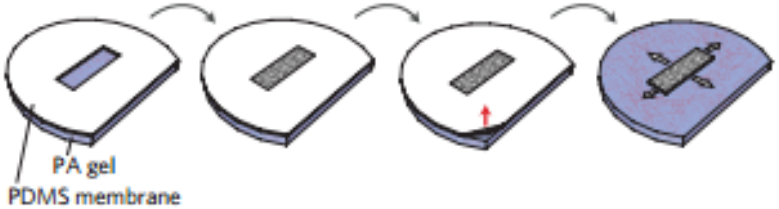
Measuring traction force, stress and velocity

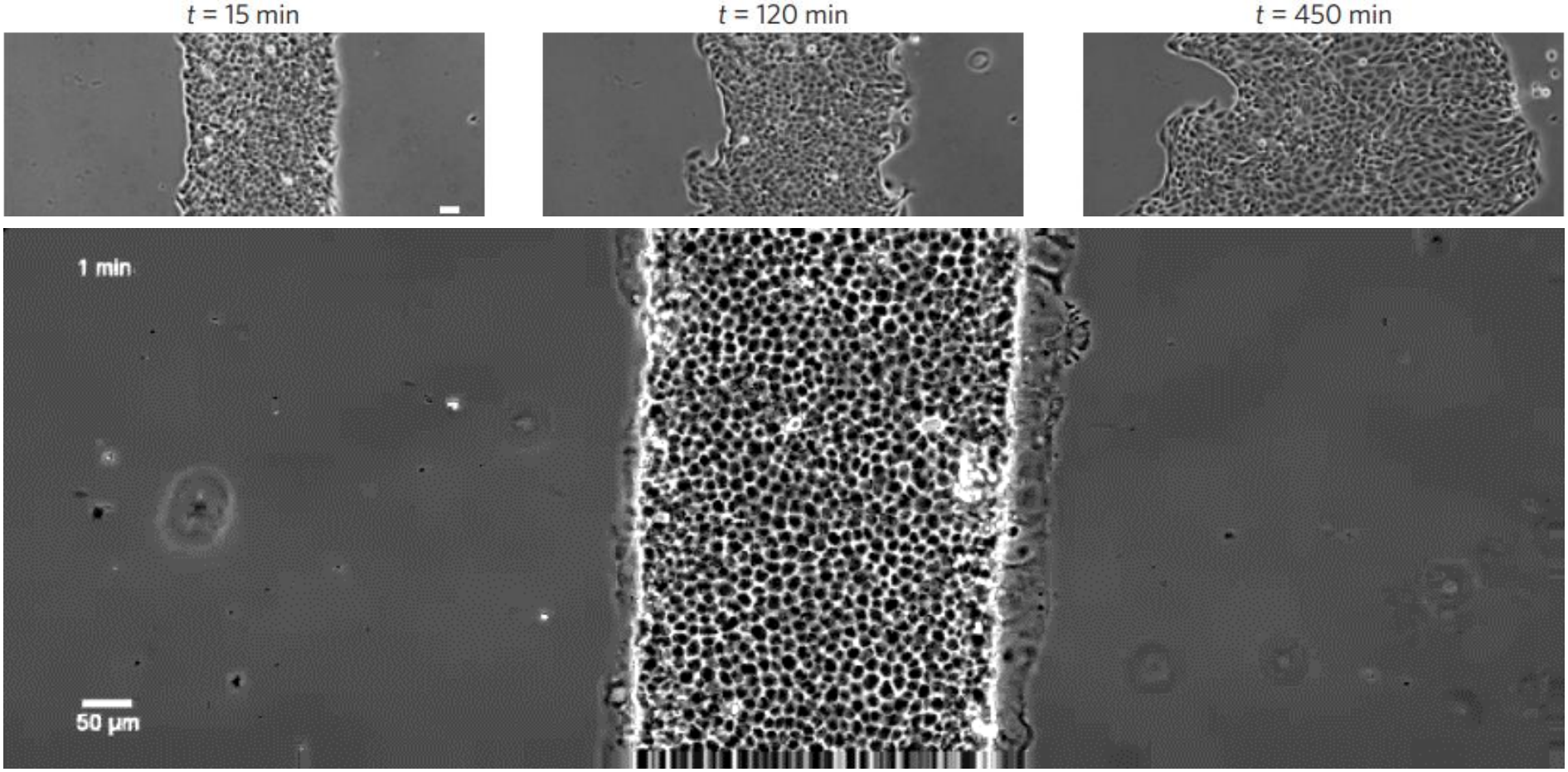


Trepat et al. (2009)
Tambe et al. (2011)
Serra-Picamal & Conte et al. (2012)

Data from: Serra-Picamal and Conte et al.

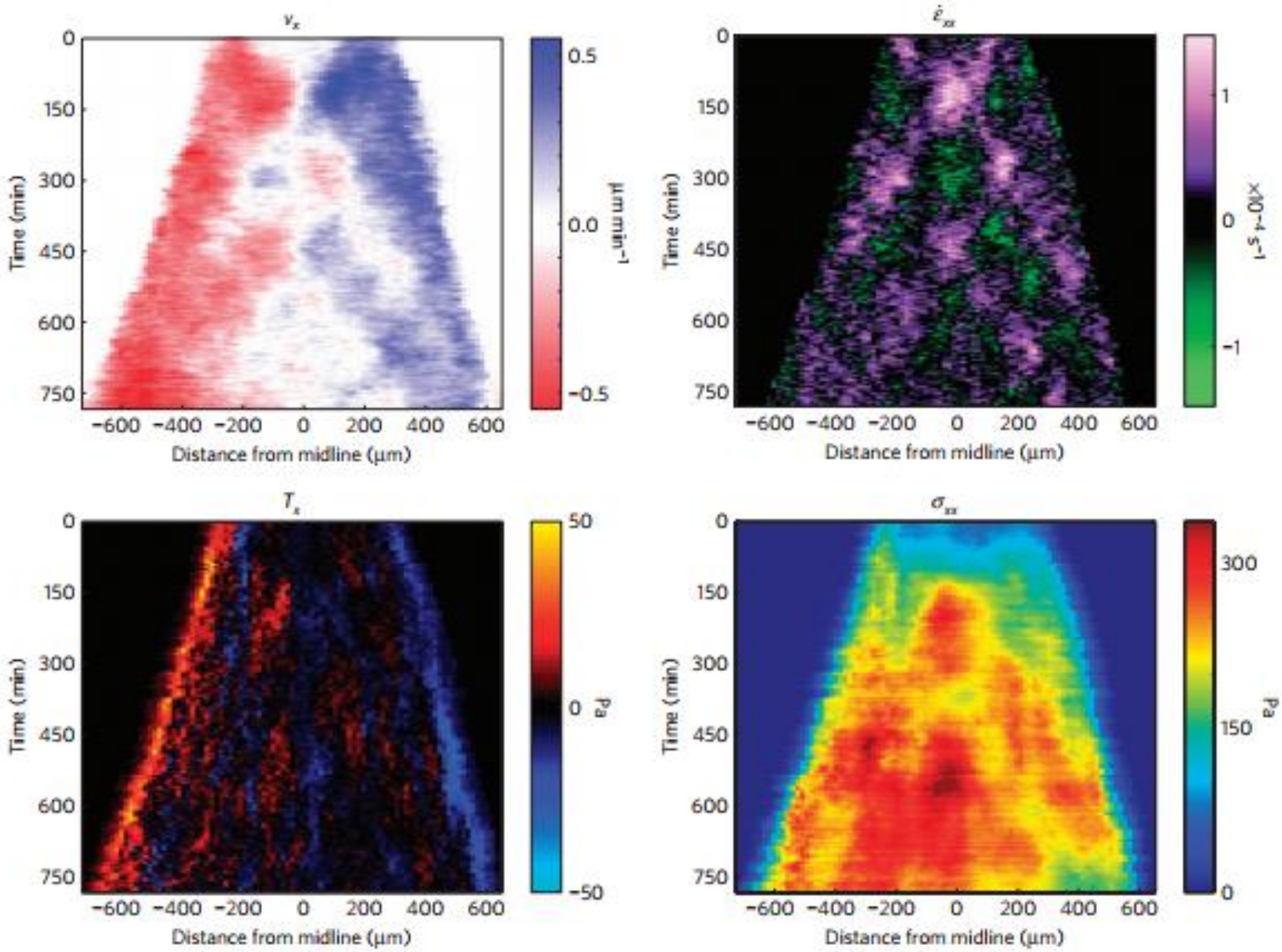
MDCK cells, N = 4 experiments





Serra-Picamal & Conte et al. (2012)

Strain rate X-waves

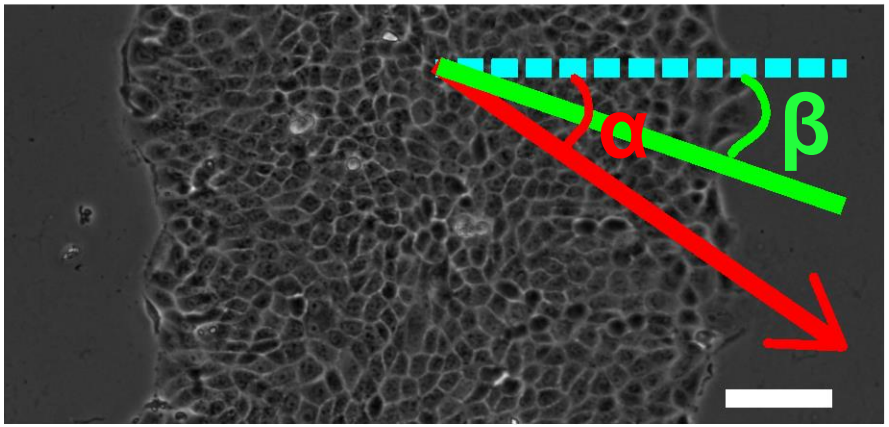


Serra-Picamal & Conte et al. (2012)

Alignment of motion and stress

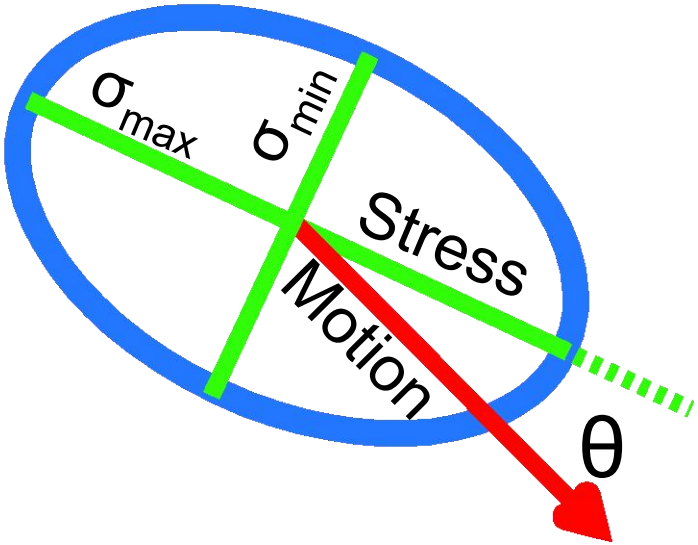
Motion-stress alignment

Velocity angle,
stress orientation



$$-90 \leq \propto, \beta \leq 90$$

Motion-stress
alignment

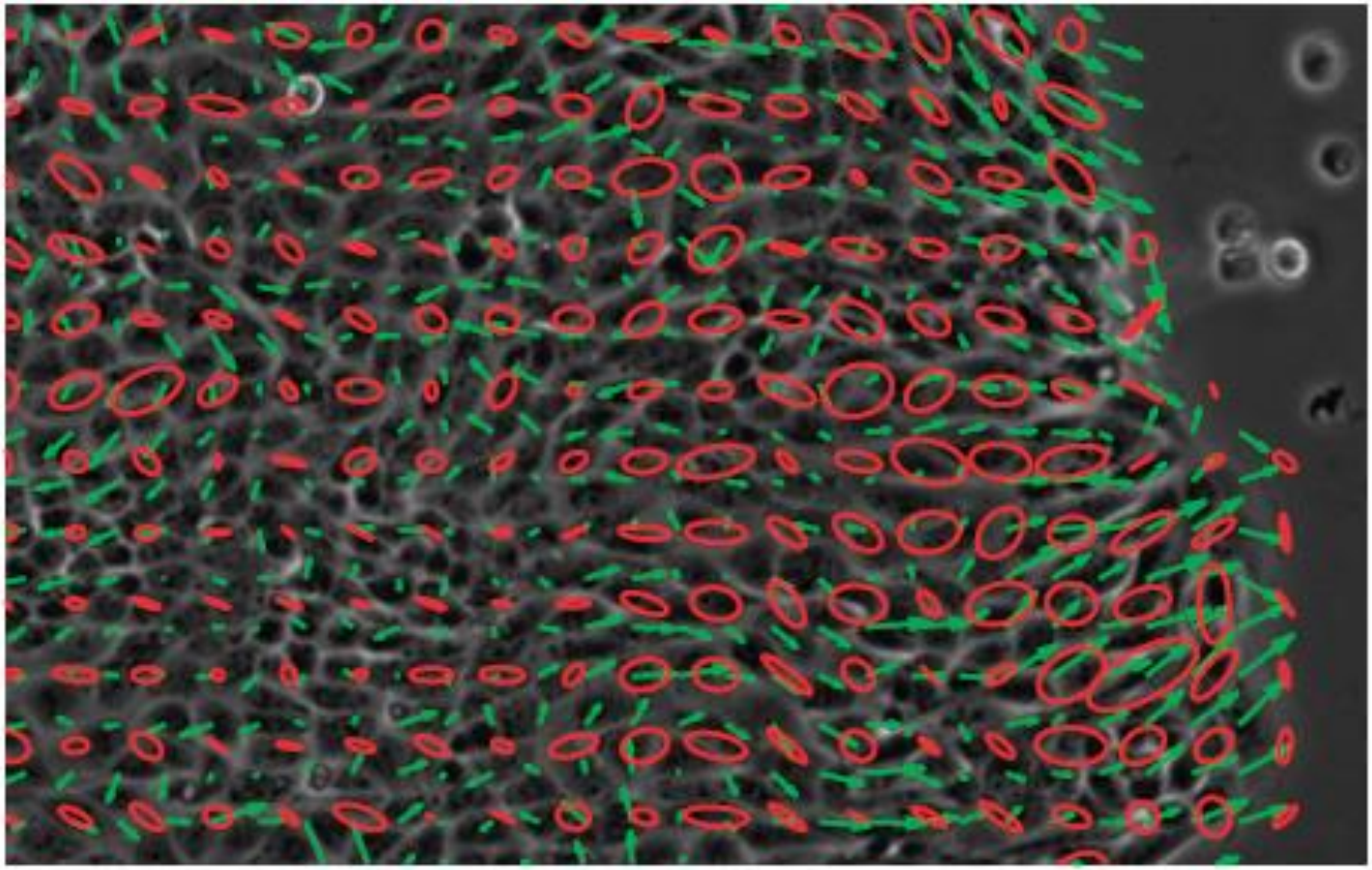


$$0 \leq \theta \leq 90$$

Tambe et al. (2011)
Trepat & Fredberg. (2011)

Plithotaxis

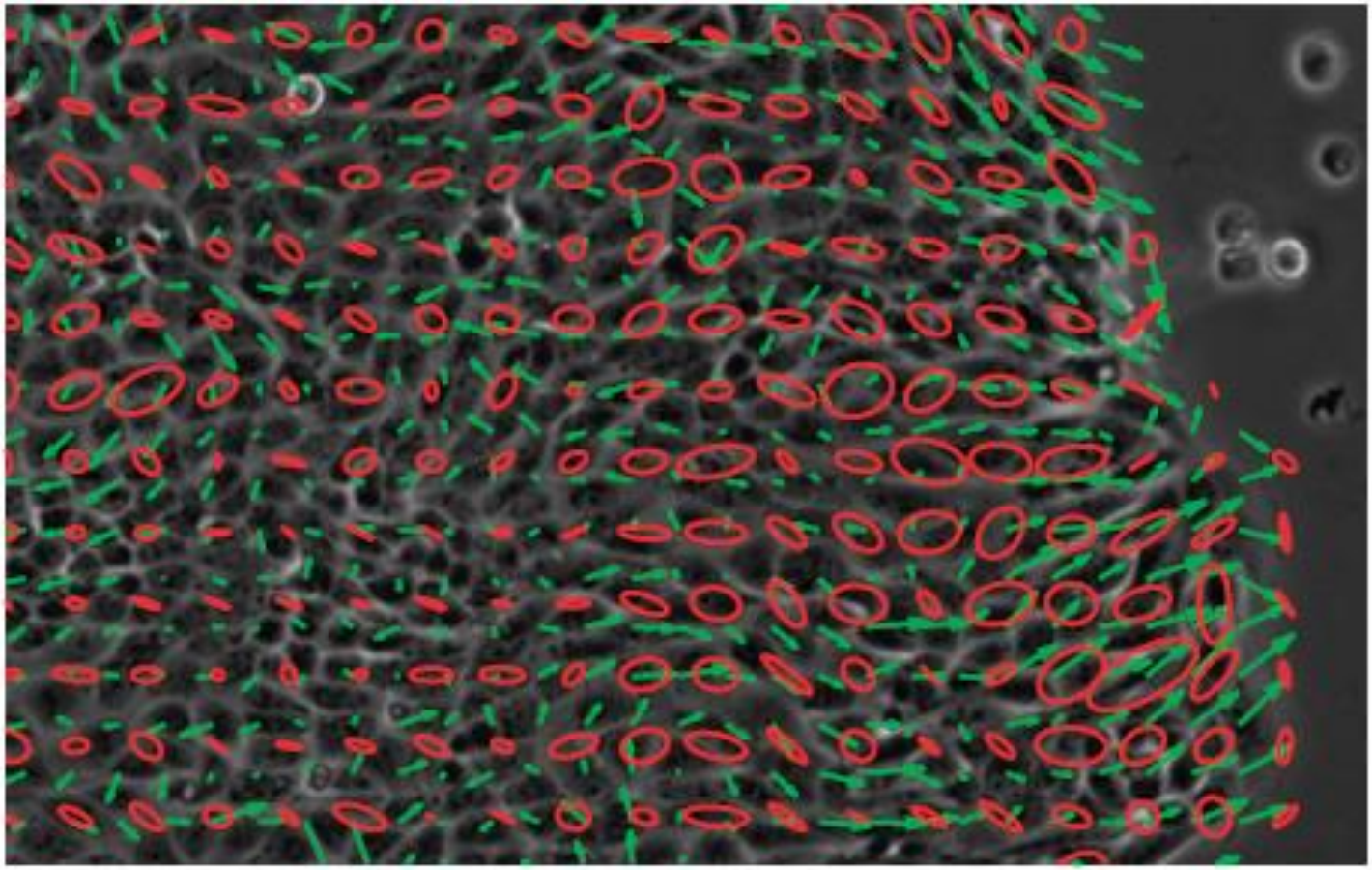
“tendency for each individual cell within a monolayer to migrate along the local orientation of the maximal principal stress.”



Tambe et al. (2011)
Trepat and Fredberg (2011)
Serra-Picamal and Conte et al. (2012)

Plithotaxis

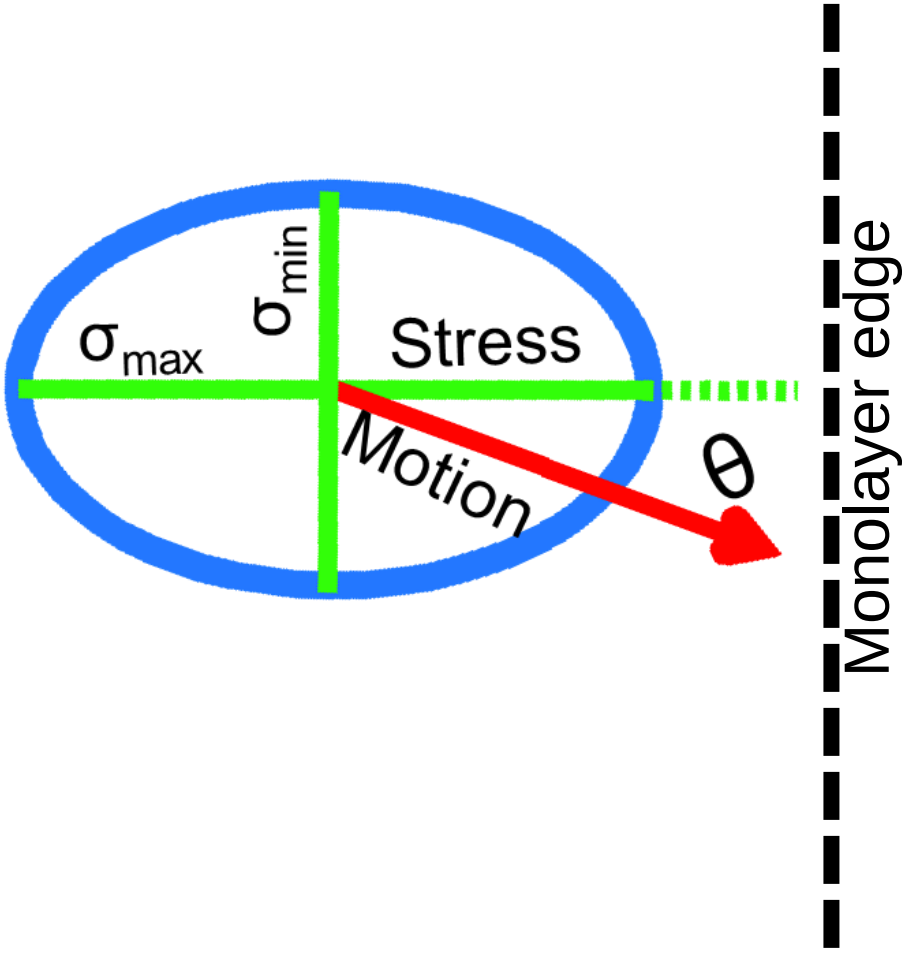
"tendency for **each individual** cell within a monolayer to migrate along the **local orientation** of the maximal principal stress."



Tambe et al. (2011)
Trepat and Fredberg (2011)
Serra-Picamal and Conte et al. (2012)

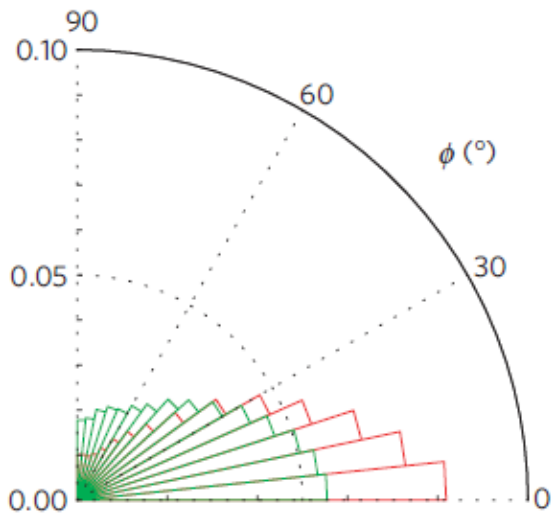
Plithotaxis?

"tendency for **each individual** cell within a monolayer to migrate along the **local orientation** of the maximal principal stress."

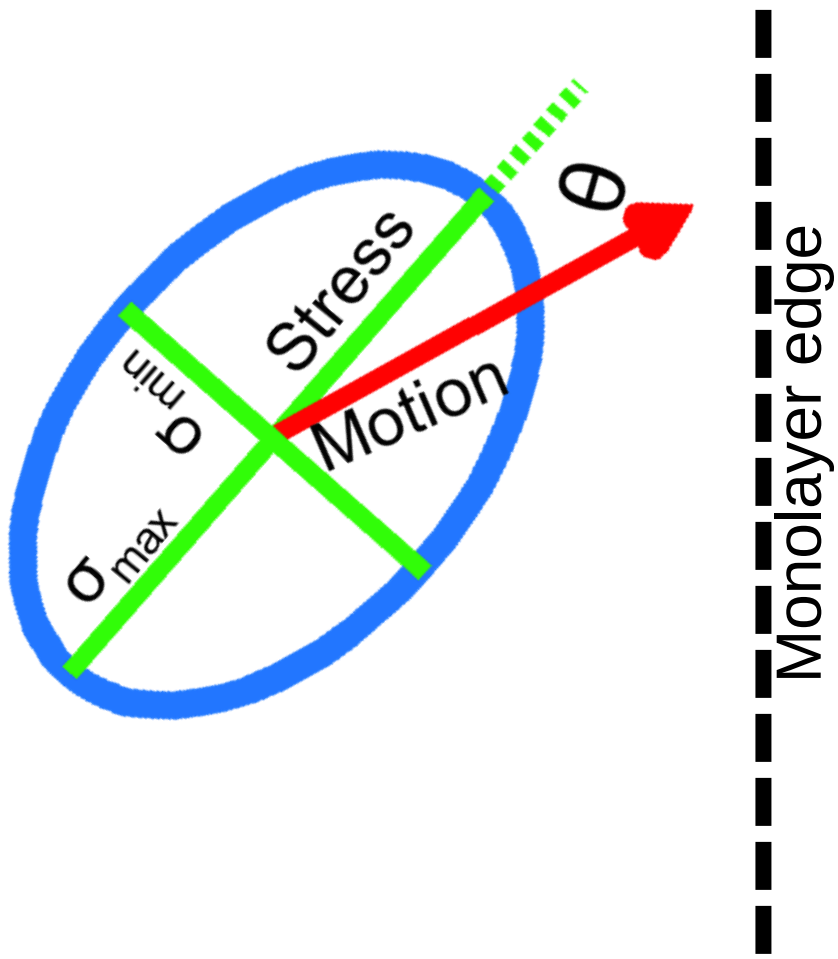


Plithotaxis?

"tendency for each **individual** cell within a monolayer to migrate along the **local orientation** of the maximal principal stress."



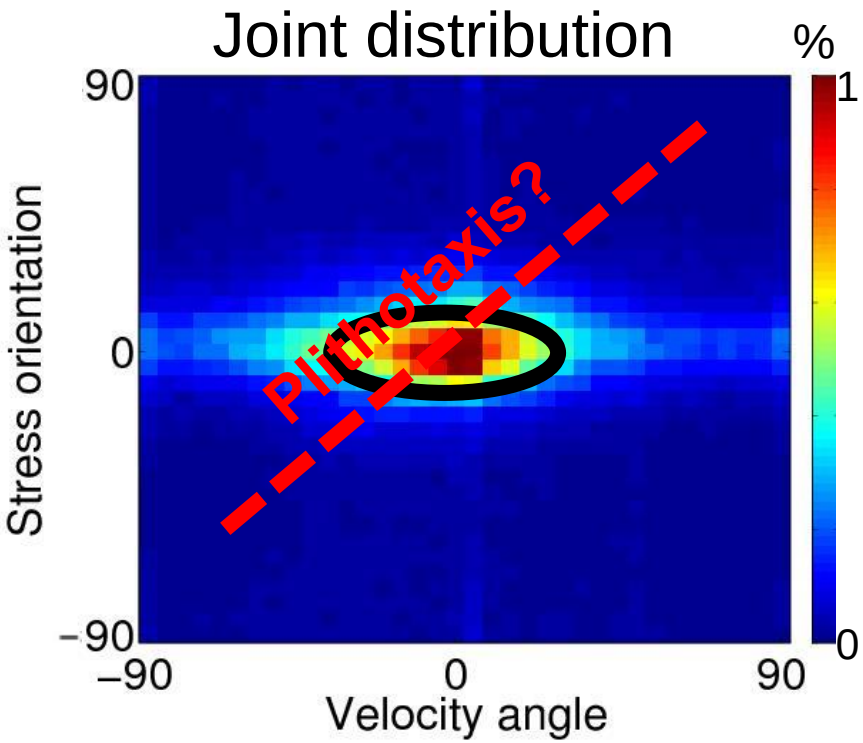
Serra-Picamal and Conte et al. (2012)

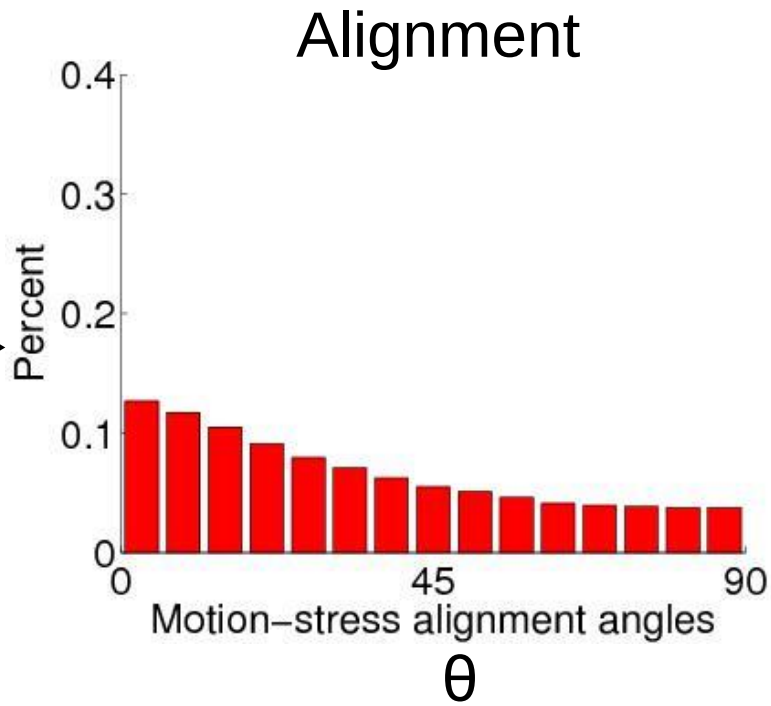


Part I

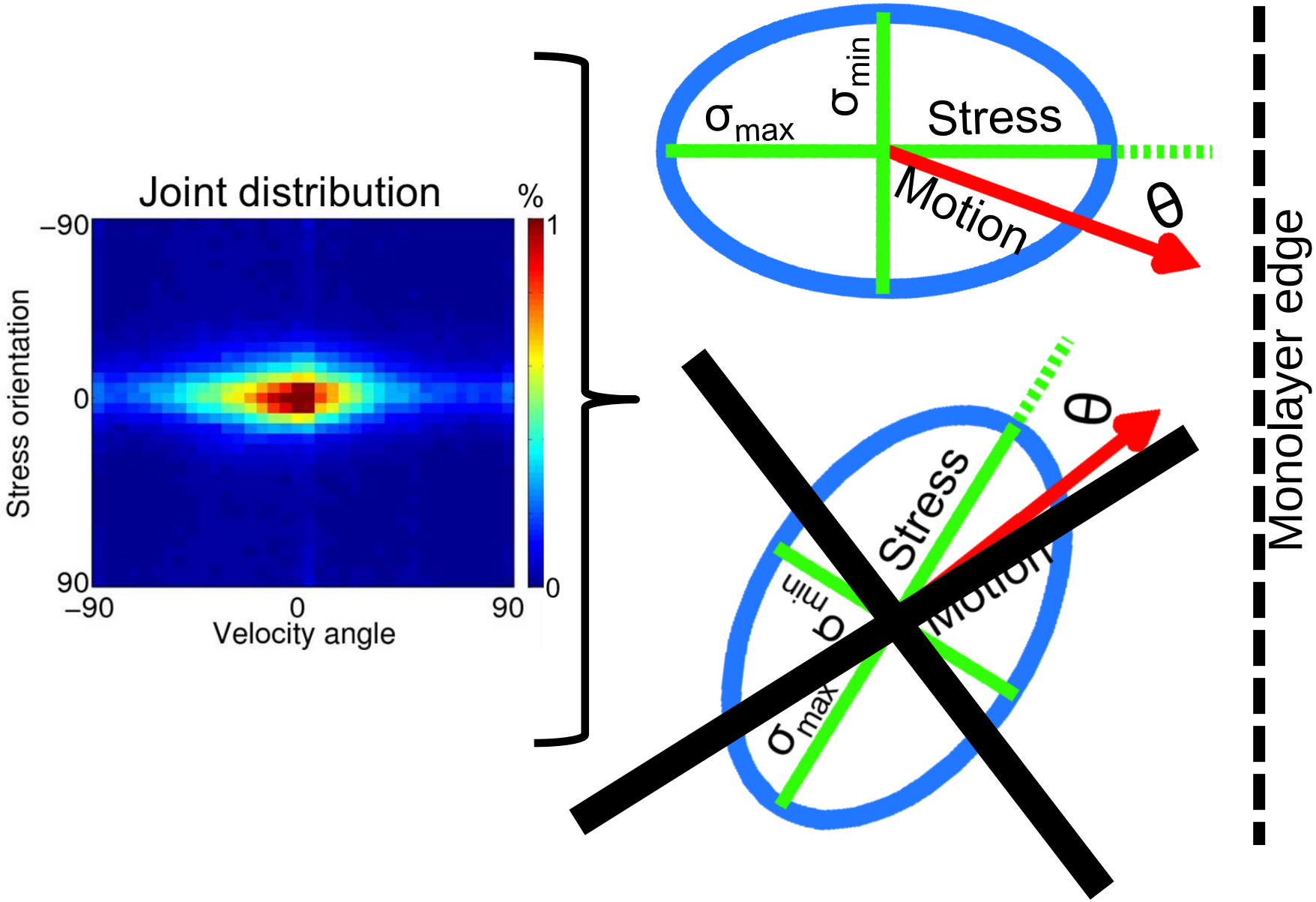
The roles of (global) monolayer geometry versus (local) plithotaxis in inducing motion-stress alignment

Motion-Stress Alignment ≠ Plithotaxis





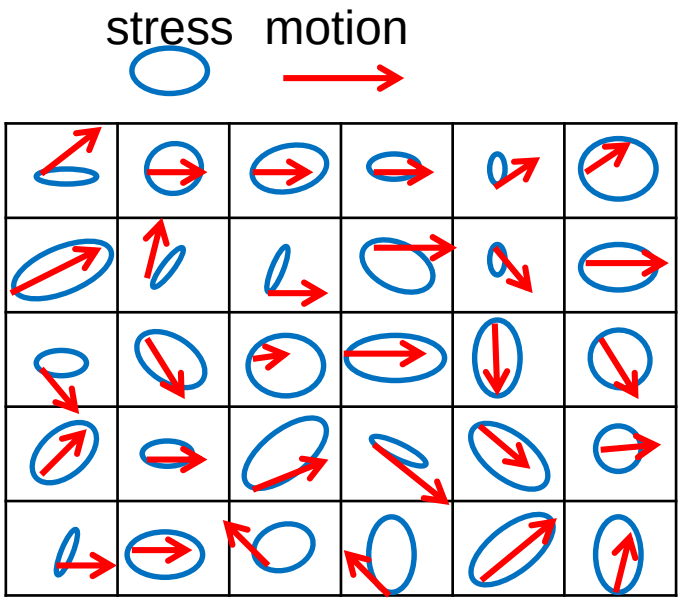
No Evidence for Plithotaxis!



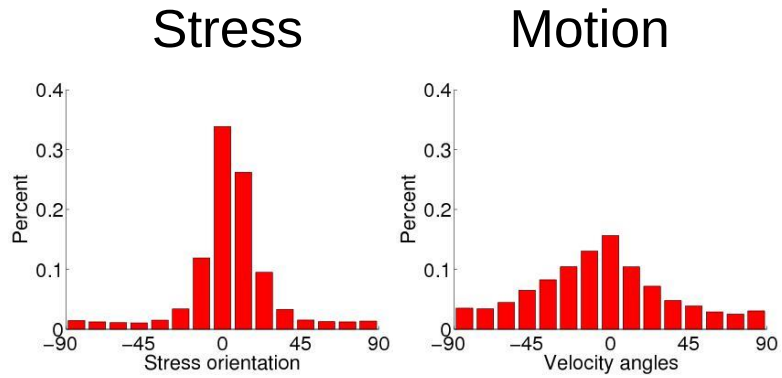
Components of motion-stress alignment

Observed motion-stress alignment = Global contribution (geometry) + ? Local contribution (plithotaxis)

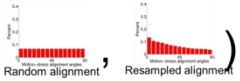
Quantifying the role of monolayer geometry & plithotaxis in motion-stress alignment



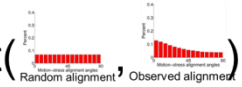
Discard pairwise orientation



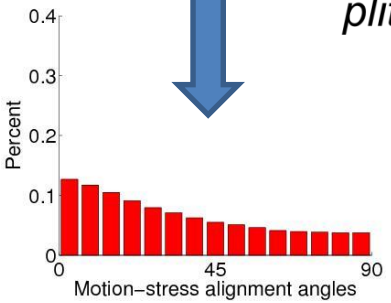
Random resampling

geometry ≤ dist(

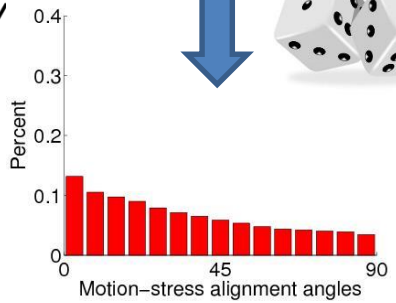
)

plithotaxis ≥ dist(

) - *geometry*

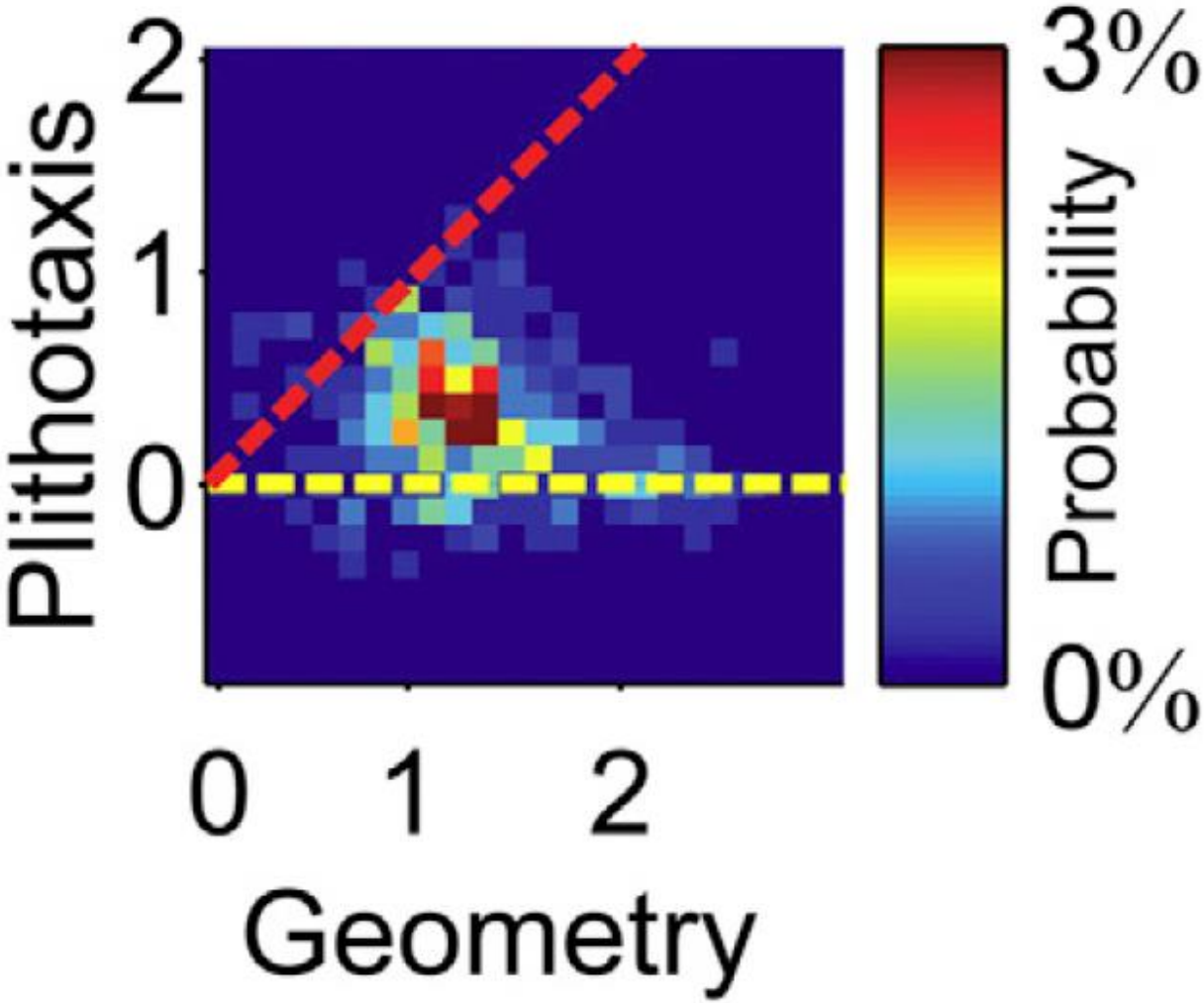


Observed alignment



Resampled alignment

Plithotaxis exists!



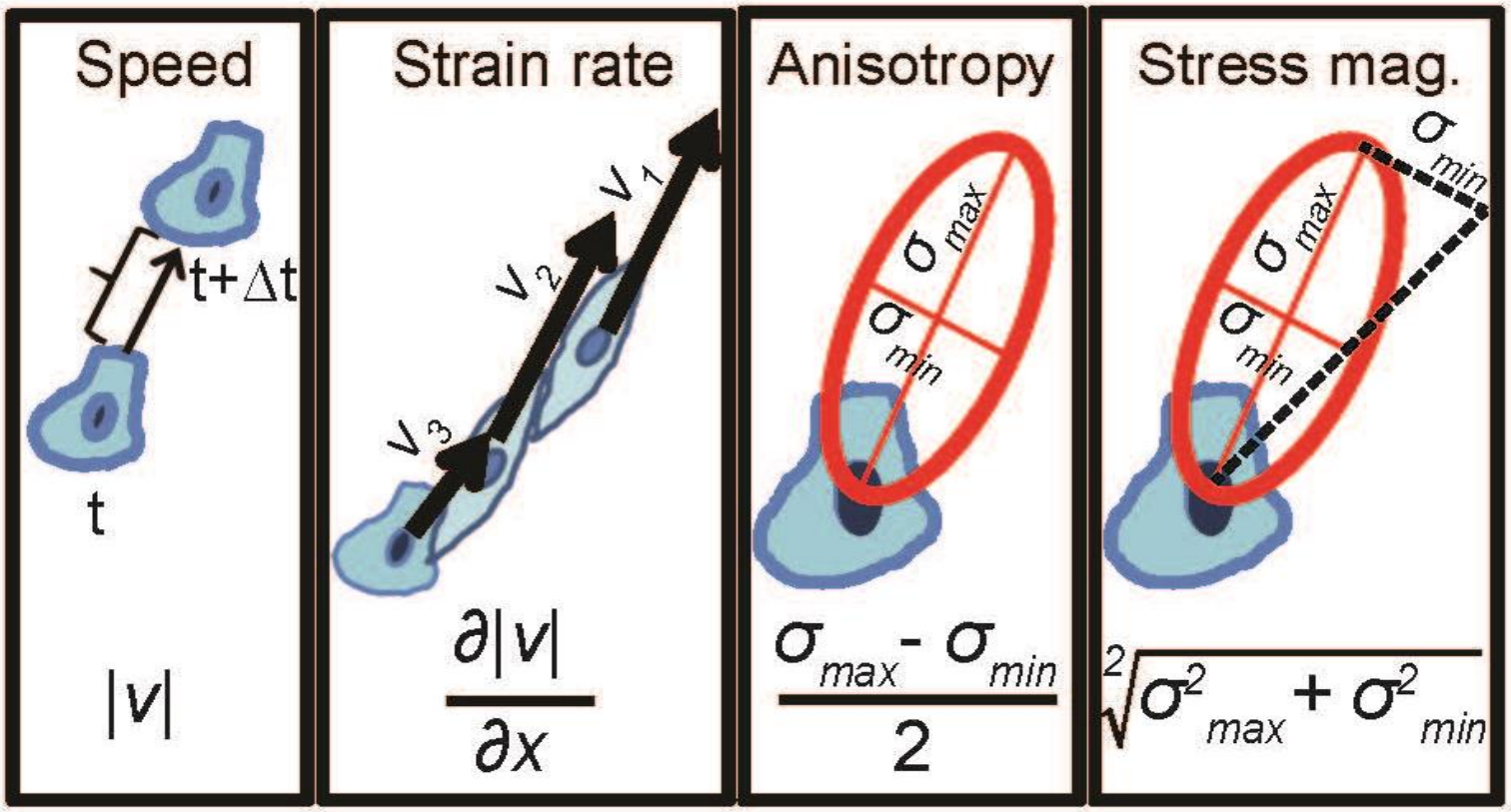
(N = 4 independent experiments, n = 96 frames in each)

Geometry > Plithotaxis

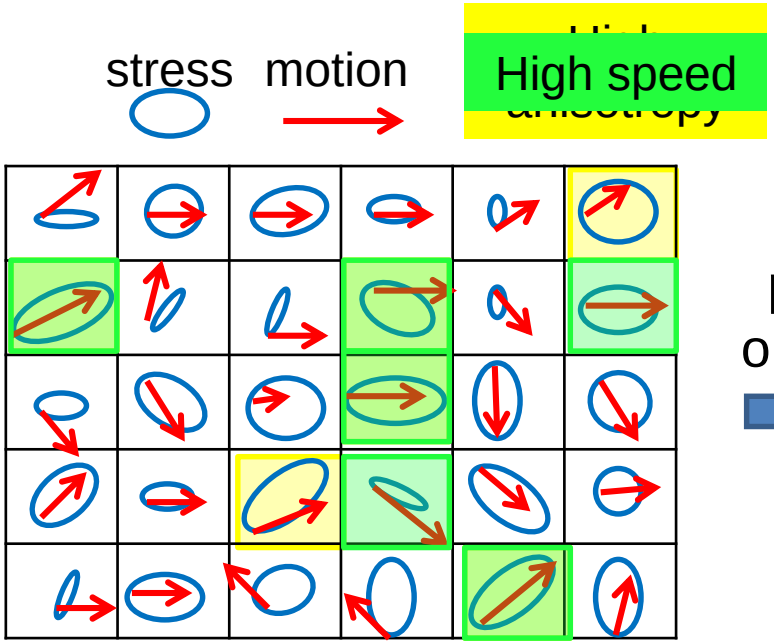
Part II
Properties of cells exhibiting plithotaxis and motion-stress alignment

<u>Working hypothesis</u>: enhanced plithotaxis and motion-stress alignment enables more efficient migration during monolayer expansion

Physical attributes considered

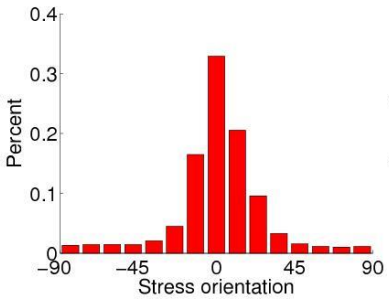


Quantifying plithotaxis for subgroups of cells

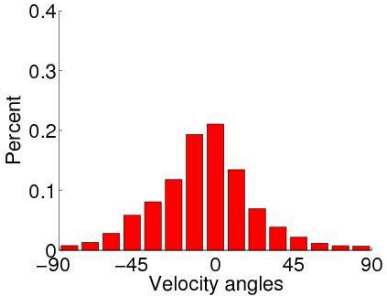


Discard pairwise orientation

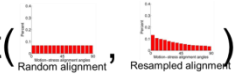
Stress



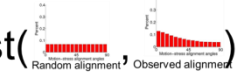
Motion



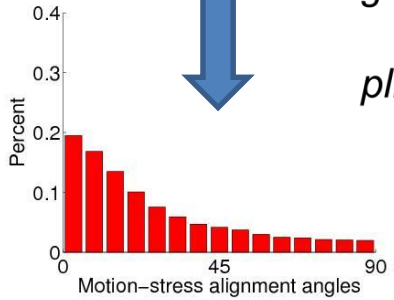
Random resampling

geometry ≤ dist(

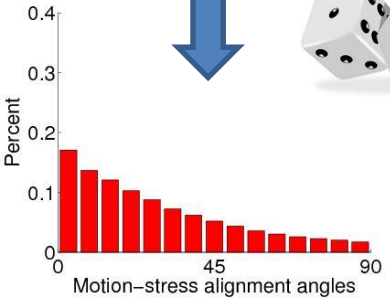
)

plithotaxis ≥ dist(

) - *geometry*



Observed alignment



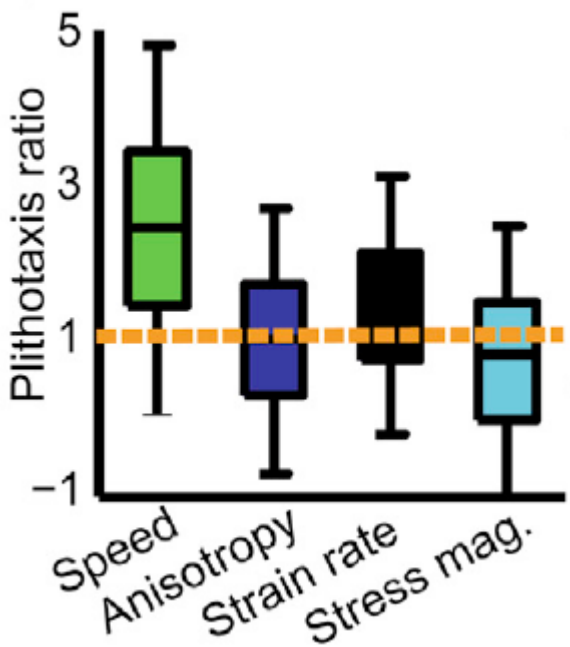
Resampled alignment

Fast cells exhibit elaborated plithotaxis

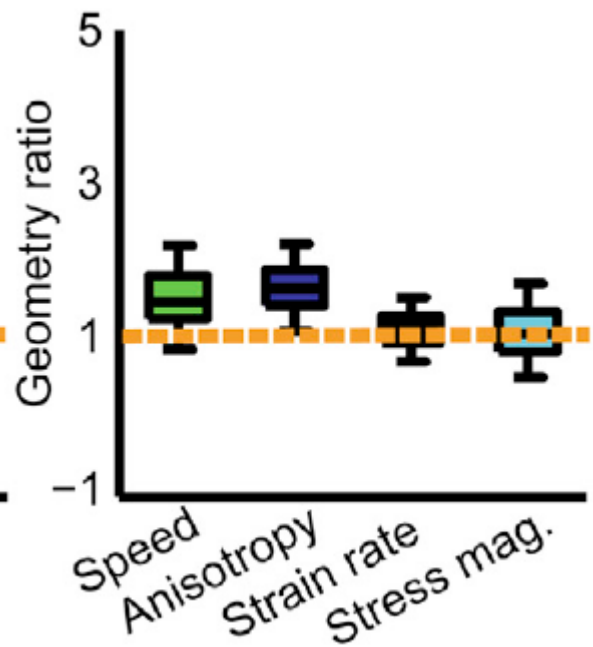
$$Ratio_i = \frac{plithotaxis_i}{plithotaxis_{all}},$$

i = Speed, Anisotropy, Strain rate, Stress magnitude

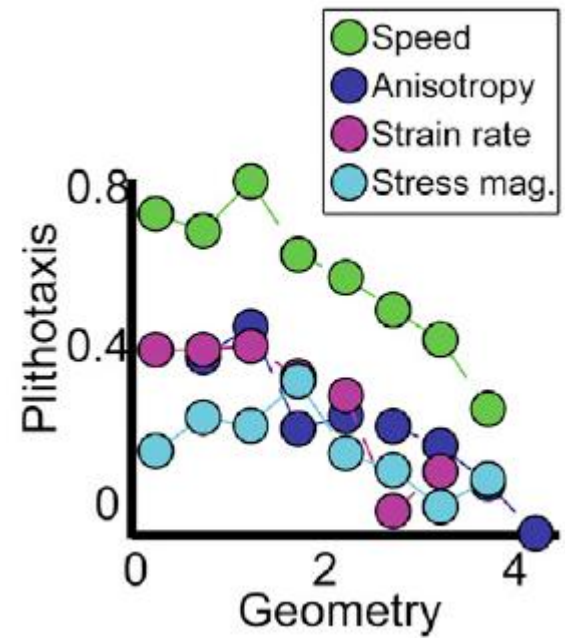
Plithotaxis



Geometry

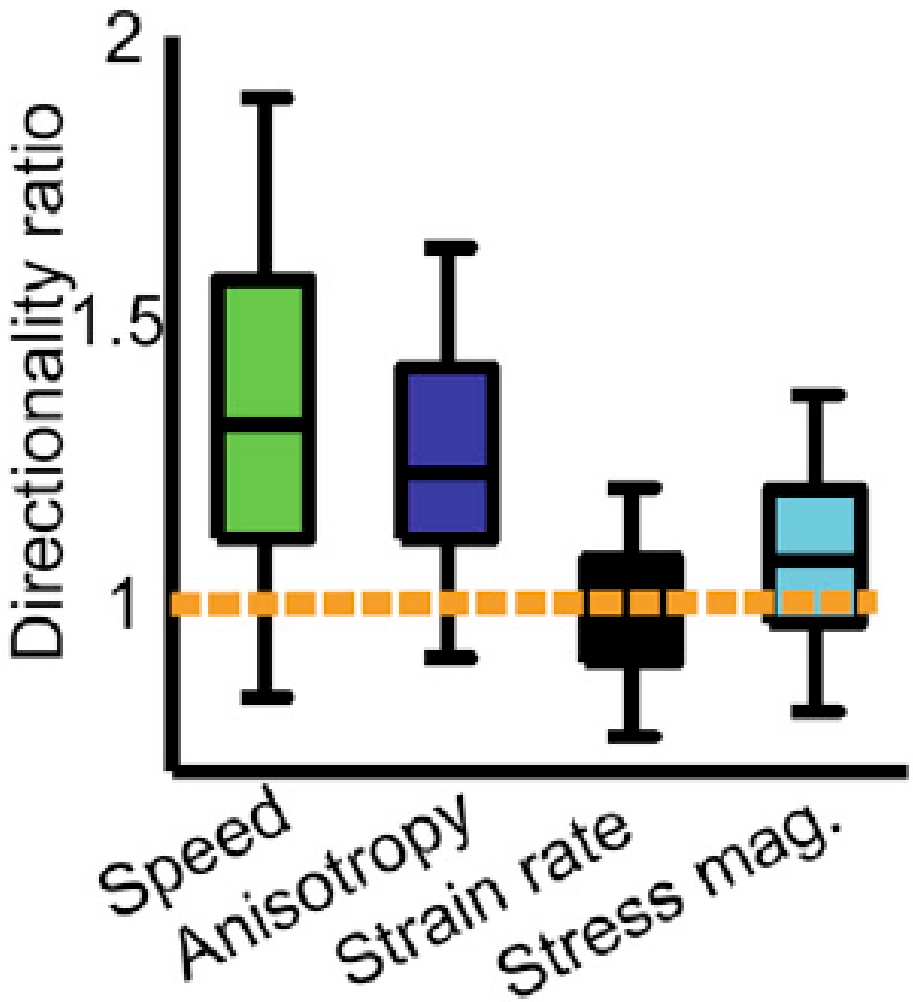


Regardless of geometry index



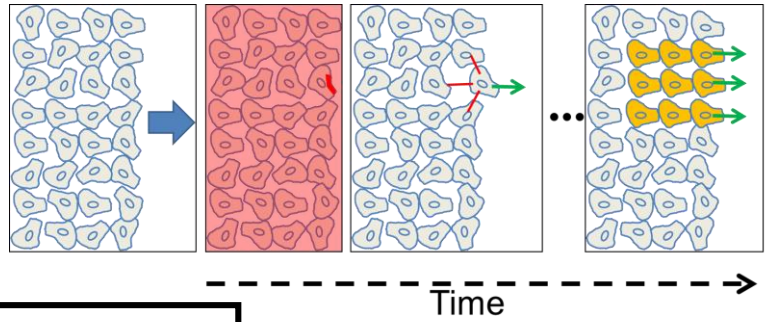
$plithotaxis_{all}$ is calculated for all cells. The ratio is calculated for each frame independently (N = 4 independent experiments, n = 96 frames each)

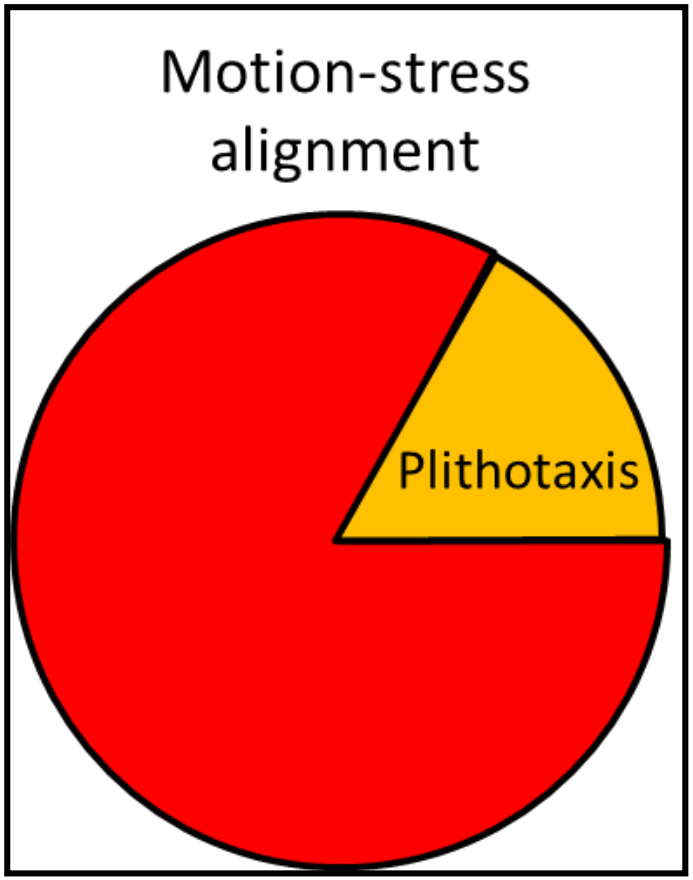
Fast cells exhibit elevated directionality

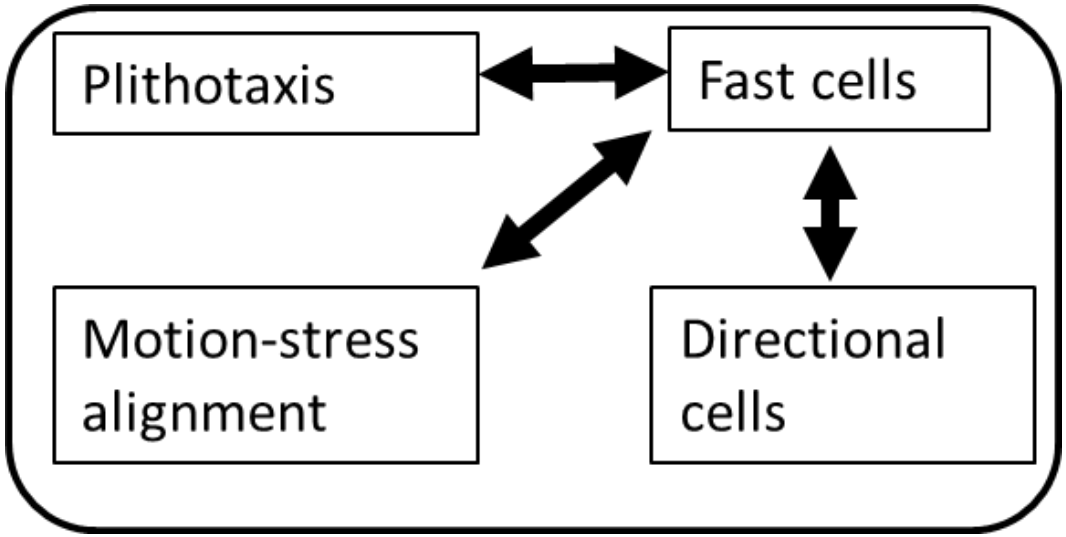


Part I & II: Conclusions

(local alignment of motion and stress)

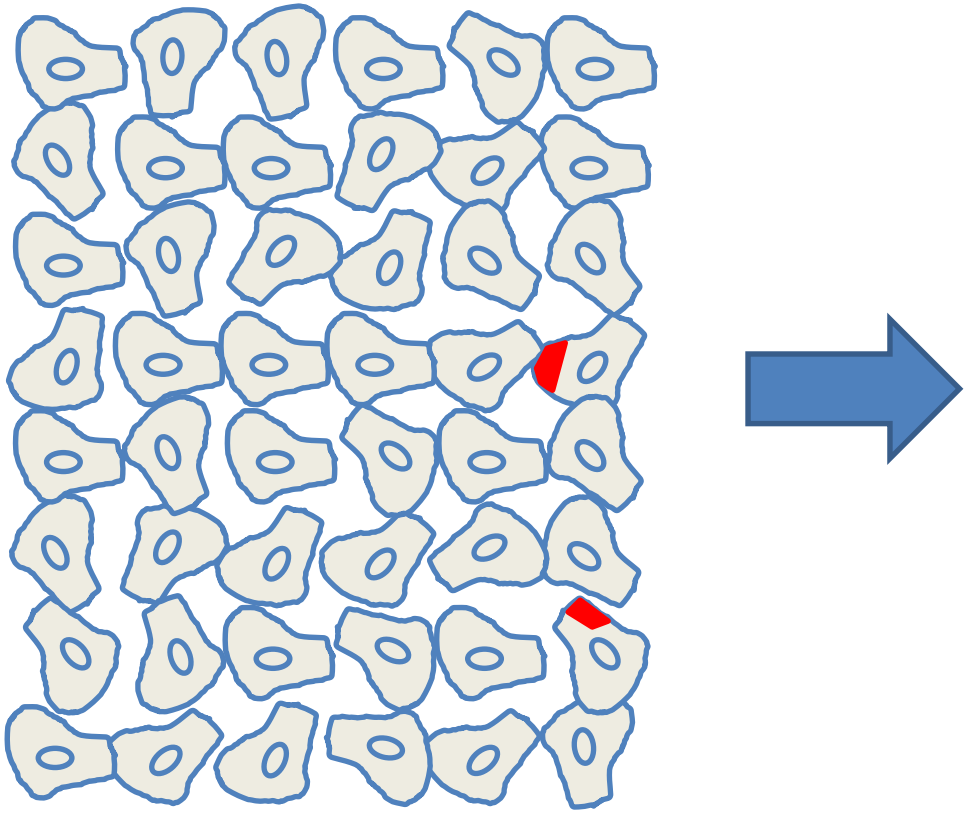




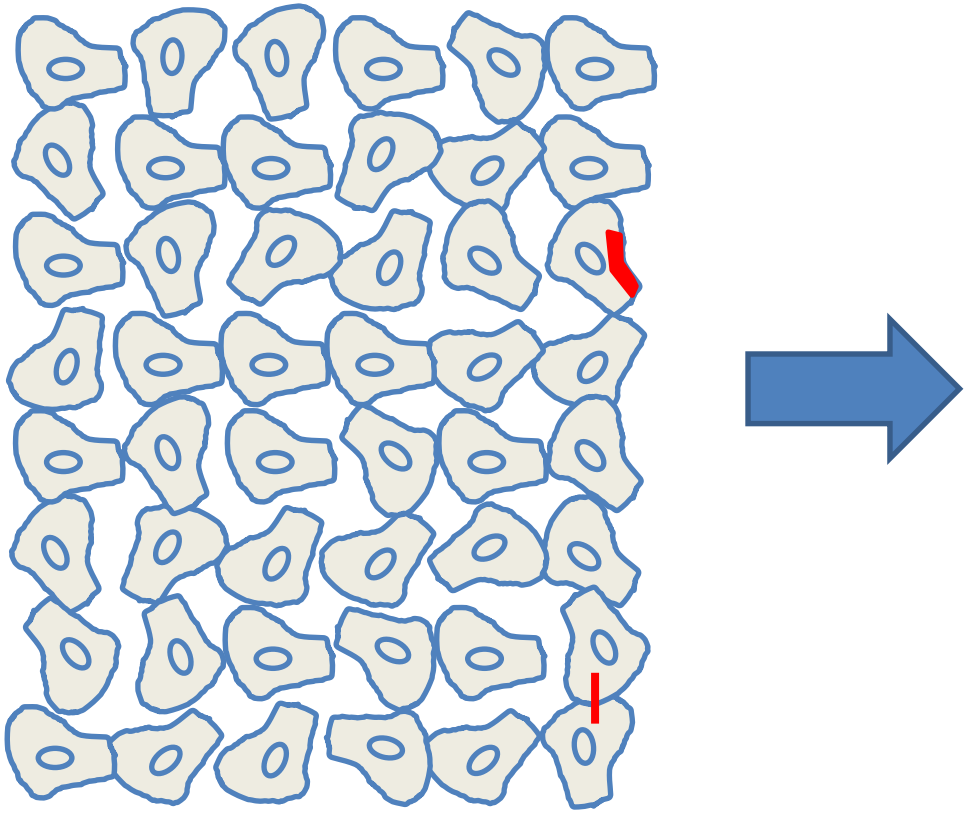


Part III
Intercellular coordination, global alignment of stress and velocity

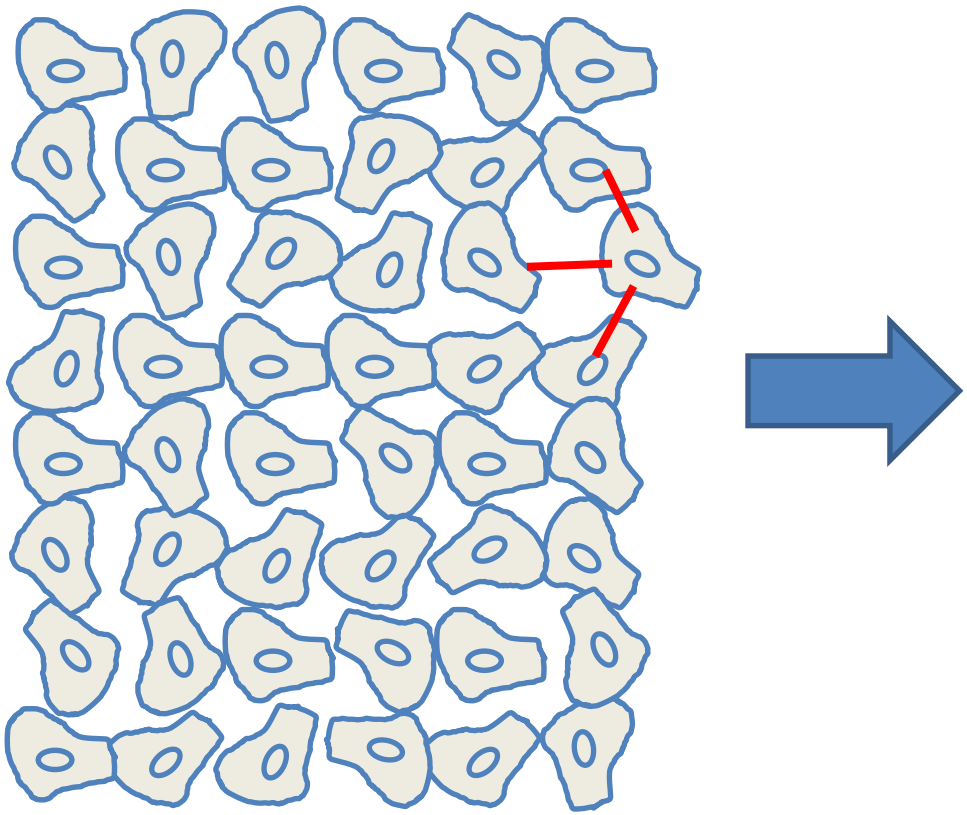
Hypothetical model



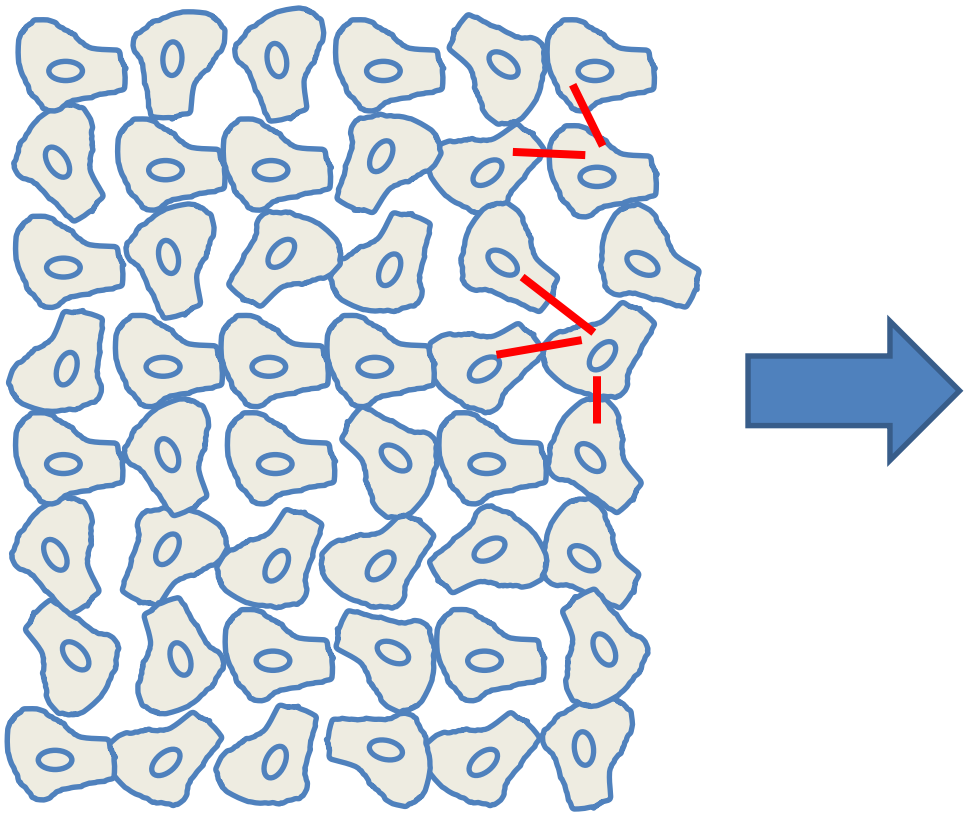
Hypothetical model



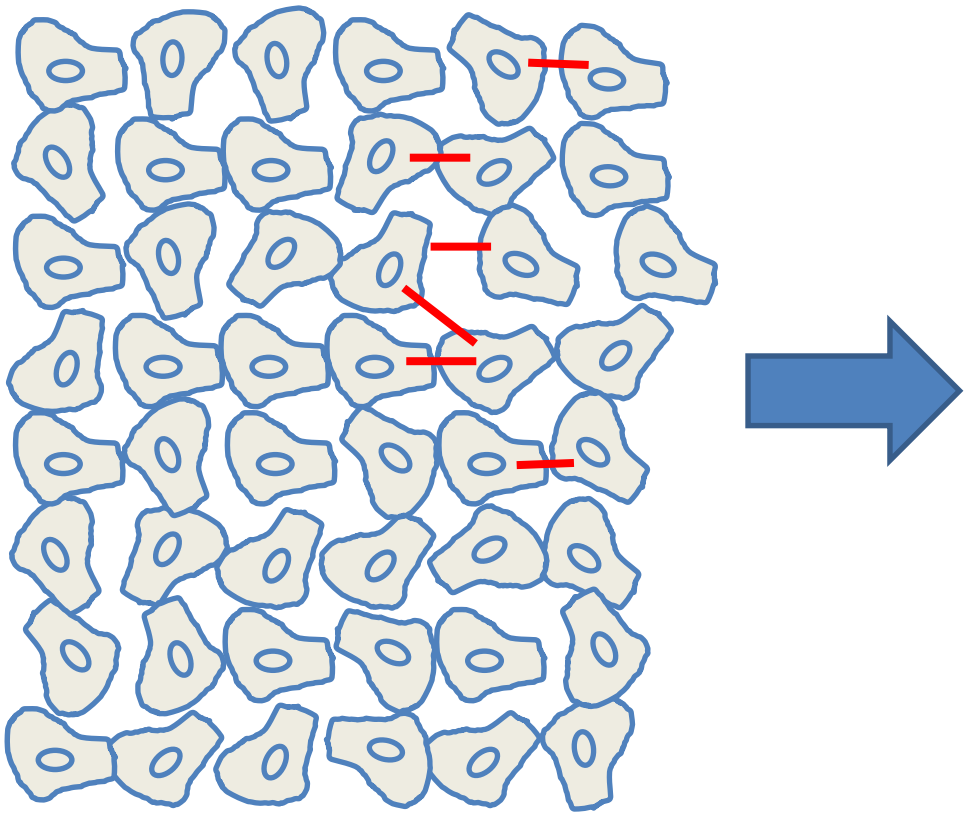
Hypothetical model



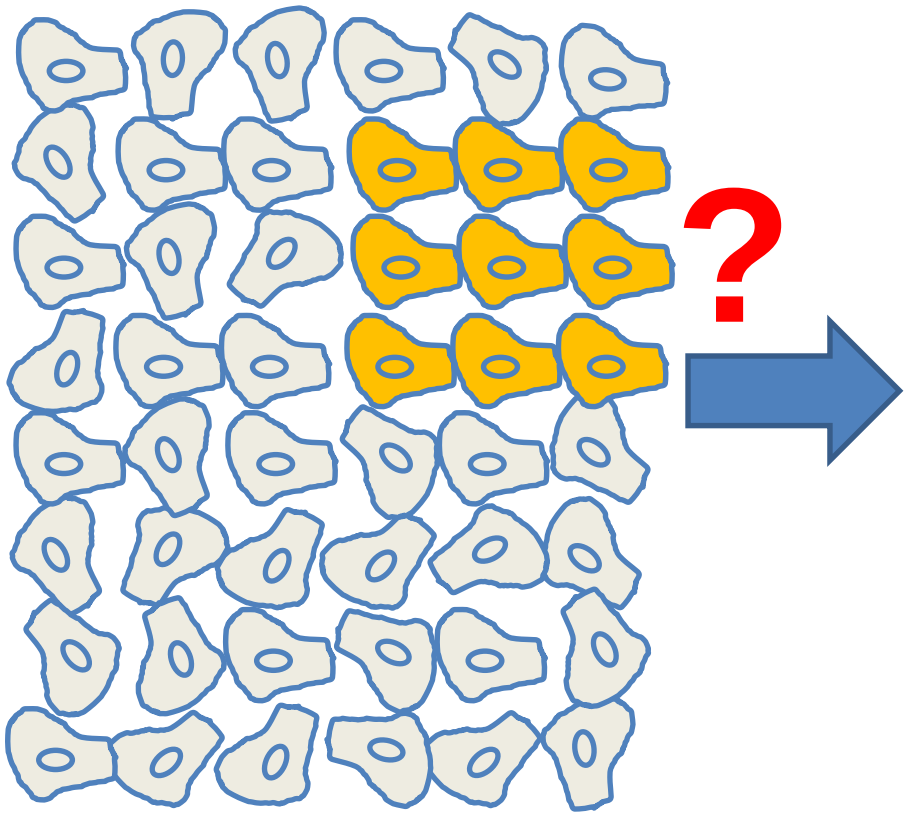
Hypothetical model



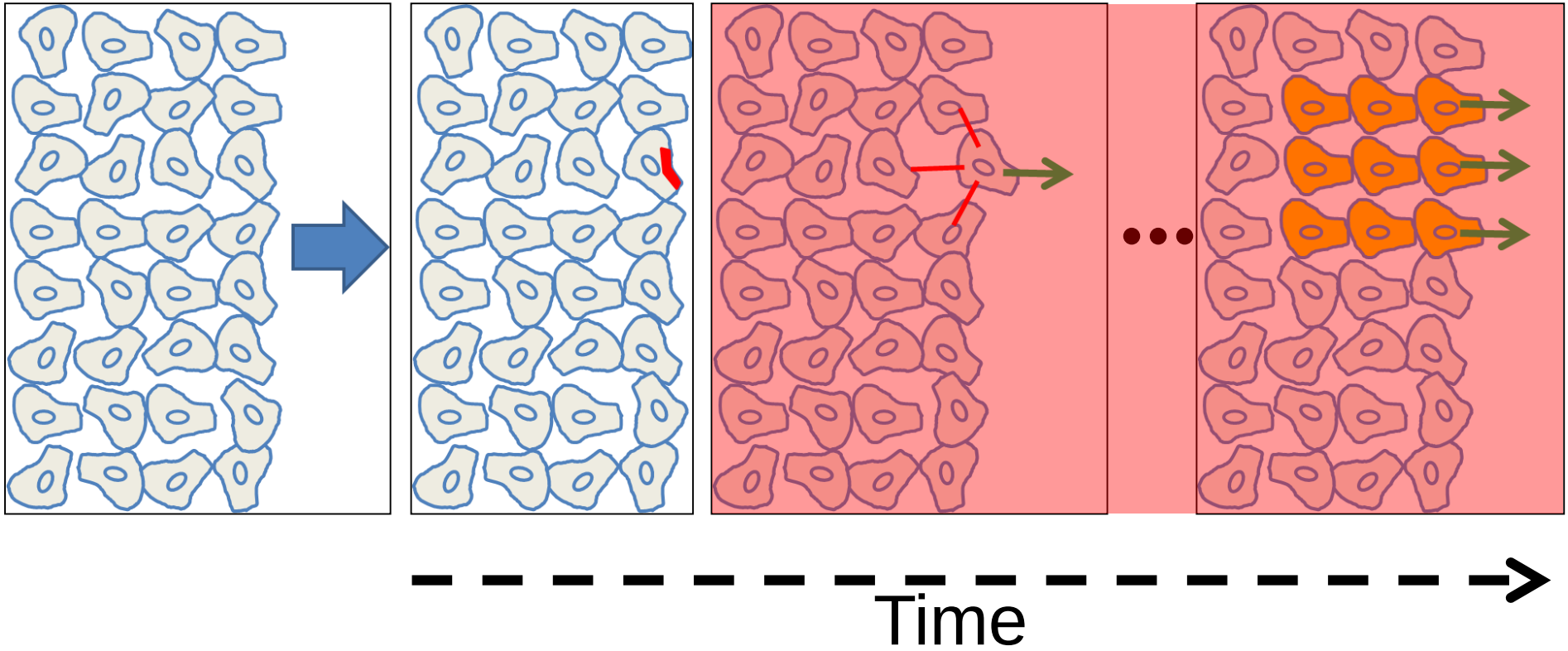
Hypothetical model



From local alignment of motion and stress to intercellular coordination?



Suggested model

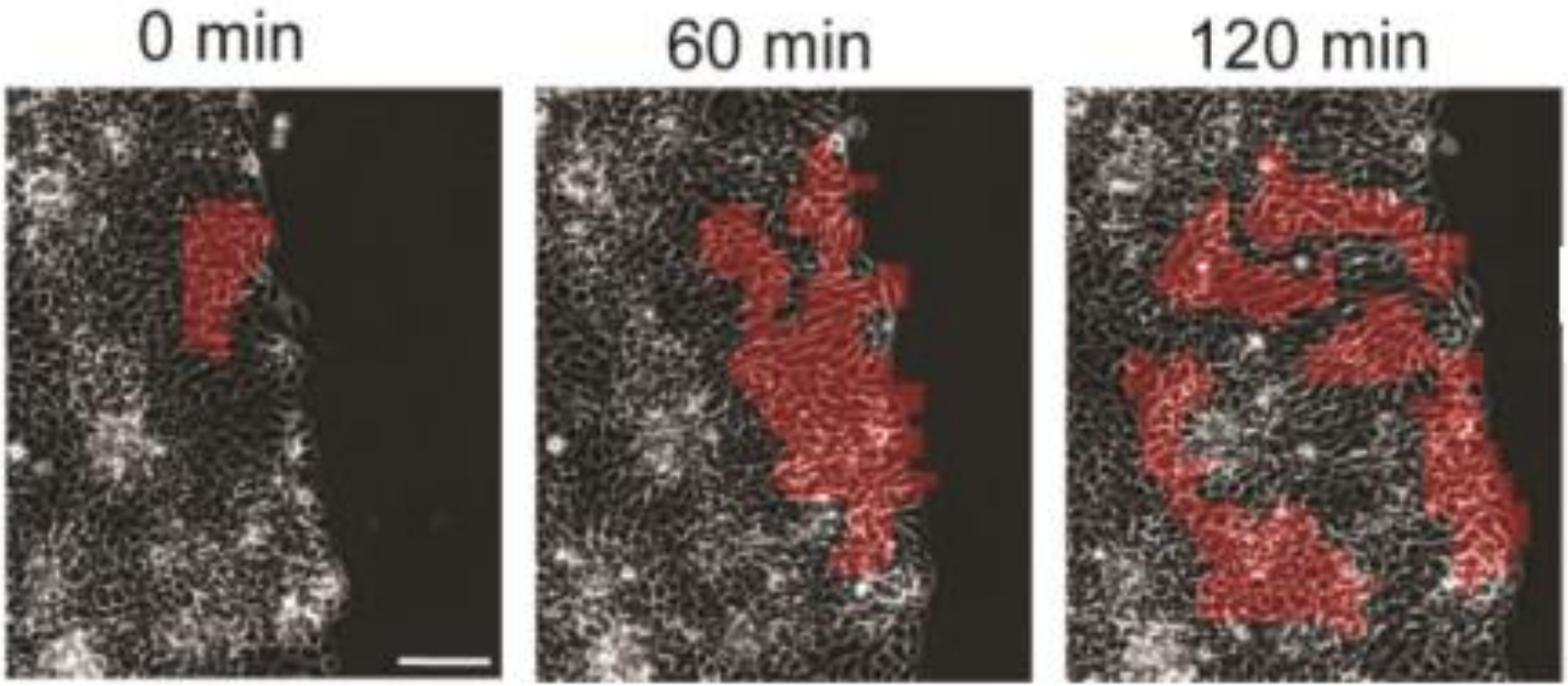


Stochastic force exertion transform to directional migration

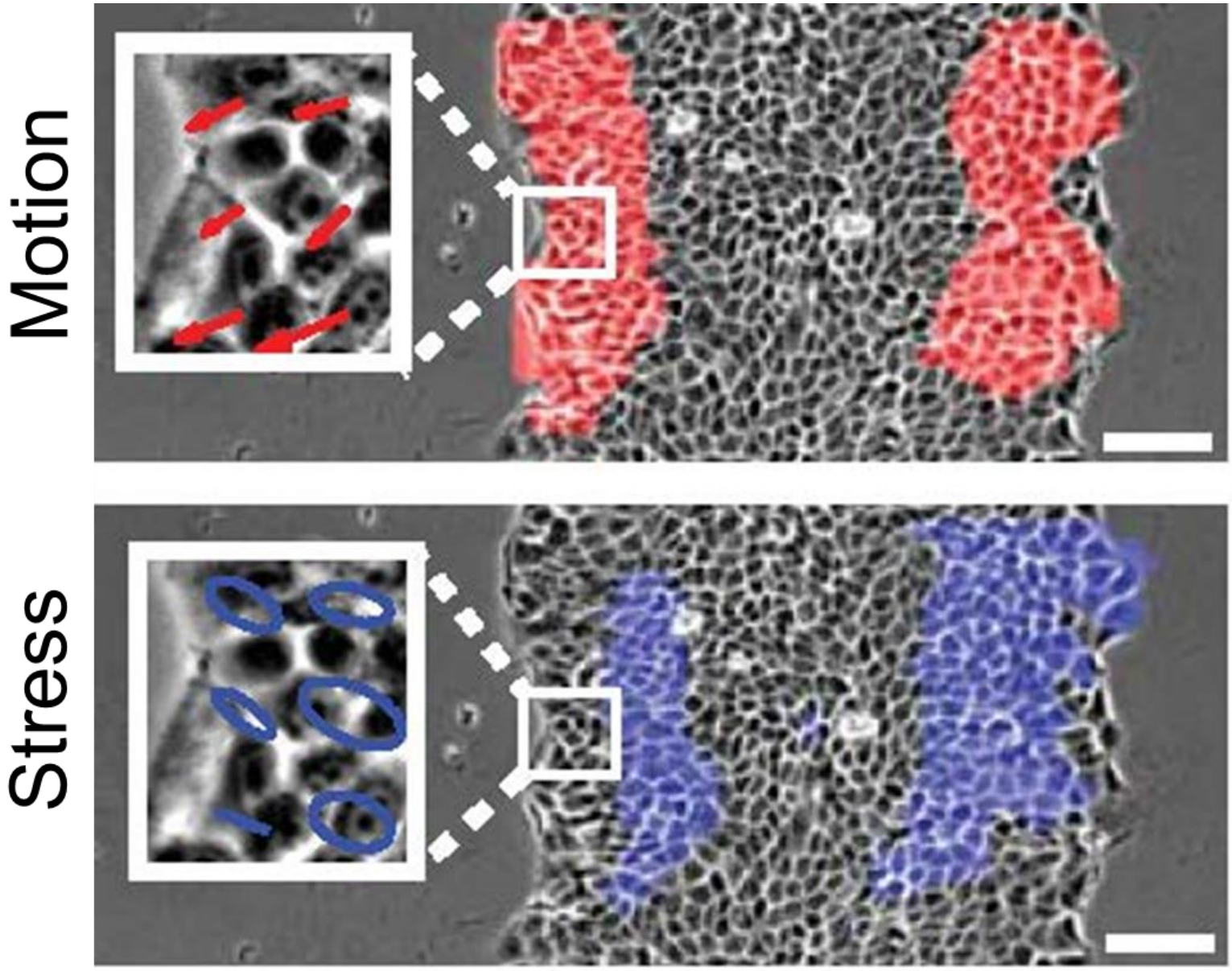
Strain on neighbors coordinate their movement

Propagation in time and space to guide groups of cells

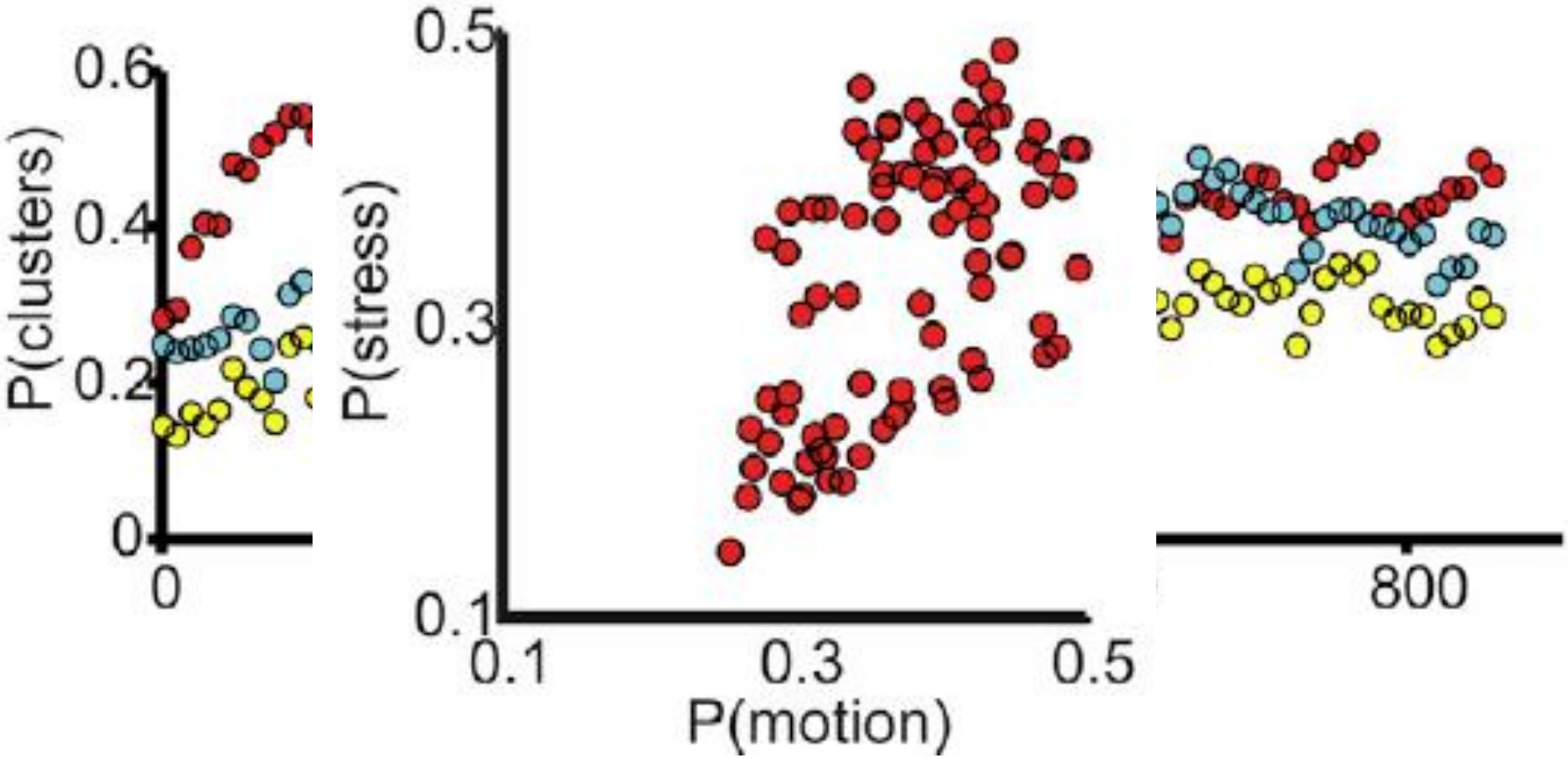
Region-growing segmentation for ***explicit*** detection of coordinated migration clusters



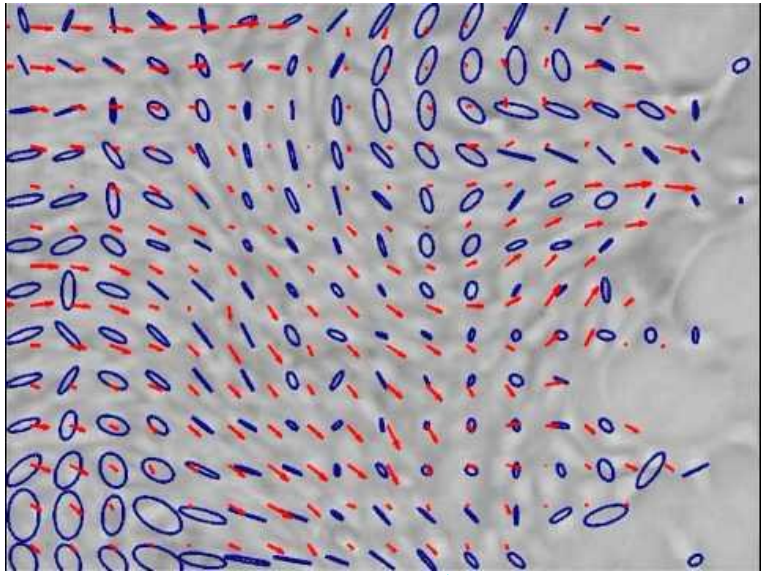
Associating coordinated stress and motion

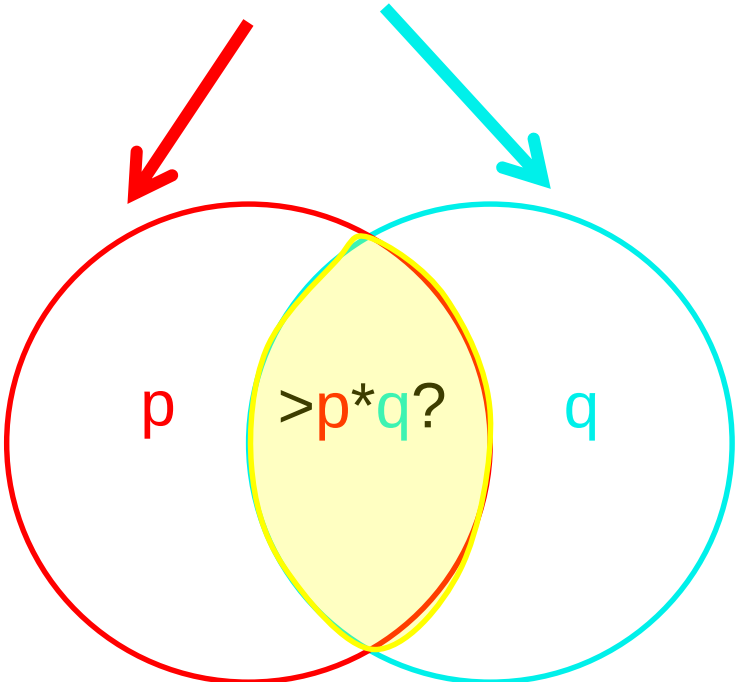


Coordinated motion is correlated to coordinated stress

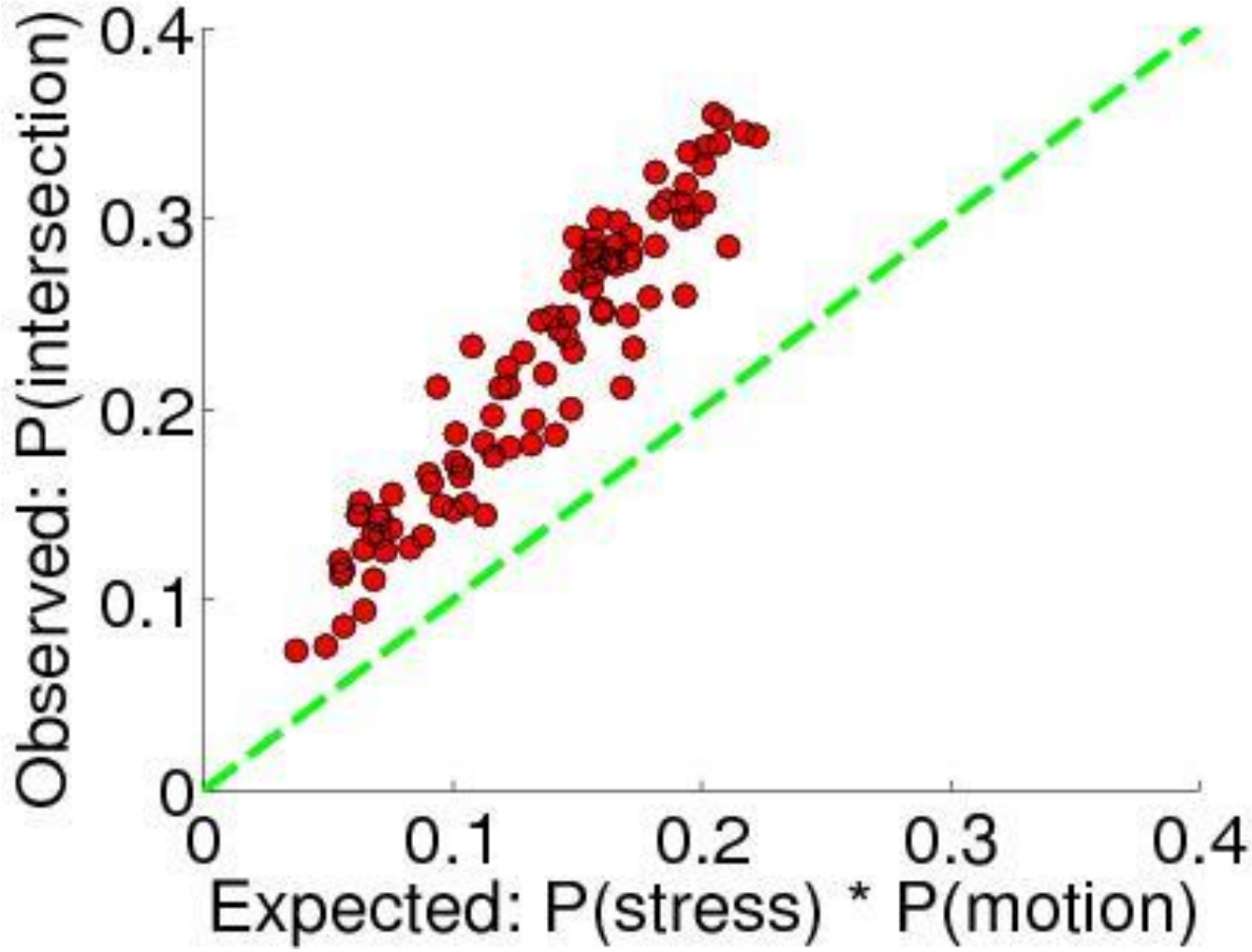


Associating coordinated stress and motion



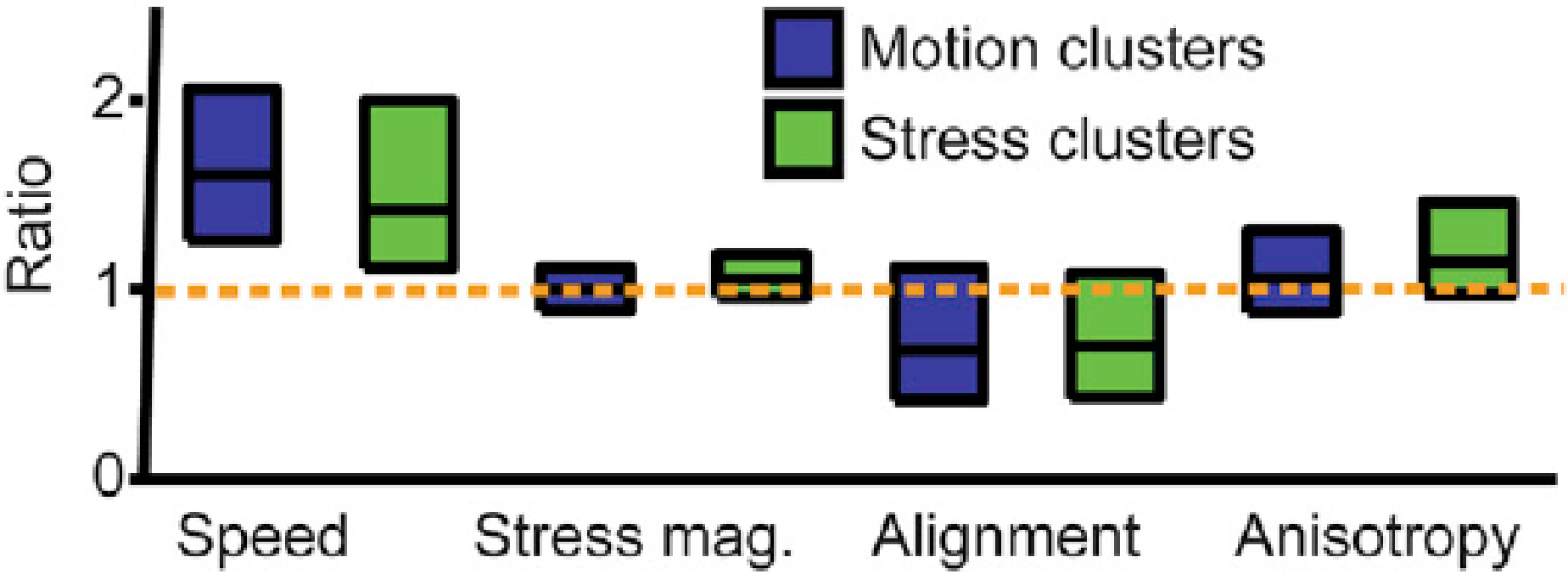


Motion- and stress- coordinated clusters are interlinked

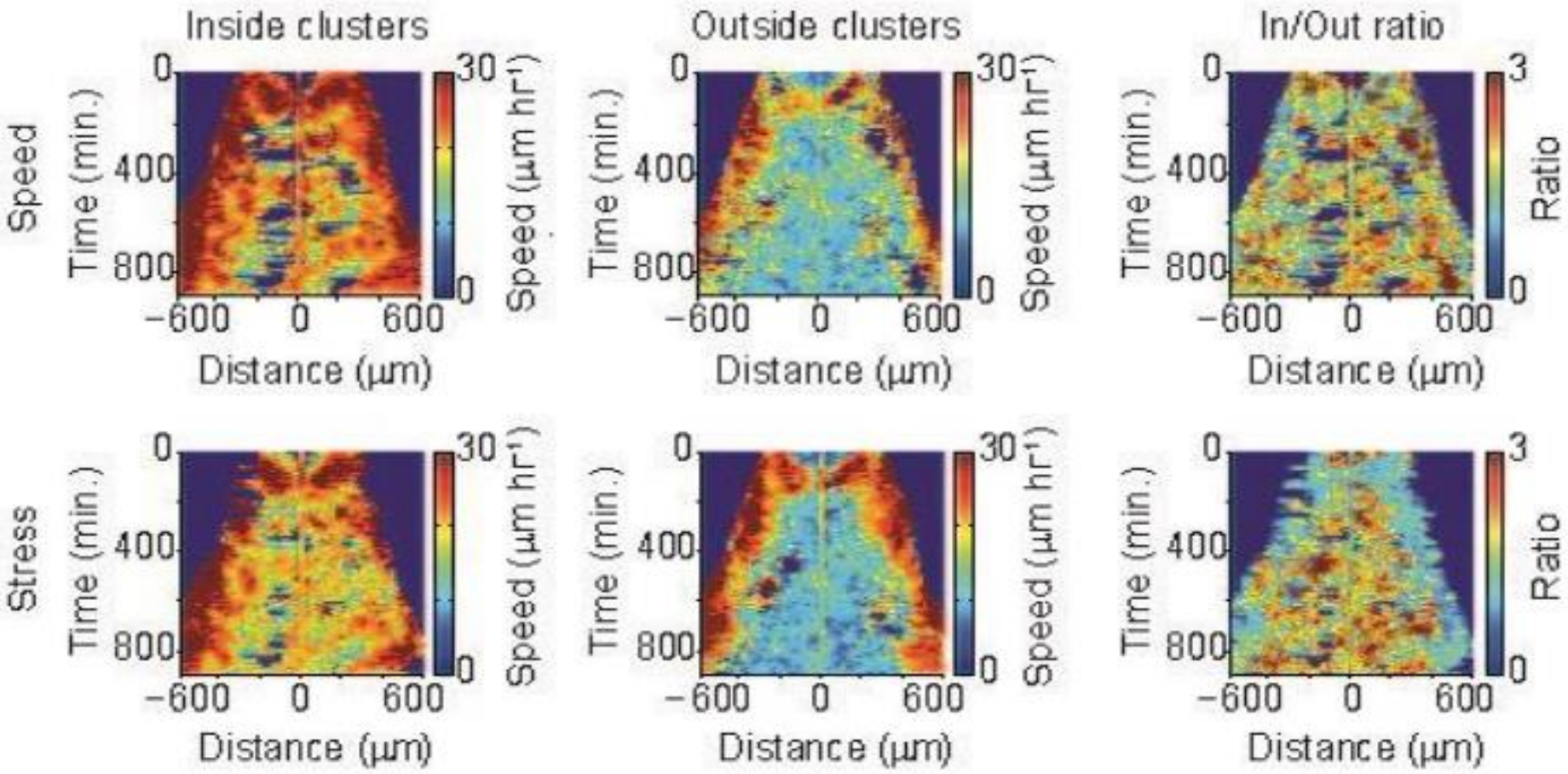


Cells in coordinated clusters move faster, with enhanced motion-stress alignment

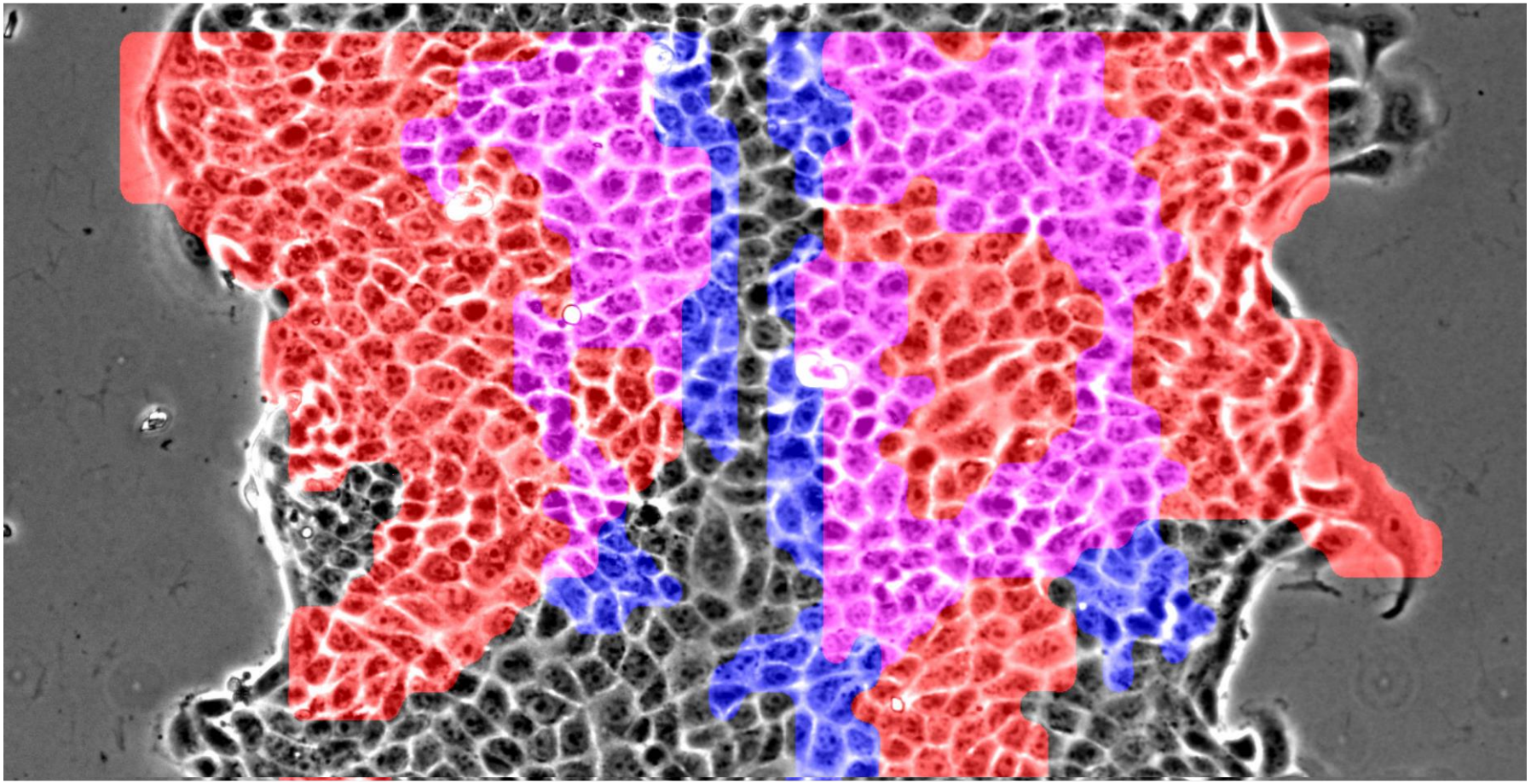
$$\mathrm{Ratio}_{\mathrm{property}} = \frac{\mathrm{property}_{\mathrm{in}}}{\mathrm{property}_{\mathrm{out}}}$$



Spatiotemporal analysis

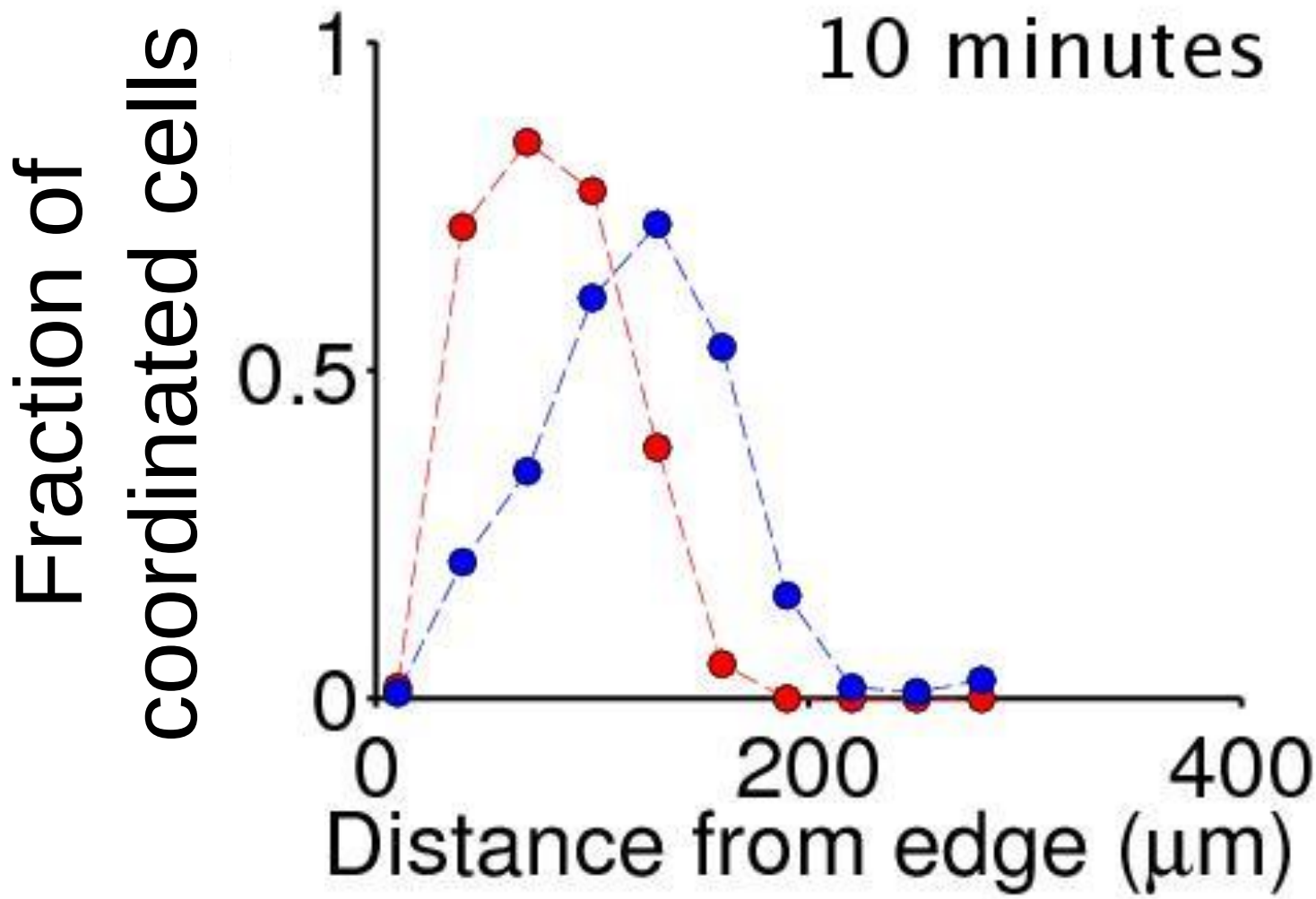


Strain-induced motion coordinates cluster's motility



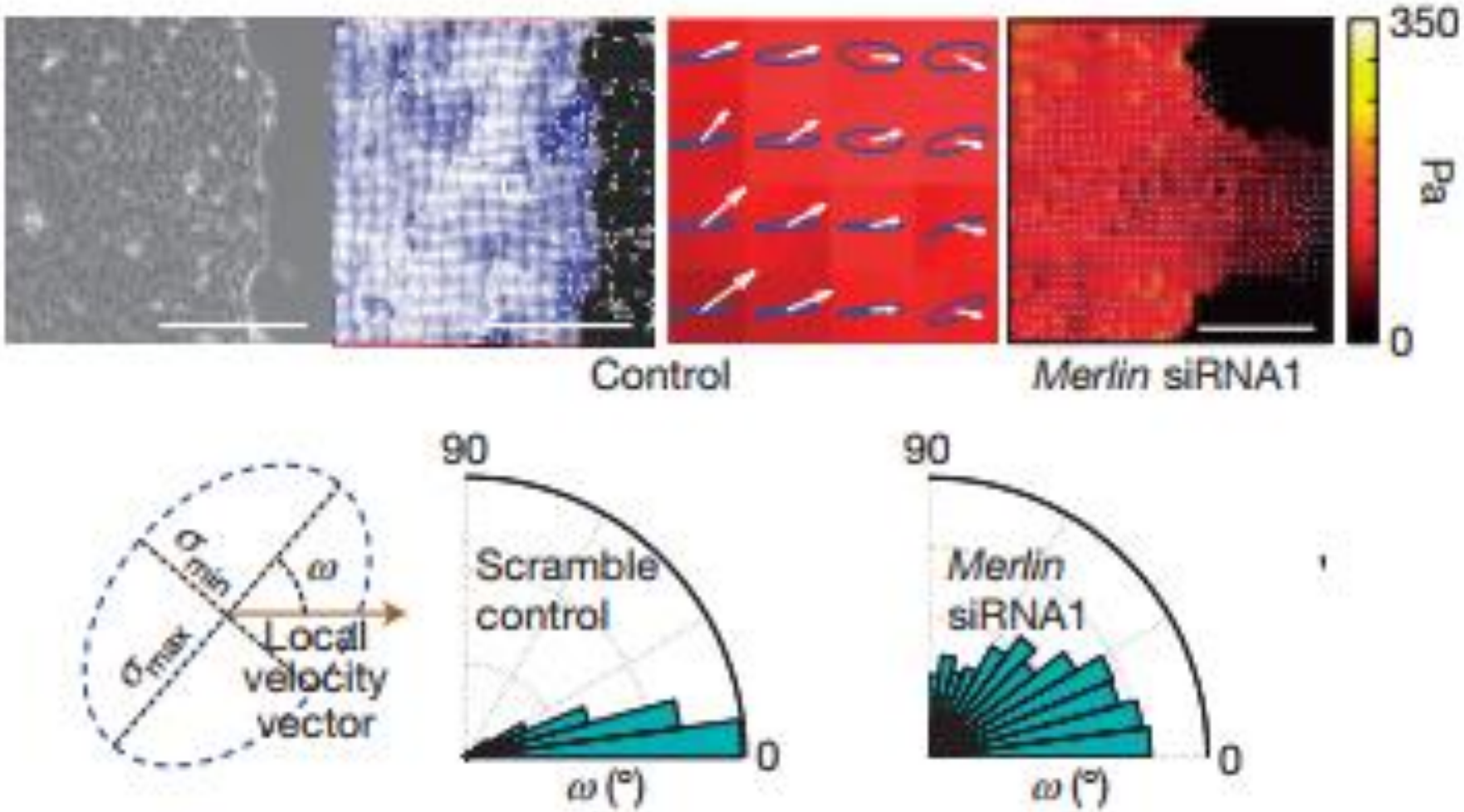
Red: motion clusters, Blue: stress, Magenta: both

Strain-induced motion coordinates cluster's motility



Red: motion clusters, Blue: stress

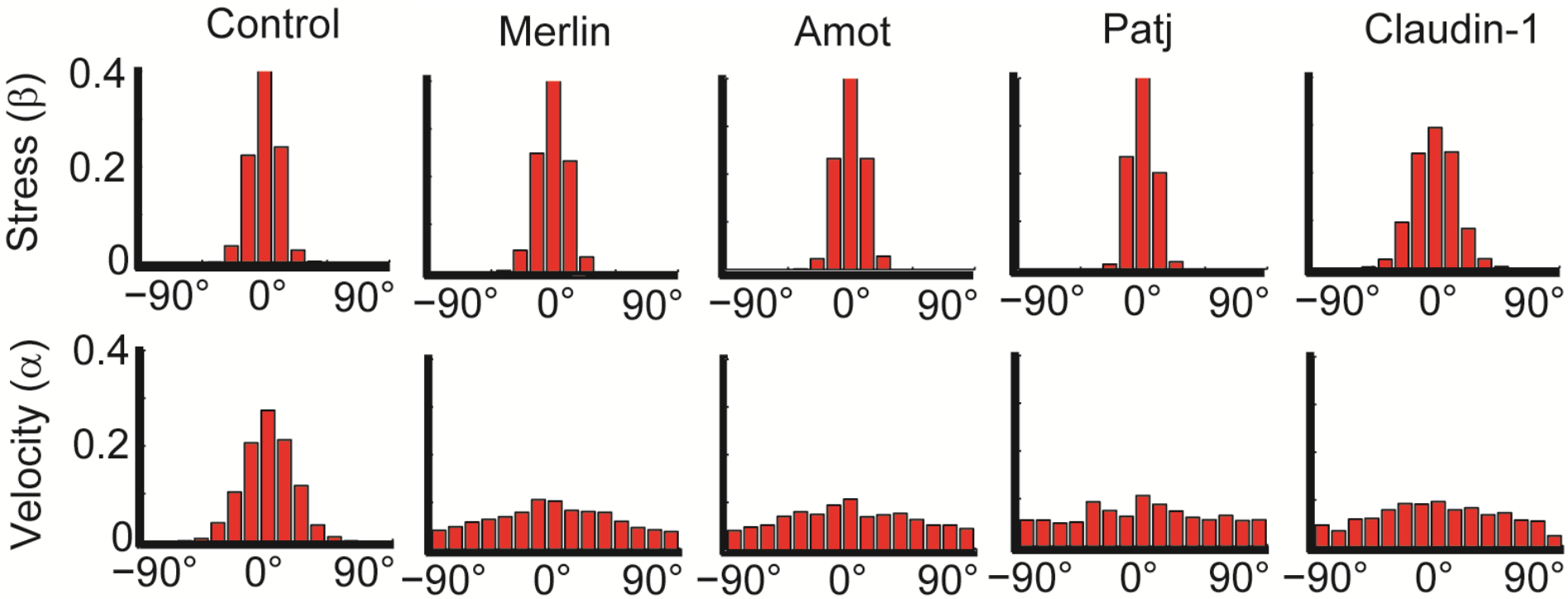
Data from: Das et al.



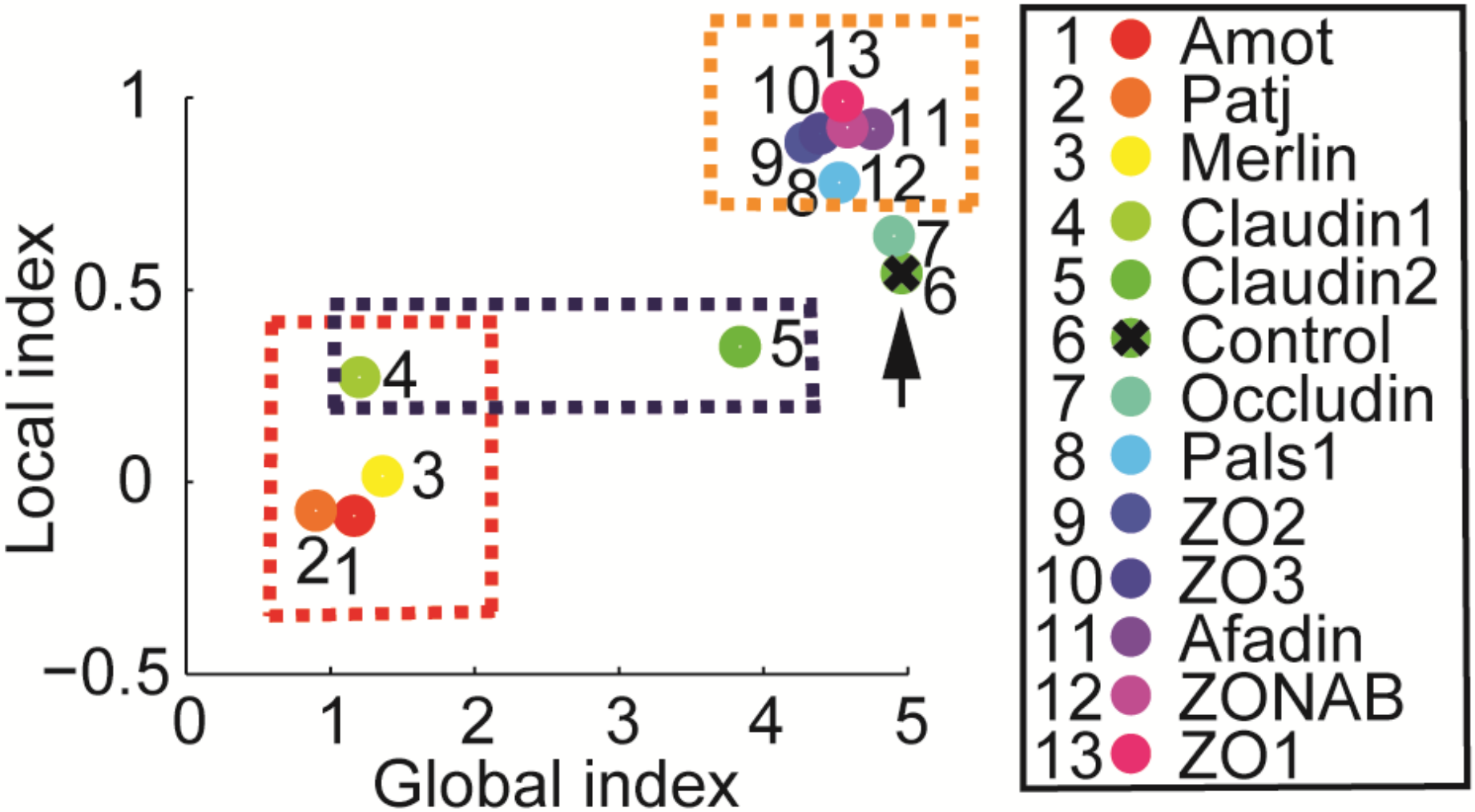
Das et al. (2015)

Stress aligns motion

Tight junction proteins play a role in effective transmission of aligned stress to aligned motion

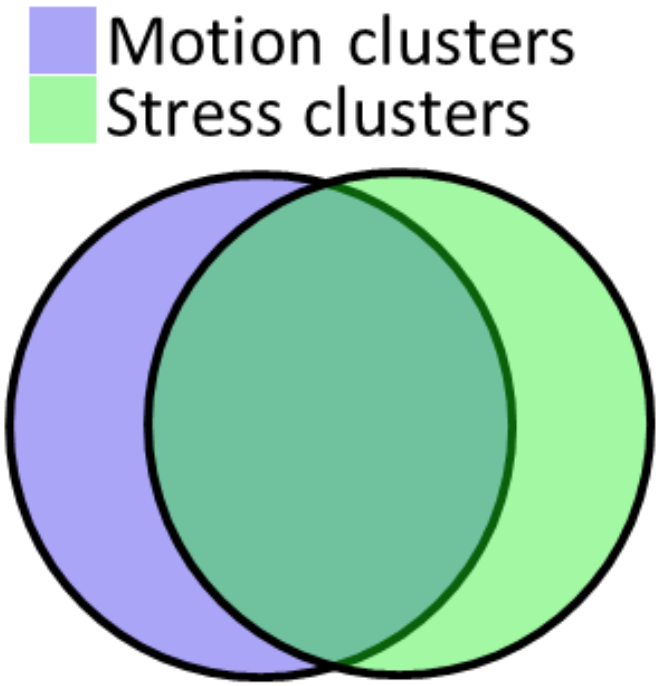


Reuse x2

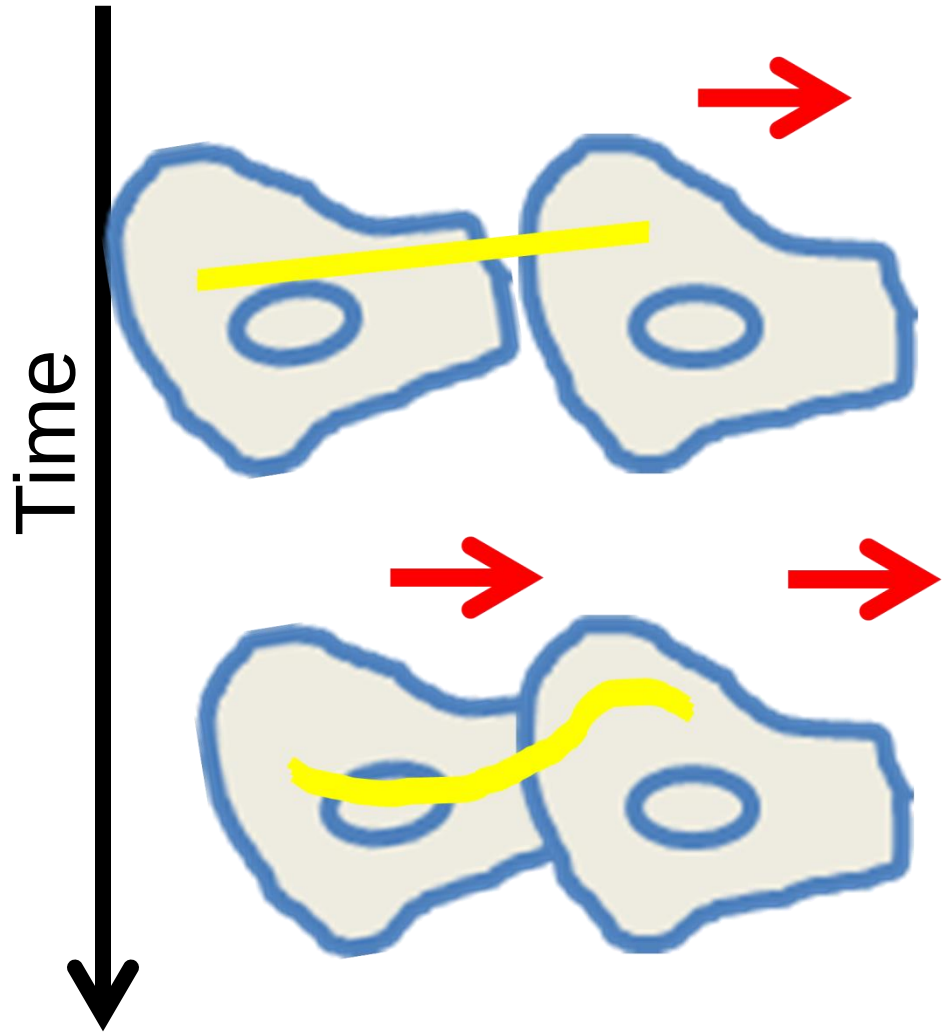

Local index
Global index
1
0.5
0
−0.5
0
1
2
3
4
5
1 Amot
2 Patj
3 Merlin
4 Claudin1
5 Claudin2
6 Control
7 Occludin
8 Pals1
9 ZO2
10 ZO3
11 Afadin
12 ZONAB
13 ZO1

Part III: Conclusions

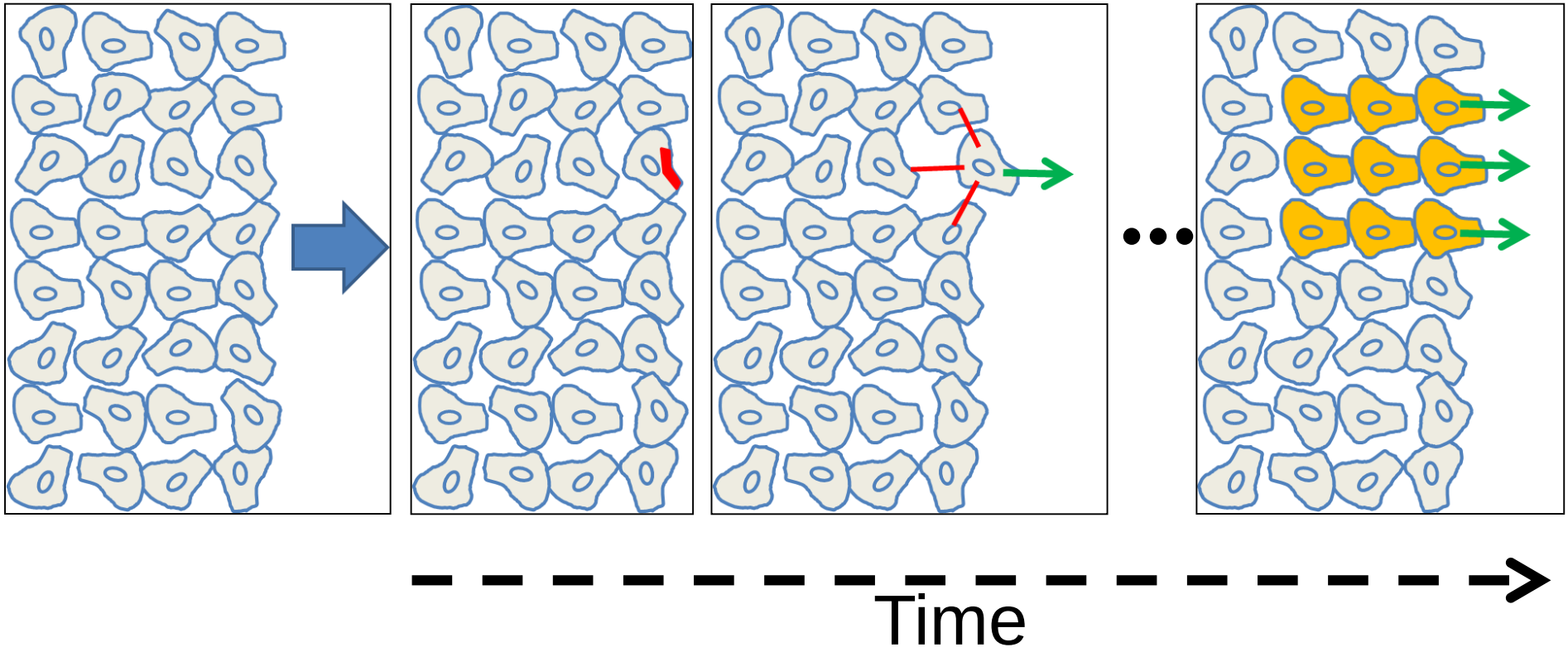
group plithotaxis



Enhanced Speed,
Motion-stress alignment,

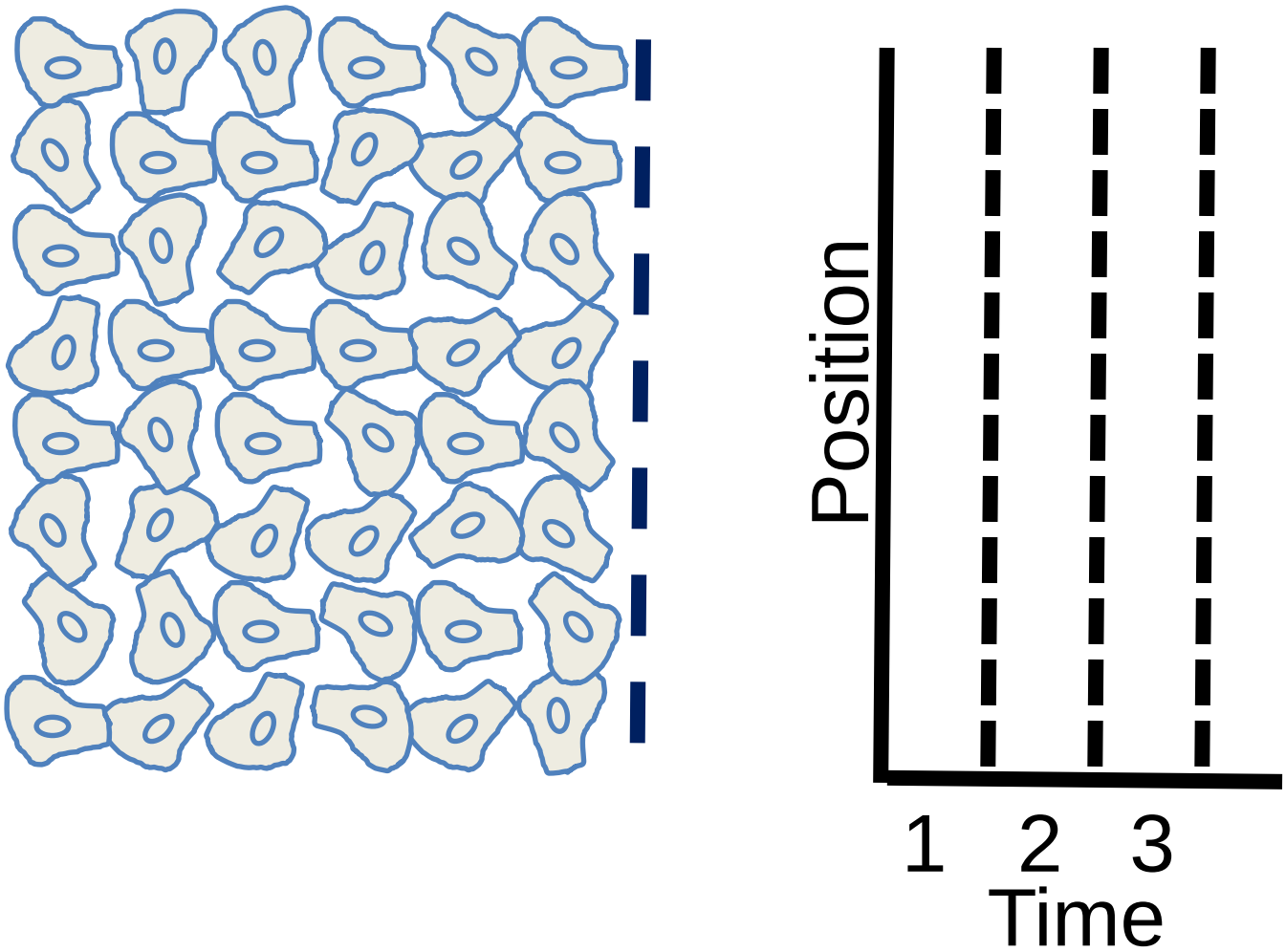


Coordination migration emerges from cell-cell junctional transmission of mechanical guidance cues

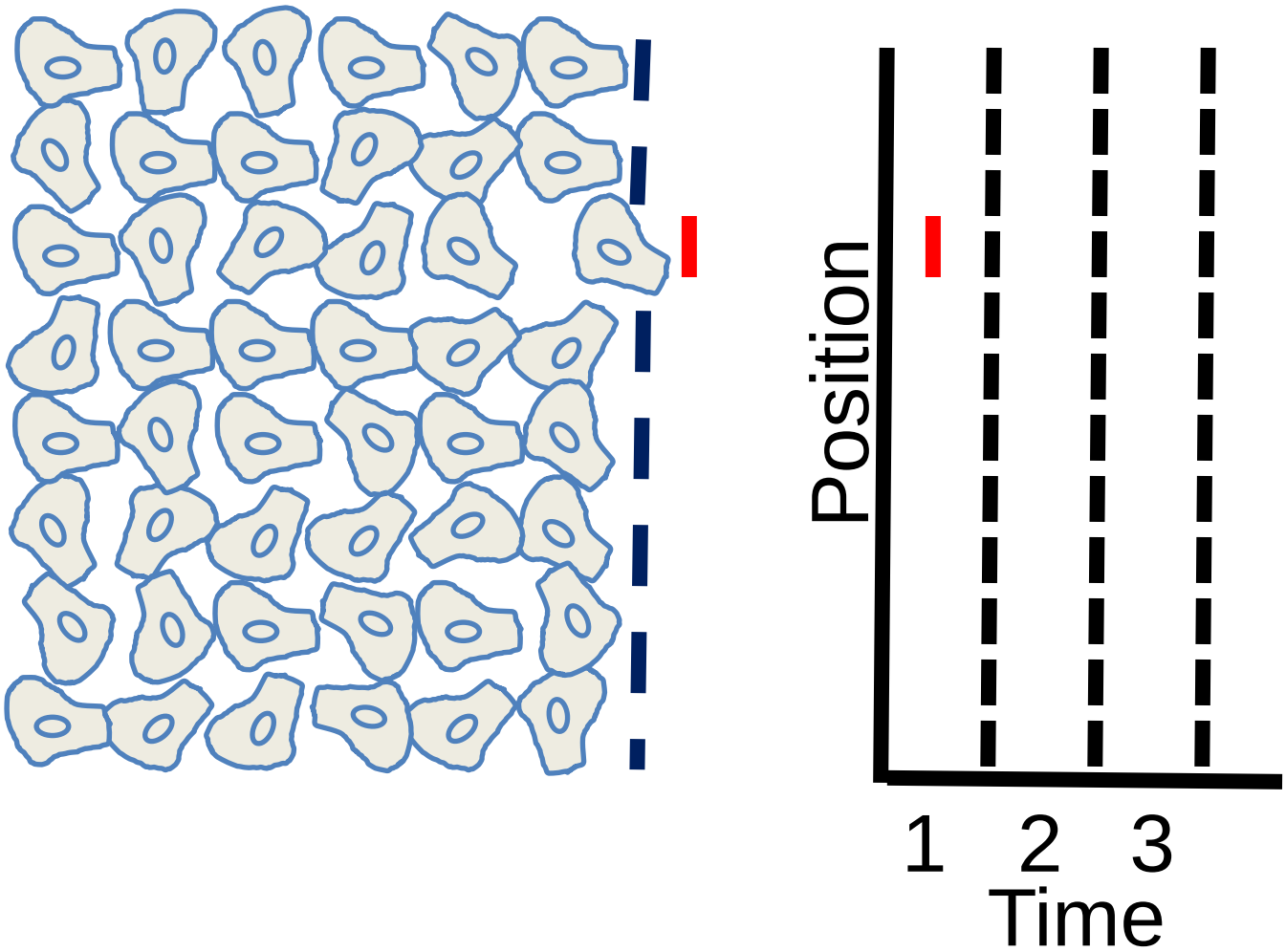


Part IV
Lateral guidance cues
(HBEC data is new)

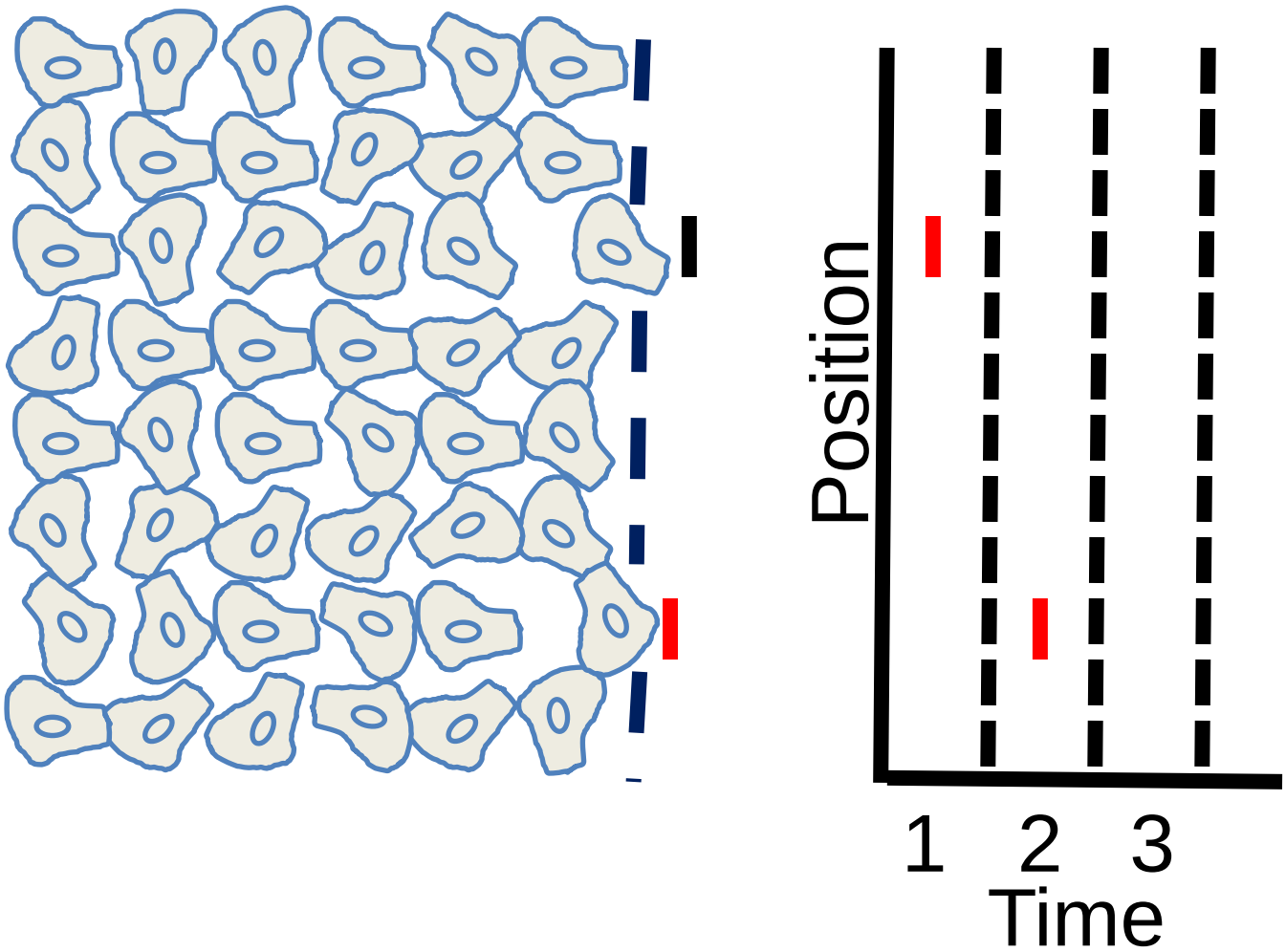
Protruding cells kymograph



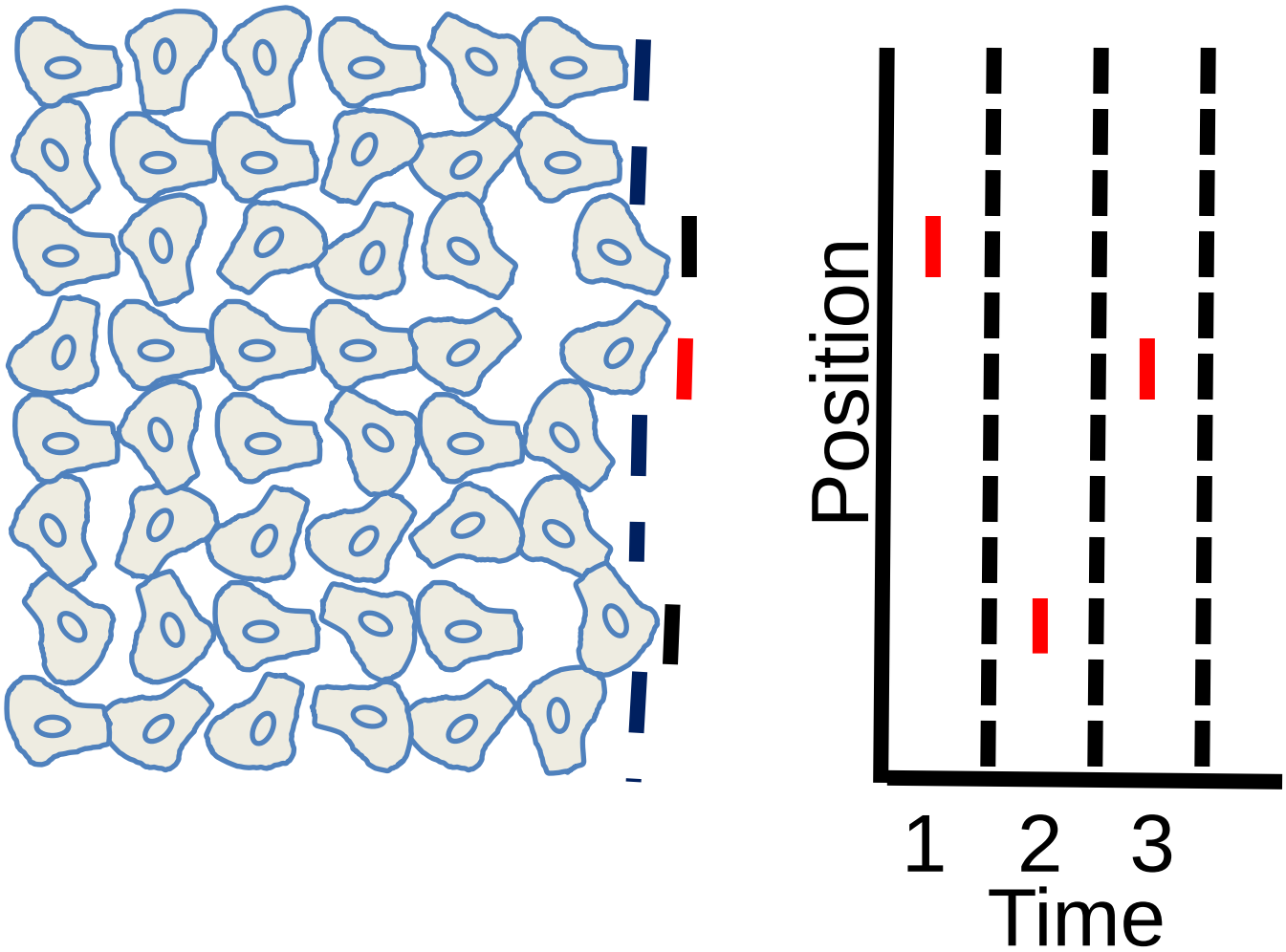
Protruding cells kymograph



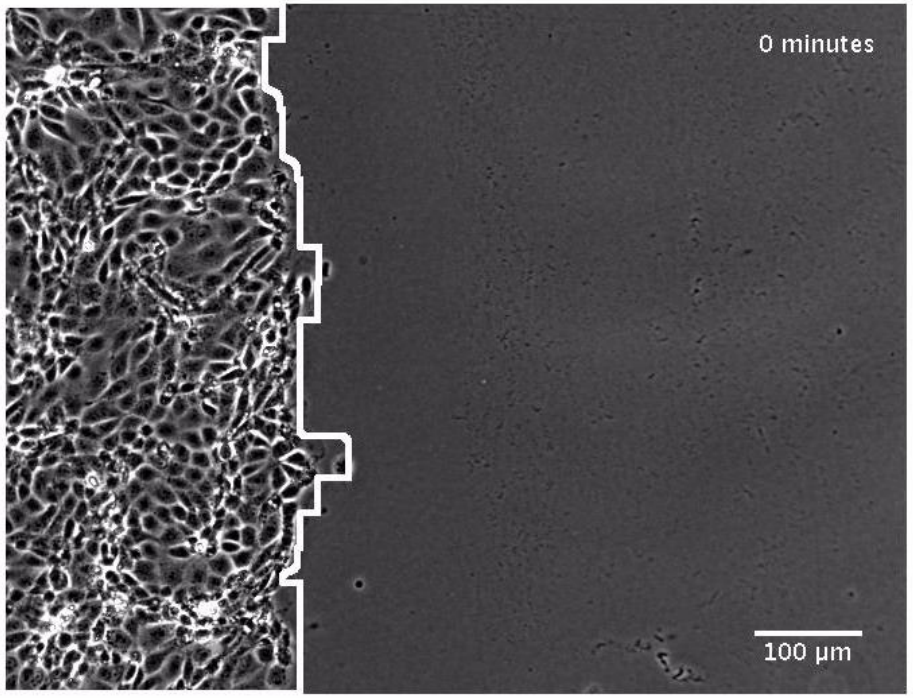
Protruding cells kymograph

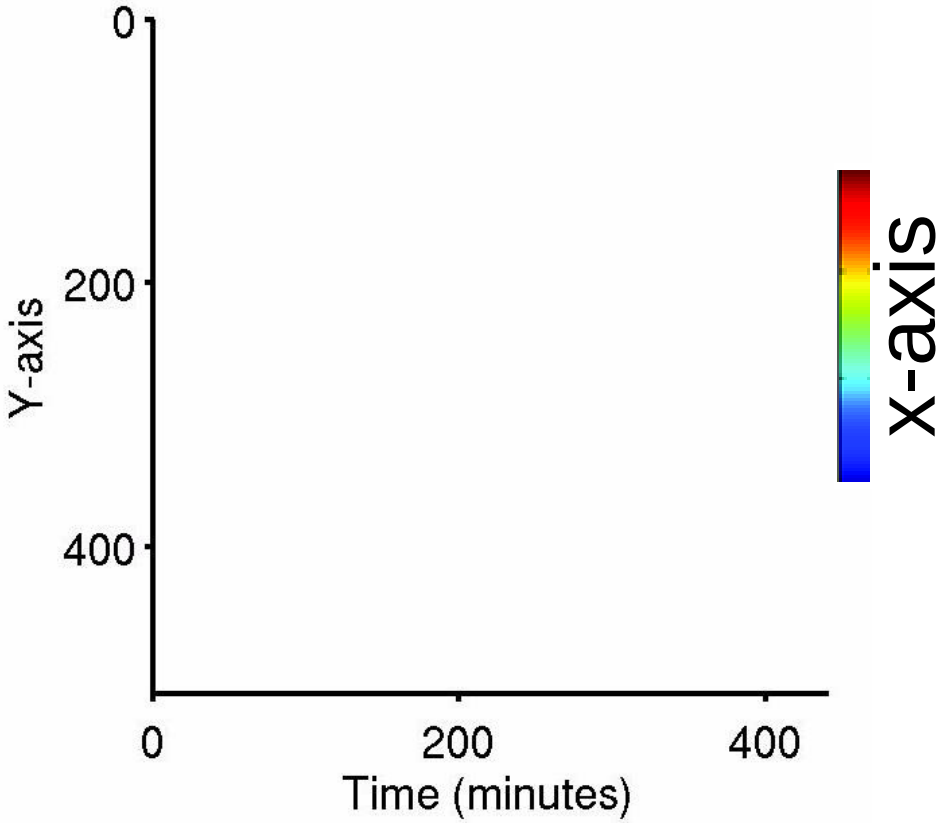


Protruding cells kymograph

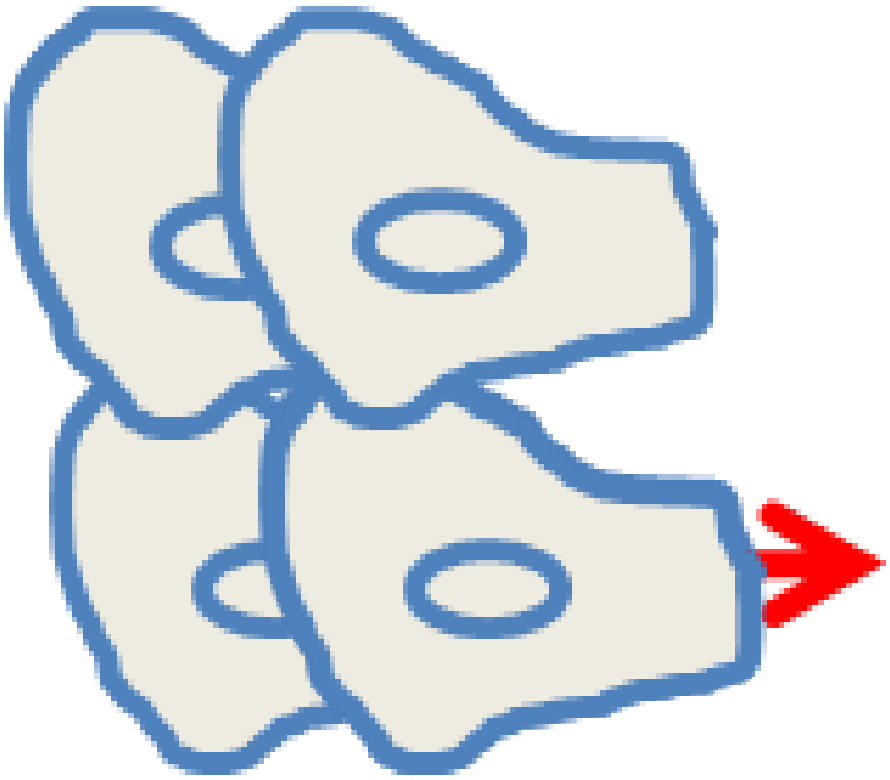


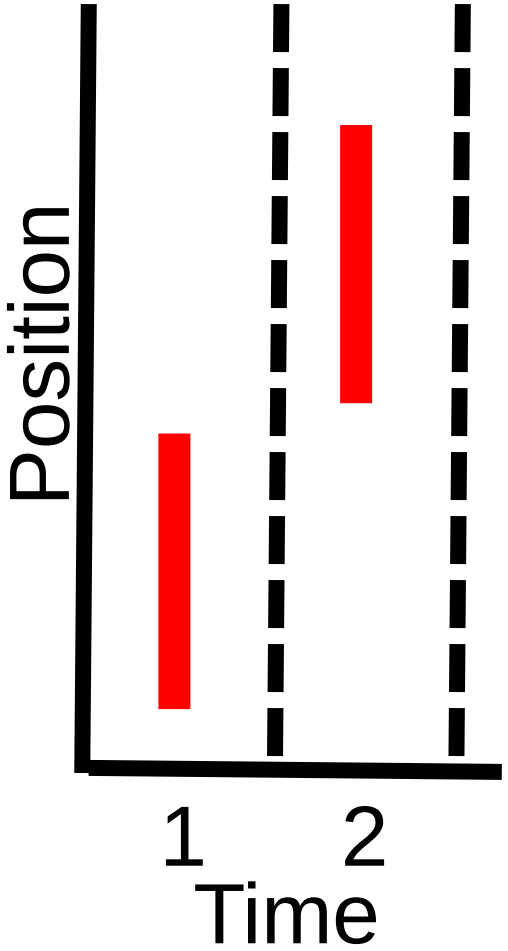
Lateral shear-strain waves



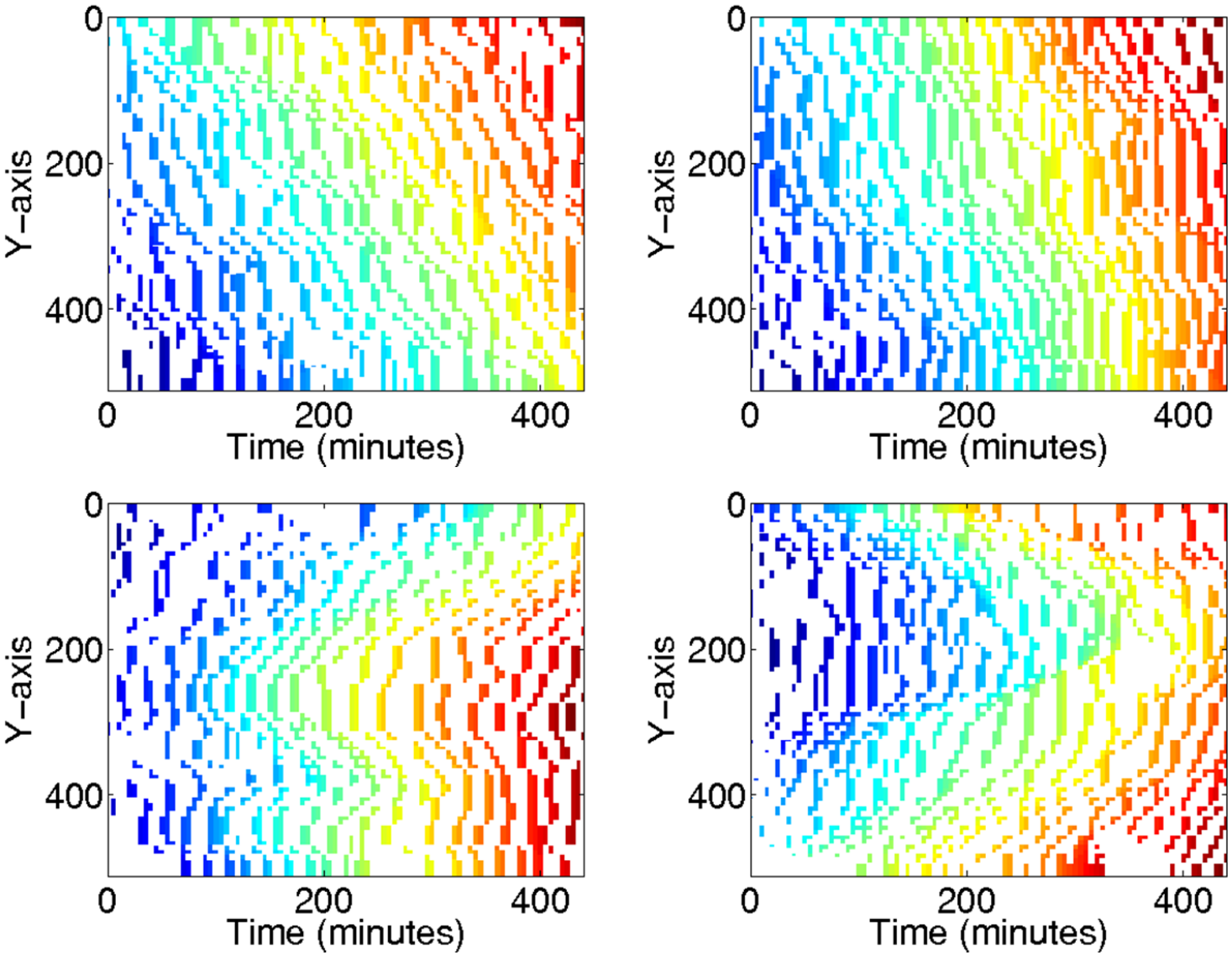


Interpretation of diagonal patterns
(simplified)

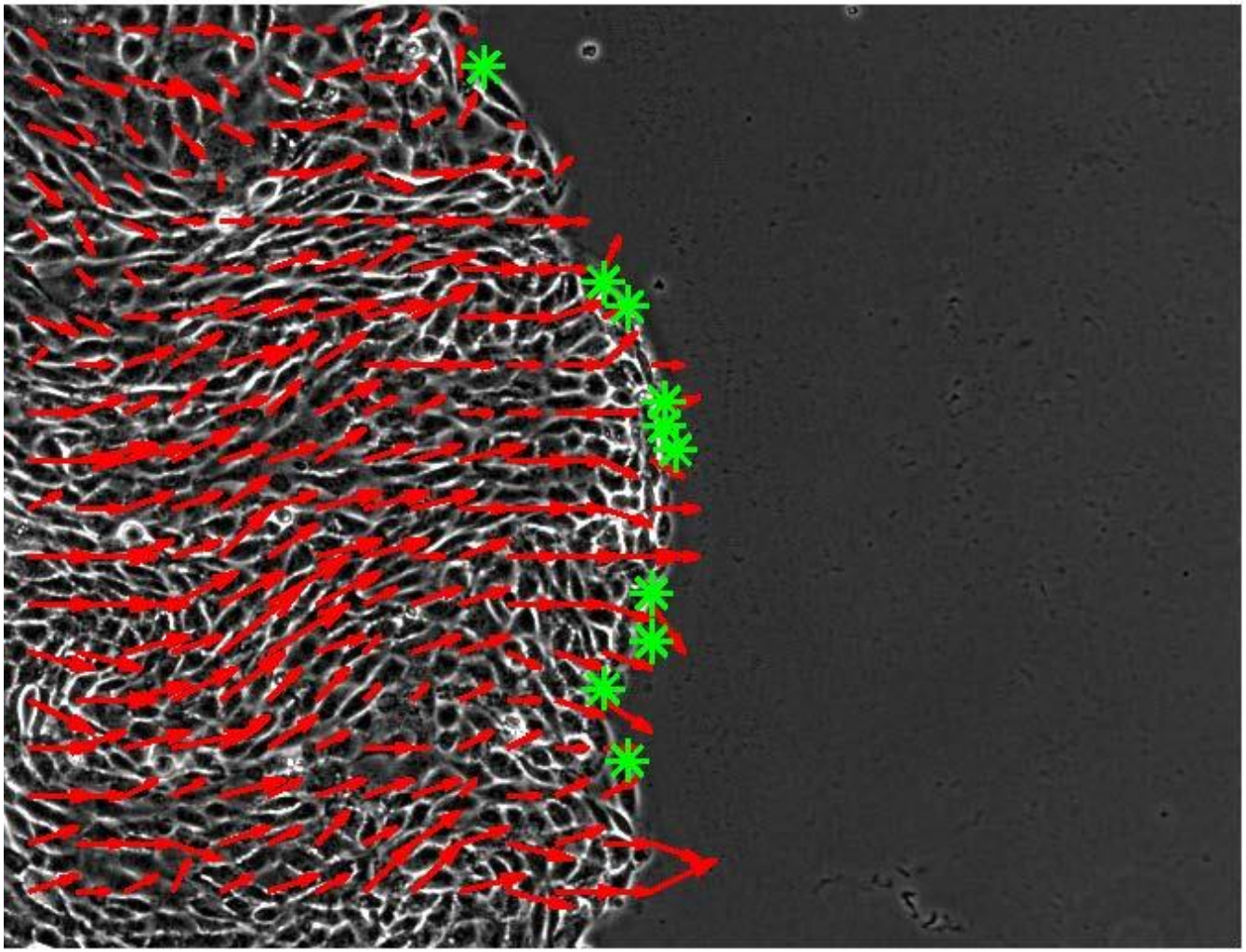




Lateral shear-strain waves

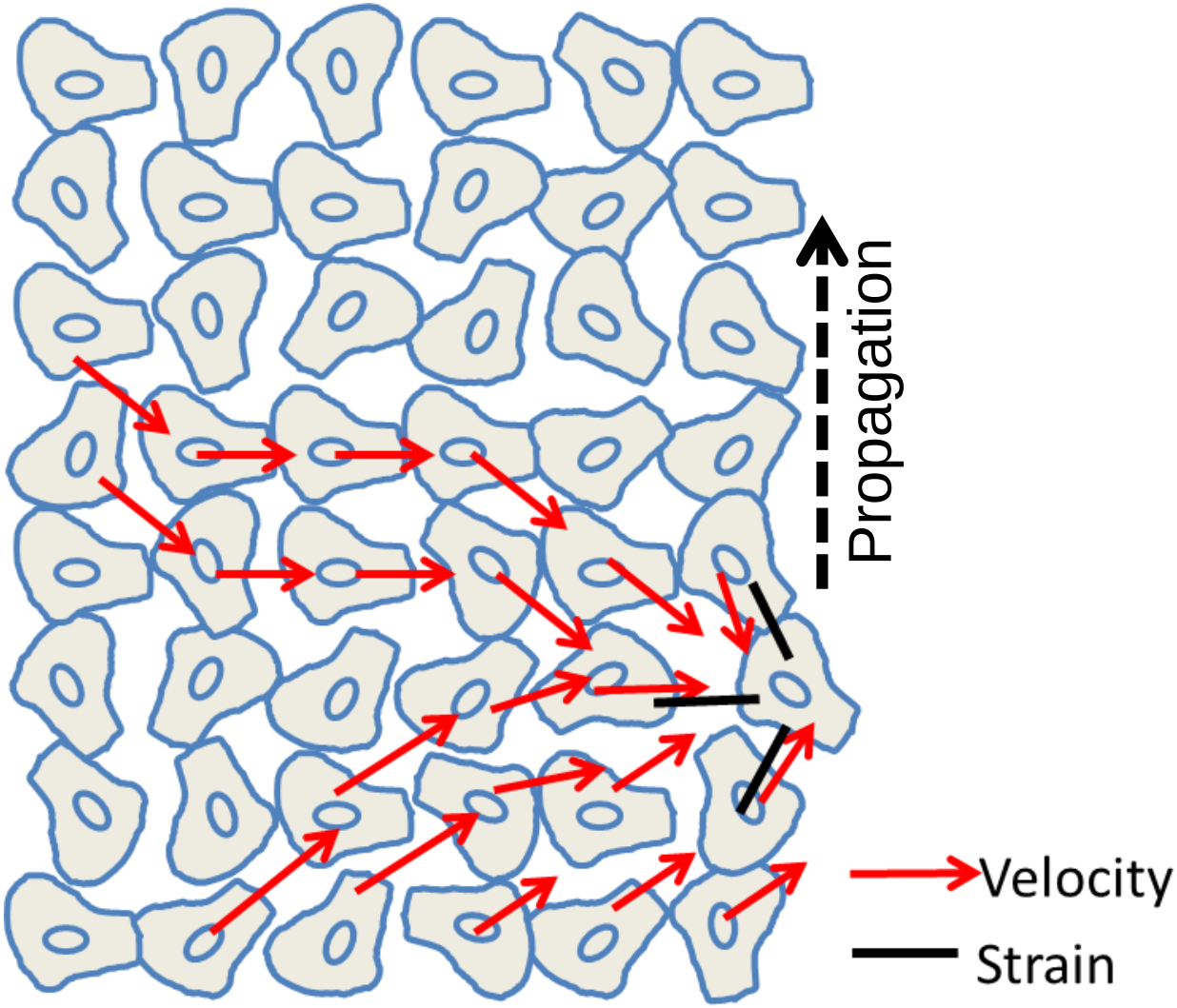


Rear cells guided to location of shear-strain events at the front



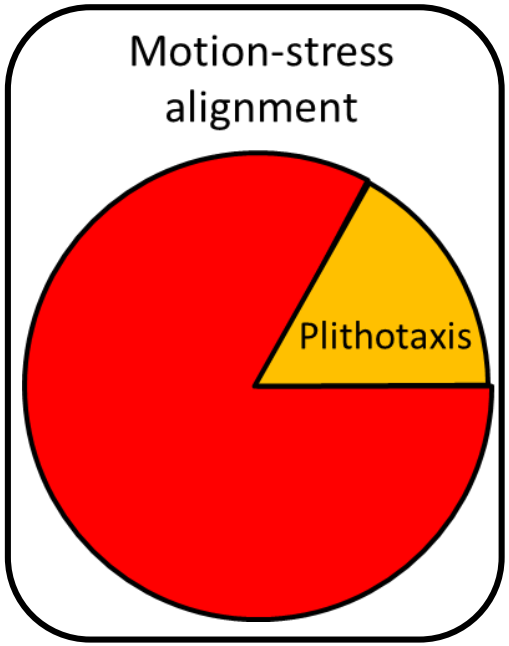
Part IV: Conclusions

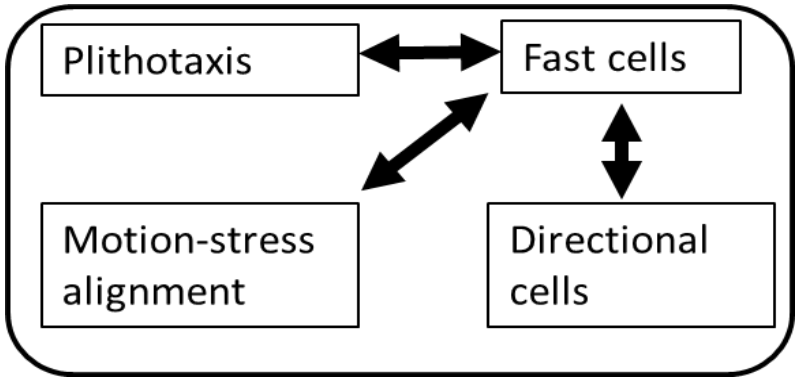
Long range guidance of by leader cells

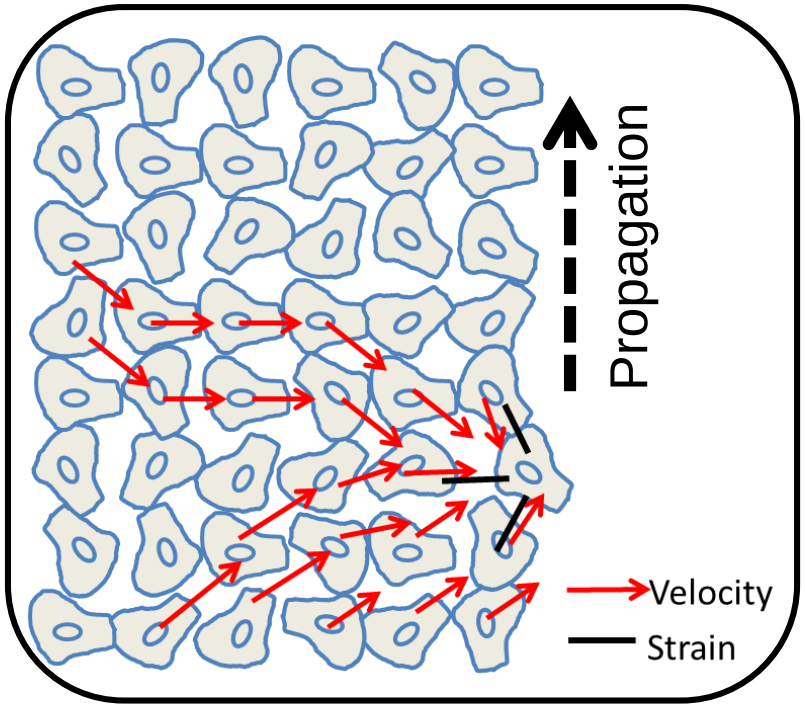


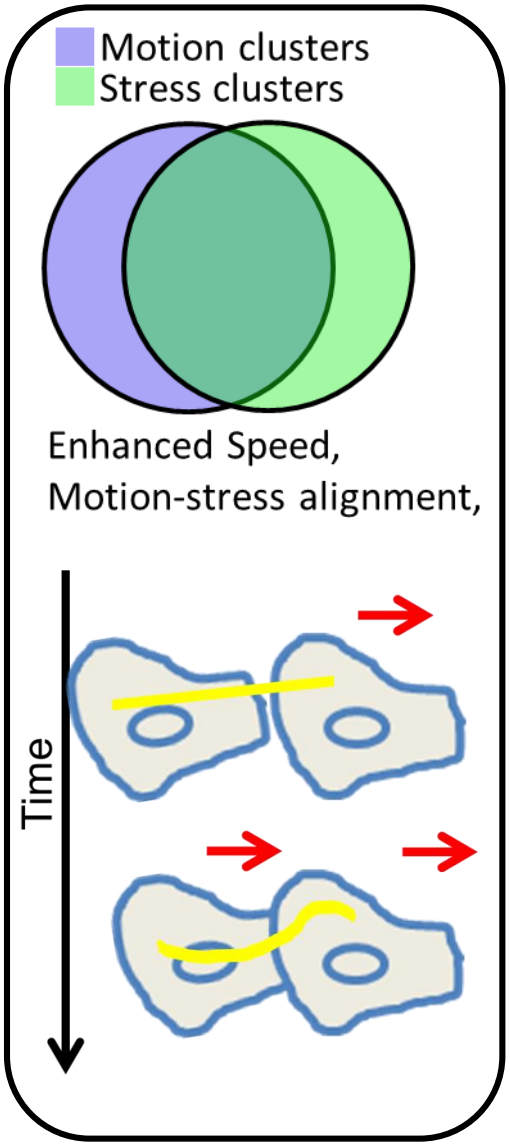
Summary

Seeds of locally aligned motion and stress coordinate collective cell migration

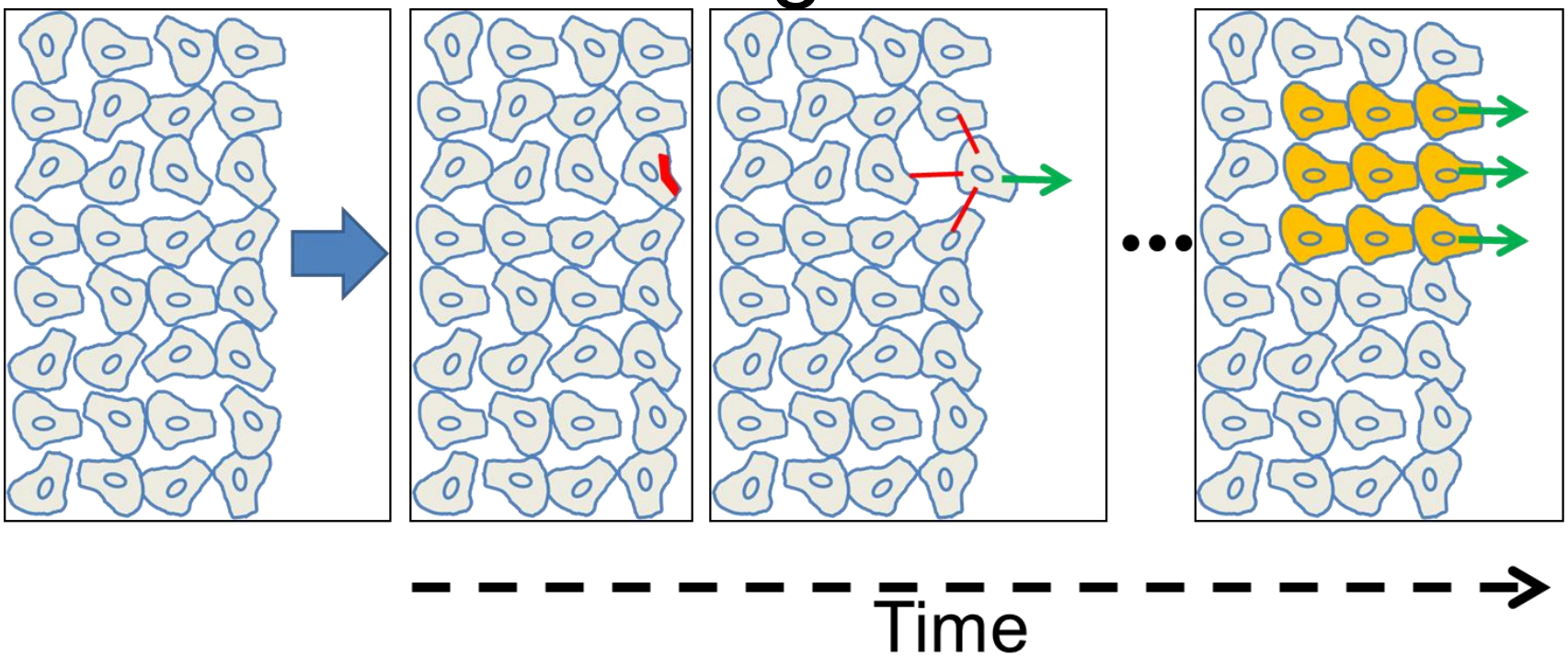








Seeds of Locally Aligned Motion and Stress Coordinate Collective Cell Migration



Stochastic force exertion transform to directional migration
Strain on neighbors coordinate their movement
Propagation in time and space to guide groups of cells

Thanks for sharing your data!

Institute for bioengineering of Catalonia

Memorial Sloan Kettering Cancer Center

Xavier Serra-Picamal

Xavier Trepat

Yun-Yu Tseng

Angeles Rabadan

Joachim Spatz

Tamal Das