

The Intersection of Physics And Biology

Jan Liphardt
UC Berkeley
Liphardt@berkeley.edu

physics  berkeley

BA PS-OC Bay Area
Physical Sciences-Oncology Center



qb3
ucb·ucsc·ucsf

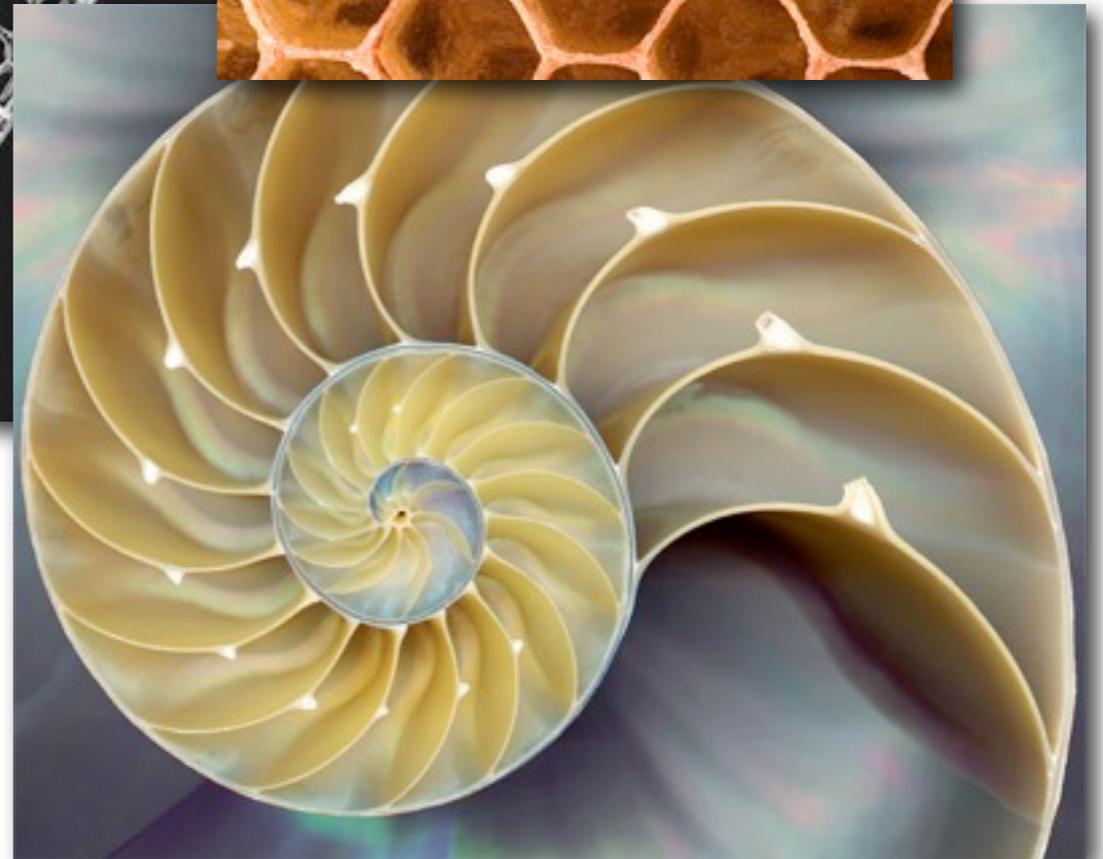
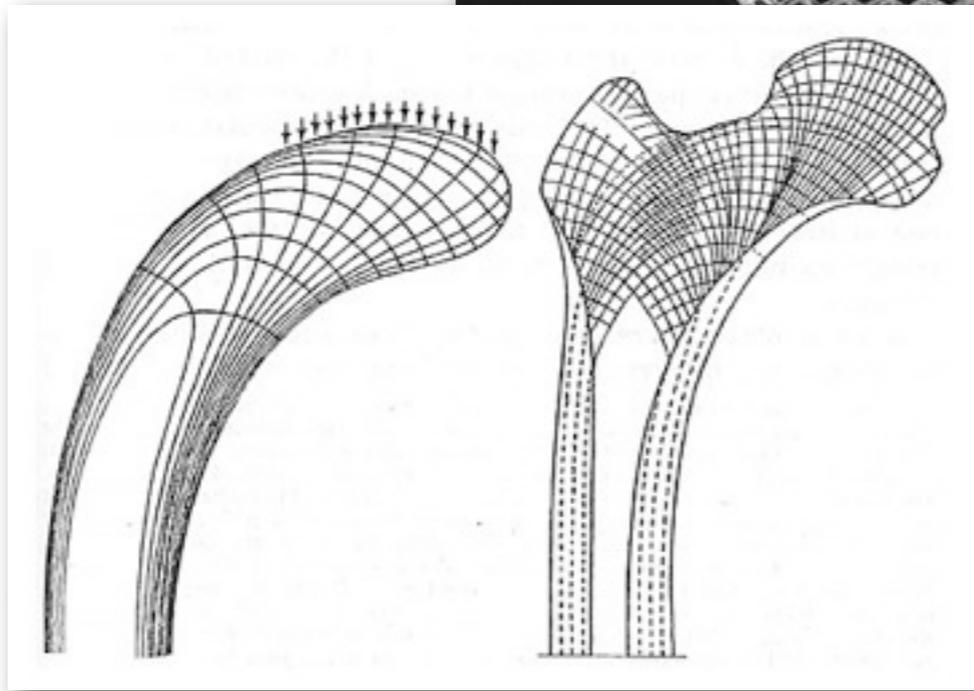
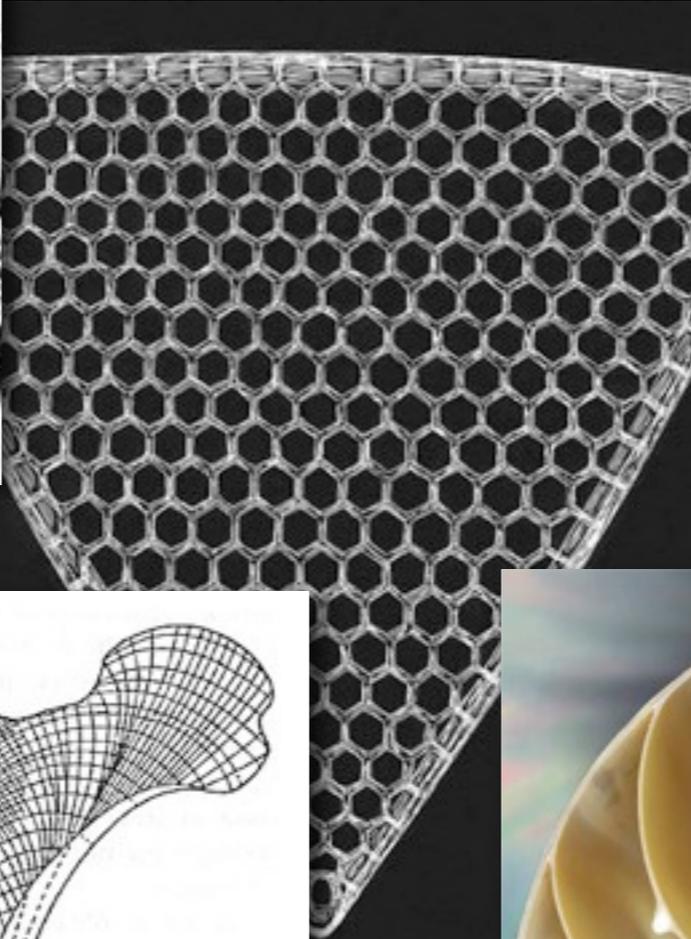
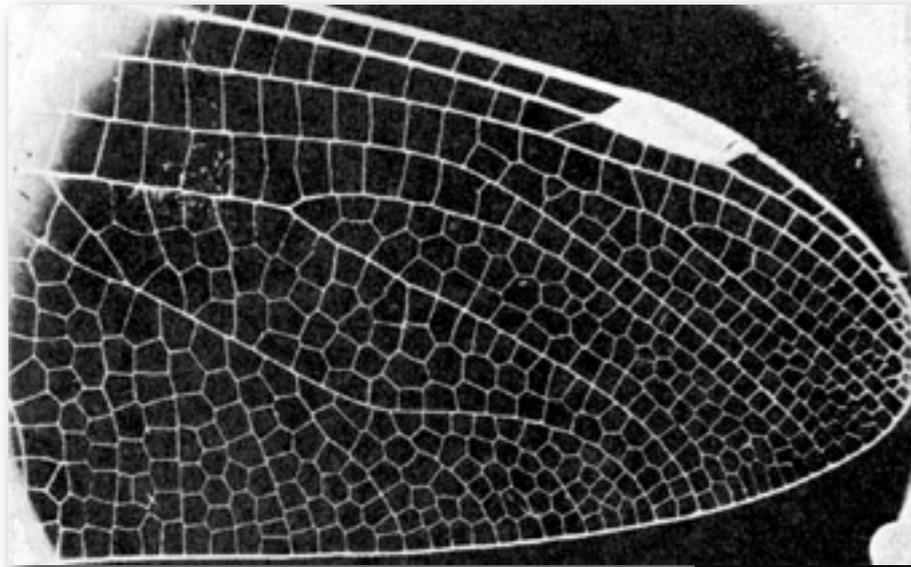
California Institute for
Quantitative Biosciences

This talk will (probably) be strange

It's about things we don't understand

Biology through the centuries

Turn of last century, biology involved chemistry, geometry, and physics



On Growth and Form,
D'Arcy Wentworth Thompson, 1917

Biology through the centuries

Fascination with patterns,
spirals,
self-similarity,
mechanics,
cells vs. inanimate objects



Romanesco broccoli

Kenneth Libbrecht, Caltech Physics

Biology in the 20th century

Massive explosion of molecular biology and biochemistry

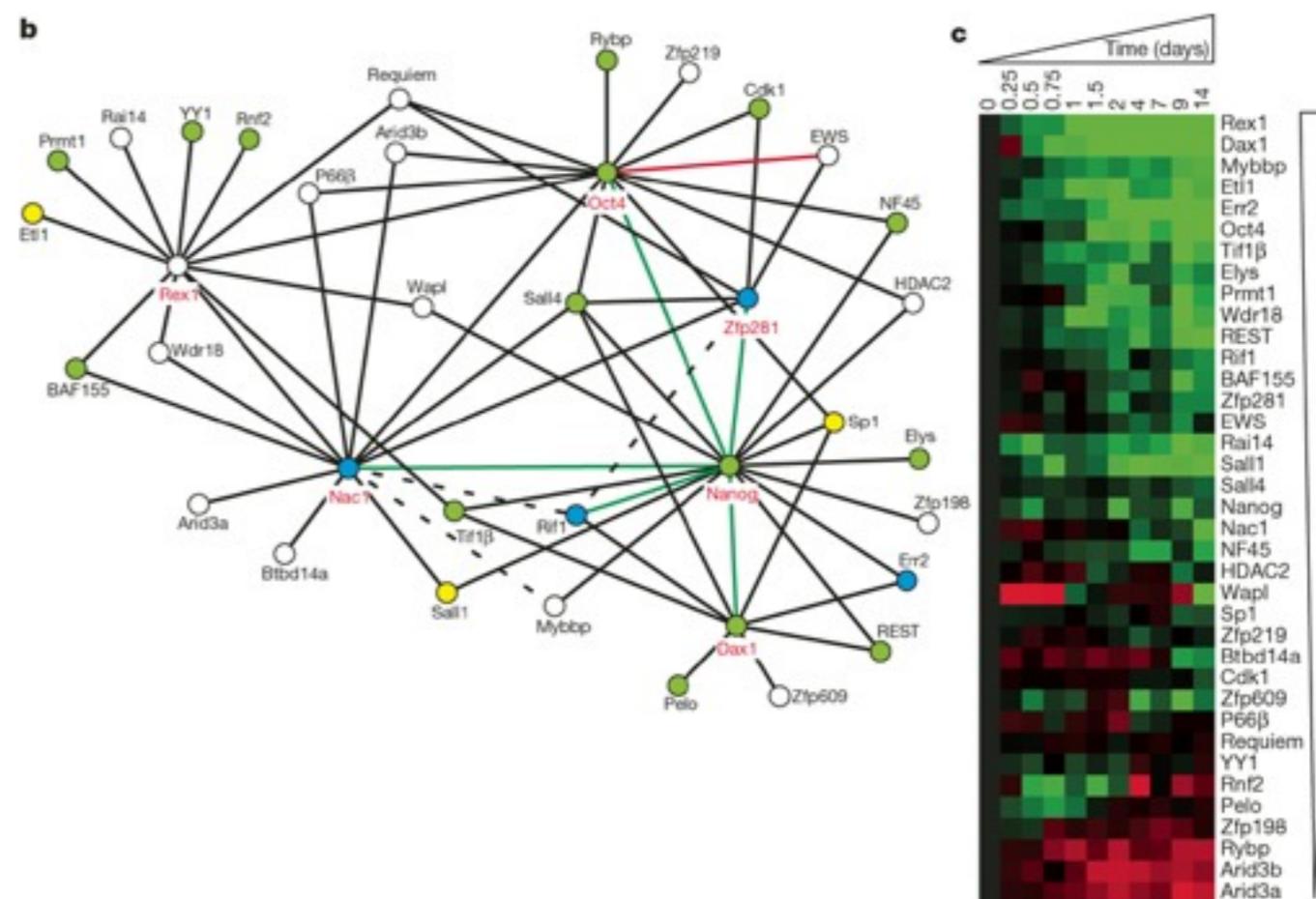
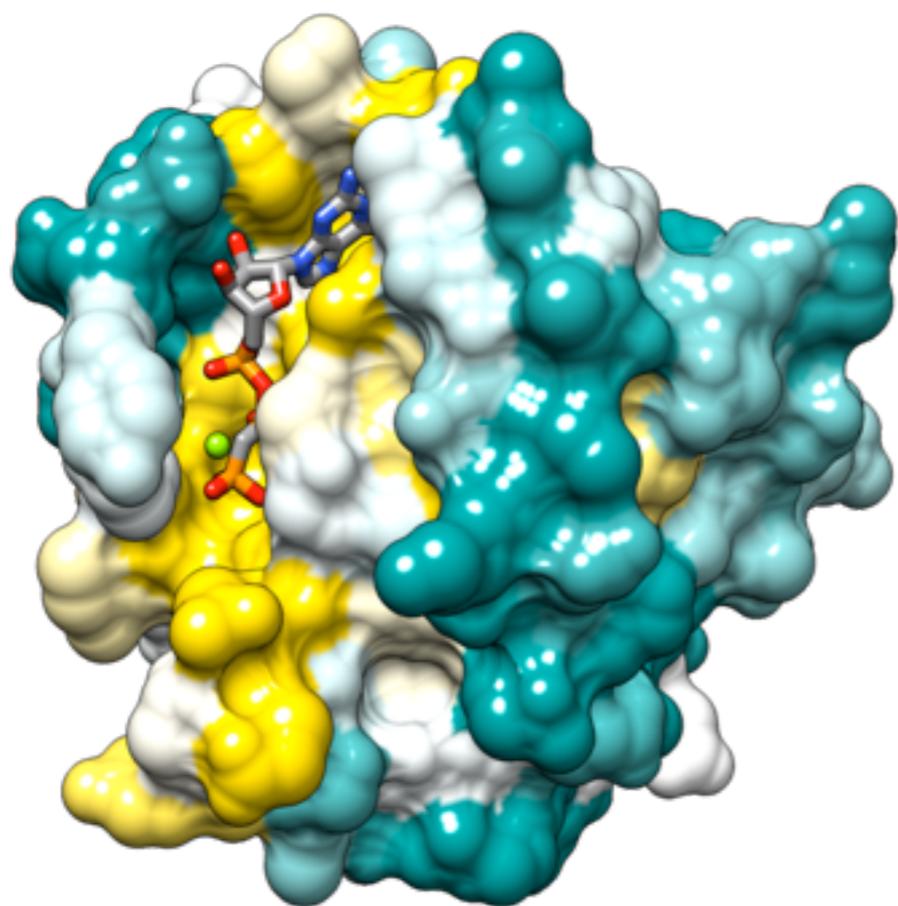
Physics largely disappears, except instrumentation/methods

Defined by the gene, culminating with the sequencing of the human genome in 2004

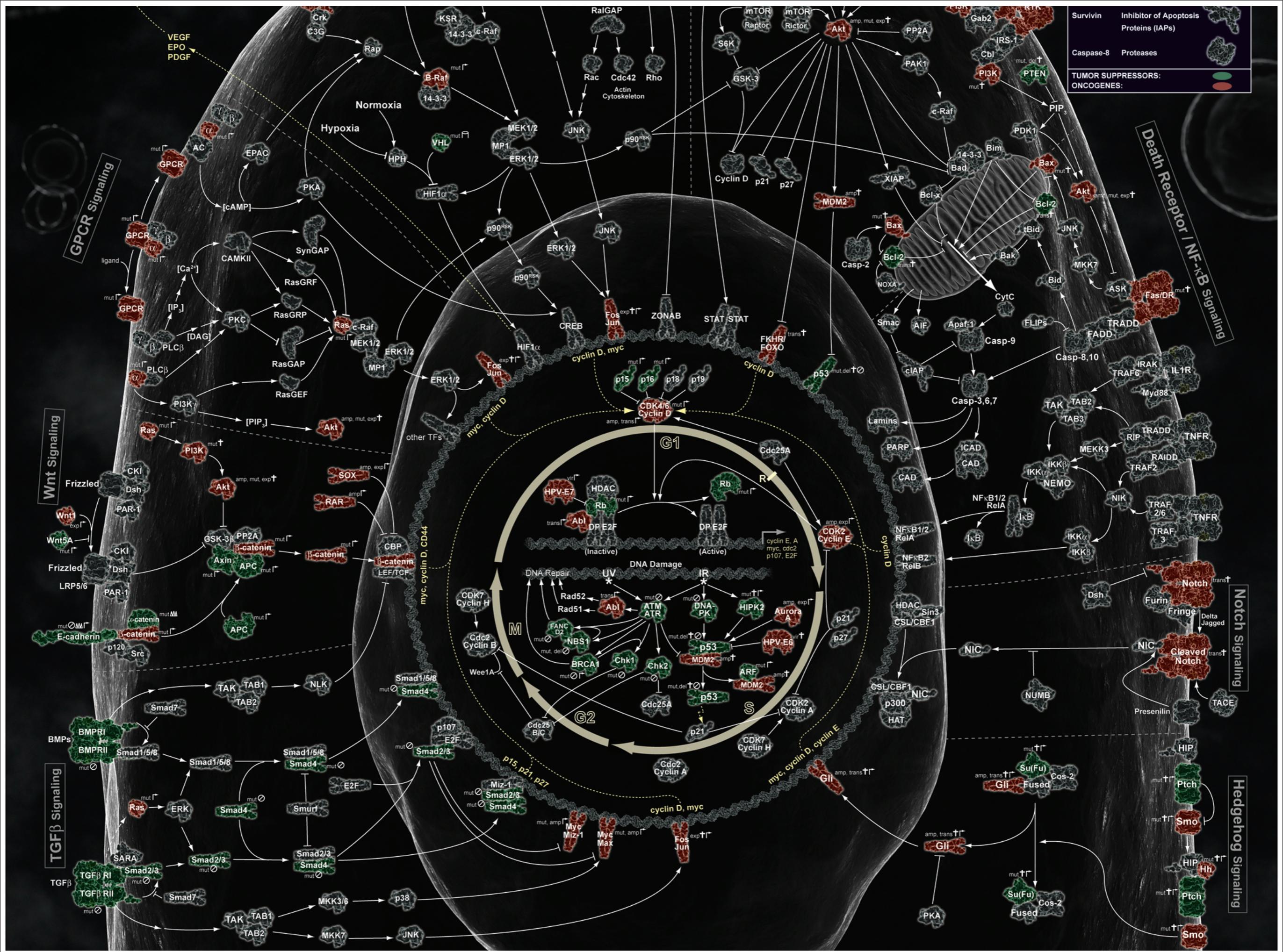
Key idea: Sequence the genome and then read the blueprint

Protein X causes disease Y

Function is defined by protein biochemistry and network connectivity



Wang et al., *Nature* 444, 2006, Interaction network for pluripotency of embryonic stem cells



Taking stock, 2010

Entire human genome sequenced



Know most of the key proteins



Know their structure and function, individually



Interaction networks



Taking stock, 2010

Entire human genome sequenced



Know most of the key proteins



Know their structure and function, individually



Interaction networks



Is biology over?

Taking stock, 2010

Entire human genome sequenced



Know most of the key proteins



Know their structure and function, individually



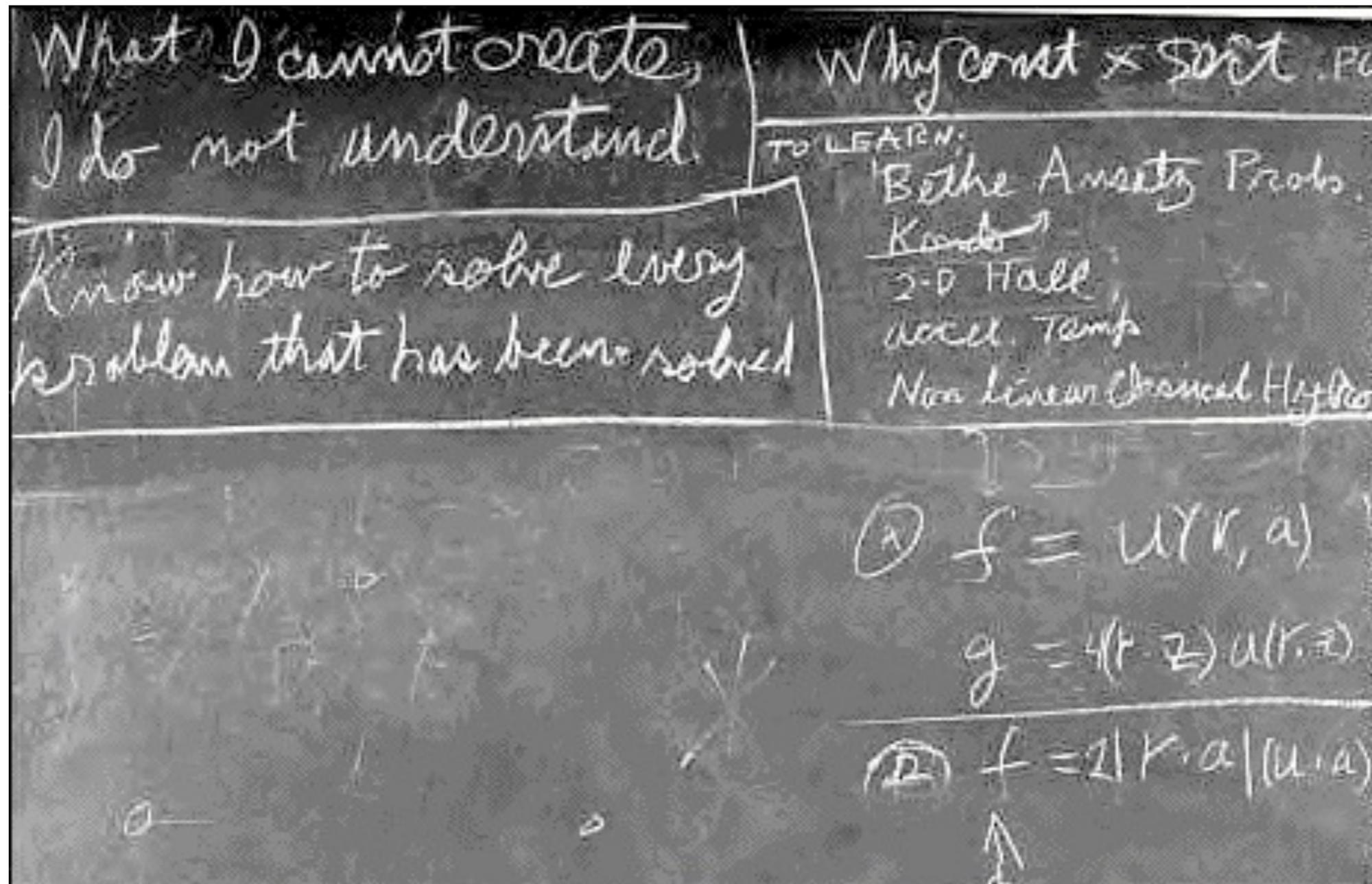
Interaction networks



Is biology over?

What's the grading scheme, how do we know where we're at?

R. P. Feynman's Blackboard



Scorecard

Ok - so how much of the biological world around us can we recreate, from first principles?

Effectively, zero!

Not a simple matter of scale, cost, infrastructure, dedication, lack of interest...

Reflects a significant lack of knowledge of fundamental biological organizing principles and mechanisms

There is no “Standard Model” of biology



Problem I: Genes \Leftrightarrow Cell Types

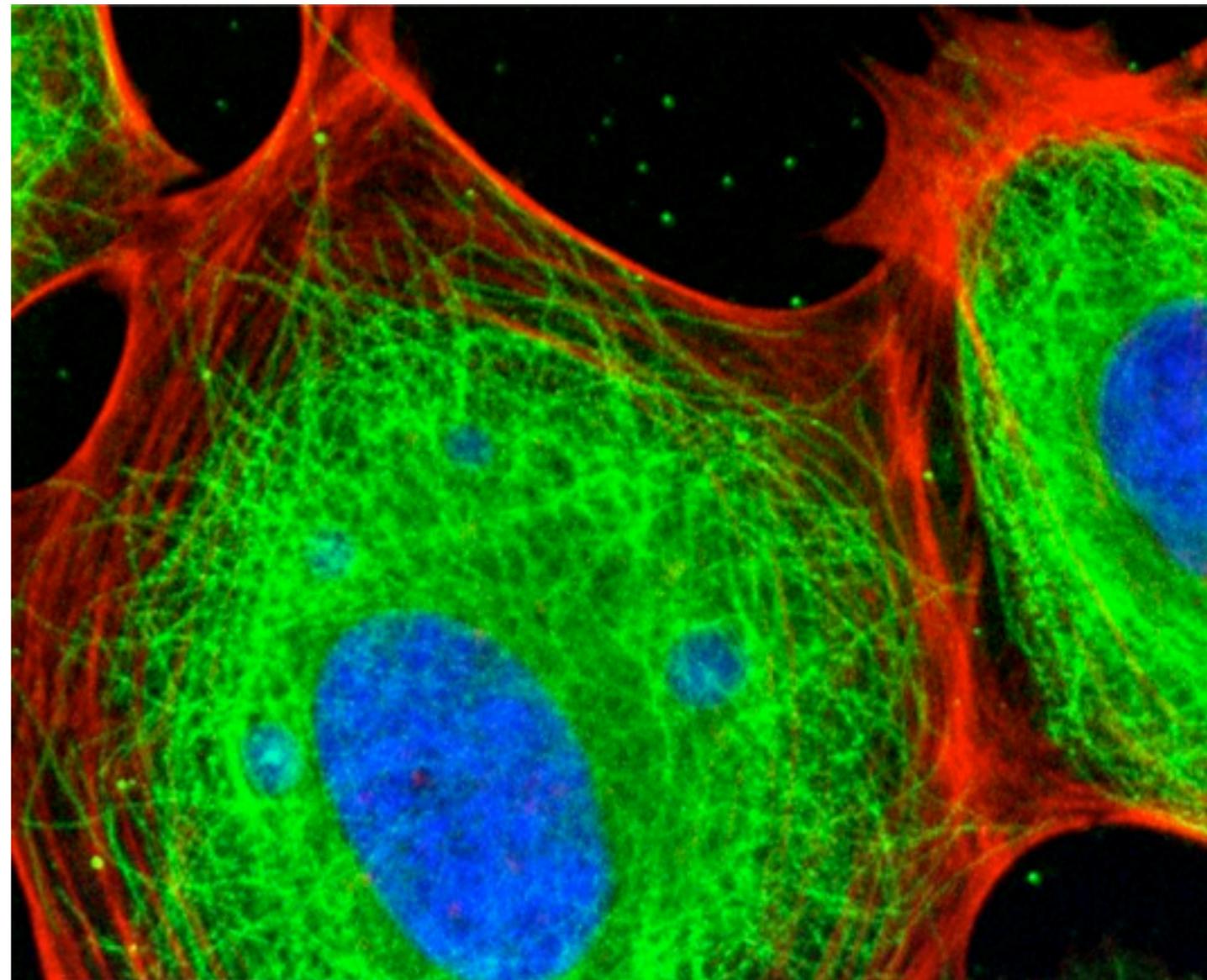
Human genome has ~ 23,000 protein coding genes.

1.5% of the genome codes for proteins

The rest: non-coding RNA genes, regulatory sequences, introns, and “other” DNA

Each protein is subject to exquisite regulation of concentration, activity, and location

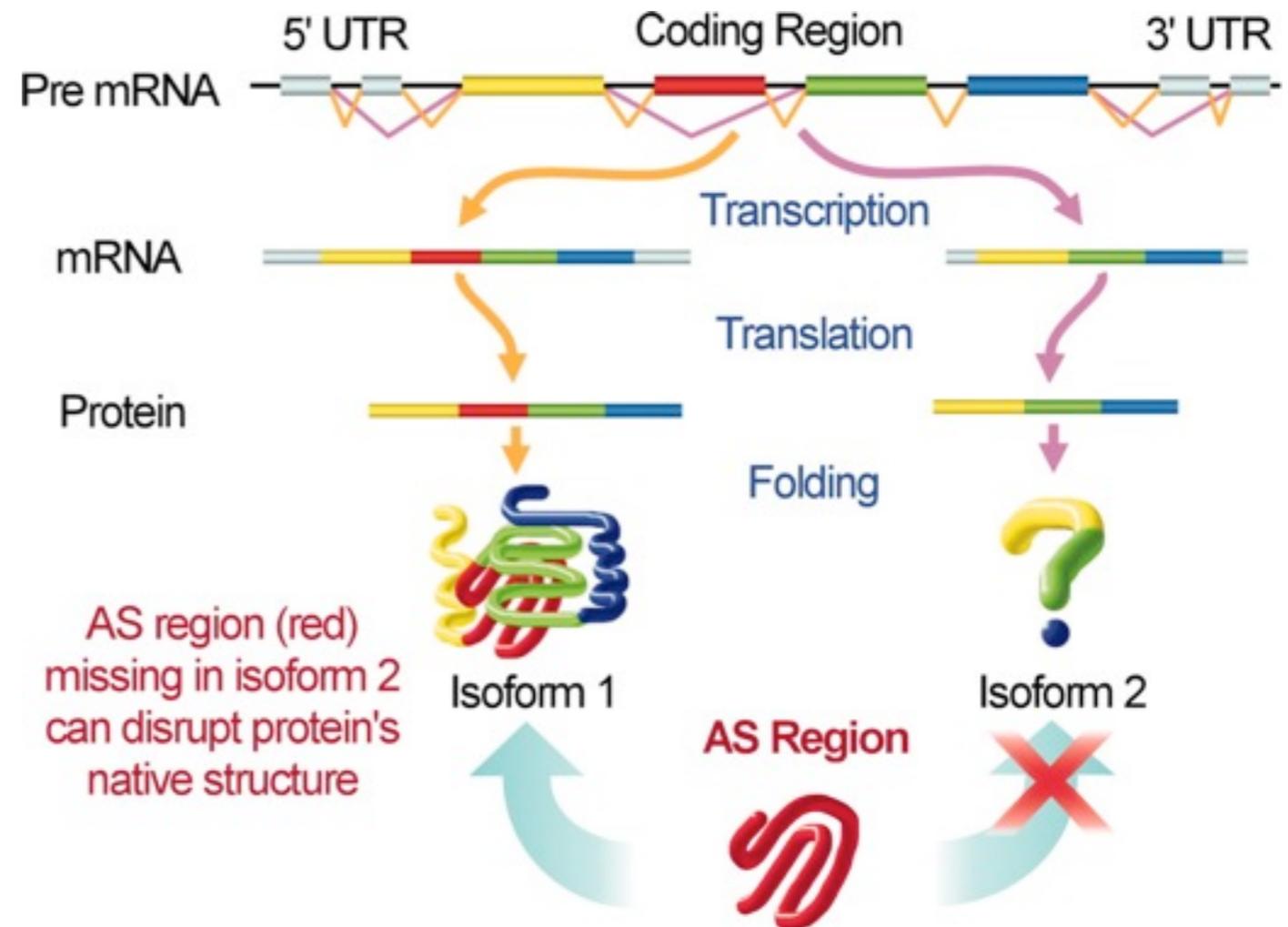
- via pre- and post-translational, control of nuclear export, alternative splicing, phosphorylation, SUMOlation



Problem 1: Genes \Leftrightarrow Cell Types

Alternative splicing: combinatorics are absolutely overwhelming

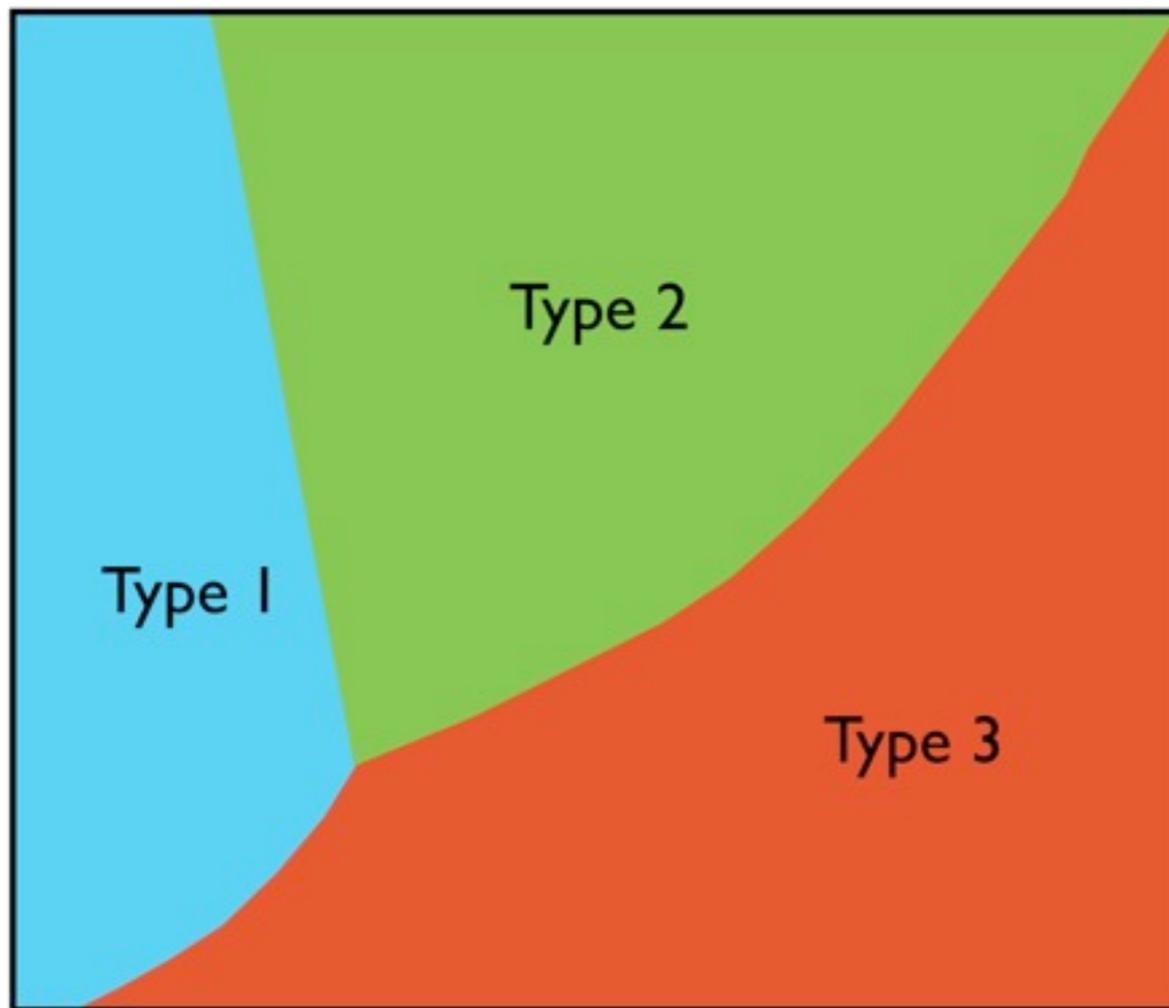
Thousands of genes have more than 2 isoforms, some as many as 12



Number of theoretically possible distinct mRNA/protein sets is staggering

Problem I: Genes \Leftrightarrow Cell Types

And yet, there are only about 210 distinct, discrete human cell types.

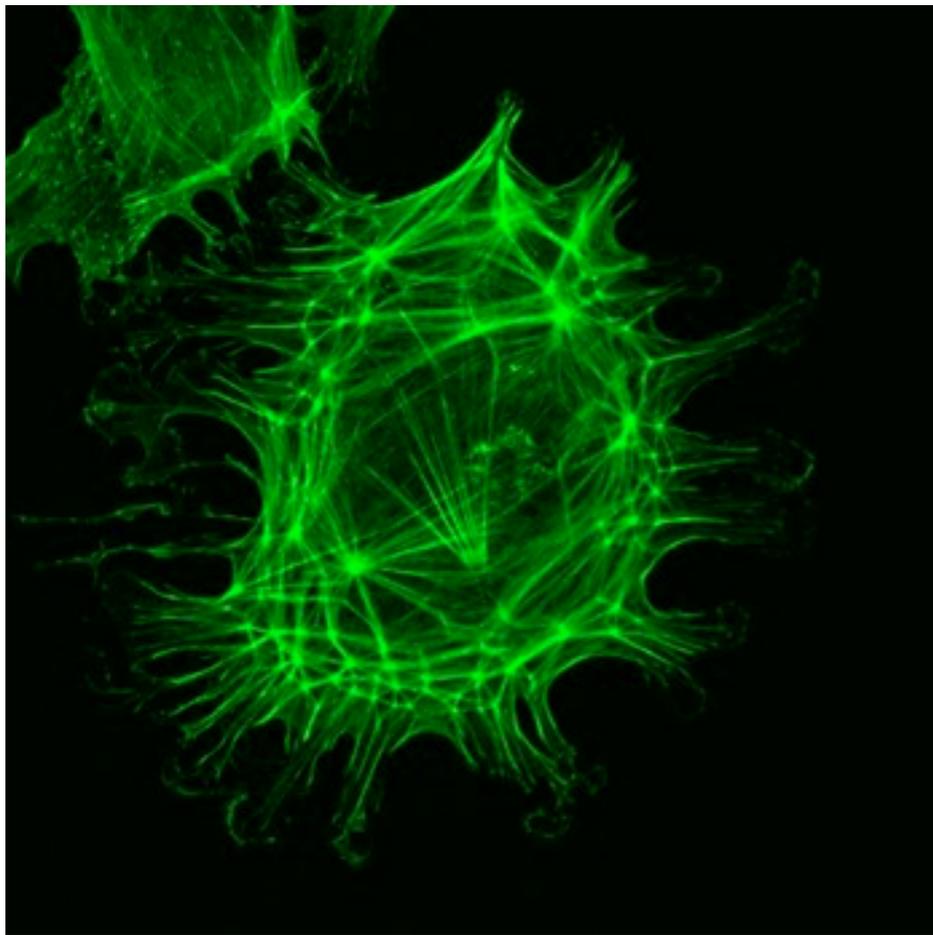
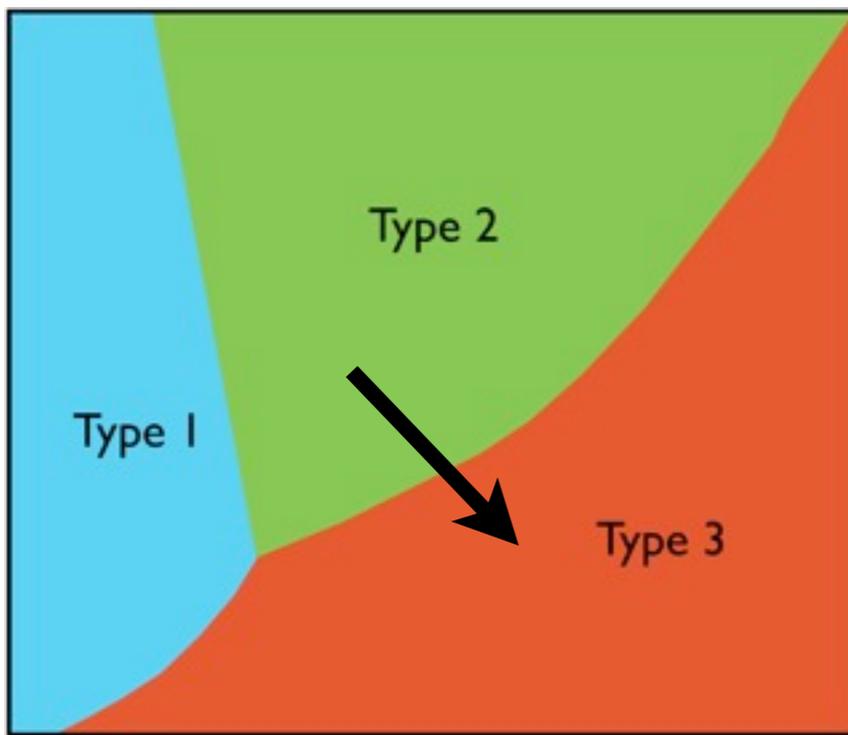


System is highly degenerate -

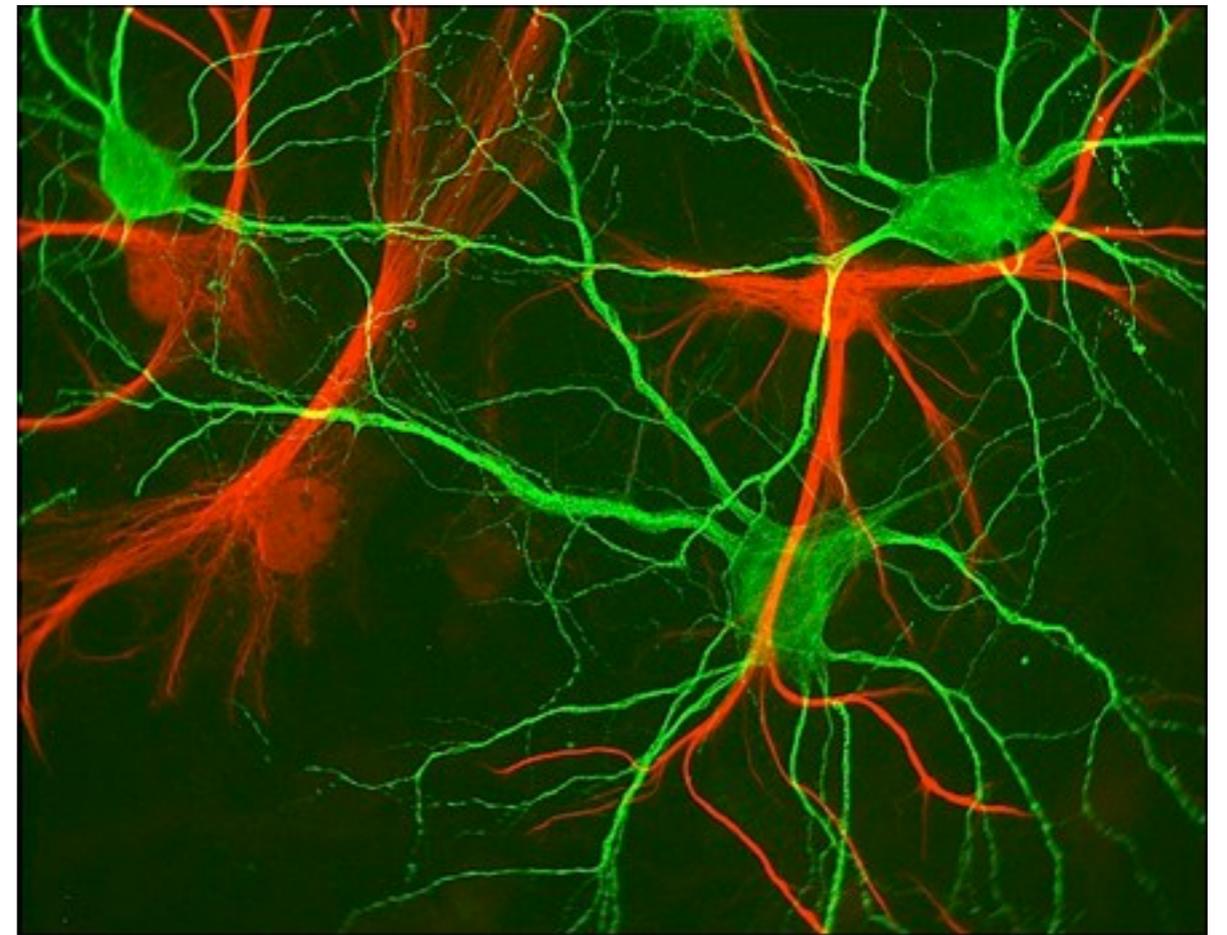
many different (combinations) of pathways,
genes, “microstates” lead to same cell type

Schematic of the cellular type space
210 patches - I'm showing 3

Problem 1: Genes \Leftrightarrow Cell Types



Ascl1
Brn2
Myt1l



Direct conversion of fibroblasts to functional neurons by defined factors
doi:10.1038/Nature 08797

Problem I: Genes \Leftrightarrow Cell Types

- Astronomical number of biochemical “microstates”
- Thousands of positive and negative feedback loops
- All of this is taking place in thermal bath, energies on order of a few $k_B T$
- 10 trillion cells in the body

And yet

- Small number of cell types
- Cell types tend not to change suddenly
- Development remarkably reliable

Then again, cells are not irreversibly locked into a particular cell type

- If needed (wound healing, differentiation), cells are plastic and can switch type
- Plasticity/type switching involves a few genes

Largely unclear how this works - some type of statistical mechanics approach probably needed

Problem II: Robust Development of an Animal from one Cell



Problem II: Robust Development of an Animal from one Cell



Michael Klymkowsky, *Xenopus* embryos (frog)

Problem II: Robust Development of an Animal from one Cell



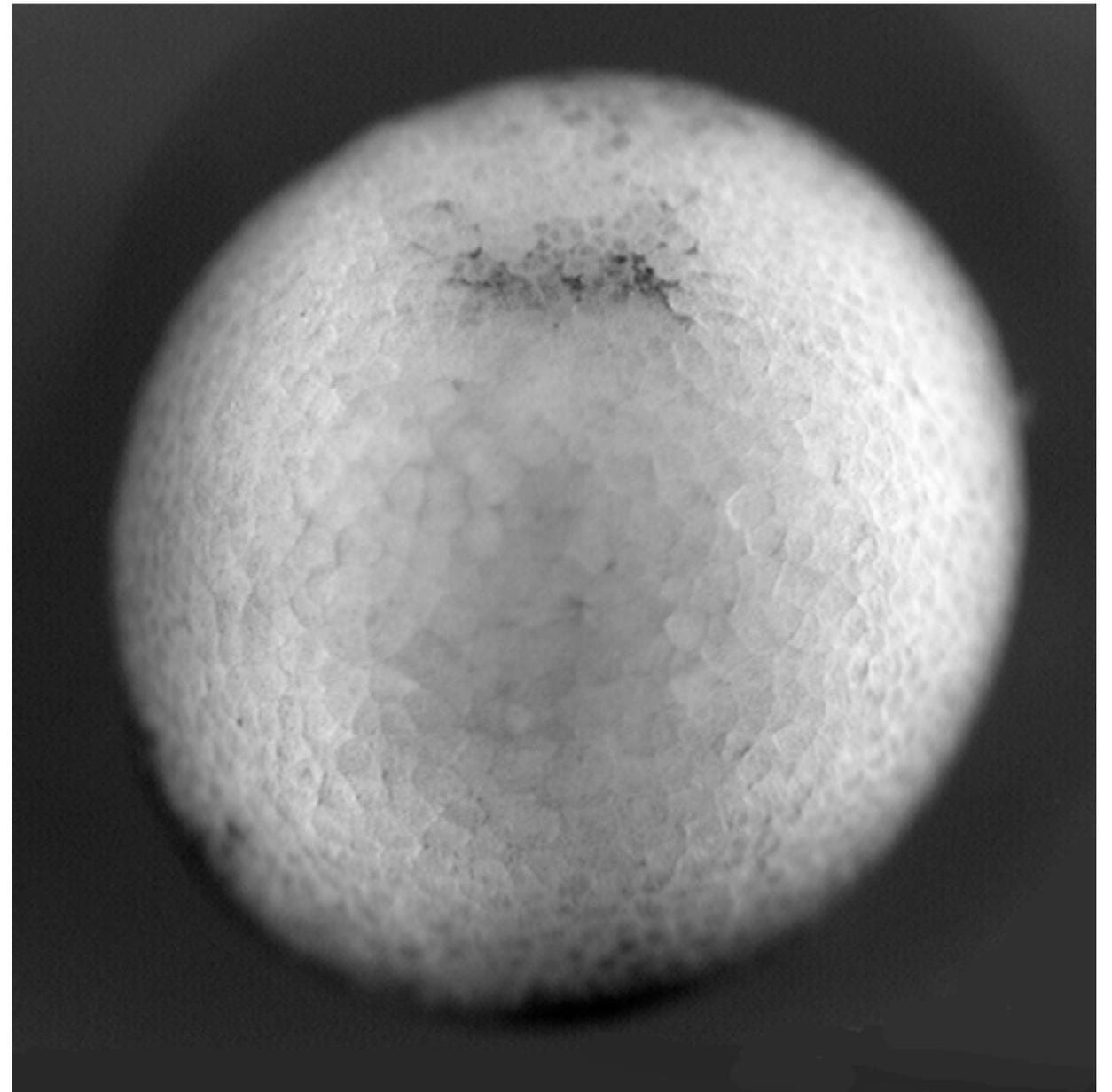
Requires the coordinated, error-correcting, self-replication and self-organization of trillions of cells

Xenopus laevis Vegetal View of Gastrulation & Neurulation: (15.0 hours elapsed, 48 minutes/second)
<http://www.gastrulation.org/>

Problem II: Robust Development of an Animal from one Cell



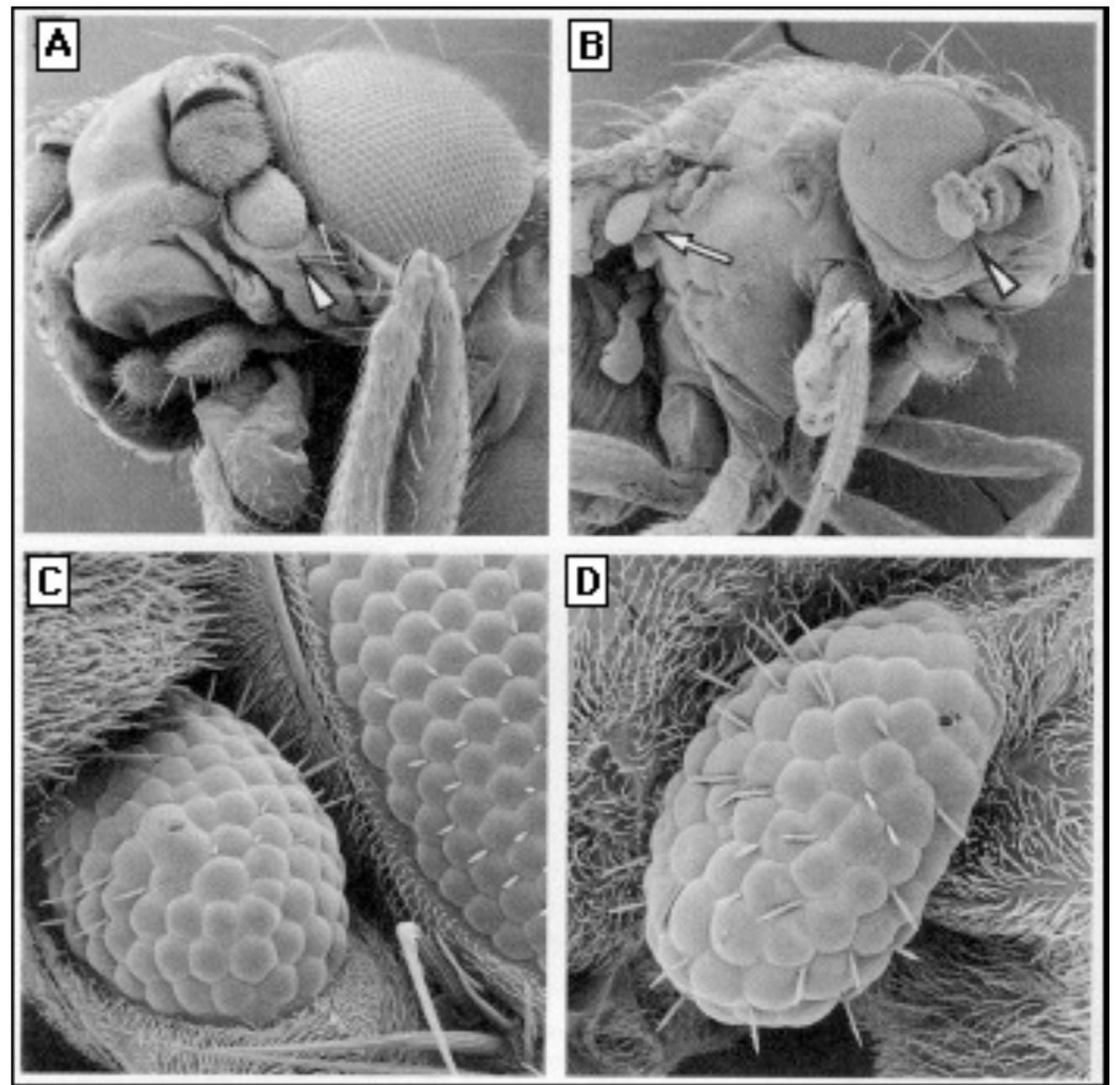
Requires the coordinated, error-correcting, self-replication and self-organization of trillions of cells



Xenopus laevis Vegetal View of Gastrulation & Neurulation: (15.0 hours elapsed, 48 minutes/second)
<http://www.gastrulation.org/>

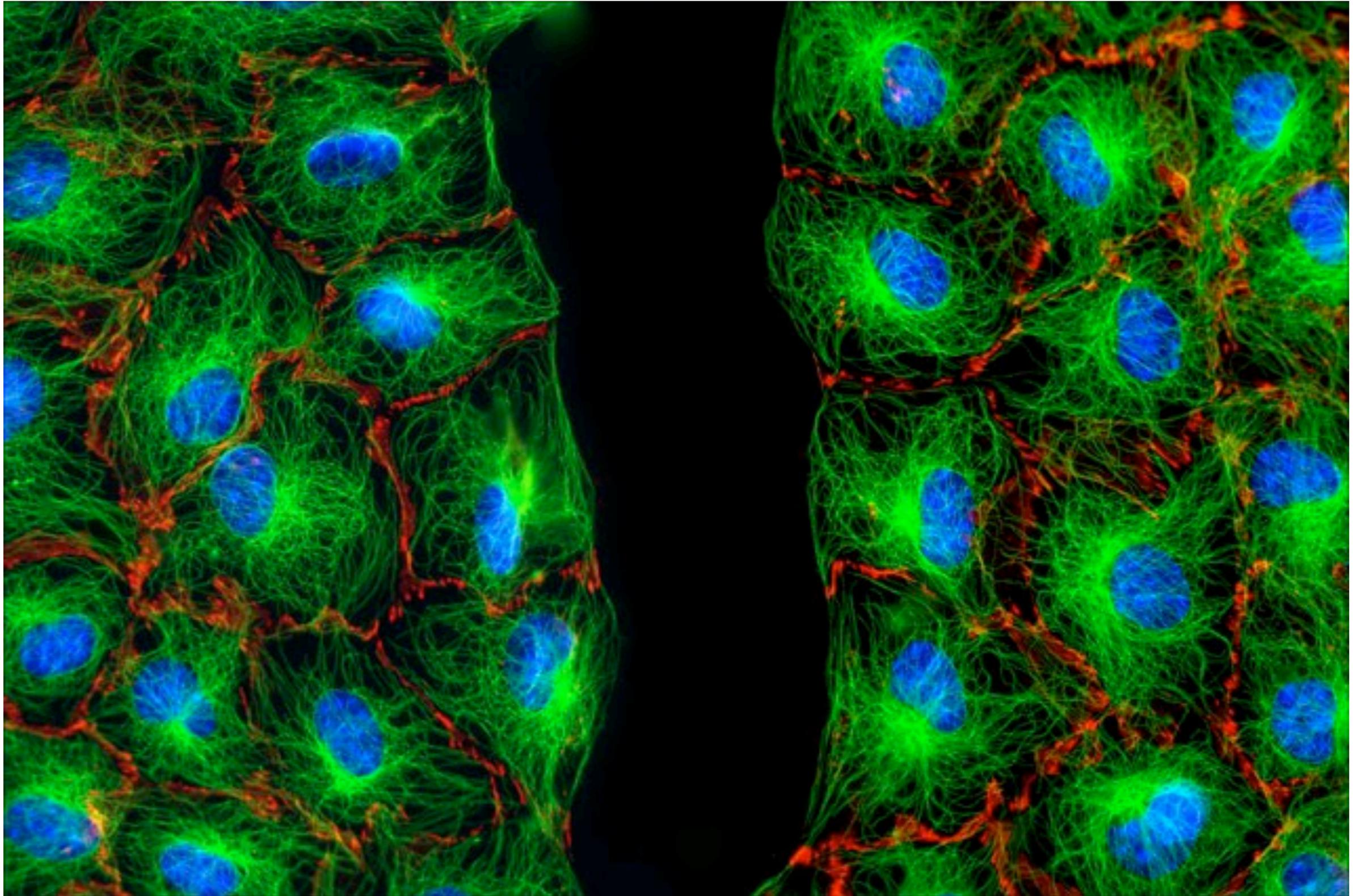
Problem II: Robust Development of an Animal from one Cell

- Reliable self-replication and self-organization of trillions of cells into an animal
- We are good at breaking development, but by Feynman's standard, have next to zero ability to reprogram development in a controlled manner.



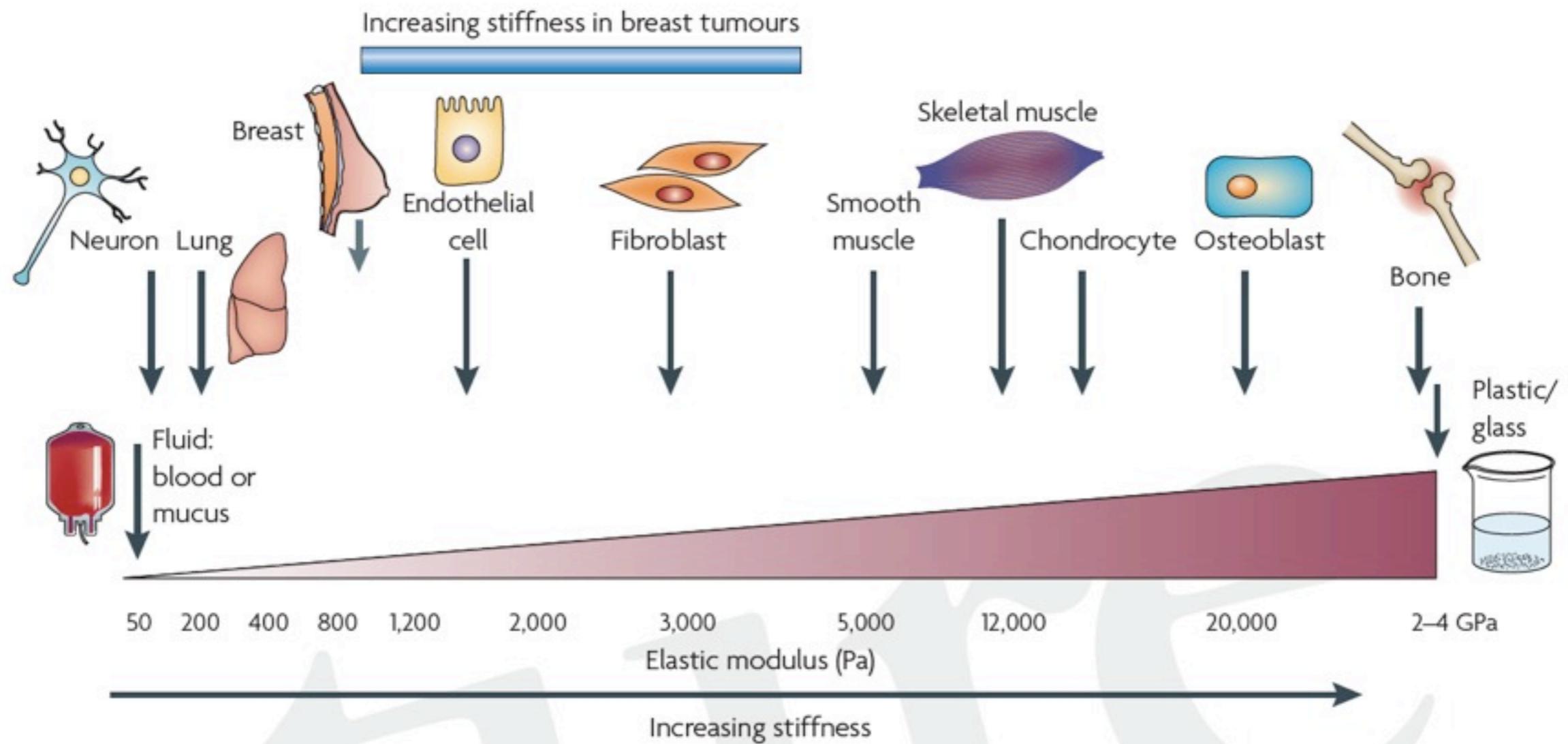
Largely unclear how this works - local and global mechanical and chemical cues are used

Problem III: Integrating Genes with Mechanical Forces



Jan Schmoranzler, Wounded monolayer of fibroblast cells in culture, Columbia University

Matrix materials properties, cellular tension and normal tissue behavior

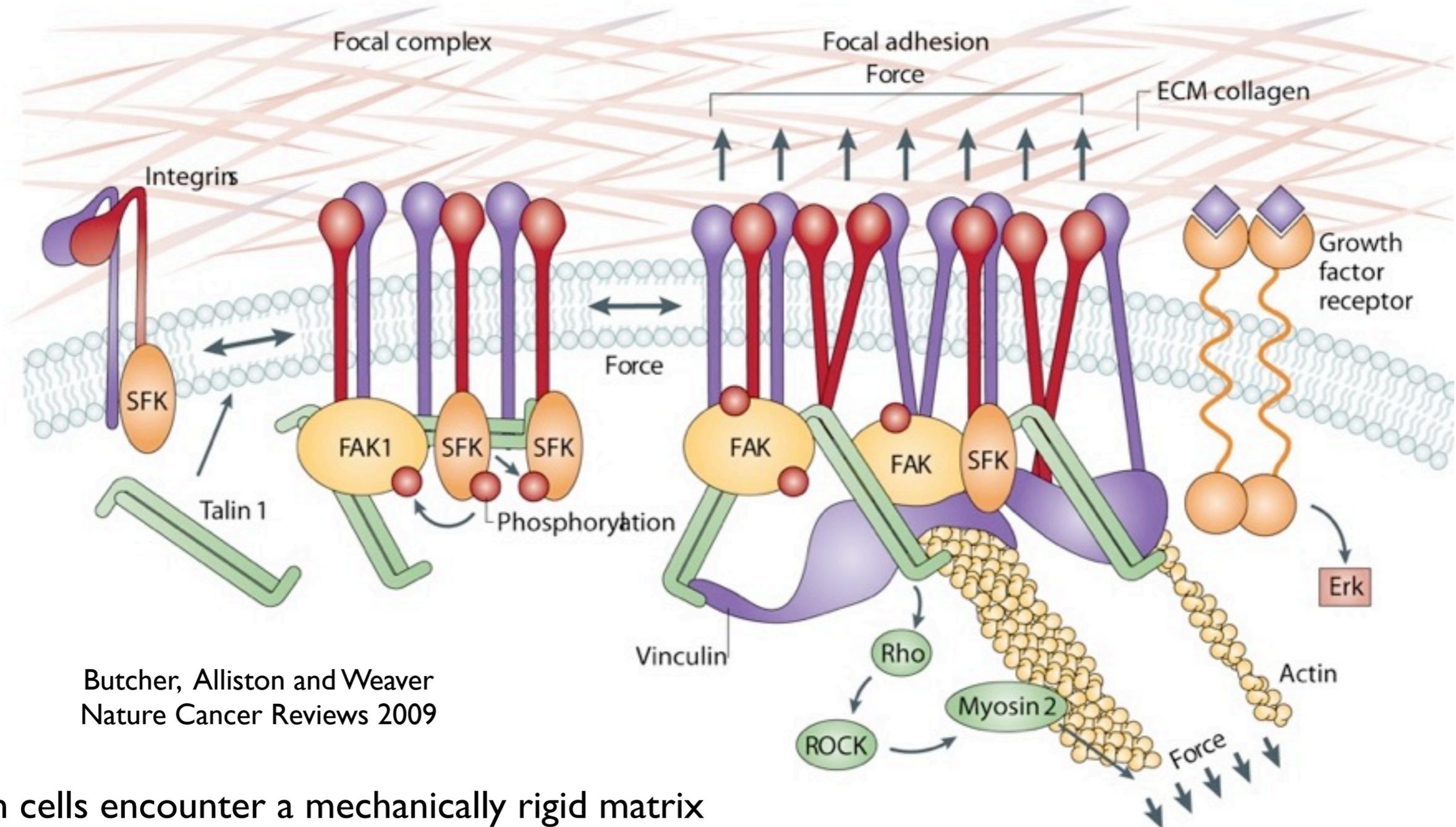


Butcher, Alliston and Weaver
Nature Cancer Reviews 2009

Normal mammary epithelial cell growth, survival, differentiation and morphogenesis are optimally supported by a soft matrix (~200 Pa).

Following transformation, breast tissue becomes progressively stiffer.

Protein and Cell-level Force Sensing/generation



Butcher, Alliston and Weaver
Nature Cancer Reviews 2009

When cells encounter a mechanically rigid matrix (or are exposed to an exogenous force) integrins are activated, favoring

- integrin oligomerization or clustering, talin I and p130Cas protein unfolding, (*)
- vinculin–talin association, (*)
- Src and focal adhesion kinase (FAK) stimulation of RhoGTPase-dependent actomyosin contractility
- actin remodeling. (*)

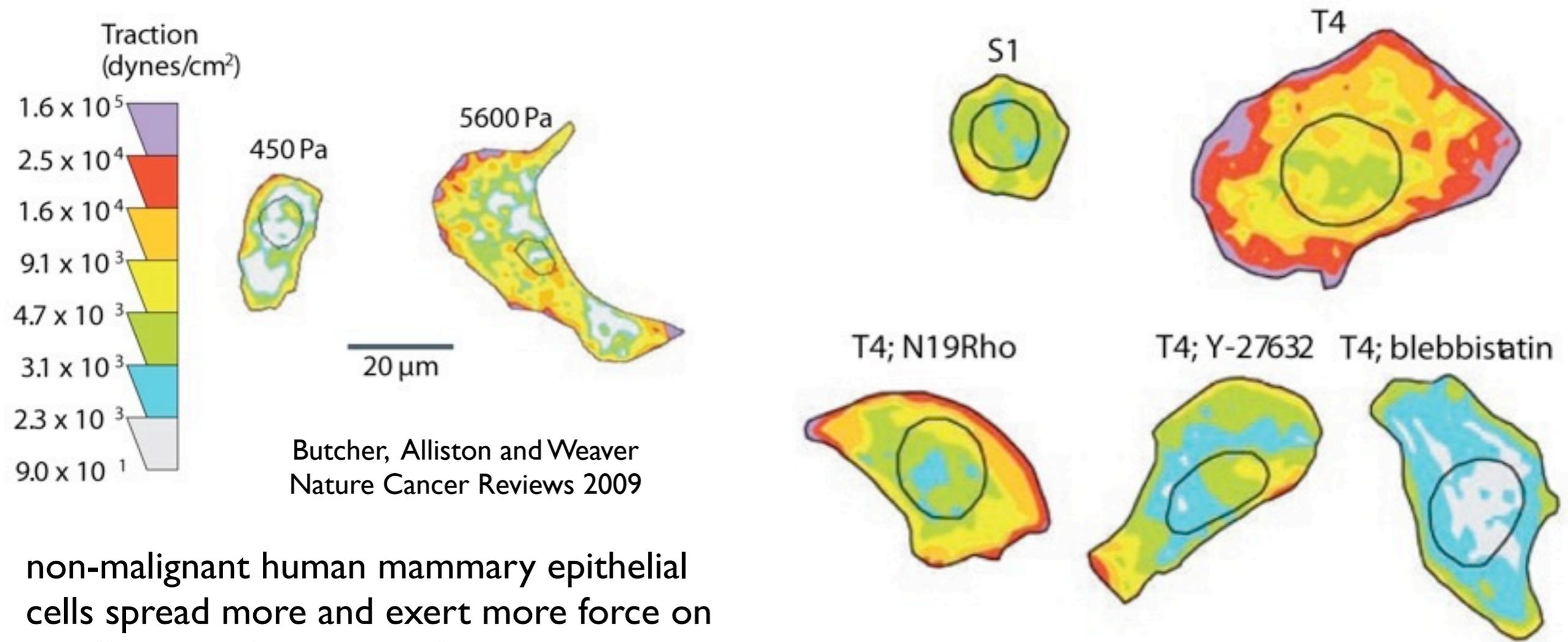
Tensional homeostasis

Cells measure forces/compliance

Cells generate forces

Cells change their compliance and the compliance of the ECM

Tension sensing/actuation loop with positive and negative feedback, with a range of setpoints

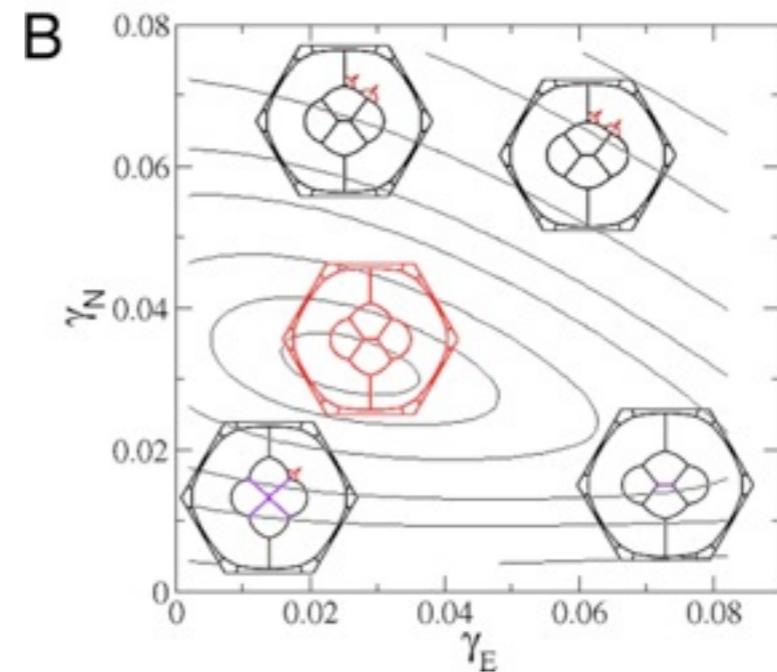
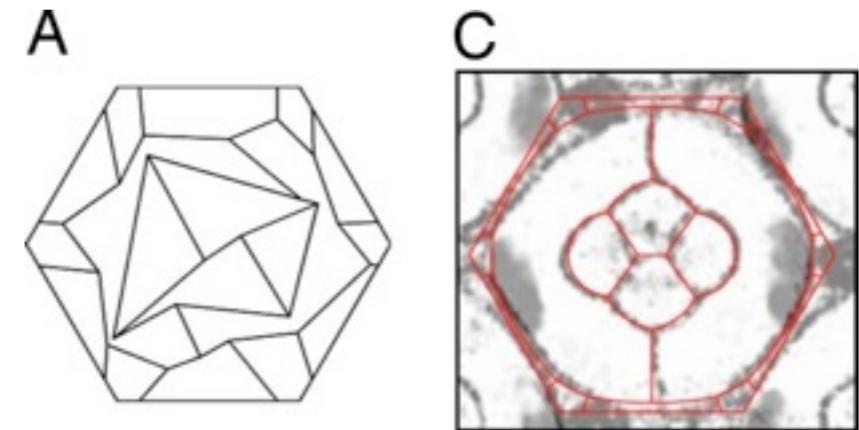
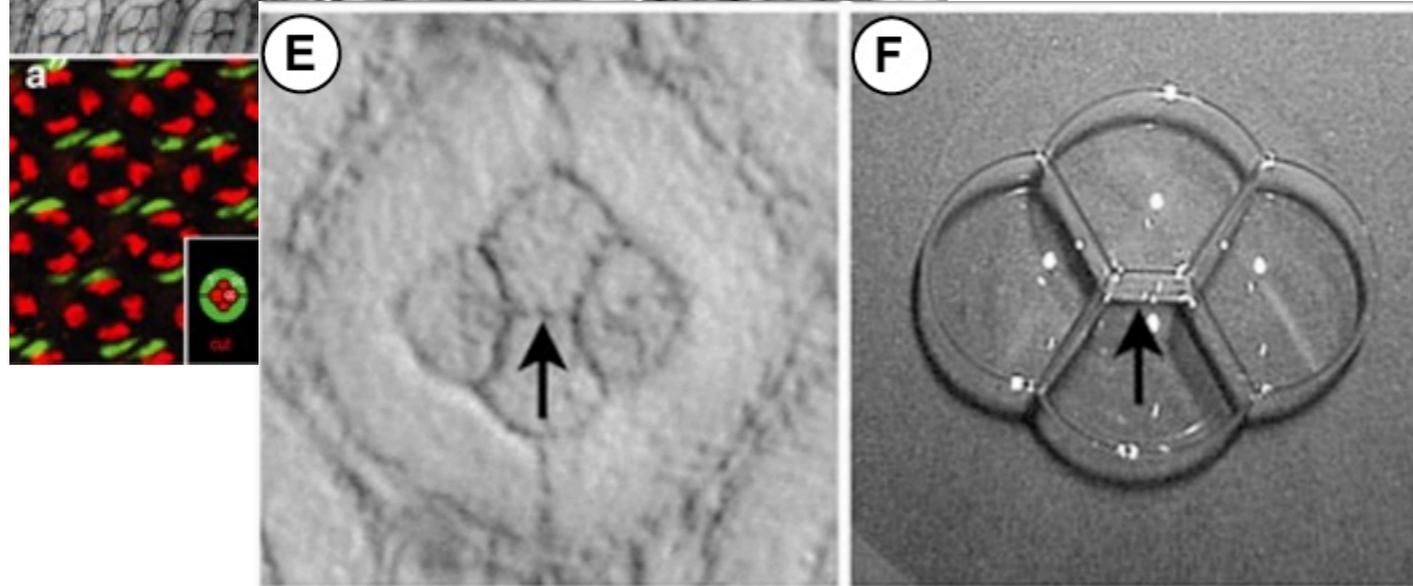
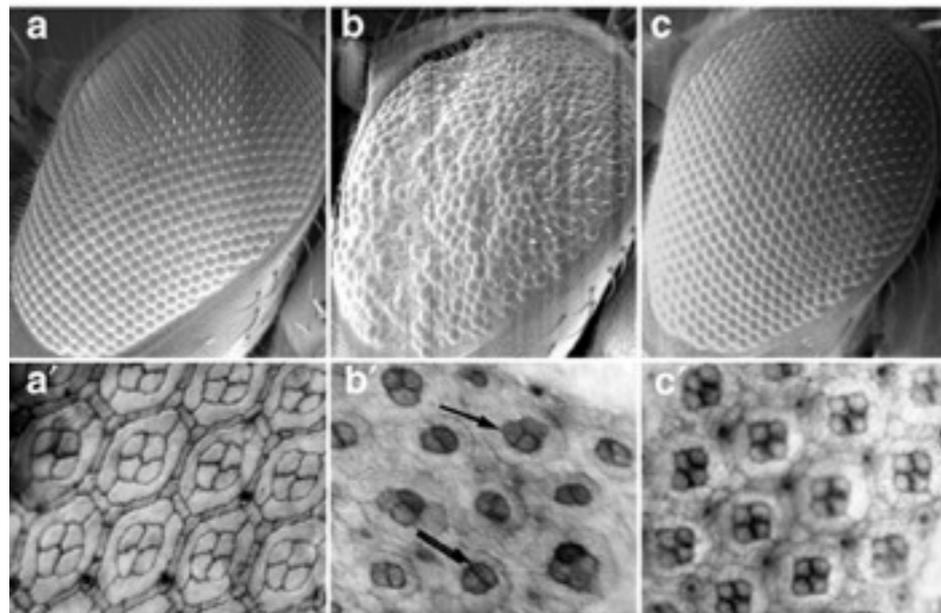


non-malignant human mammary epithelial cells spread more and exert more force on a stiff matrix than on a soft matrix

$$F_{\text{generated}} \propto \gamma \cdot EM_{\text{matrix}}$$

$$\gamma_{\text{normal}} < \gamma_{\text{cancer}}$$

Problem III: Integrating Genes with Mechanical Forces

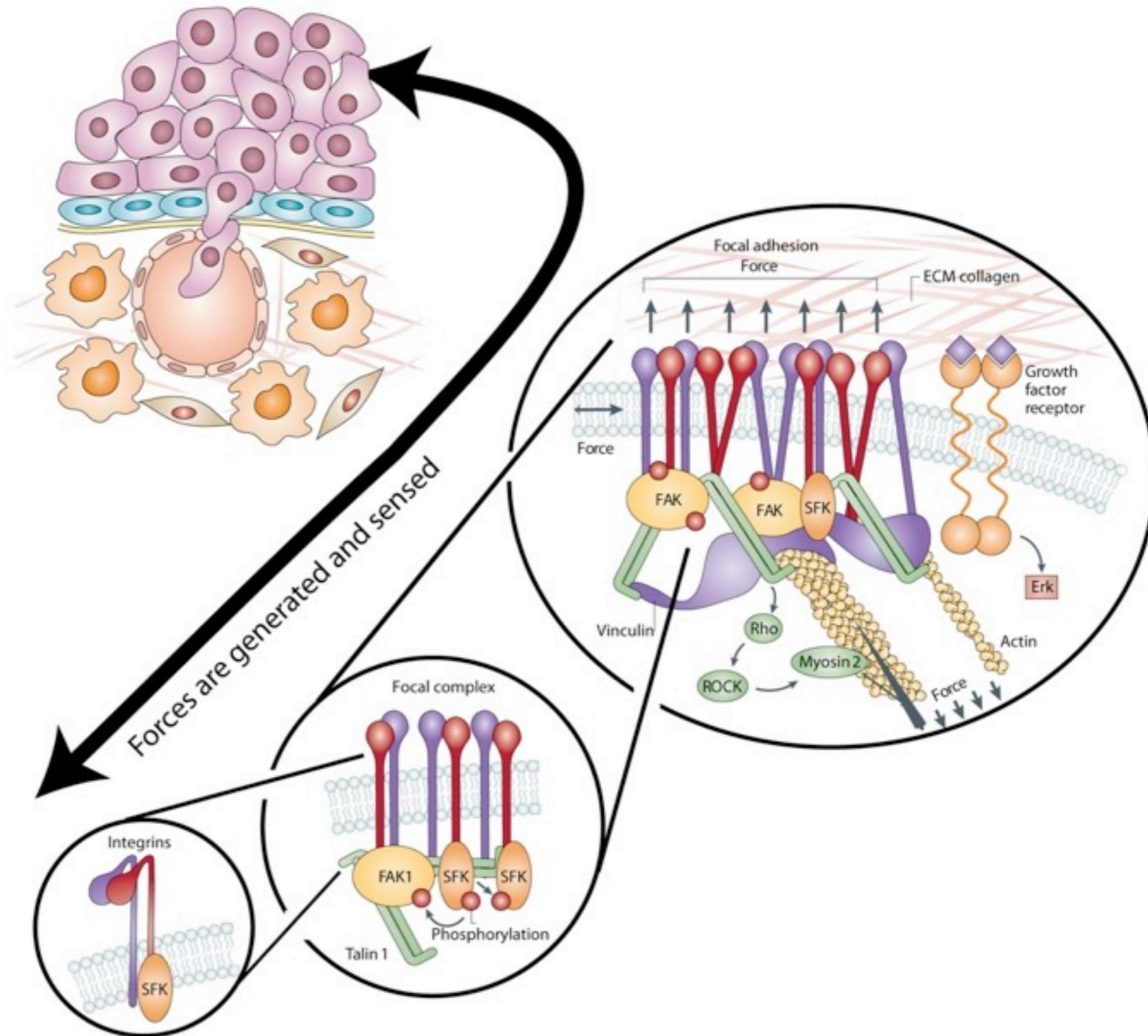


$$\mathcal{E} = \sum_i \frac{1}{2} \Delta_i^2 L_{0i} - \sum_{i,j} L_{ij} \gamma_E \delta_{i,E} \delta_{j,E} - \sum_{i,j} L_{ij} \gamma_N \delta_{i,N} \delta_{j,N}$$

Physical modeling of cell geometric order in an epithelial tissue, Hilgenfeldt, Erisken, and Carthew, PNAS 2008.

Mechanobiology ->

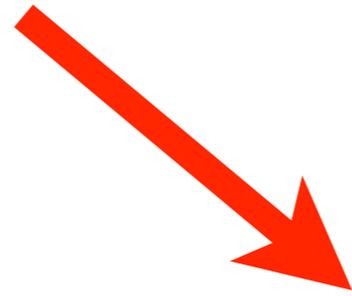
How are mechanical cues generated, transmitted, and integrated with chemical cues?



Problem I: Genes \Leftrightarrow Cell Types

Problem II: Robust Development of an Animal from one Cell

Problem III: Integrating Genes with Mechanical Forces



Single-molecule Biophysics

Mechanochemistry/
Molecular machines

Biological Form and
Function

3D arrangement vs.
function

Oncology

Mechanobiology

Precision control and measurement

quantum dots, nanowires, plasmonics

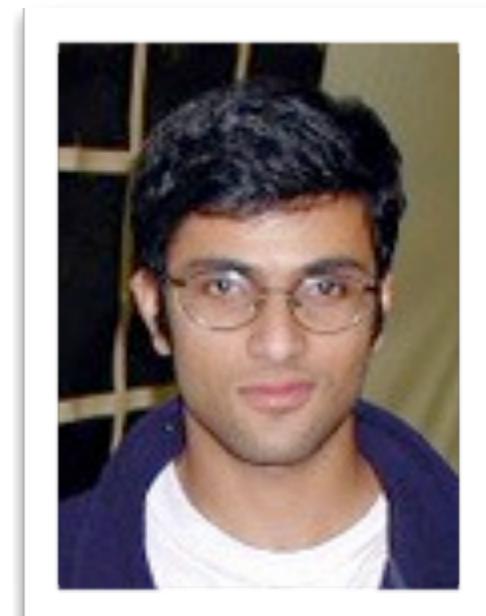
Key people:

Eric Betzig (HHMI, Janelia Farm)

Hari Shroff (HHMI, Janelia Farm)

Ann McEvoy and Derek Greenfield

Ned Wingreen (Princeton)



Spatial cell biology:

Parts, and their relative organization, determine function

Spatial cell biology:

Parts, and their relative organization, determine function

Key: reliable counting plus relative organization

Spatial cell biology:

Parts, and their relative organization, determine function

Key: reliable counting plus relative organization

Does this cell have 37 polymerases?

Does this cell have 8, 93, 764, or 8092 copies of Protein X?

How are those 8092 copies organized in space, relative to one-another?

Who cares?

What possible utility could there be to know 3 vs. 36 molecules?
5435 vs. 8735?

Who cares?

What possible utility could there be to know 3 vs. 36 molecules?
5435 vs. 8735?

Precision counting certainly relevant when something is rare
⇒ big difference between 0 and 1

⇒ fluxes... rate of production versus time

⇒ thermodynamics...

⇒ essential for robust model testing...

Who cares?

What possible utility could there be to know 3 vs. 36 molecules?
5435 vs. 8735?

Precision counting certainly relevant when something is rare
⇒ big difference between 0 and 1

⇒ fluxes... rate of production versus time

⇒ thermodynamics...

⇒ essential for robust model testing...

What you really want is precision counting + relative location
⇒ special combination

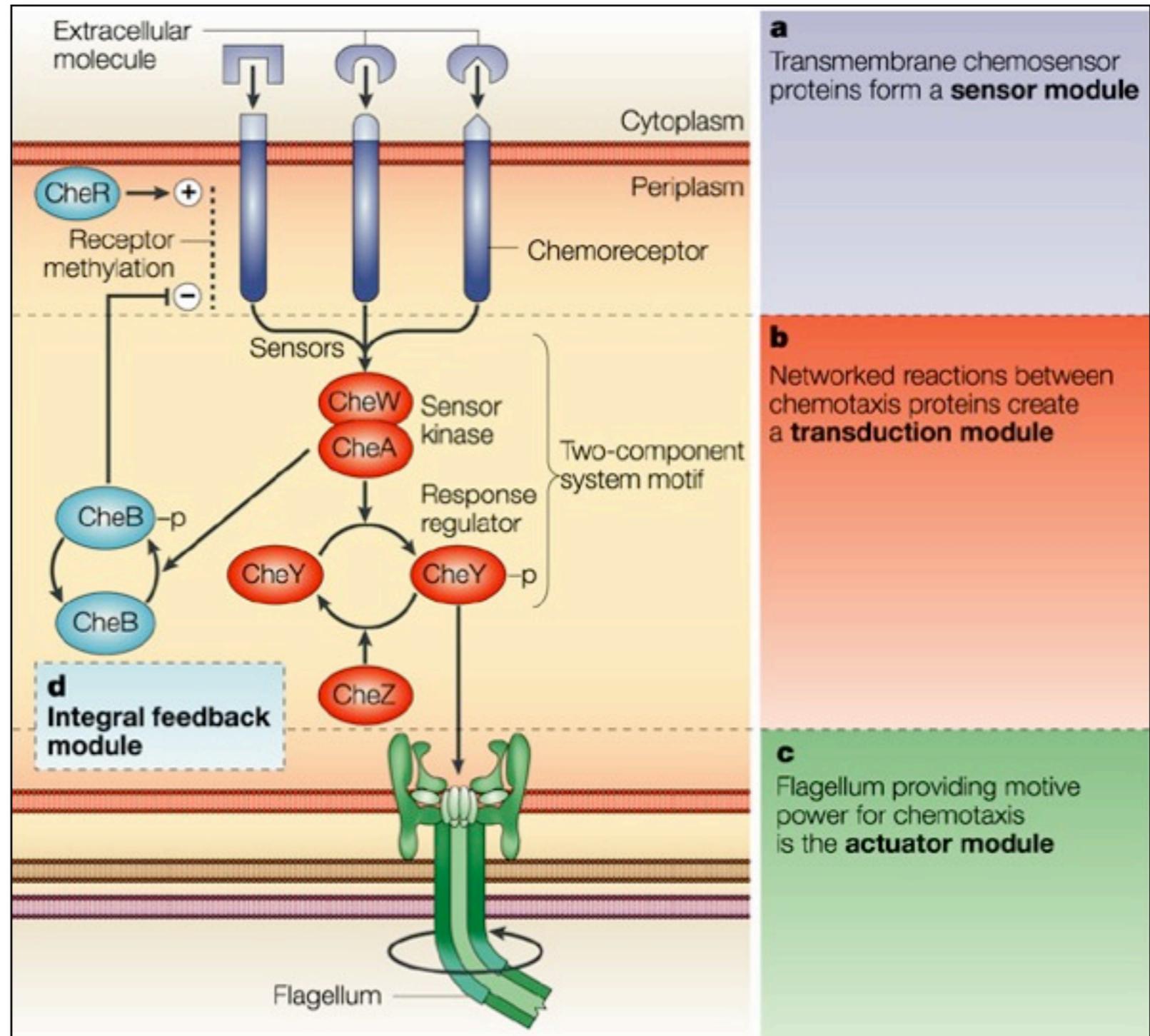
Numbers + relative position - why special?



- The exponentially distributed sizes of rain drops reflect their spontaneous aggregation and growth
- The Gaussian distribution of cell length in *E. coli* reflects the tightly regulated process of growth and division
- The relative spatial positioning of clusters and the precise distribution of cluster sizes contains information about the mechanism of cluster formation

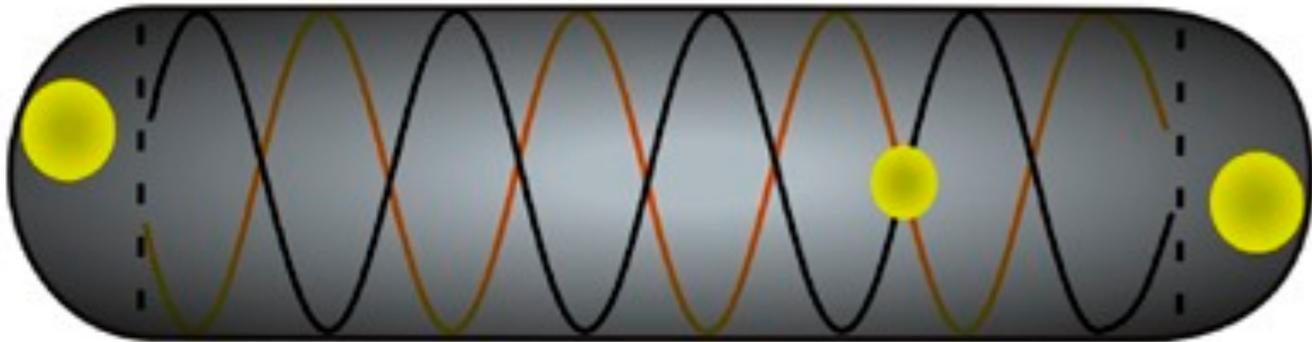
E. coli chemotaxis network

- The bacterial chemotaxis network directs movement toward or away from chemicals.
- Extremely well studied model system.
- Five types of transmembrane receptors form trimers of dimers.
- These cluster in large complexes that transduce signals to flagellar motors.



How do chemotaxis receptors get to specific locations?

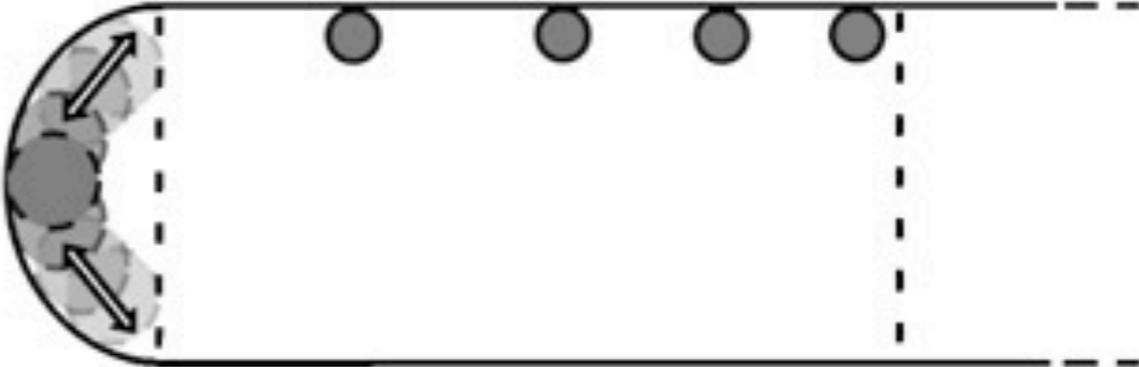
Active



Thiem, Kentner, Sourjik, *EMBO*, **26** (2007)

- Regulation of positions and sizes of clusters
- Positioned relative to unknown cellular structure
- Most receptors at specific sites

Passive

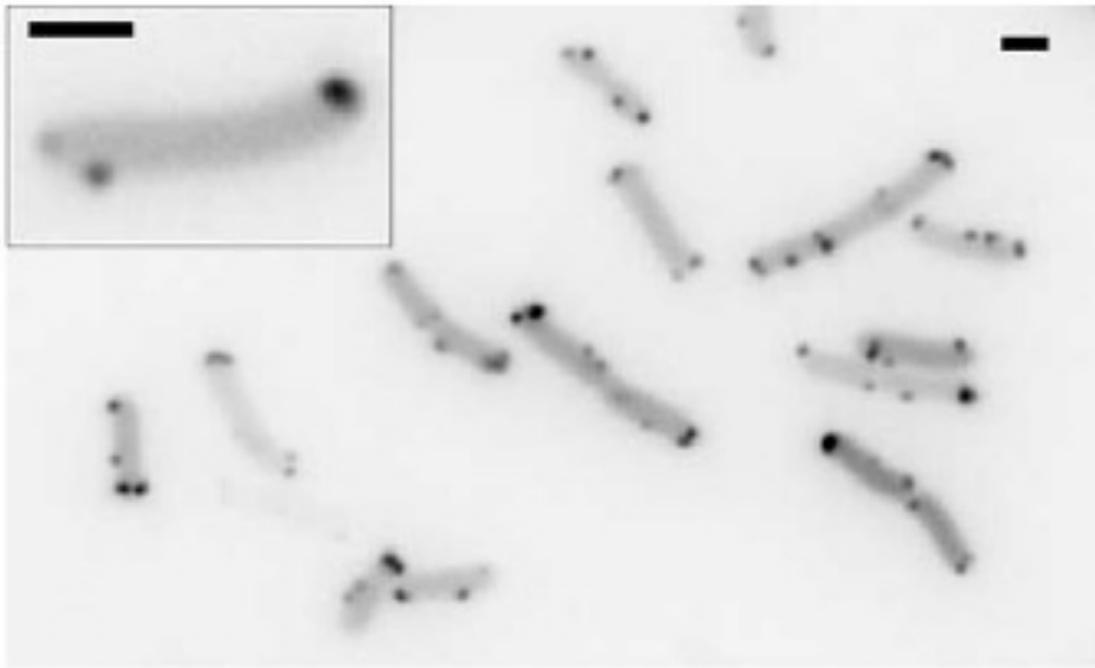


Shiomi, Yoshimoto, Homma, Kawagishi, *Mol Microbio*, **60** (2006)

- Diffusion and capture
- No active transport
- Receptors throughout membrane

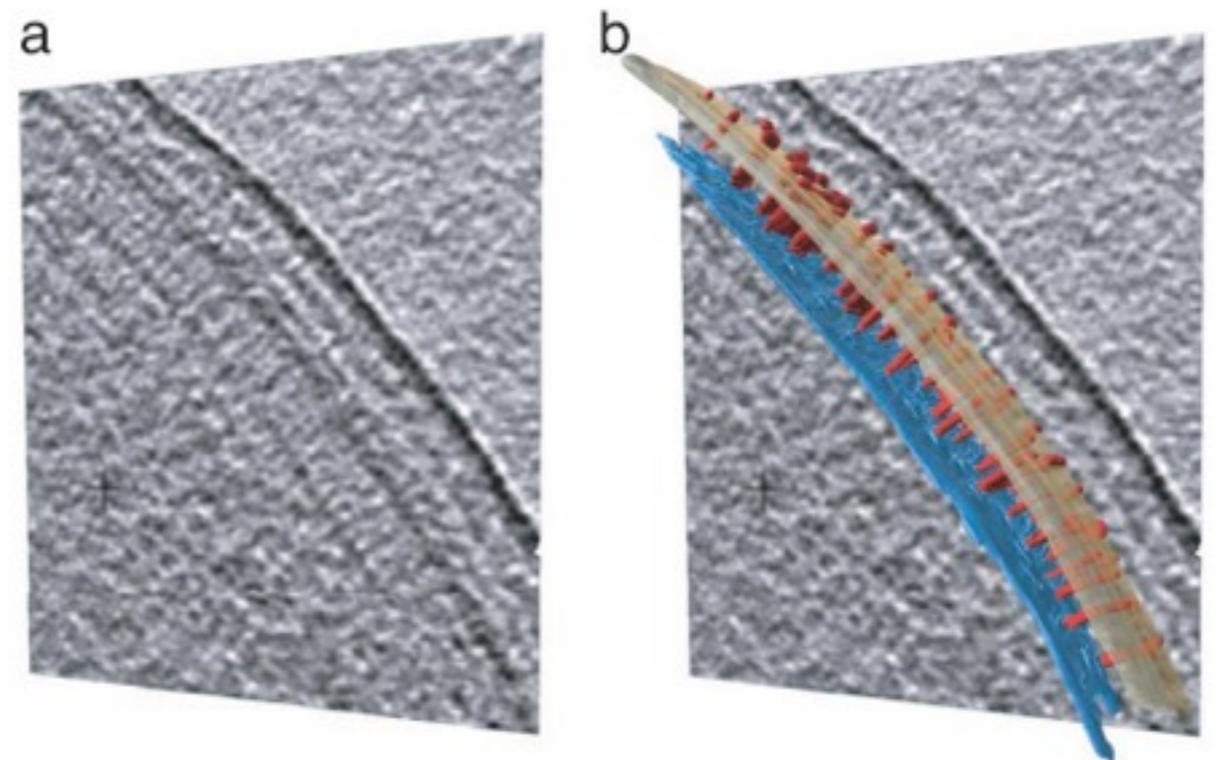
Answer by looking...

Fluorescent Microscopy
Excellent specificity but limited resolution



YFP-CheR
Thiem, Kentner, Sourjik, *EMBO*, **26** (2007)

Cryo-EM Tomography
High resolution but little specificity



Zhang, et al., *PNAS*, **104** (2007)

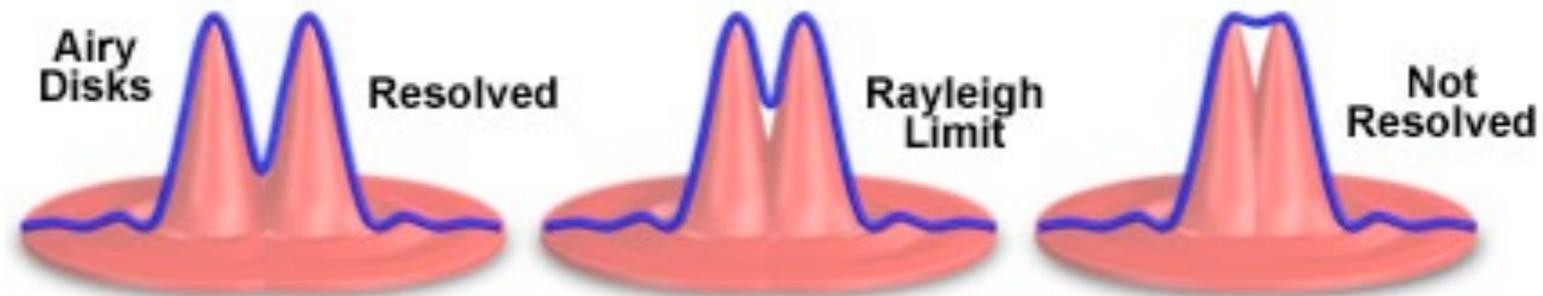
E. coli chemotaxis network

- (How do the parts work?)
- How do chemotaxis clusters form? What controls cluster size and density?
- How is the cellular location of clusters robustly maintained in growing and dividing cells?
- How does spatial organization of receptors influence signal processing?
- How does the network function?

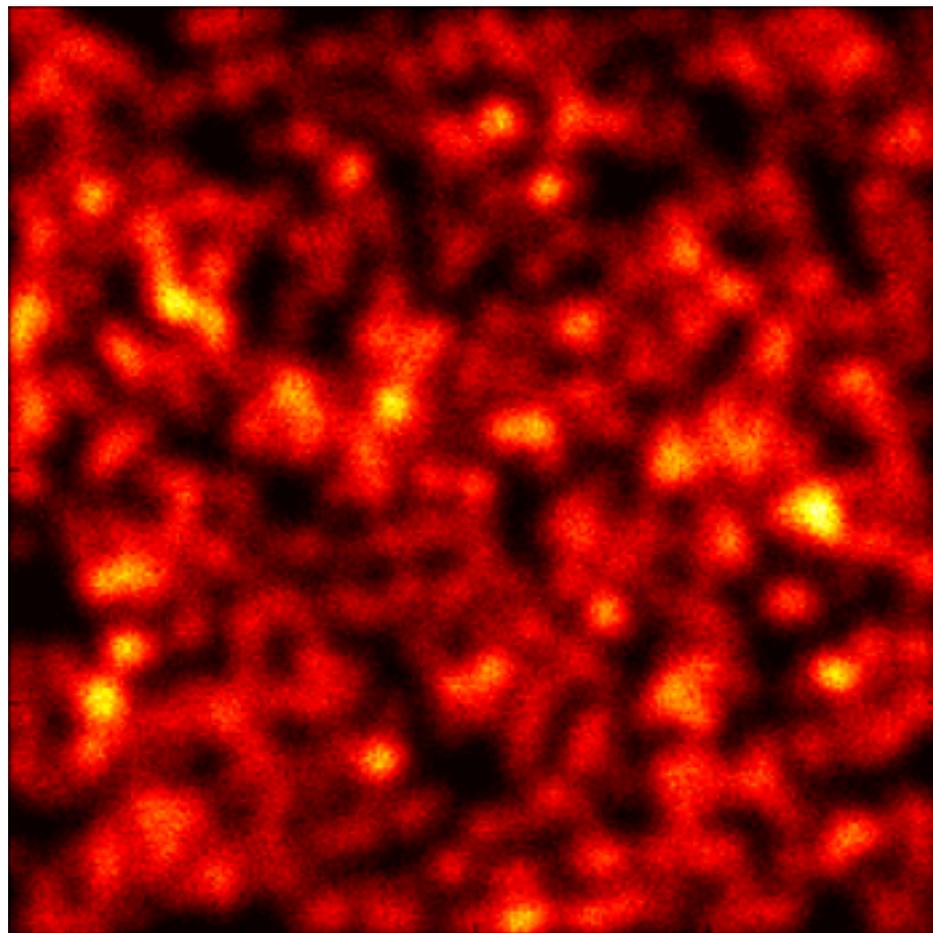
Idea: Many of these questions can be answered by counting and localization

Diffraction Limited Imaging

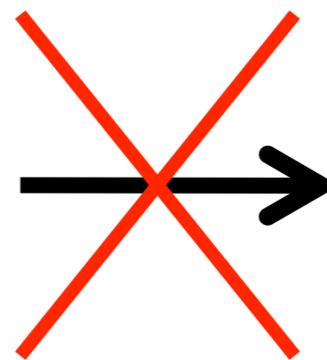
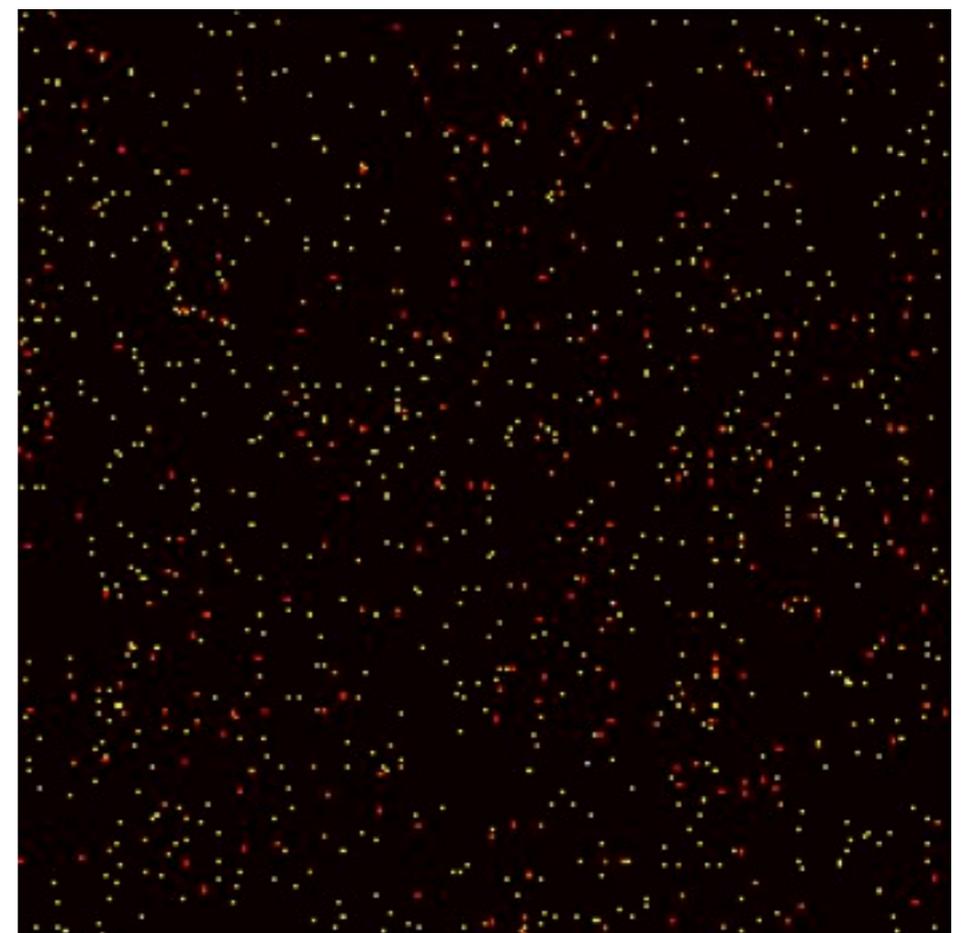
Airy Disk Separation and the Rayleigh Criterion



simulated diffraction limited image

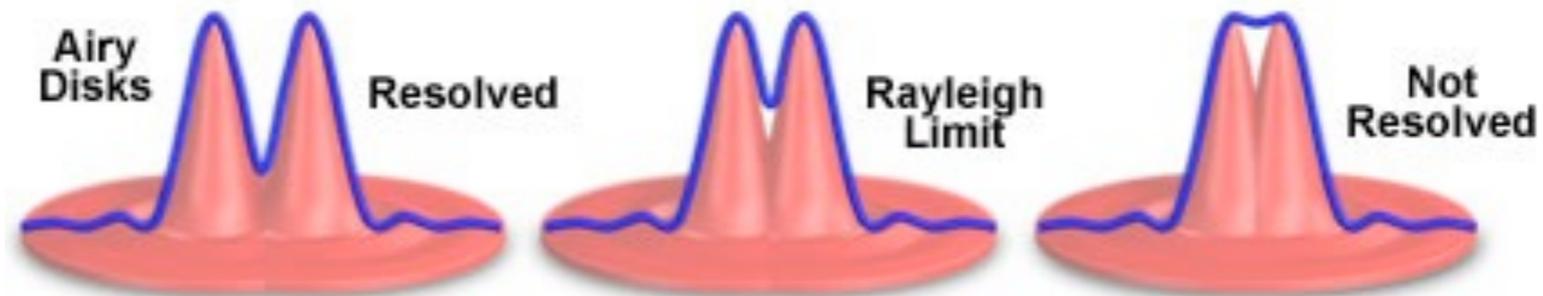


simulated molecules at a surface

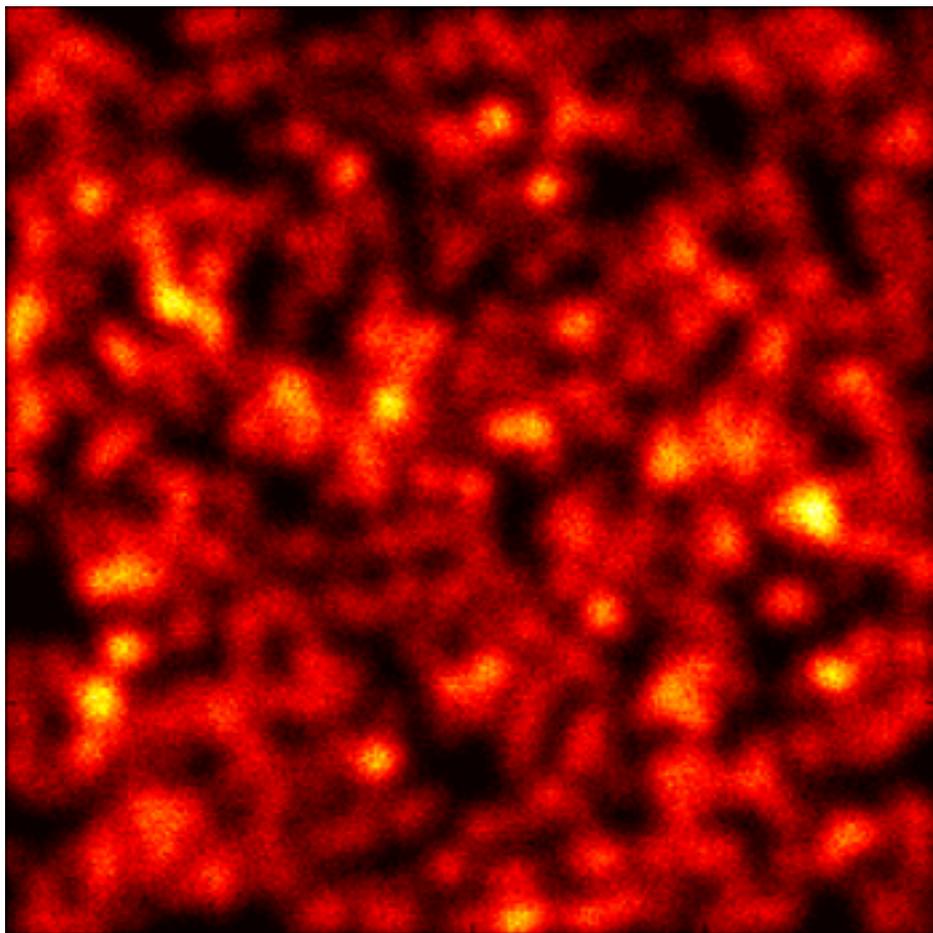


Diffraction Limited Imaging

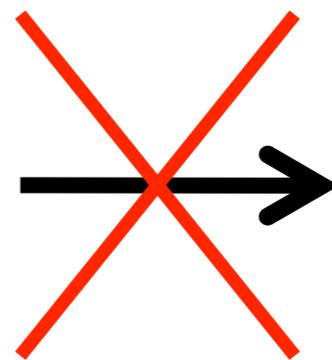
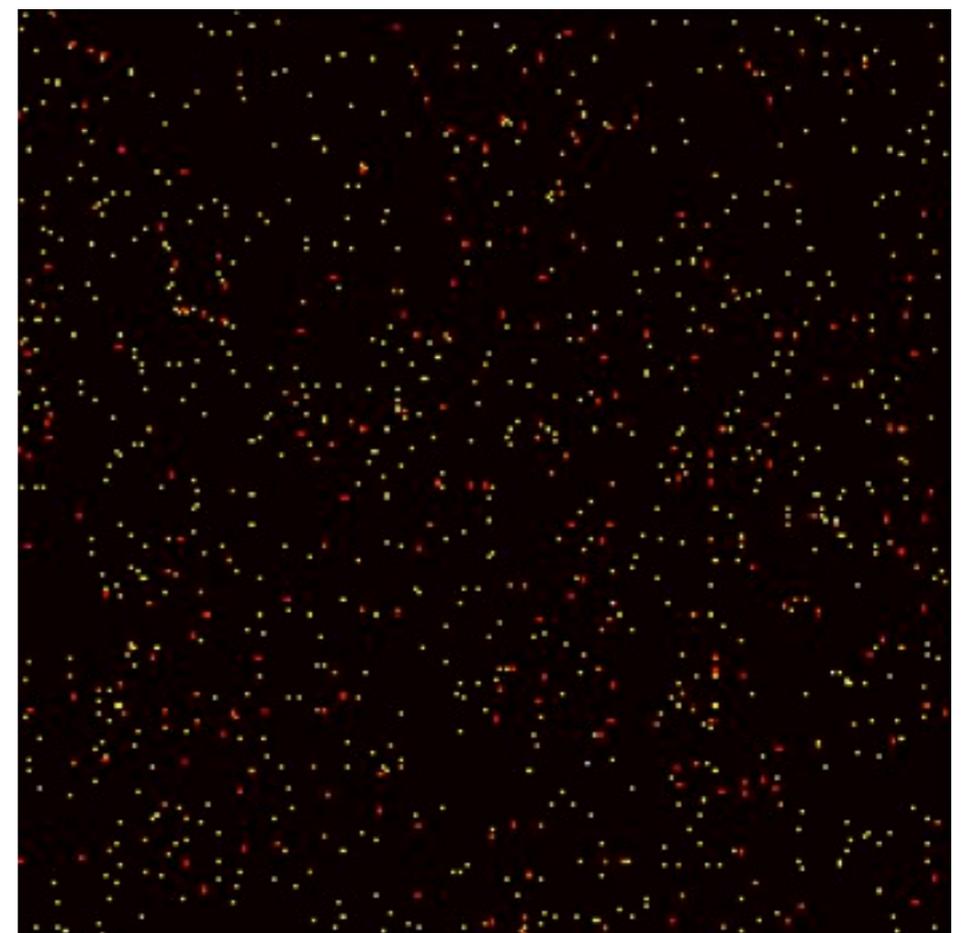
Airy Disk Separation and the Rayleigh Criterion



simulated diffraction limited image



simulated molecules at a surface



Solution: image one molecule at a time

E. Betzig, *et al.*, *Science* 313, 1642 (2006)

Proposed method for molecular optical imaging

E. Betzig

NSOM Enterprises, 17 Webster Drive, Berkeley Heights, New Jersey 07922

Received September 20, 1994

We can resolve multiple discrete features within a focal region of m spatial dimensions by first isolating each on the basis of $n \geq 1$ unique optical characteristics and then measuring their relative spatial coordinates. The minimum acceptable separation between features depends on the point-spread function in the $(m + n)$ -dimensional space formed by the spatial coordinates and the optical parameters, whereas the absolute spatial resolution is determined by the accuracy to which the coordinates can be measured. Estimates of each suggest that near-field fluorescence excitation microscopy/spectroscopy with molecular sensitivity and spatial resolution is possible.

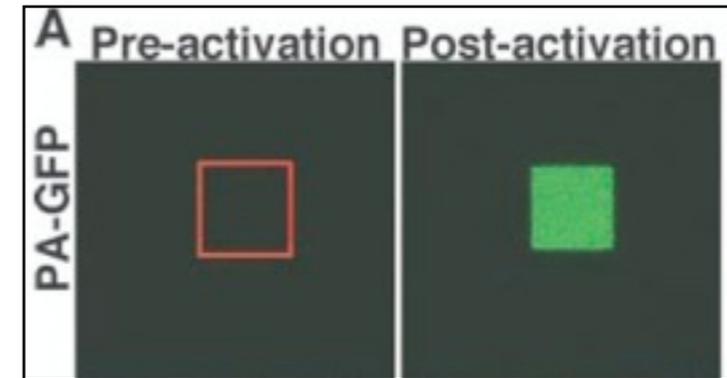
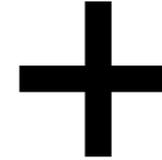
Proposed method for molecular optical imaging

E. Betzig

NSOM Enterprises, 17 Webster Drive, Berkeley Heights, New Jersey 07922

Received September 20, 1994

We can resolve multiple discrete features within a focal region of m spatial dimensions by first isolating each on the basis of $n \geq 1$ unique optical characteristics and then measuring their relative spatial coordinates. The minimum acceptable separation between features depends on the point-spread function in the $(m + n)$ -dimensional space formed by the spatial coordinates and the optical parameters, whereas the absolute spatial resolution is determined by the accuracy to which the coordinates can be measured. Estimates of each suggest that near-field fluorescence excitation microscopy/spectroscopy with molecular sensitivity and spatial resolution is possible.



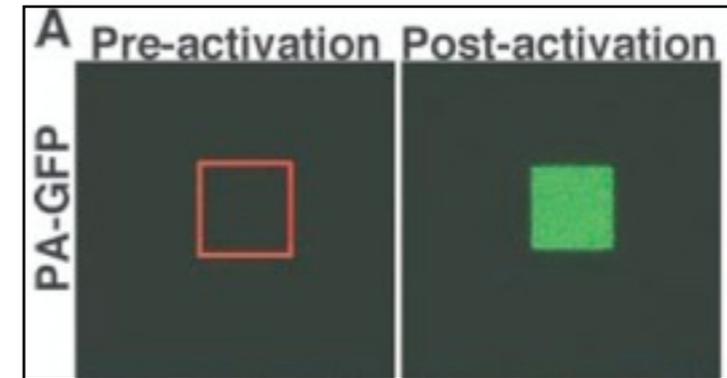
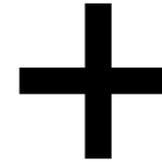
Proposed method for molecular optical imaging

E. Betzig

NSOM Enterprises, 17 Webster Drive, Berkeley Heights, New Jersey 07922

Received September 20, 1994

We can resolve multiple discrete features within a focal region of m spatial dimensions by first isolating each on the basis of $n \geq 1$ unique optical characteristics and then measuring their relative spatial coordinates. The minimum acceptable separation between features depends on the point-spread function in the $(m + n)$ -dimensional space formed by the spatial coordinates and the optical parameters, whereas the absolute spatial resolution is determined by the accuracy to which the coordinates can be measured. Estimates of each suggest that near-field fluorescence excitation microscopy/spectroscopy with molecular sensitivity and spatial resolution is possible.



PALM

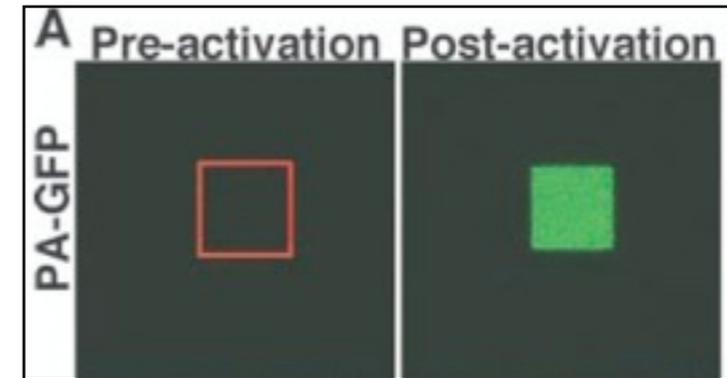
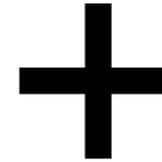
Proposed method for molecular optical imaging

E. Betzig

NSOM Enterprises, 17 Webster Drive, Berkeley Heights, New Jersey 07922

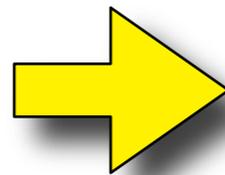
Received September 20, 1994

We can resolve multiple discrete features within a focal region of m spatial dimensions by first isolating each on the basis of $n \geq 1$ unique optical characteristics and then measuring their relative spatial coordinates. The minimum acceptable separation between features depends on the point-spread function in the $(m + n)$ -dimensional space formed by the spatial coordinates and the optical parameters, whereas the absolute spatial resolution is determined by the accuracy to which the coordinates can be measured. Estimates of each suggest that near-field fluorescence excitation microscopy/spectroscopy with molecular sensitivity and spatial resolution is possible.

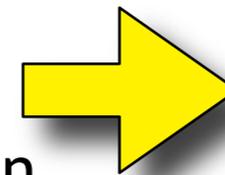


PALM

Protein initially not fluorescent at specific wavelength

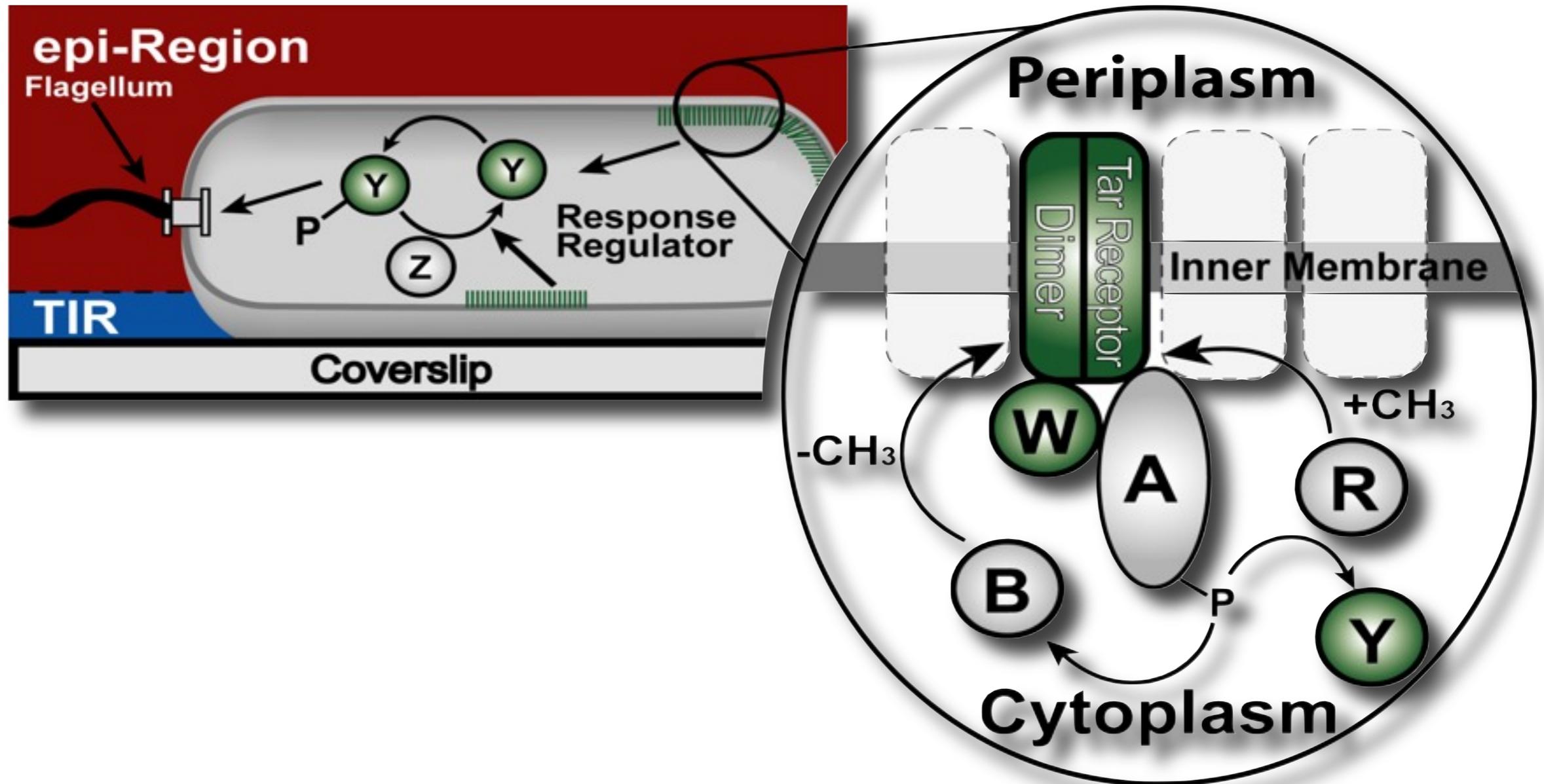


UV light induces a conformation change and "activates" protein



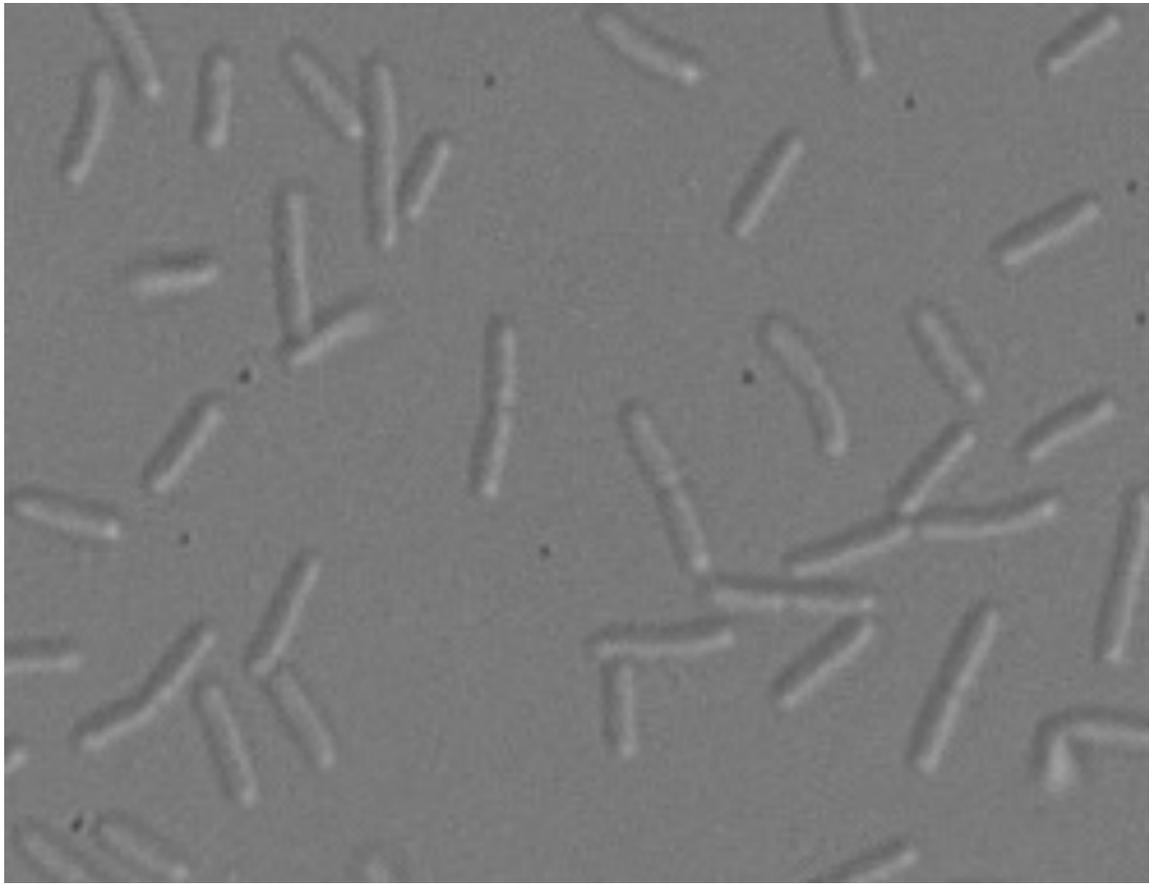
Protein becomes fluorescent and can be observed and bleached

PALMing chemotaxis receptors

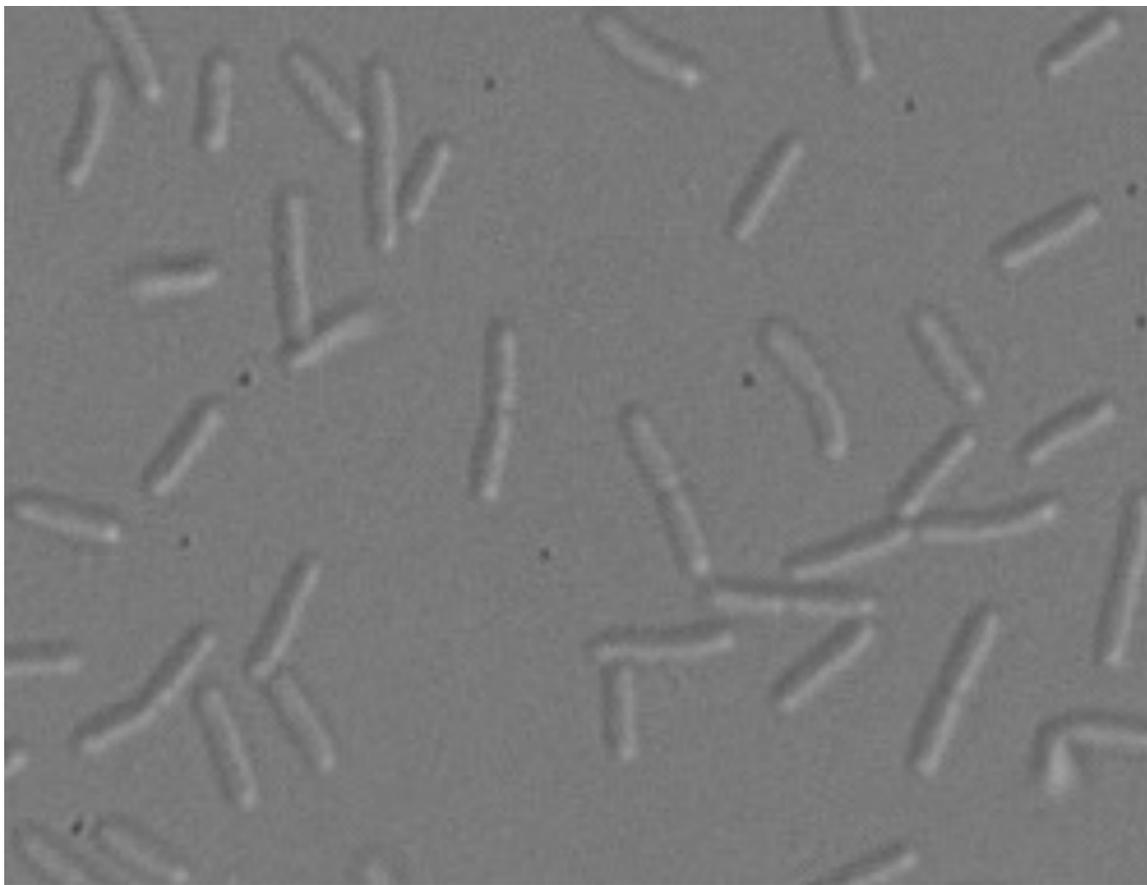


Made fusions to mEos2 (photoactivatable fluorescent protein)

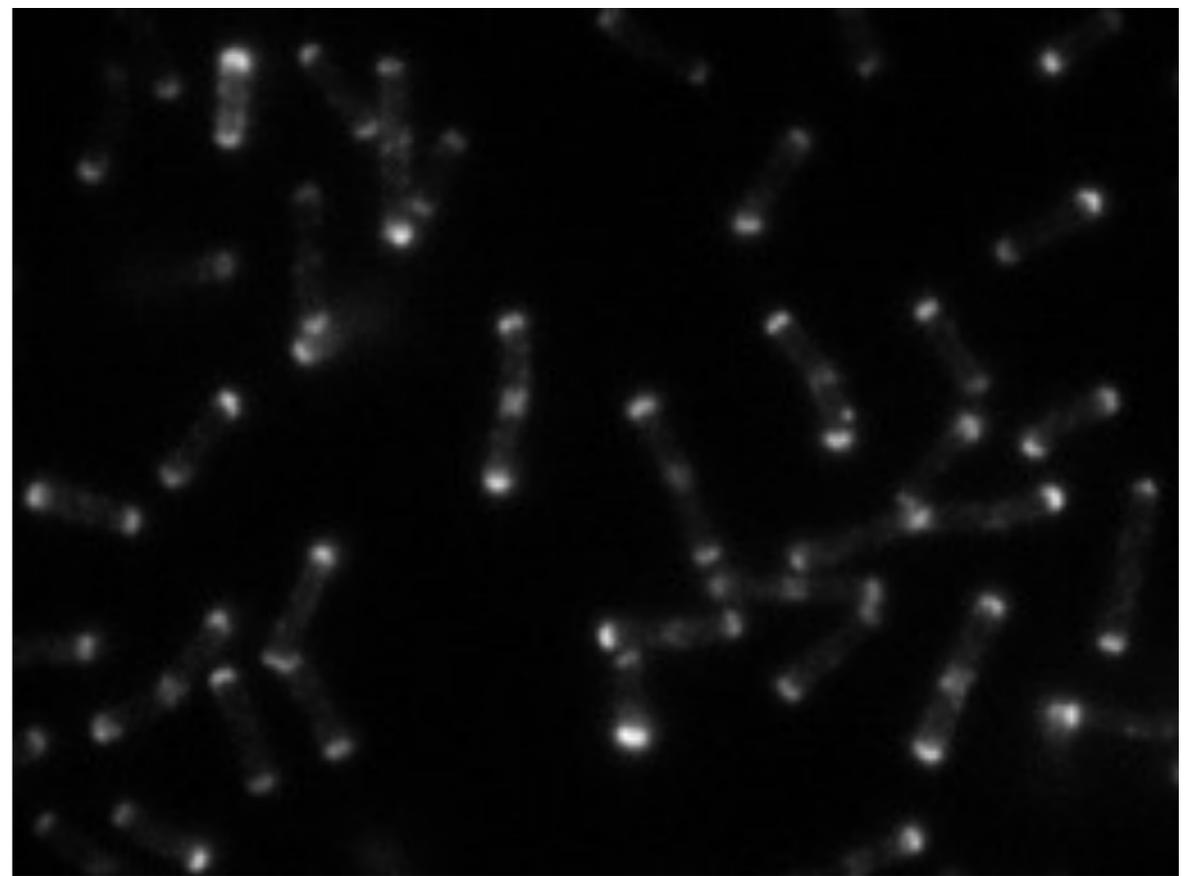
- Fixed whole cells
- Imaged in PALM (with 15 nm precision)
- Only interpretable when several types of proteins are tagged -> internal cross-validation
- most difficult part was...



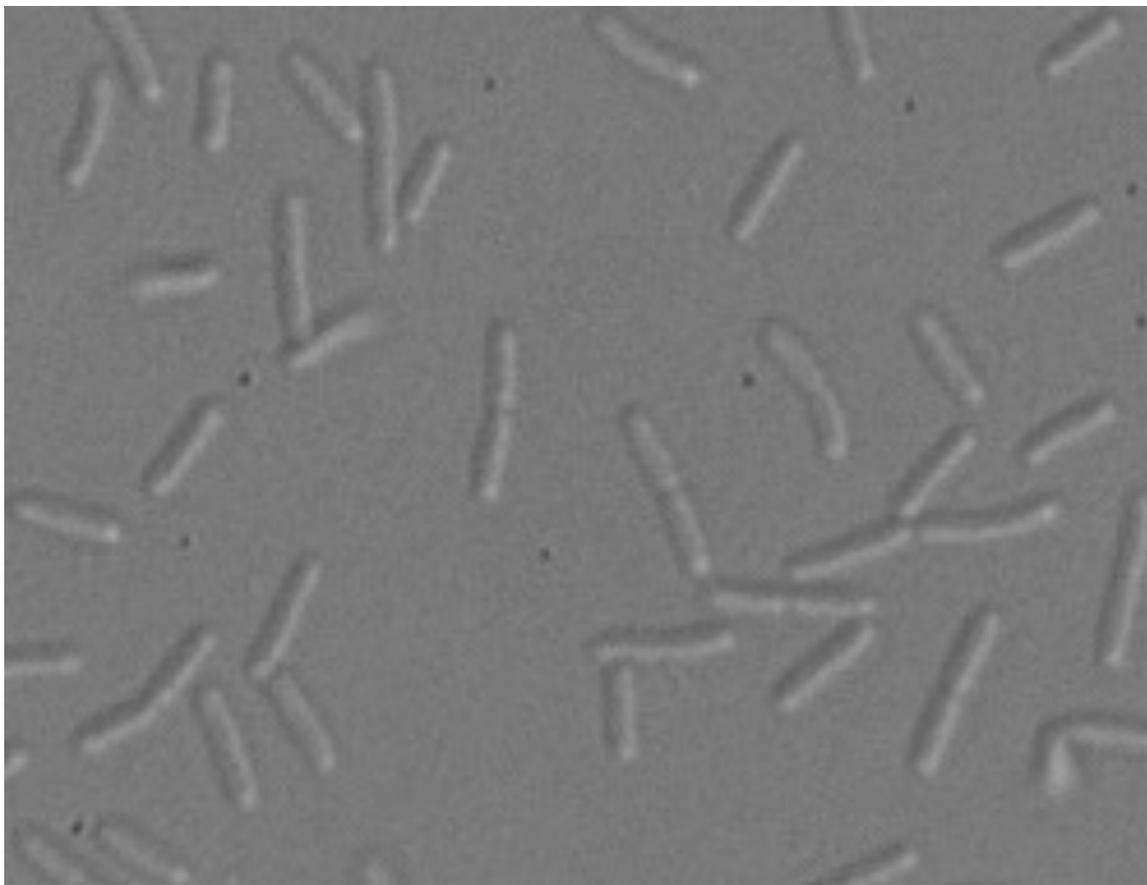
DIC



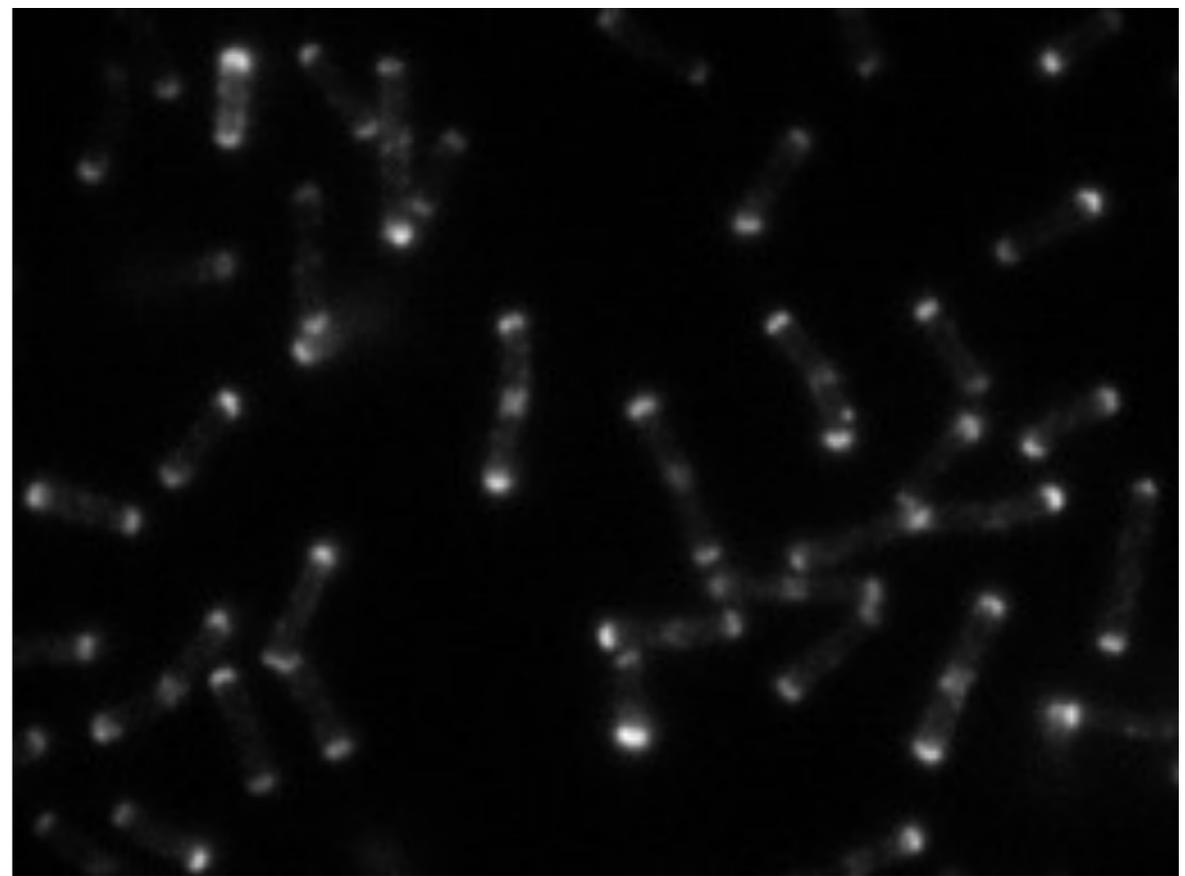
DIC



Conventional Fluorescence



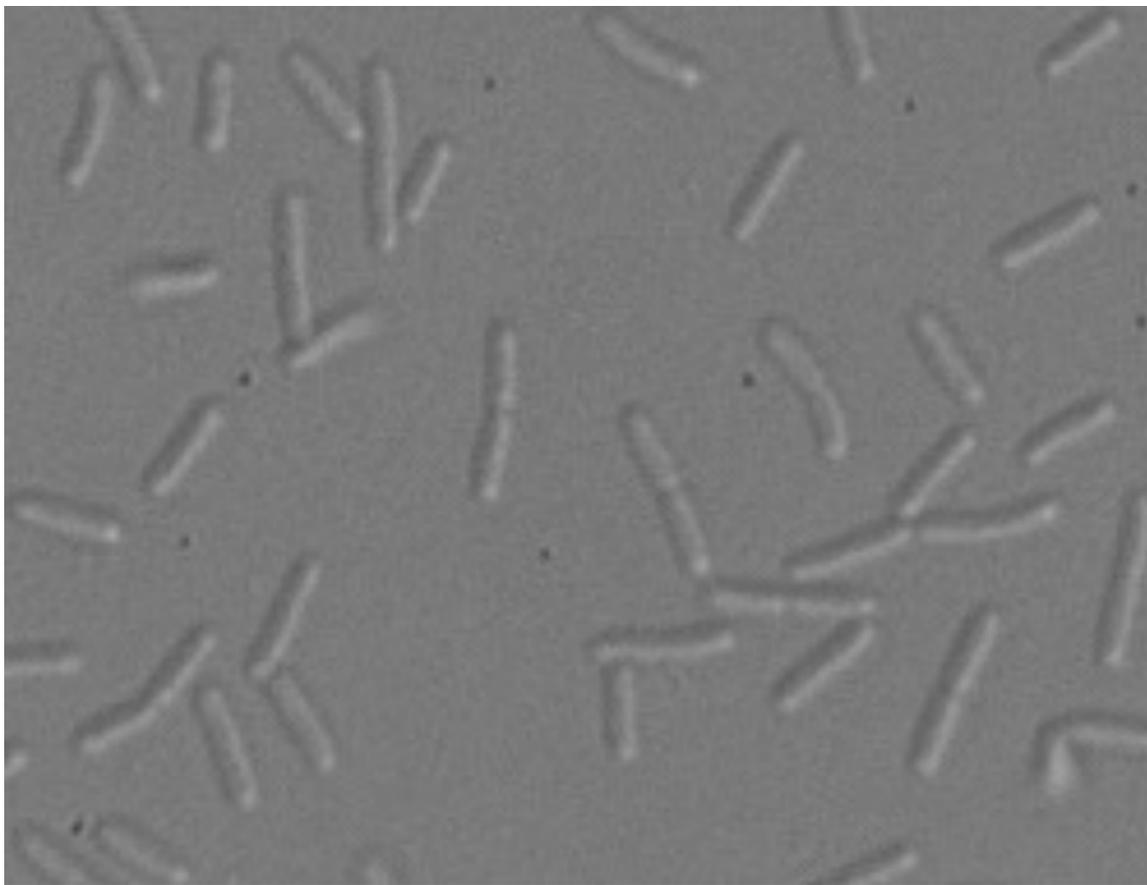
DIC



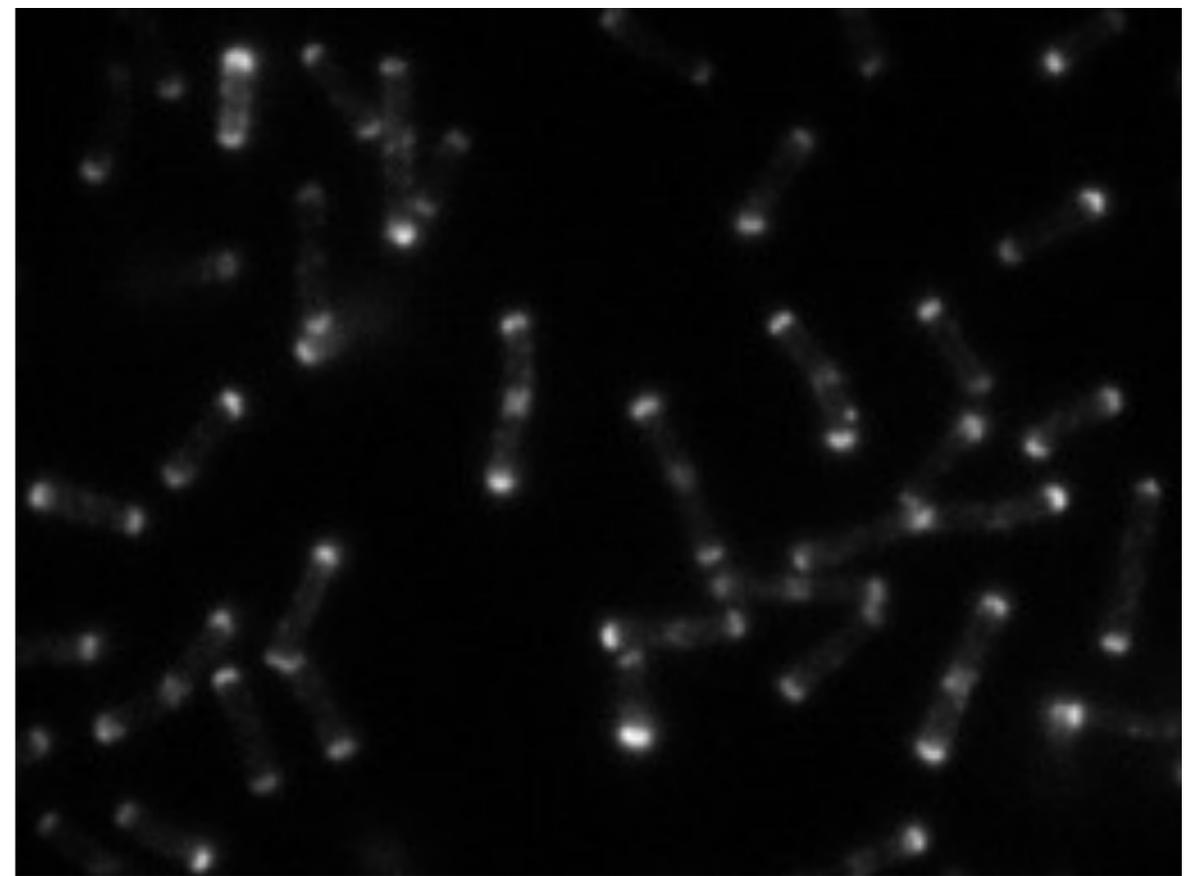
Conventional Fluorescence

PALM Acquisition



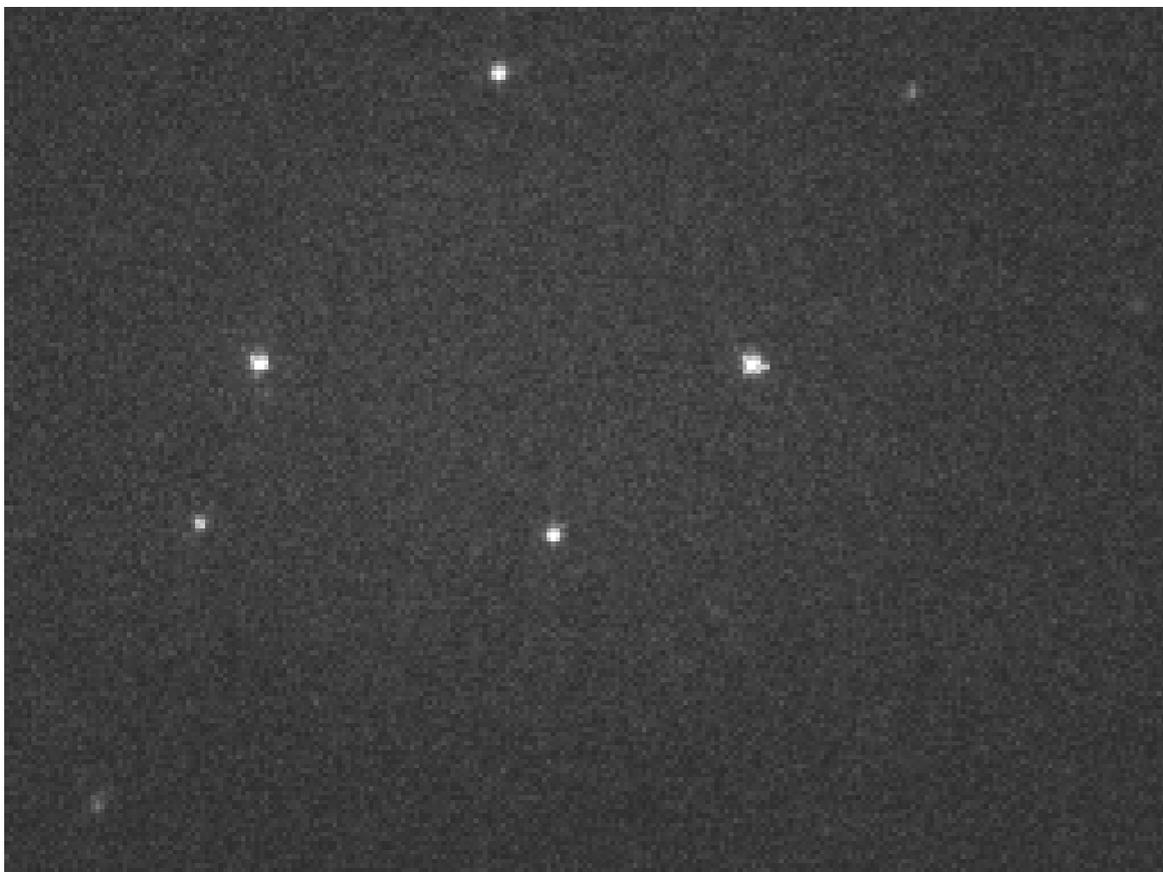


DIC

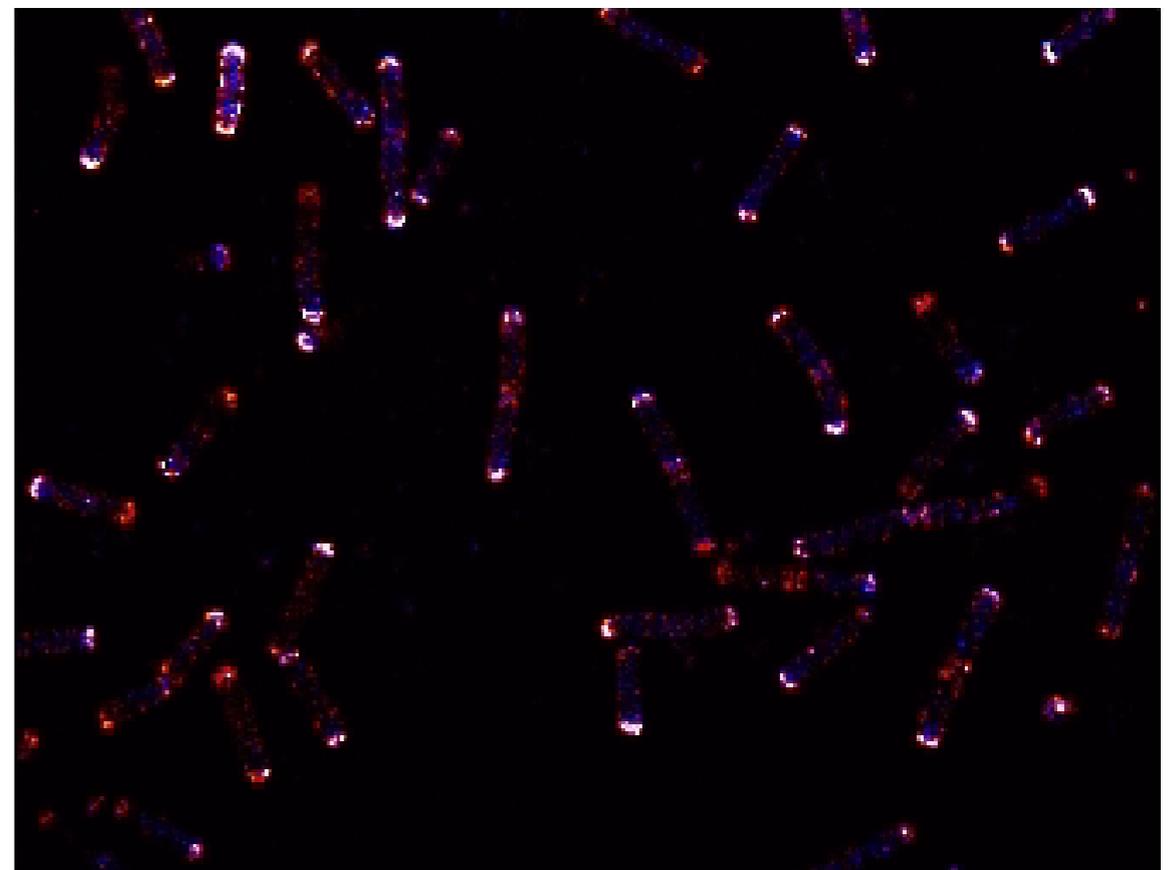


Conventional Fluorescence

PALM Acquisition

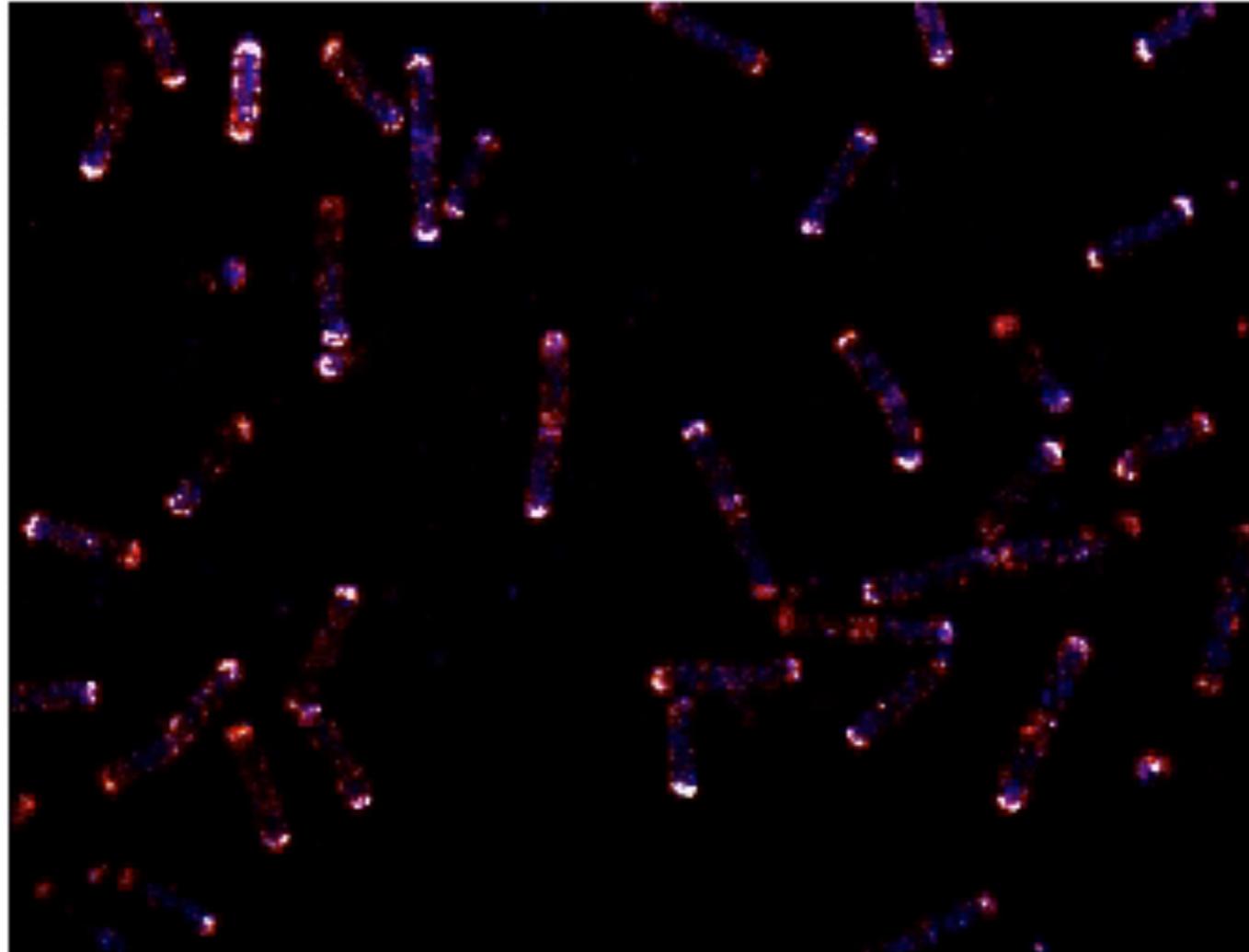


PALM

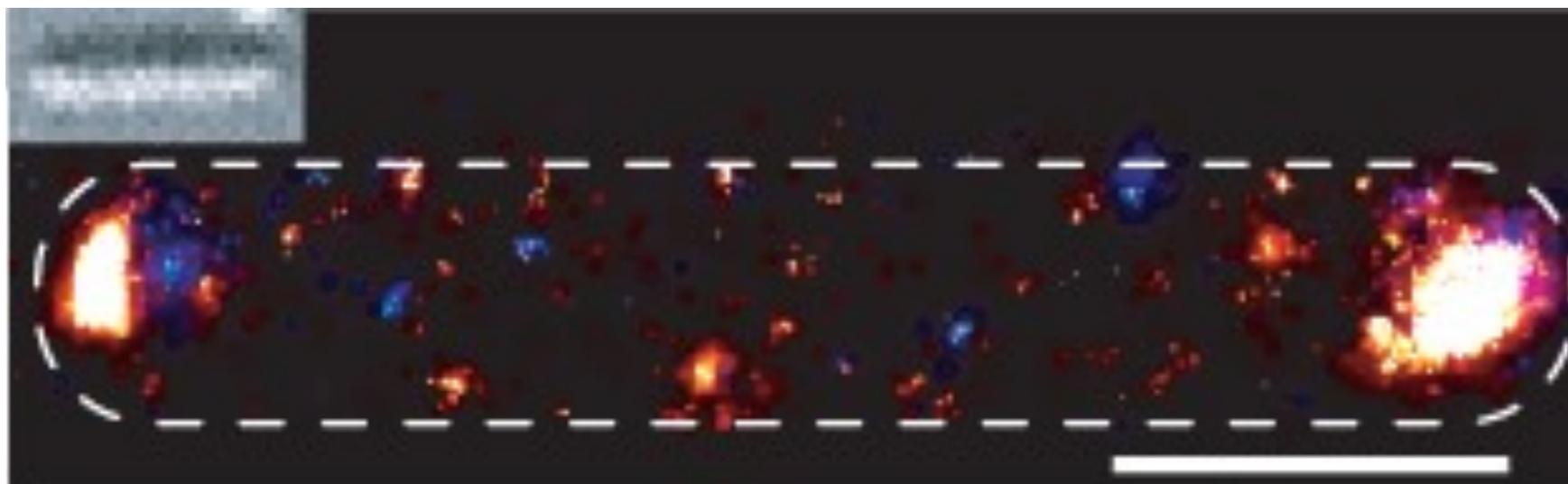


We can image many cells at once

We can image many cells at once

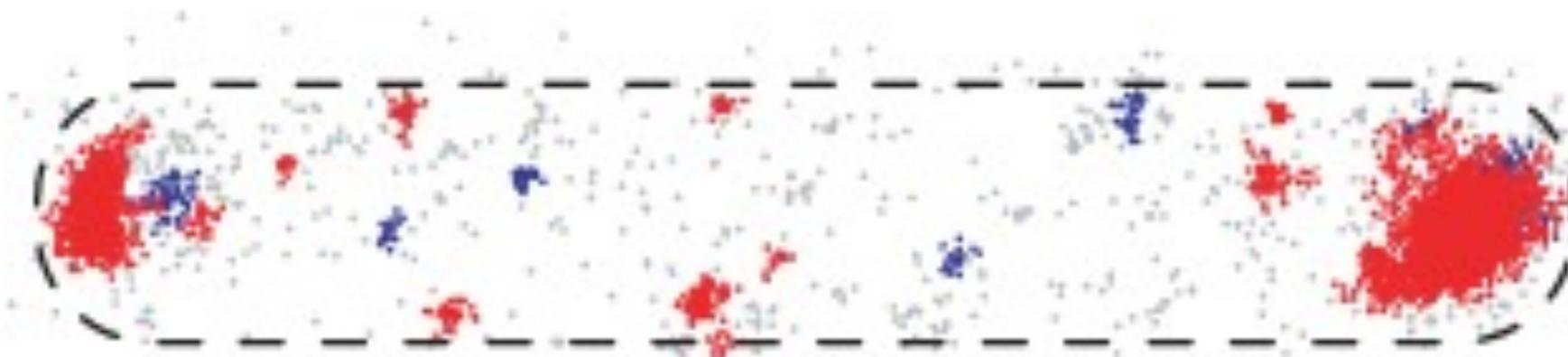


Cluster quantification



1 μm

Closely spaced proteins are part of the same cluster



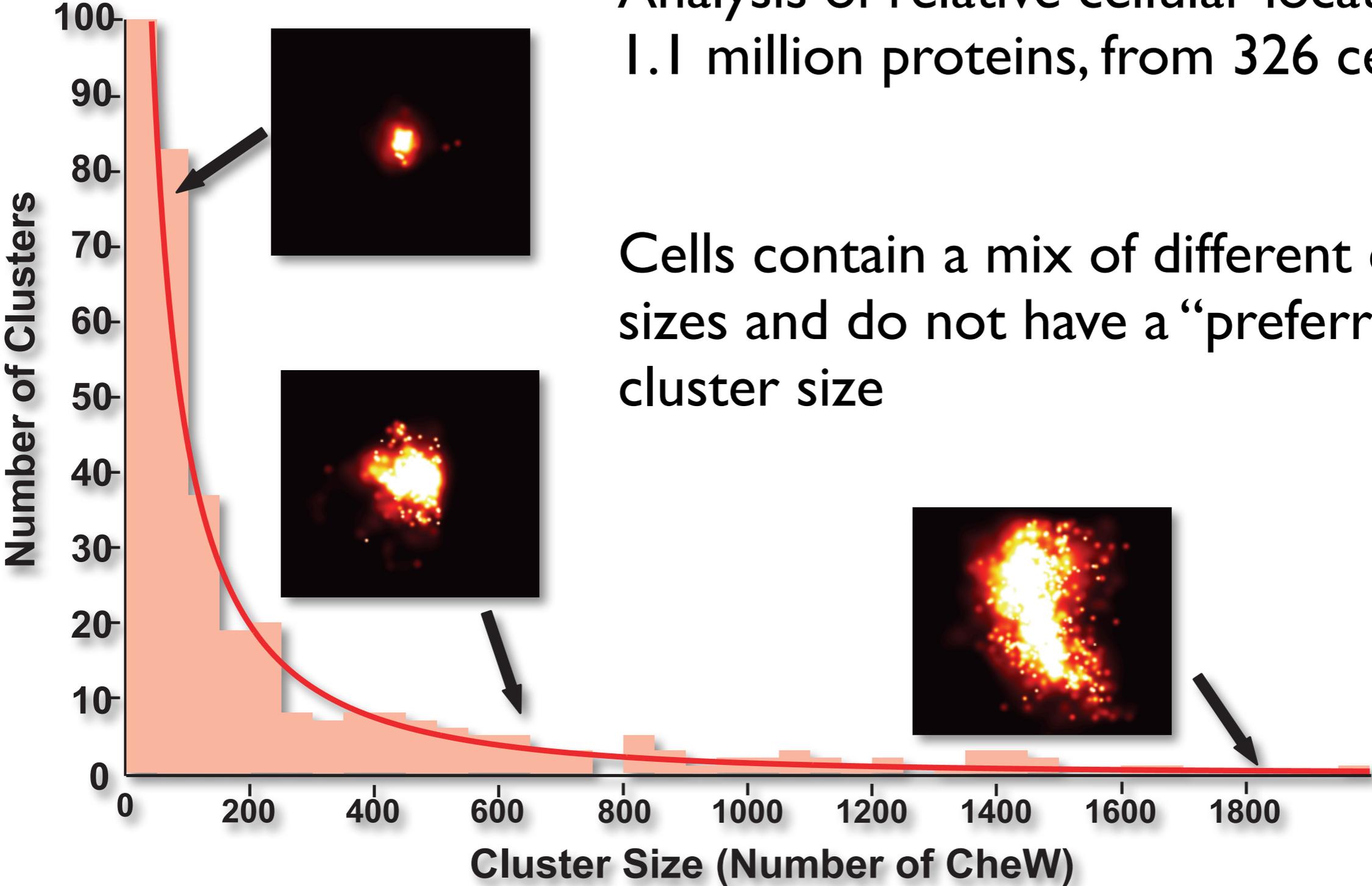
Examine 326 cells to determine the distribution of cluster sizes
(# of proteins per cluster) - 1.1 million proteins

Cluster sizes are _____ distributed

Cluster sizes are _____ distributed

Analysis of relative cellular location of 1.1 million proteins, from 326 cells

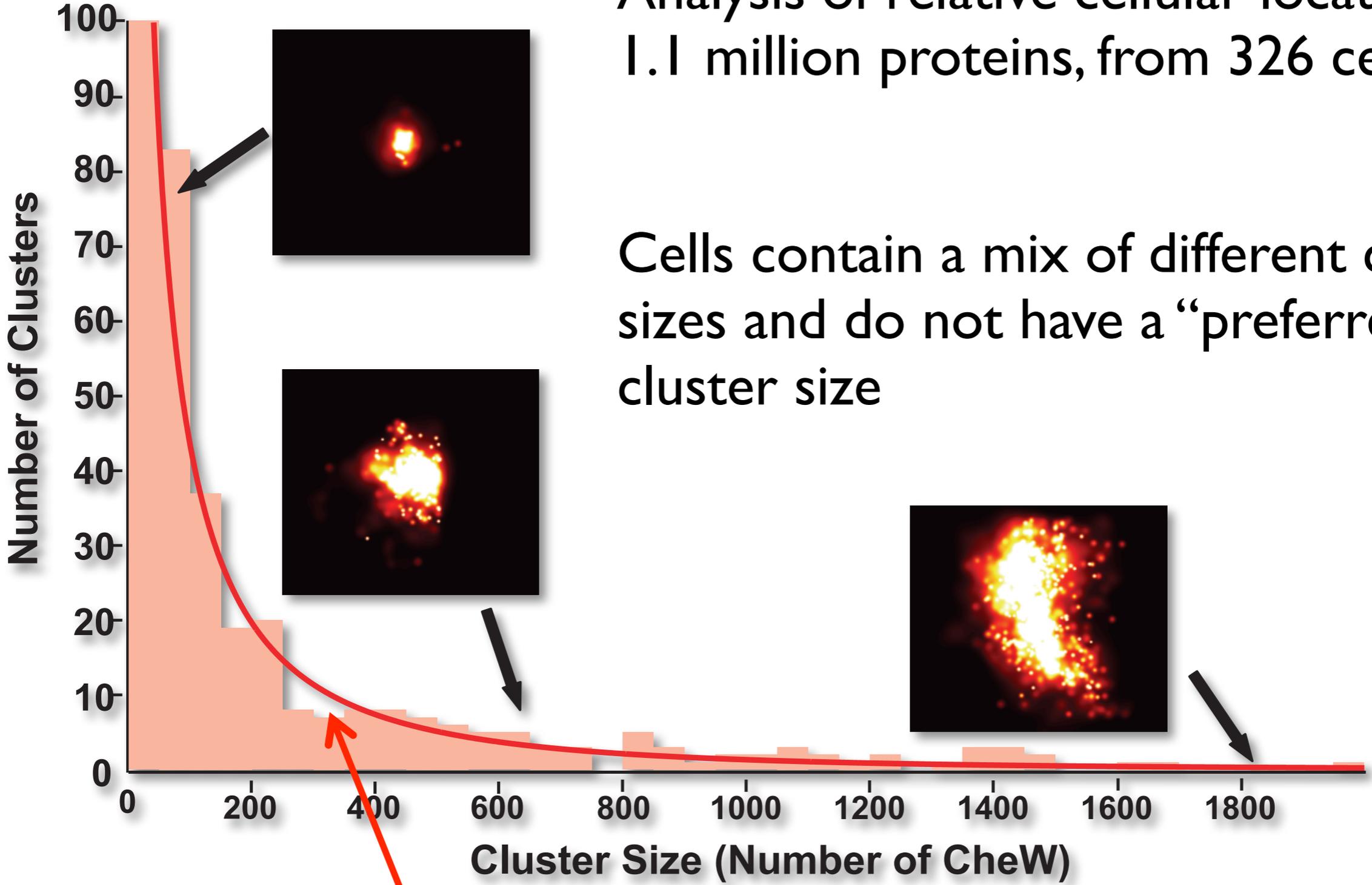
Cells contain a mix of different cluster sizes and do not have a “preferred” cluster size



Cluster sizes are _____ distributed

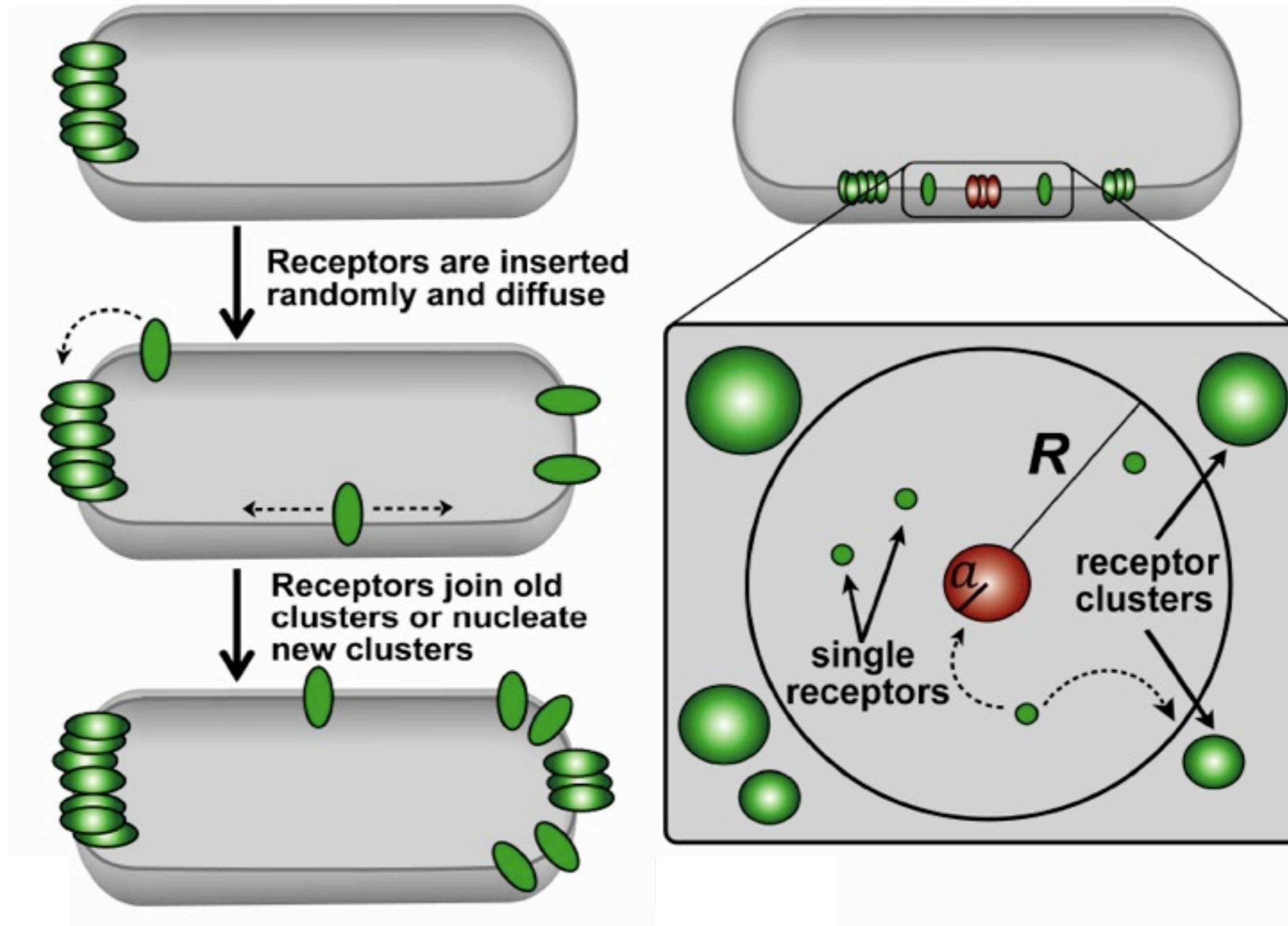
Analysis of relative cellular location of 1.1 million proteins, from 326 cells

Cells contain a mix of different cluster sizes and do not have a “preferred” cluster size



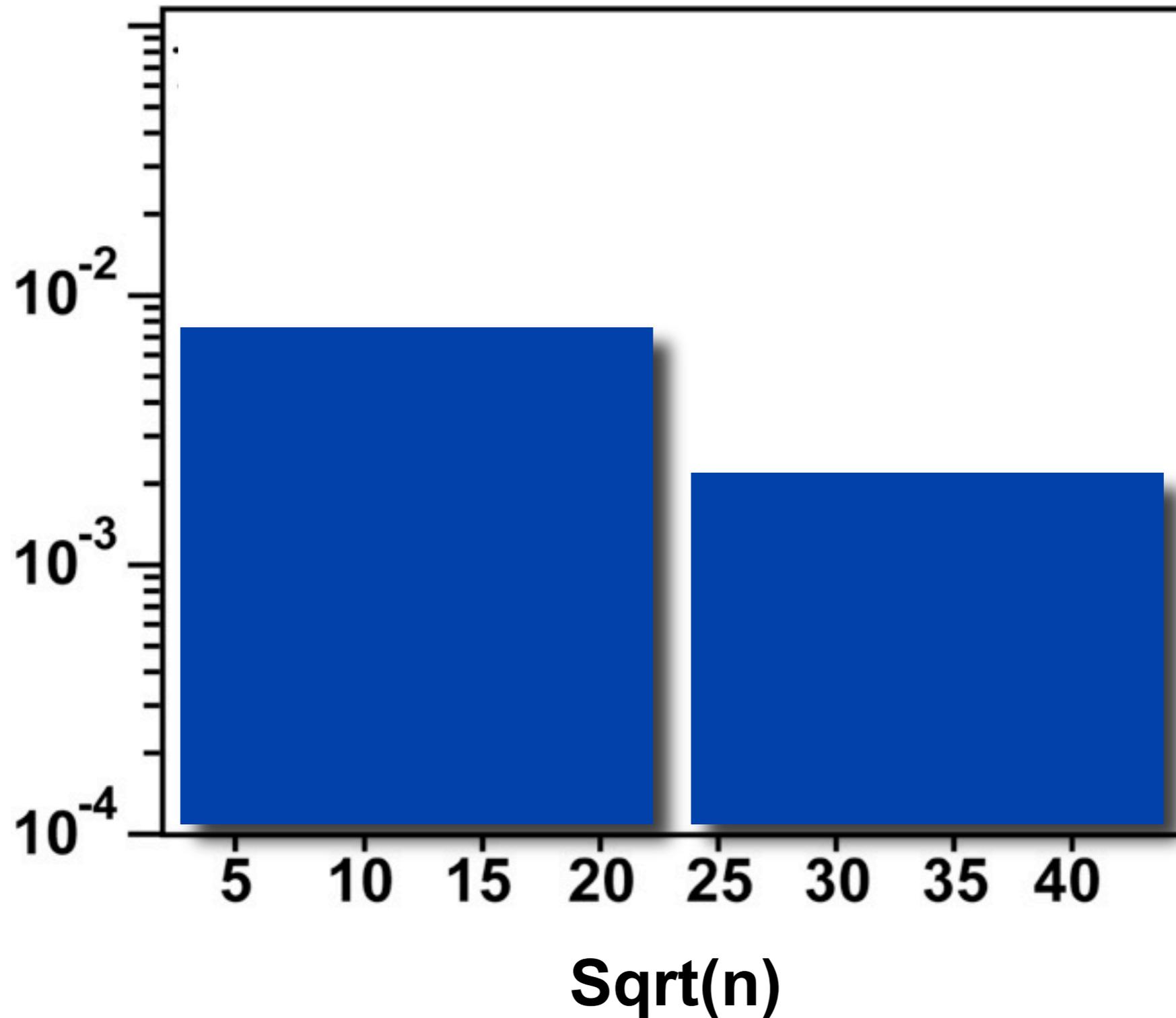
Stochastic Cluster Nucleation Model

Model: stochastic cluster nucleation

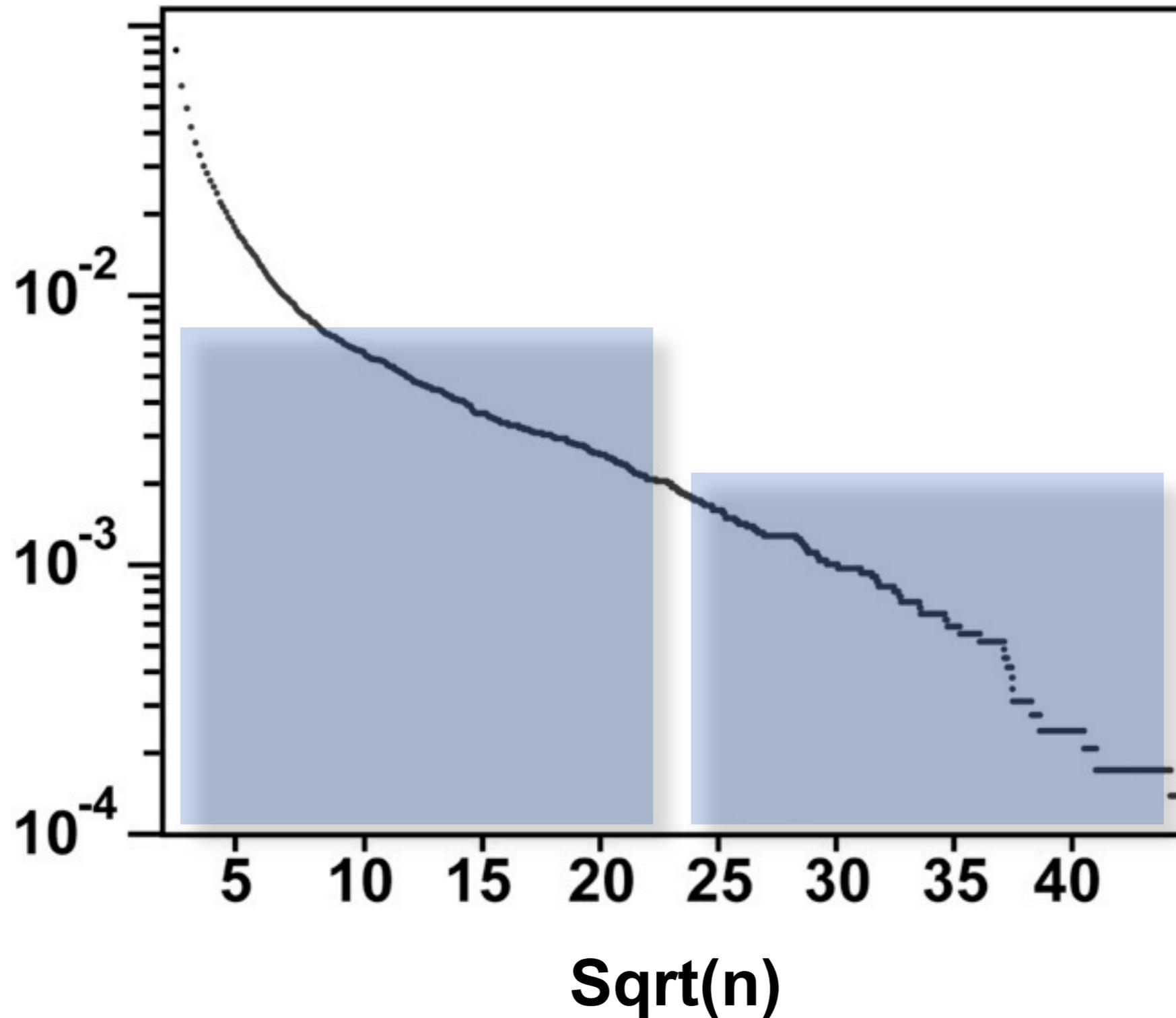


Highest density of single receptors is furthest away from existing cluster(s)

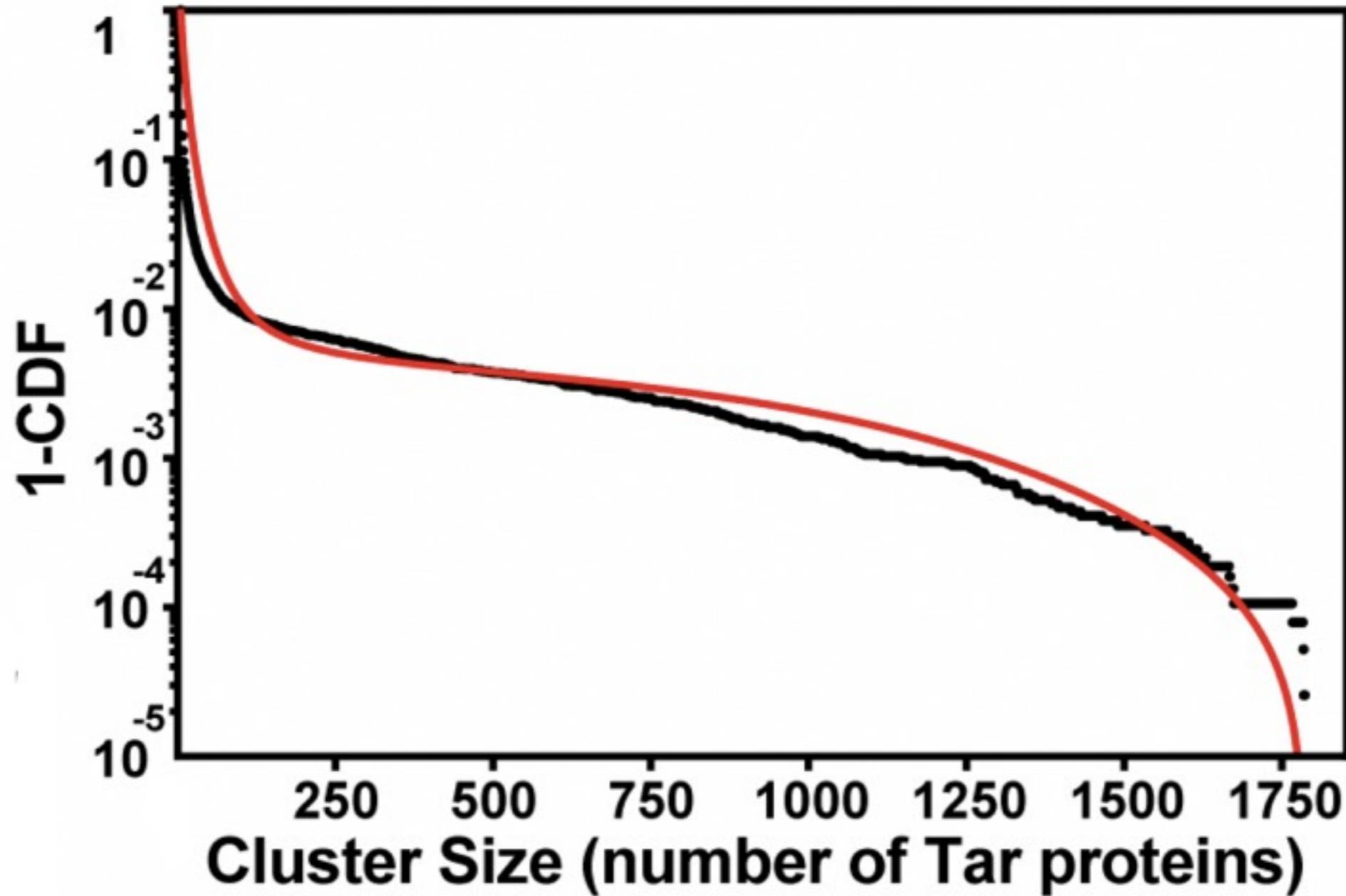
Benefits of being able to count, protein by protein



Benefits of being able to count, protein by protein



Cluster sizes are exponentially distributed

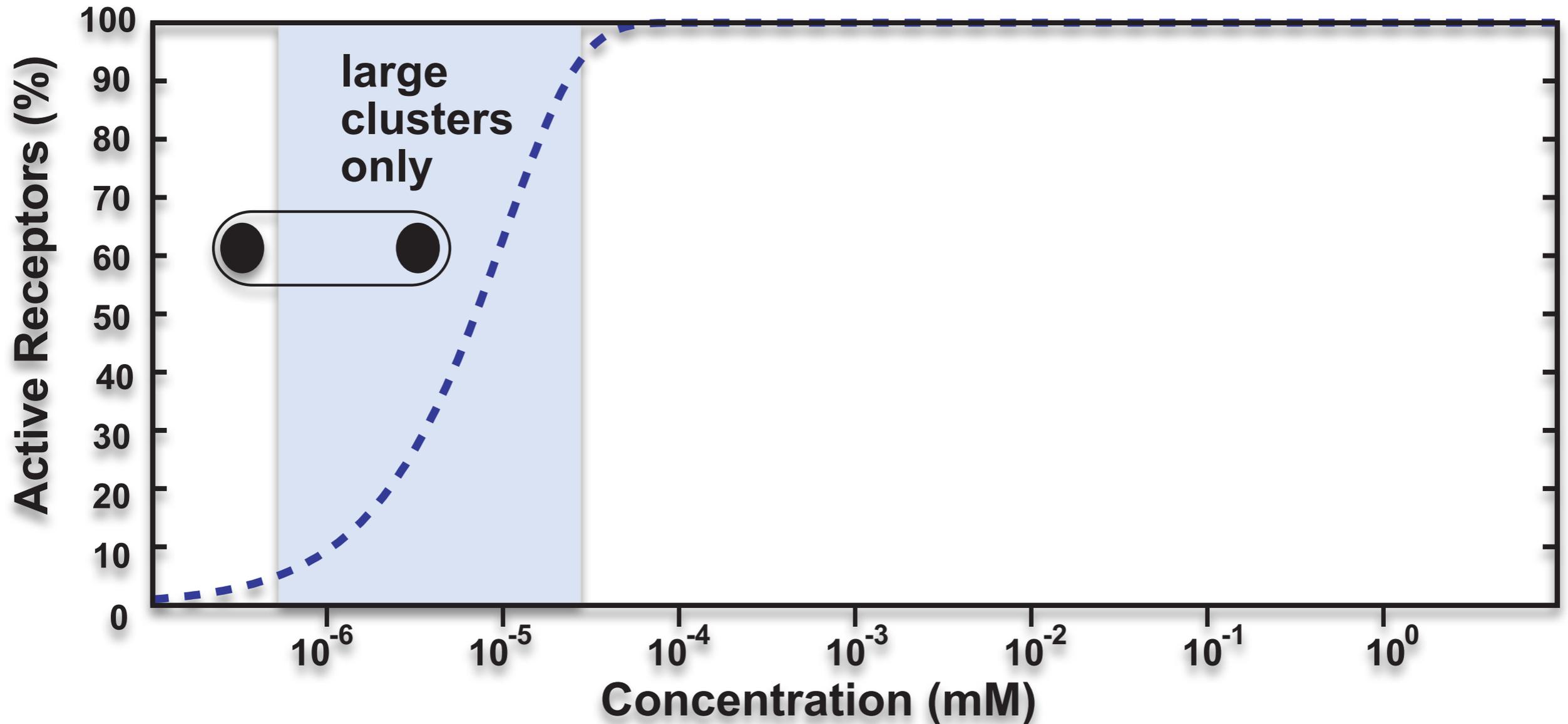


$$P_{tot}(N) \approx c_1 e^{-c_2 N + c_3 N \ln(N) - c_4 N (\ln(N))^2}$$

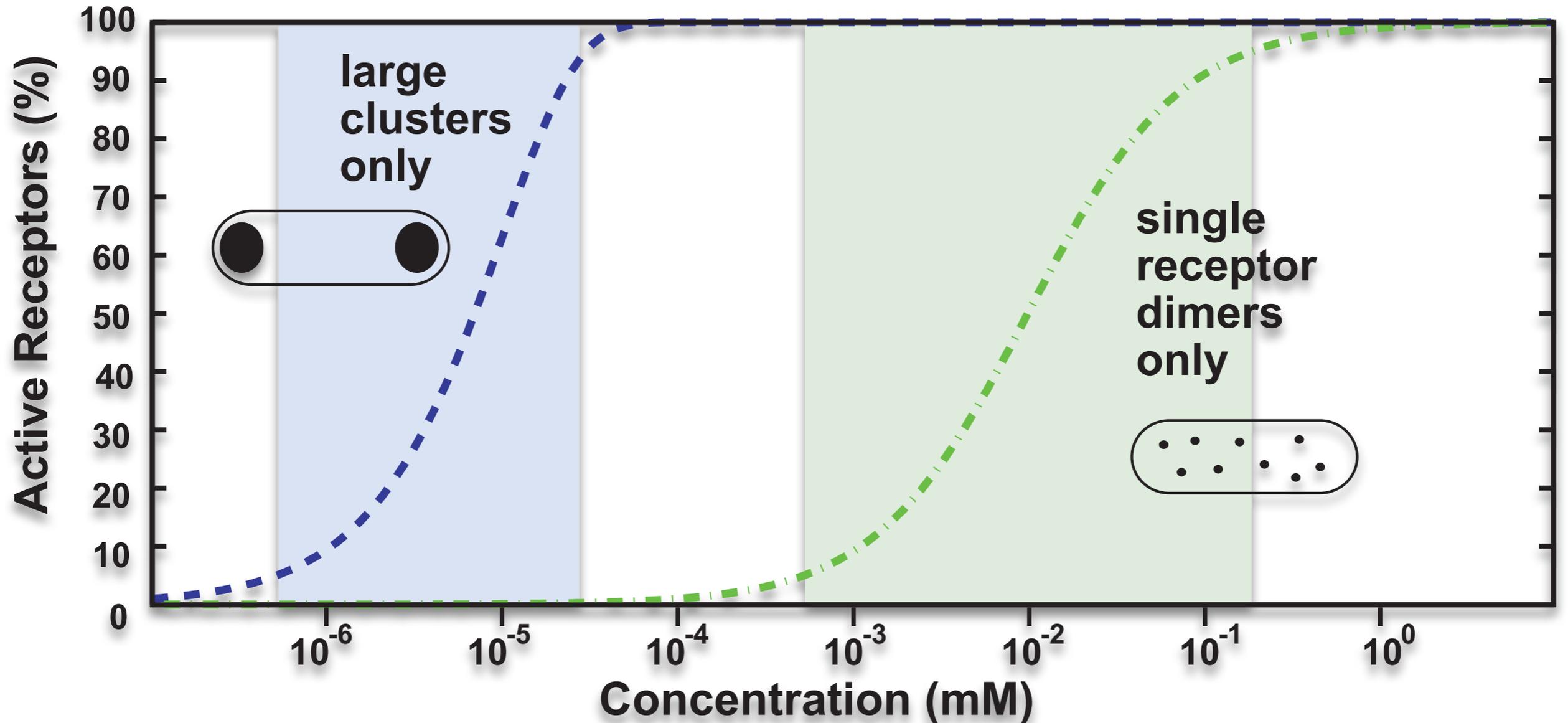
Physical picture

- Imagine you're a particle or astrophysicist, and you have a finite number of identical detectors
- Imagine that those detectors cooperate linearly (snap N of them together, N times the sensitivity, but $1/N$ range)
- Your goal is to build a detector network with high sensitivity and wide dynamic range
- What topology do you use?

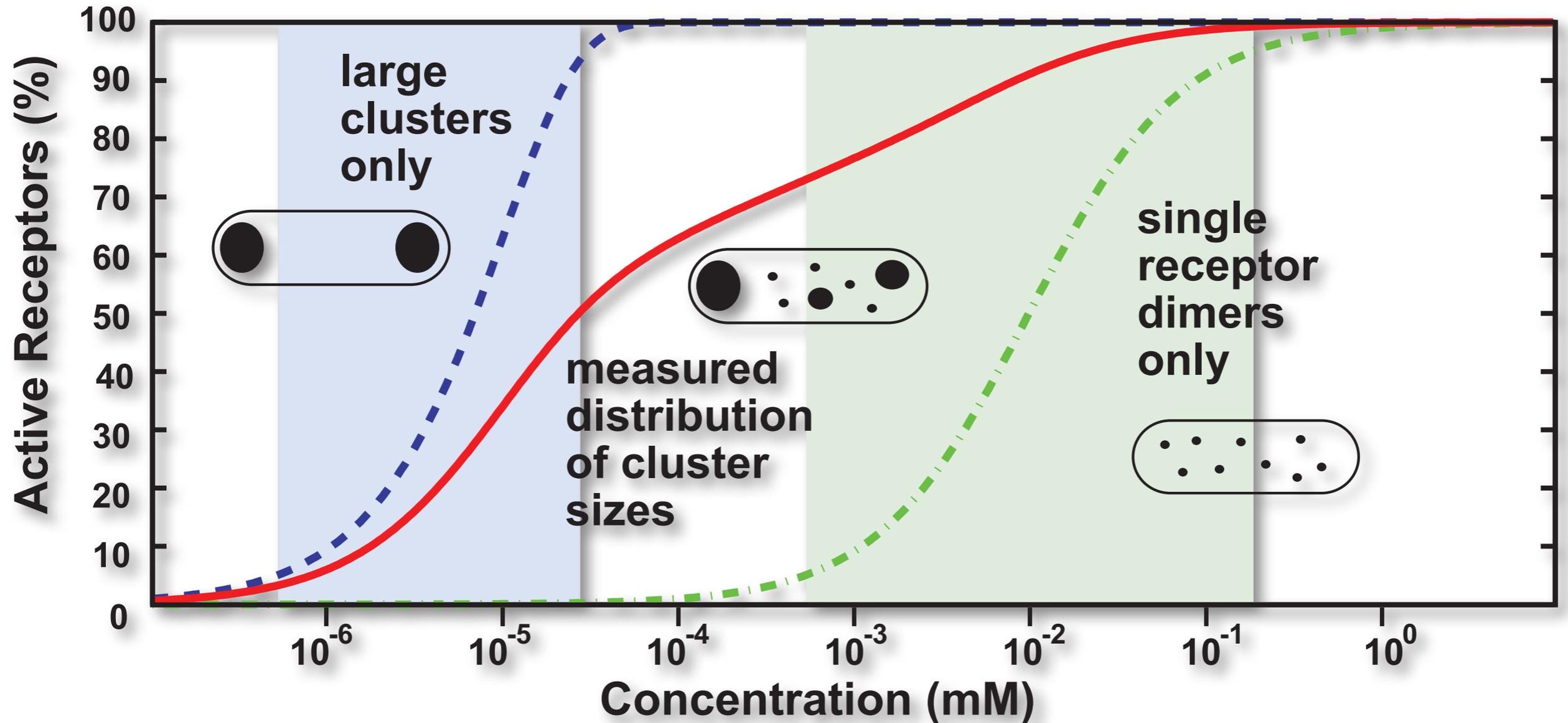
Net filter performance for different mixes



Net filter performance for different mixes

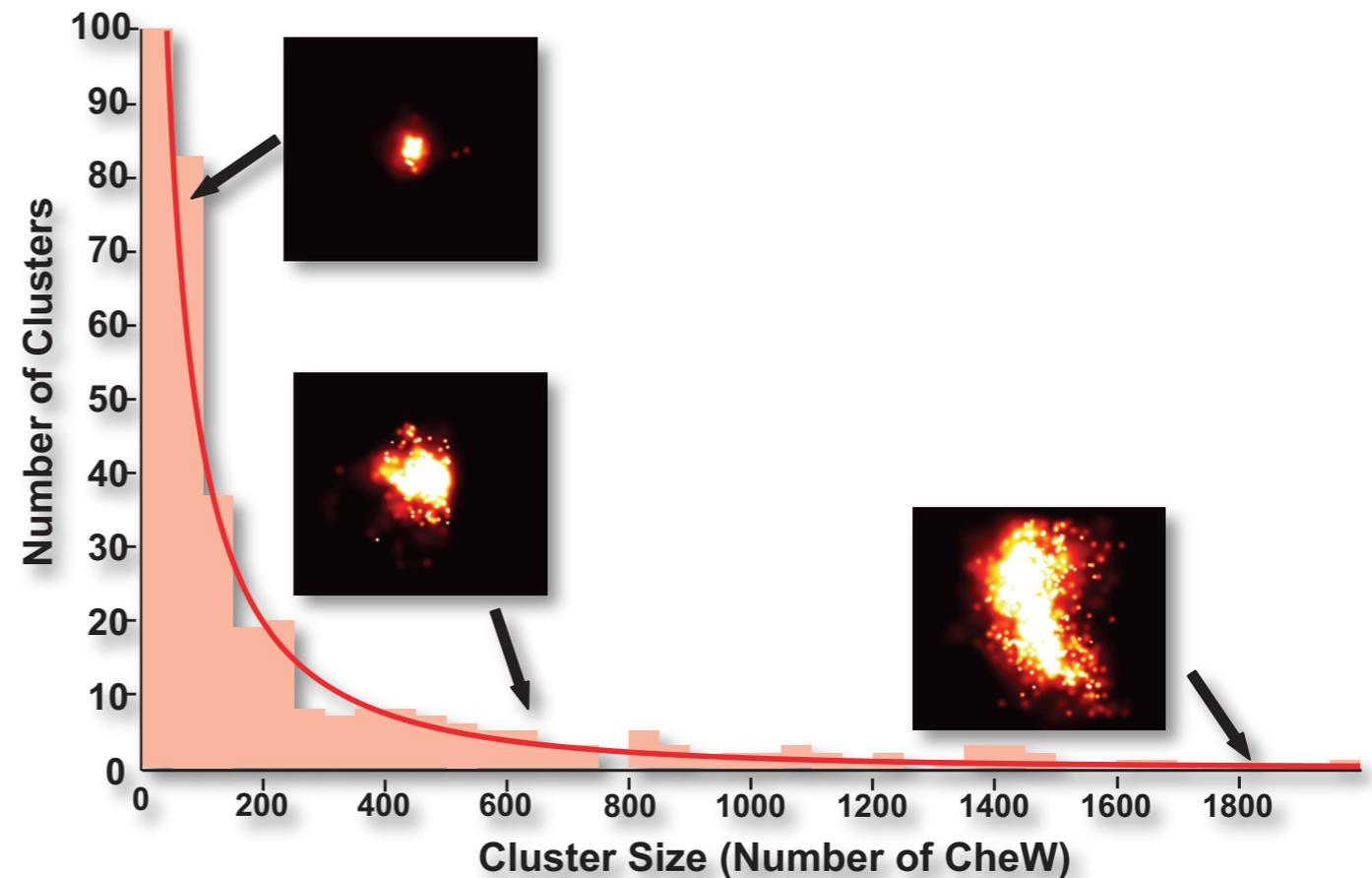


Net filter performance for different mixes

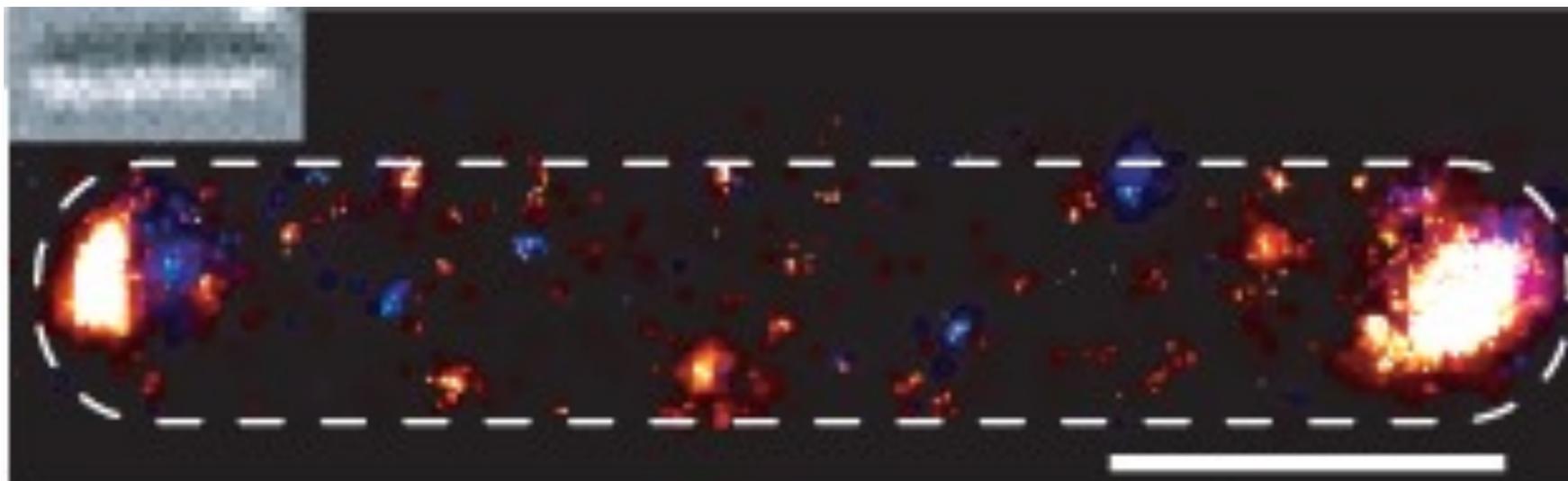


Hang on...

- Looked at cluster size distribution

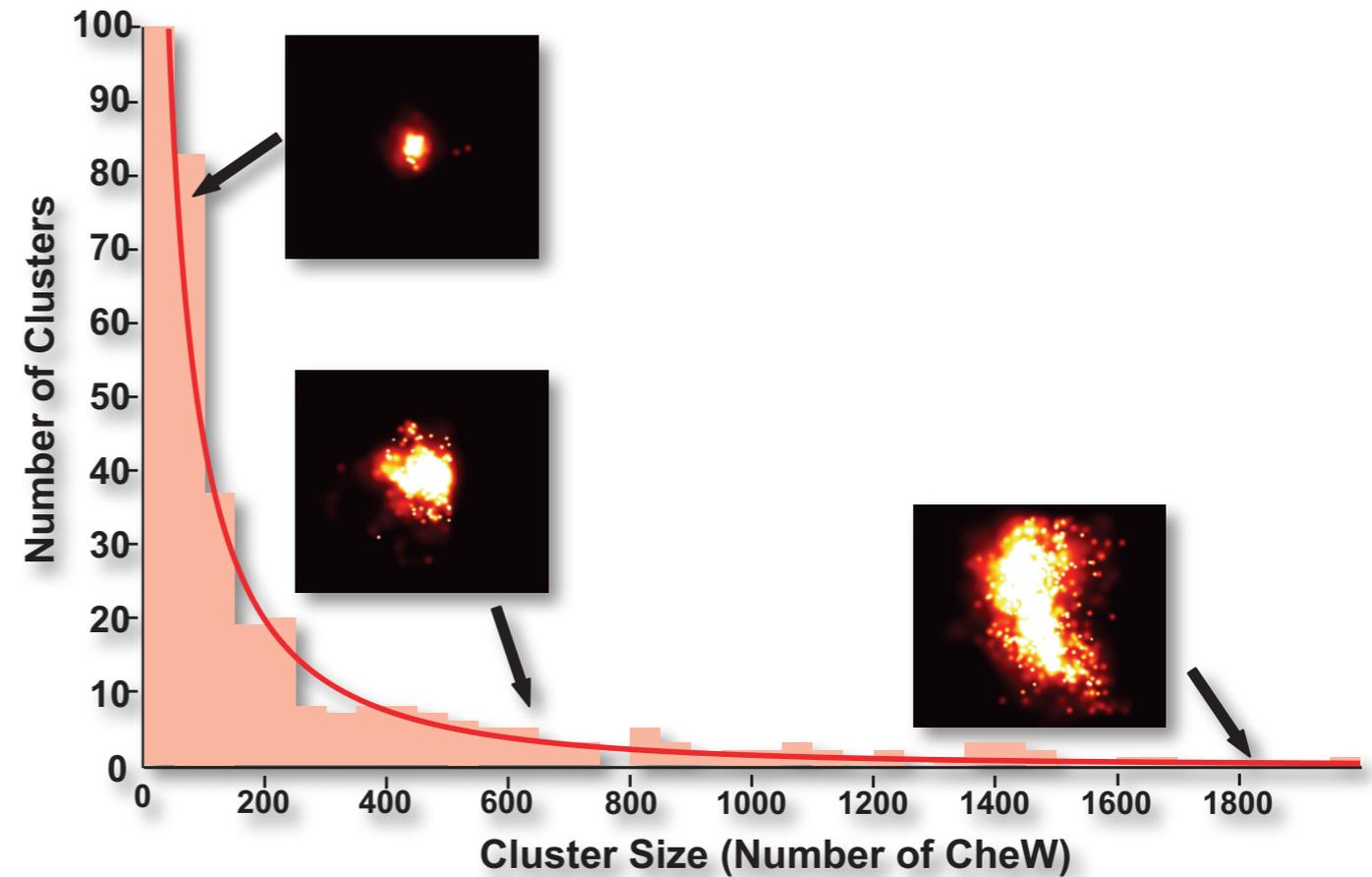


- But what about relative spatial arrangement?



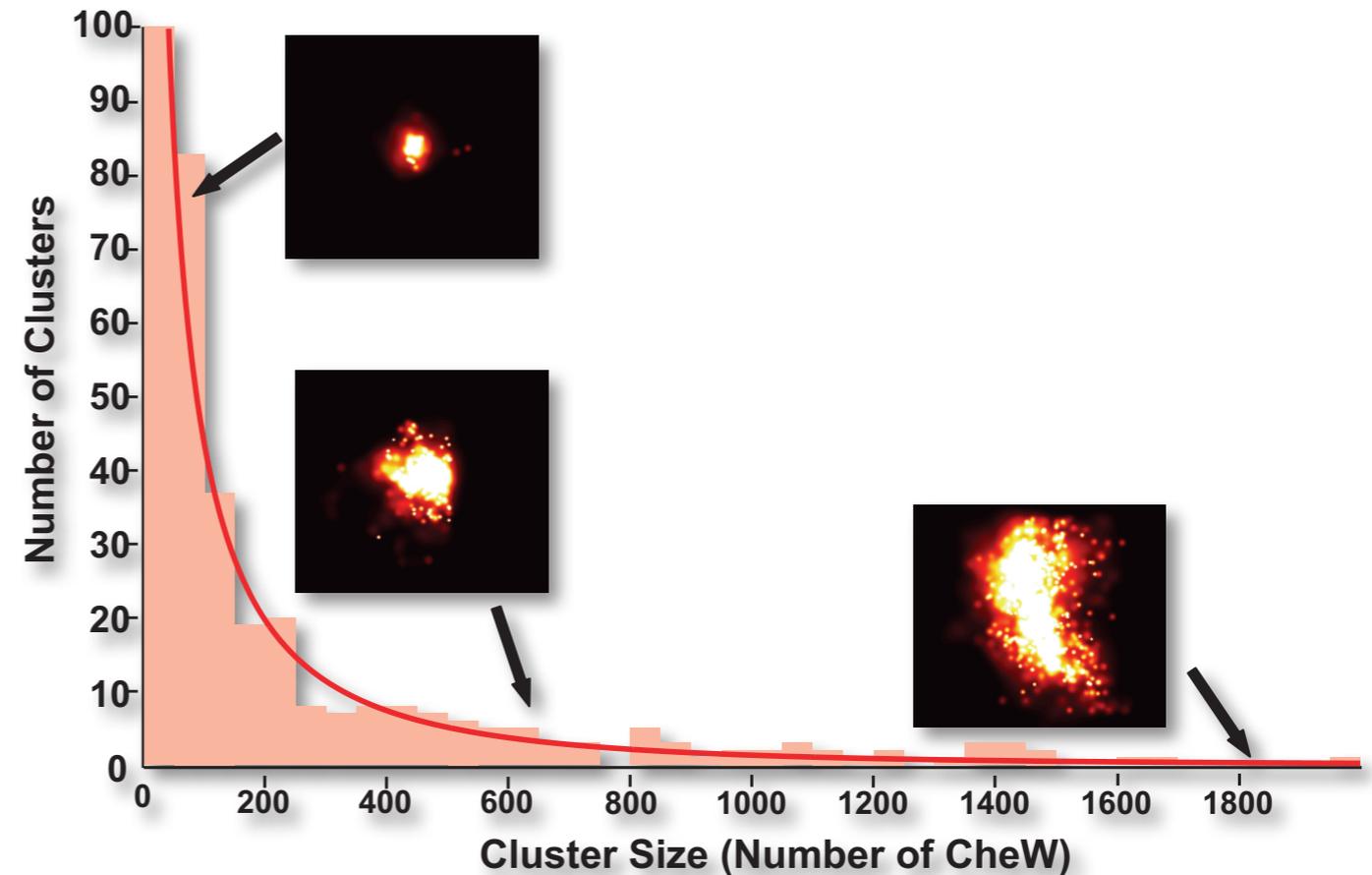
Conclusions

- Looked at cluster size distribution => direct evidence for stochastic nucleation

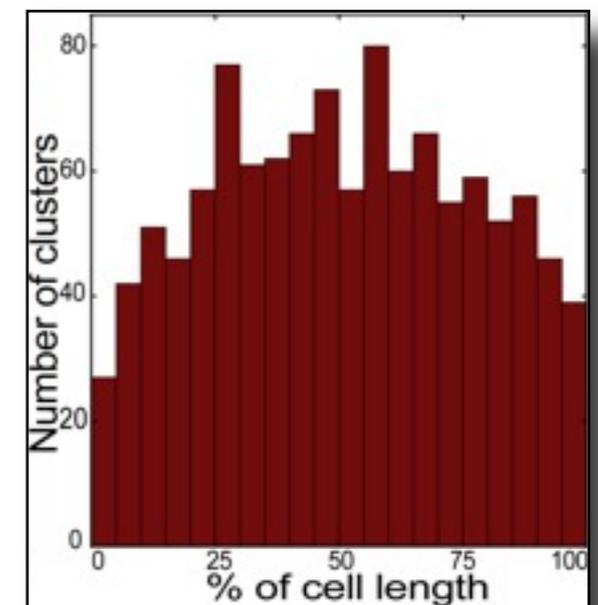


Conclusions

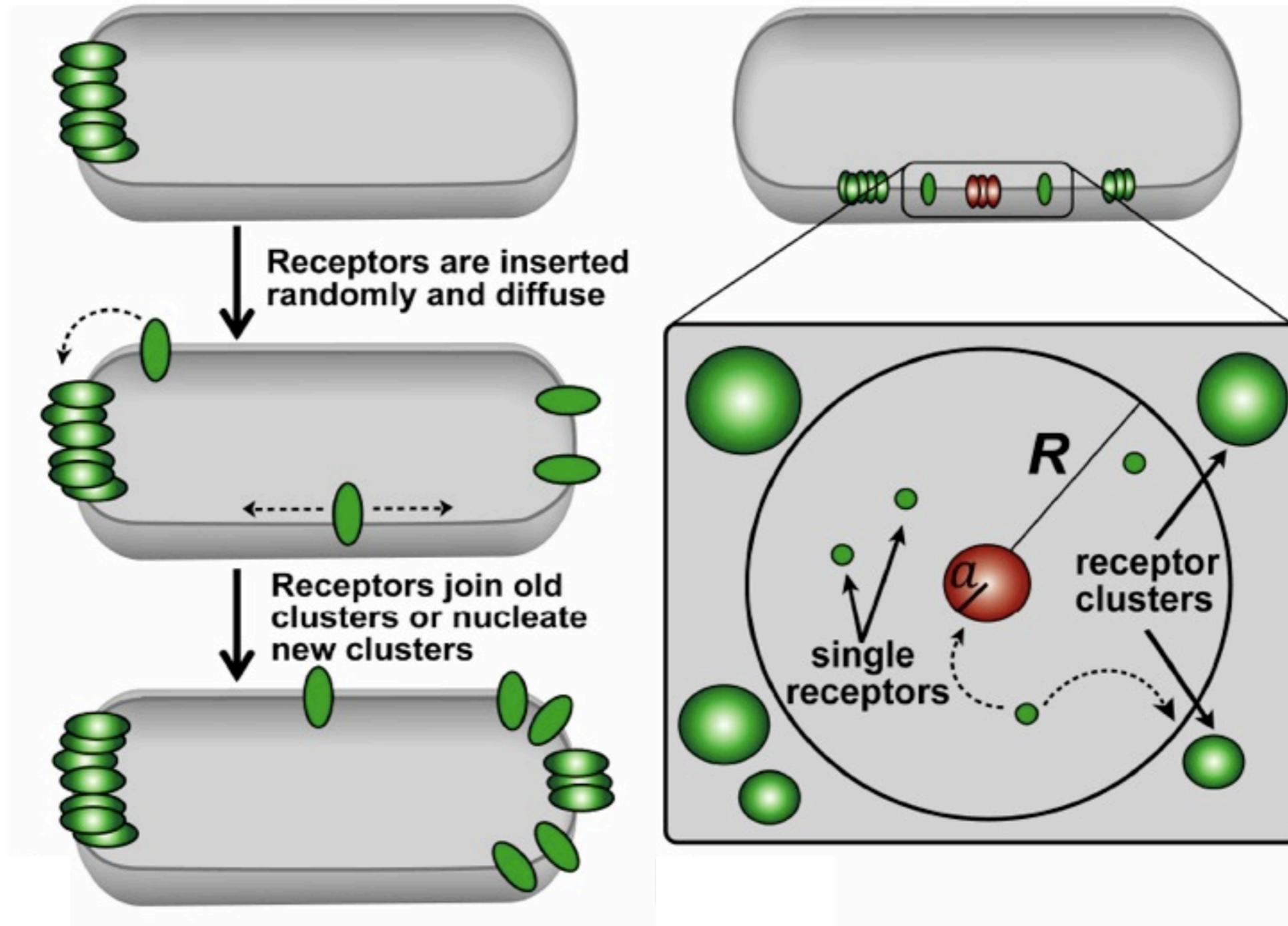
- Looked at cluster size distribution => direct evidence for stochastic nucleation



- Analysis of relative spatial arrangement reveals cluster exclusion => new clusters nucleate furthest from old clusters



Model: stochastic cluster nucleation

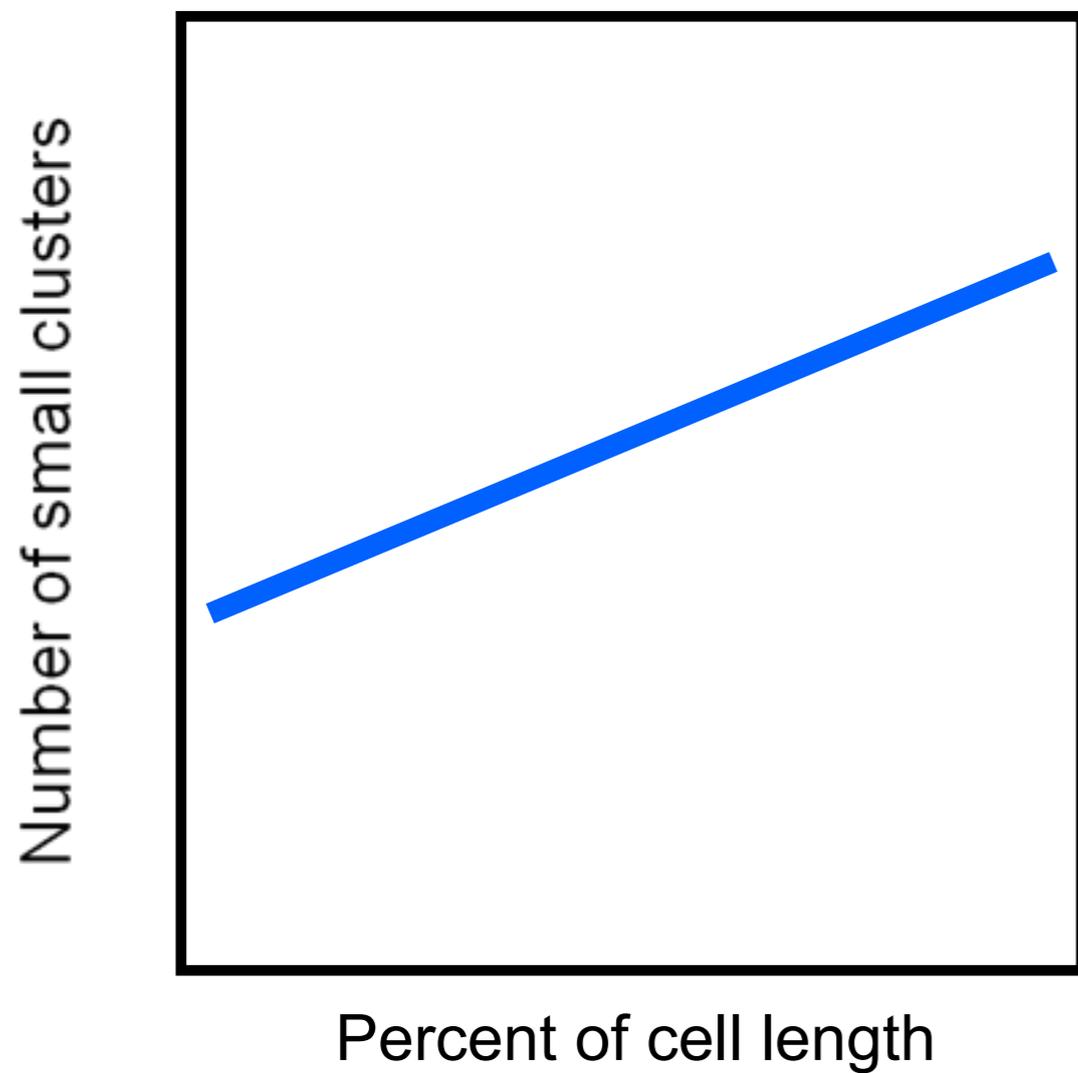


Highest density of single receptors is furthest away from existing cluster(s)

Model predictions



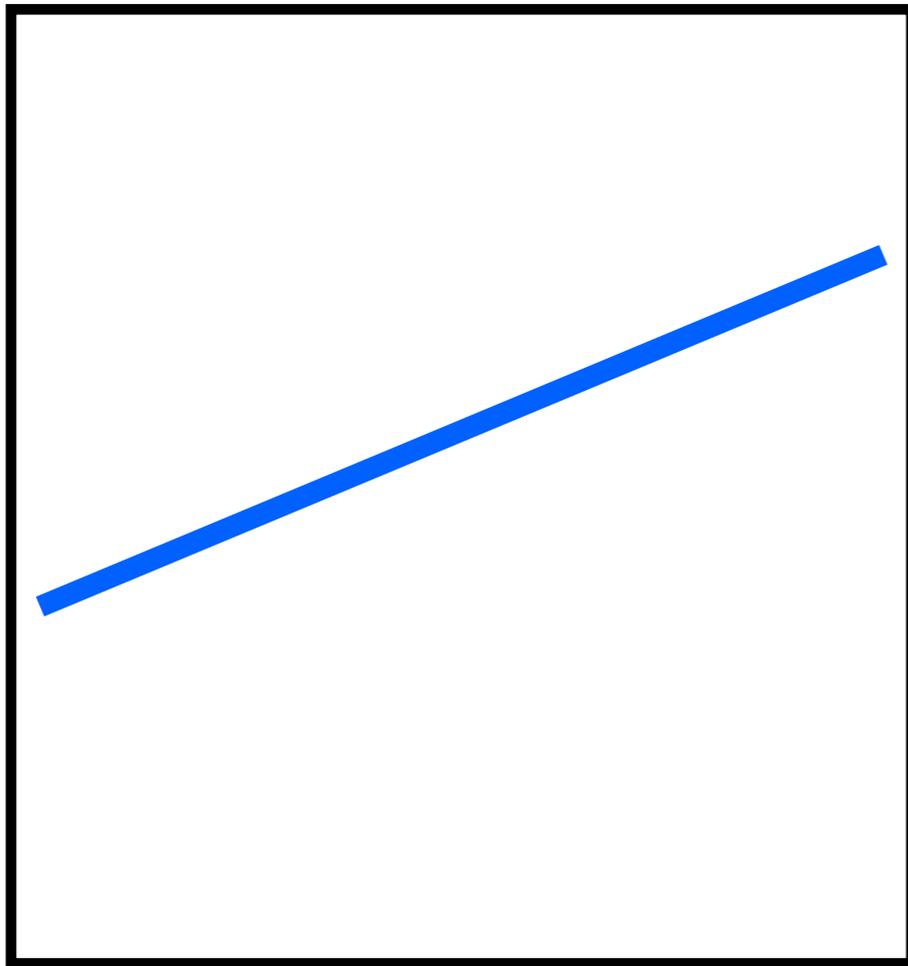
Model predictions



Model predictions

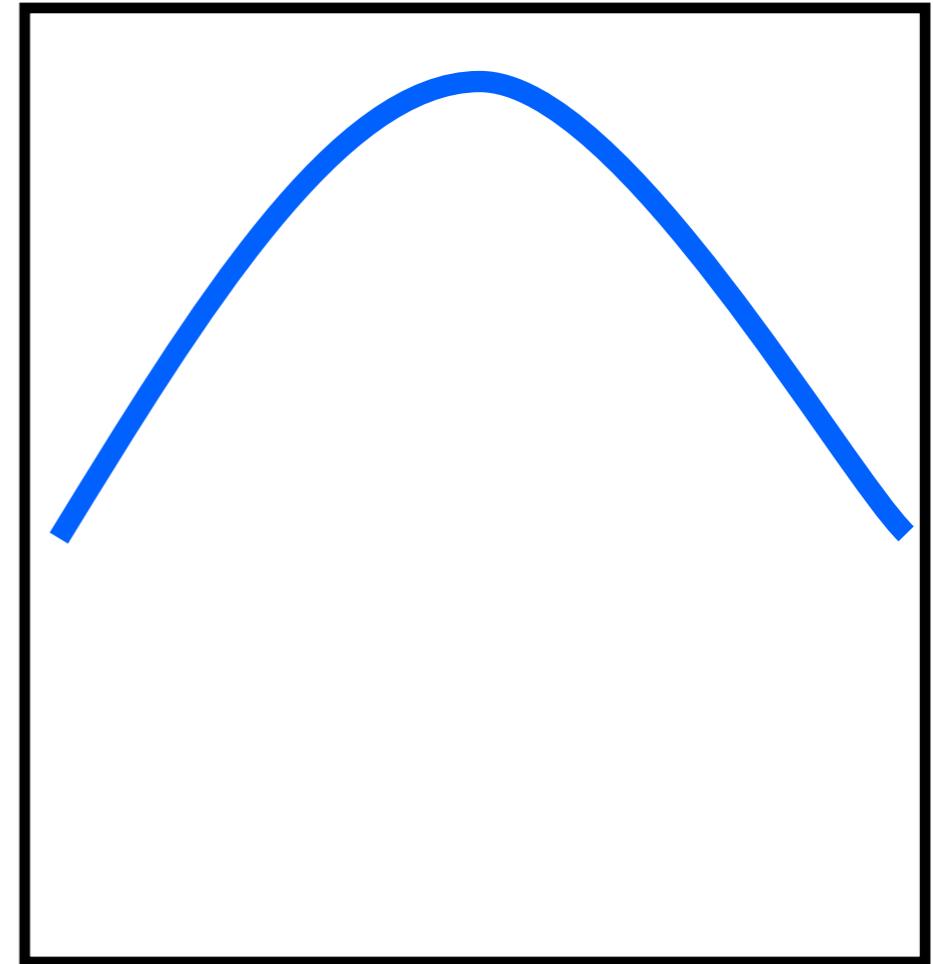


Number of small clusters



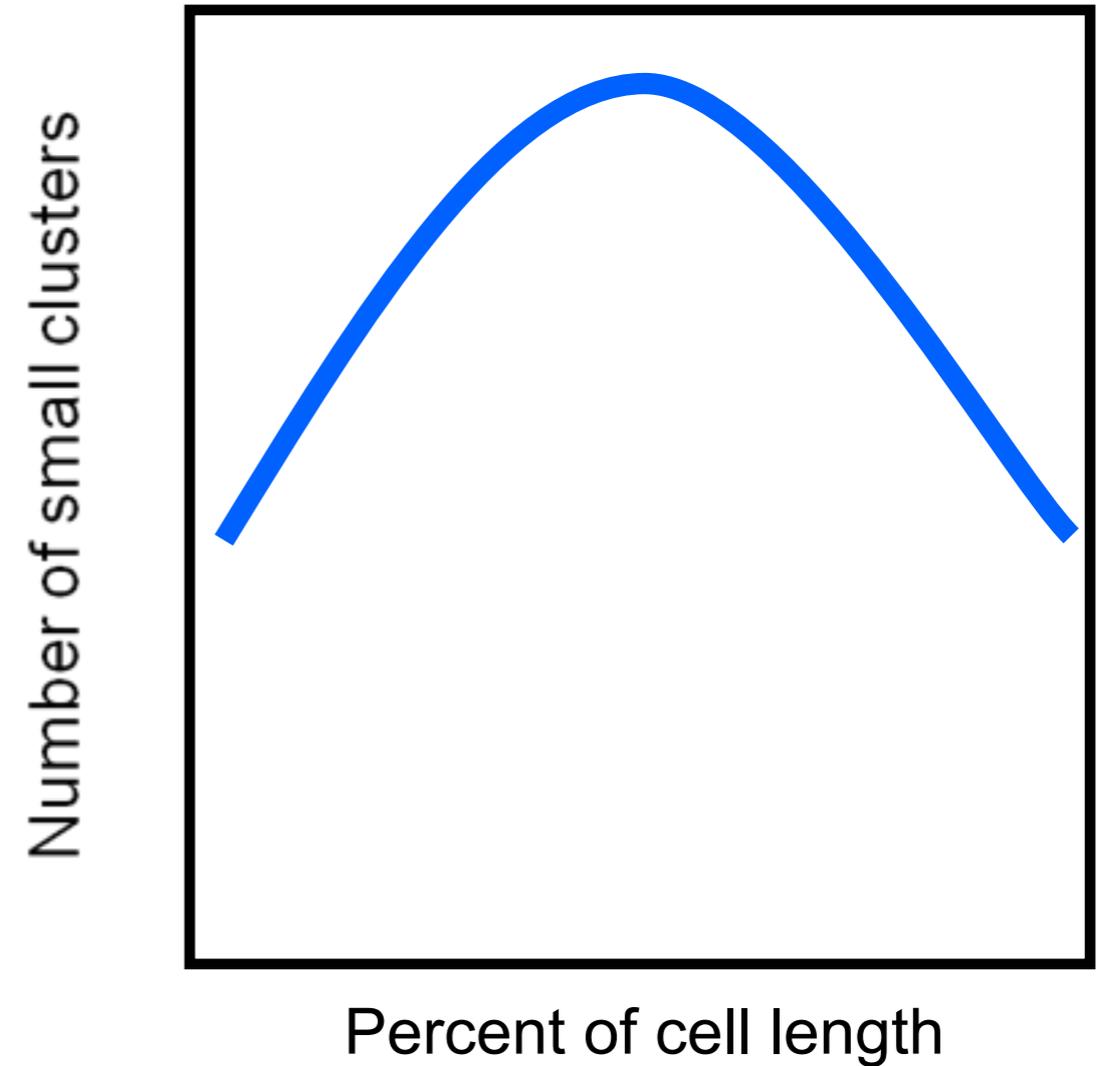
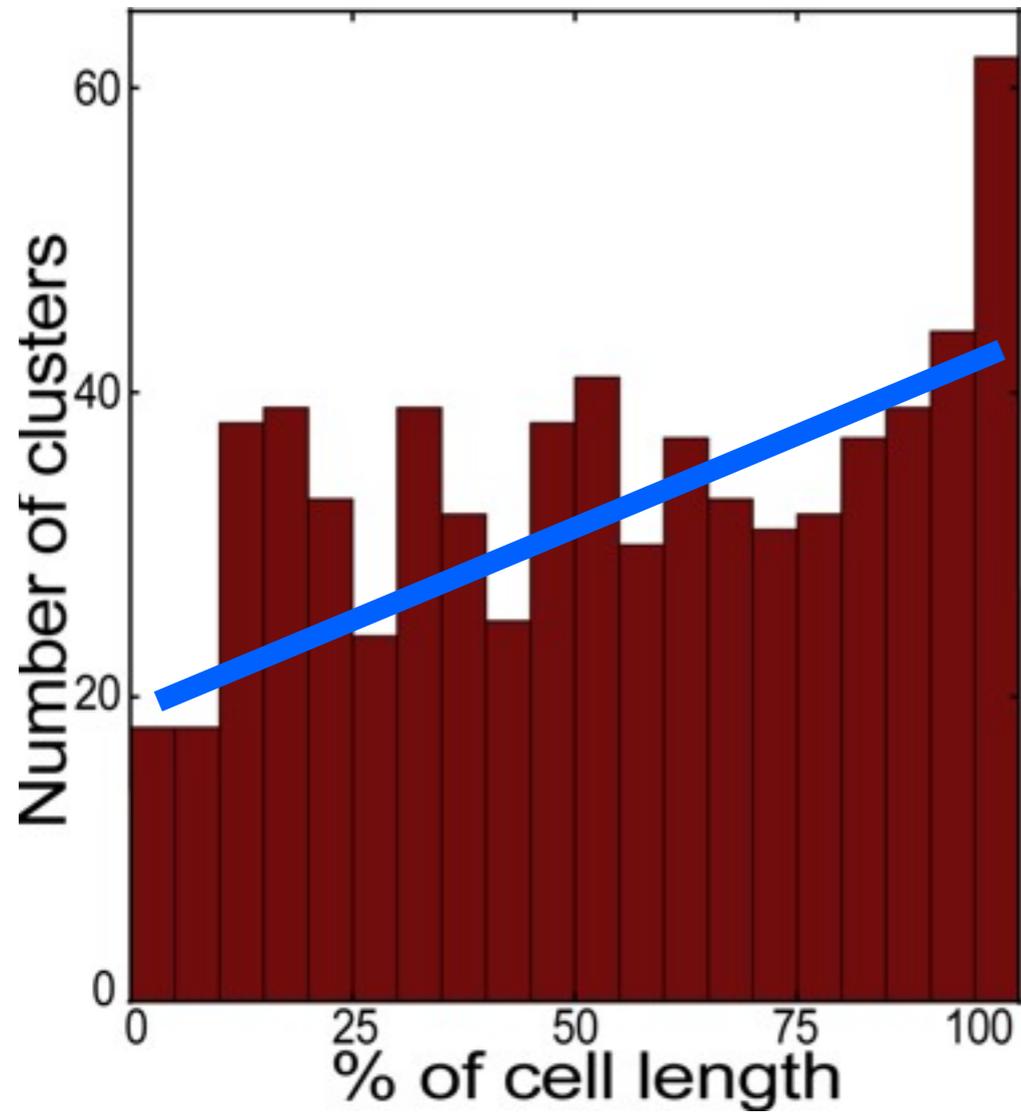
Percent of cell length

Number of small clusters

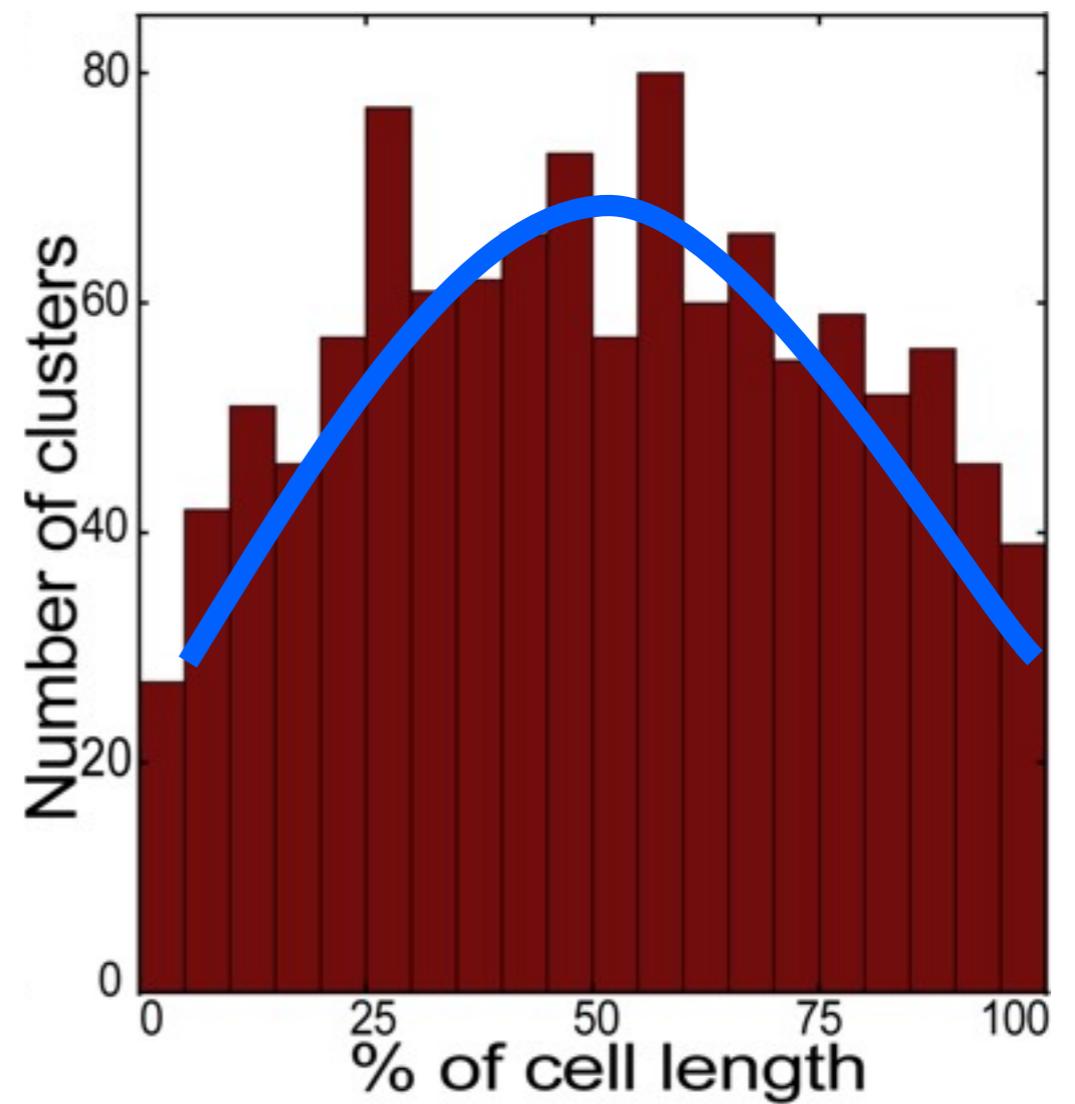
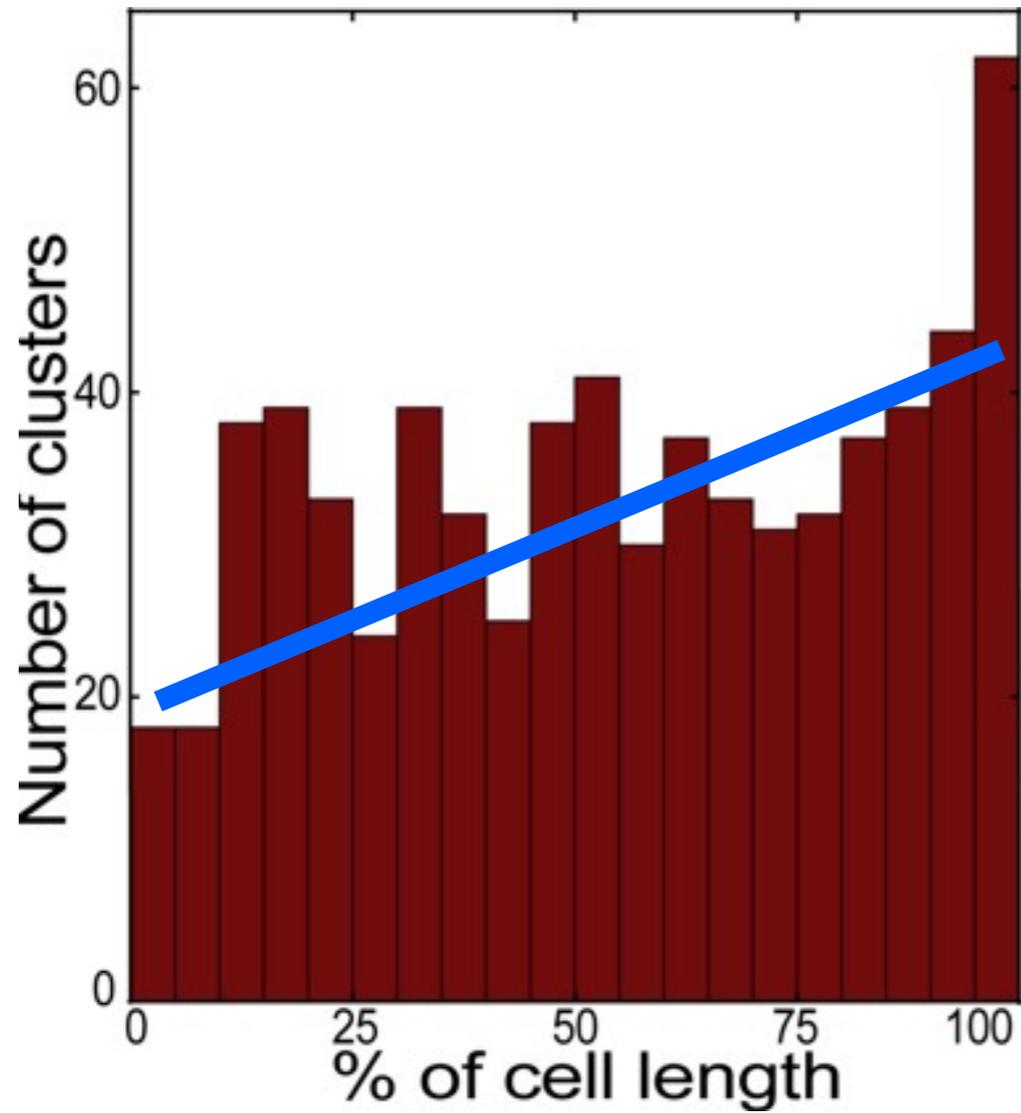


Percent of cell length

Model predictions



Model predictions

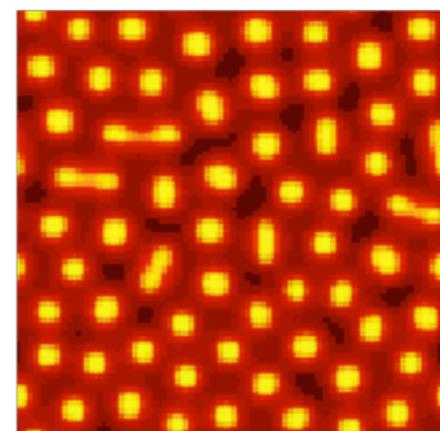
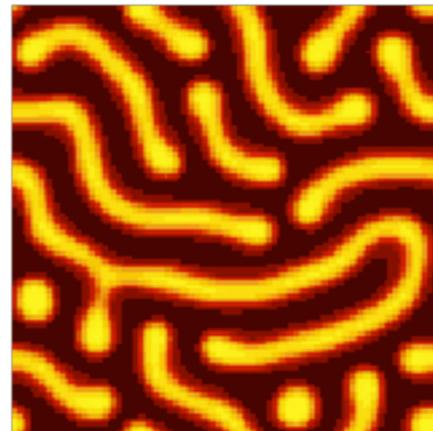
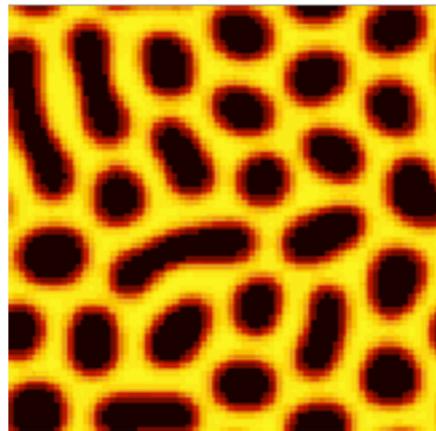


Turing patterns



Chemical morphogenesis (Turing, 1952):

- A system of reacting and diffusing chemical species can spontaneously form stationary spatial patterns given a certain set of chemically plausible mechanisms.
- Two reacting chemical species that diffuse at very different rates.
- The system is an intrinsically non-equilibrium; both substances are continuously created (by the cells) at every point in space, and also decay or are removed at specified rates.
- In these reactions the activator makes more activator and inhibitor, and the inhibitor destroys the activator.

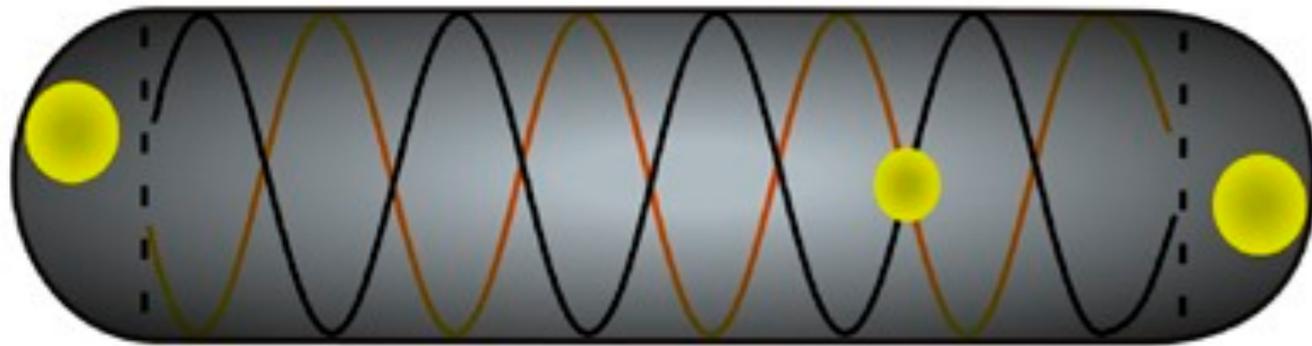


Erik M. Rauch and Mark M. Millonas, J. Theoretical Biology

How do chemotaxis receptors cluster?

How do chemotaxis receptors cluster?

Active



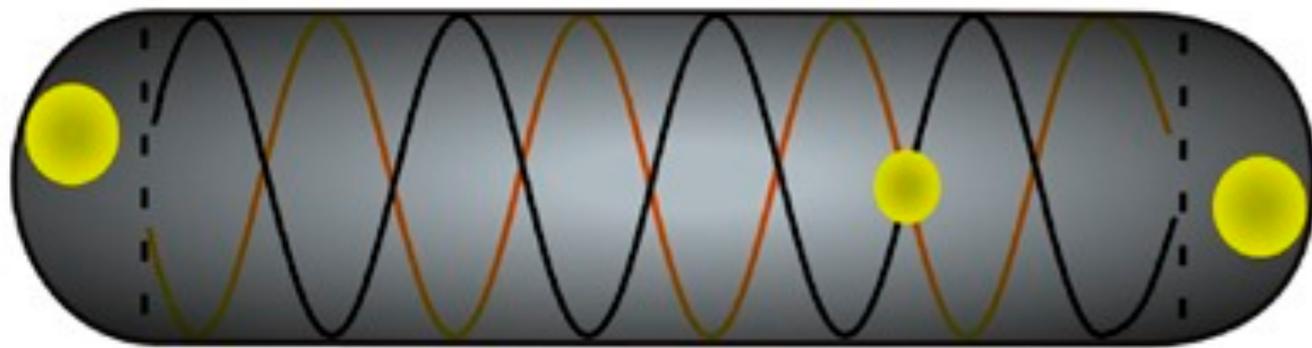
Thiem, Kentner, Sourjik, *EMBO*, **26** (2007)

Positioned relative to cytoskeleton
Transported or captured

Receptors at specific sites

How do chemotaxis receptors cluster?

Active

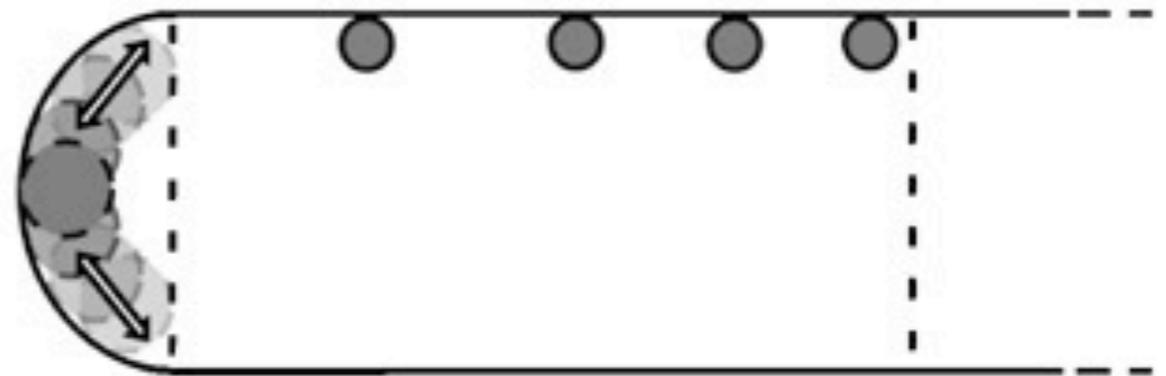


Thiem, Kentner, Sourjik, *EMBO*, **26** (2007)

Positioned relative to cytoskeleton
Transported or captured

Receptors at specific sites

Passive

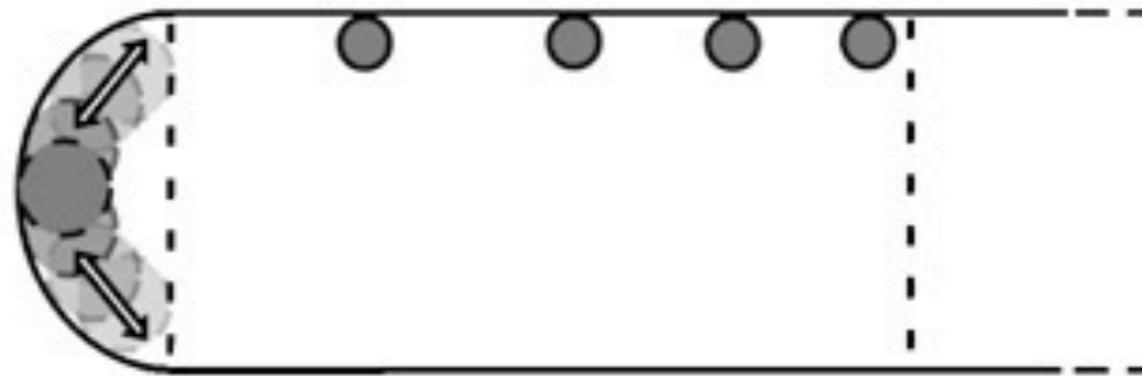


Diffusion and stochastic self-assembly
Membrane curvature as energy minimum

Receptors throughout membrane

How do chemotaxis receptors cluster?

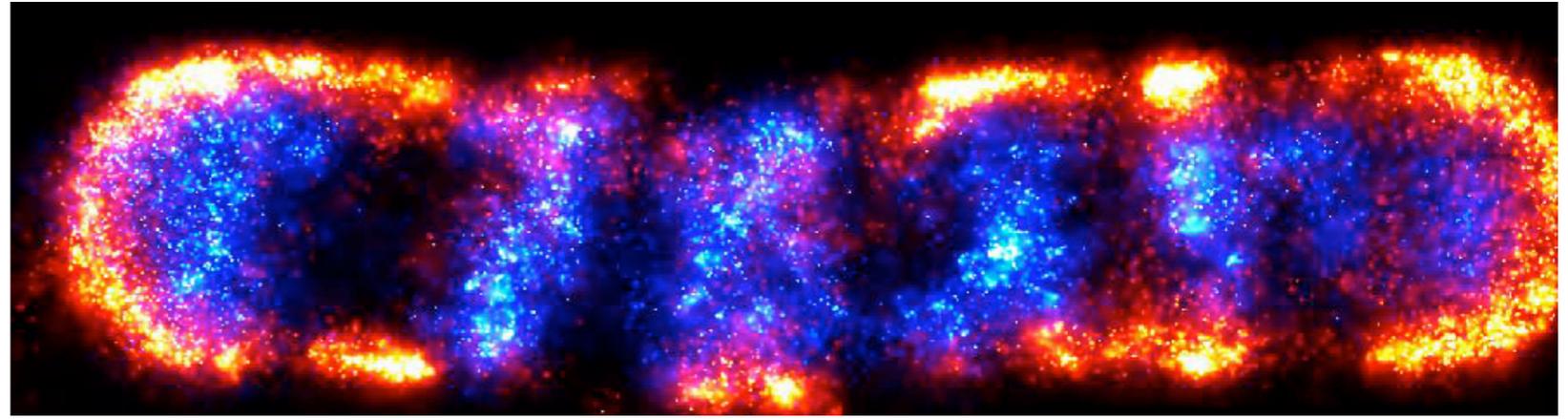
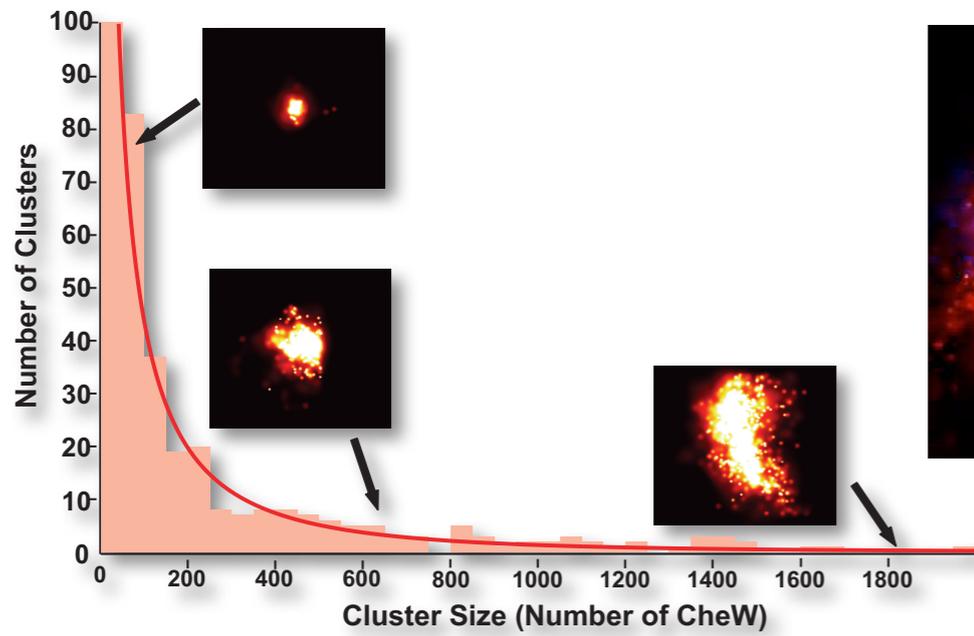
Passive



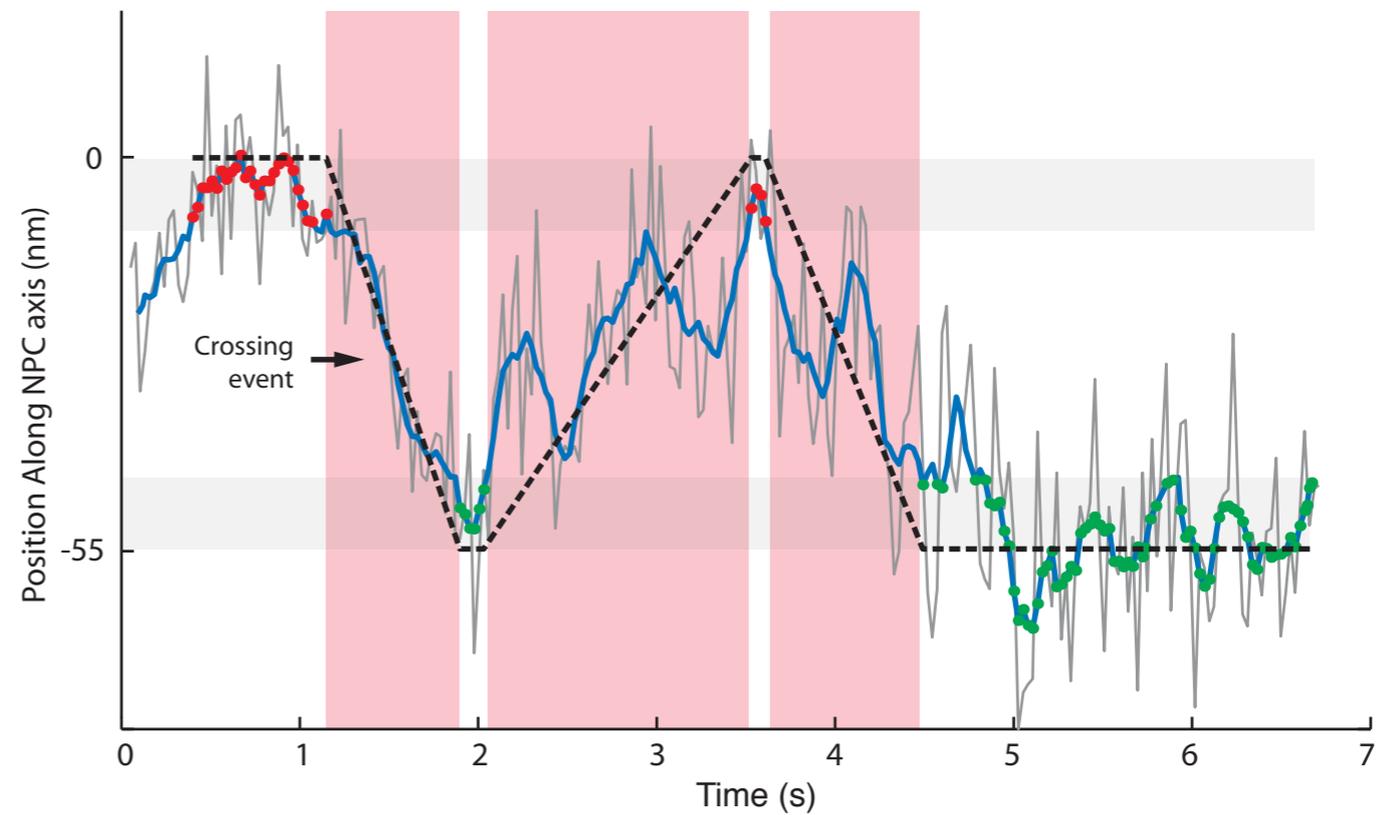
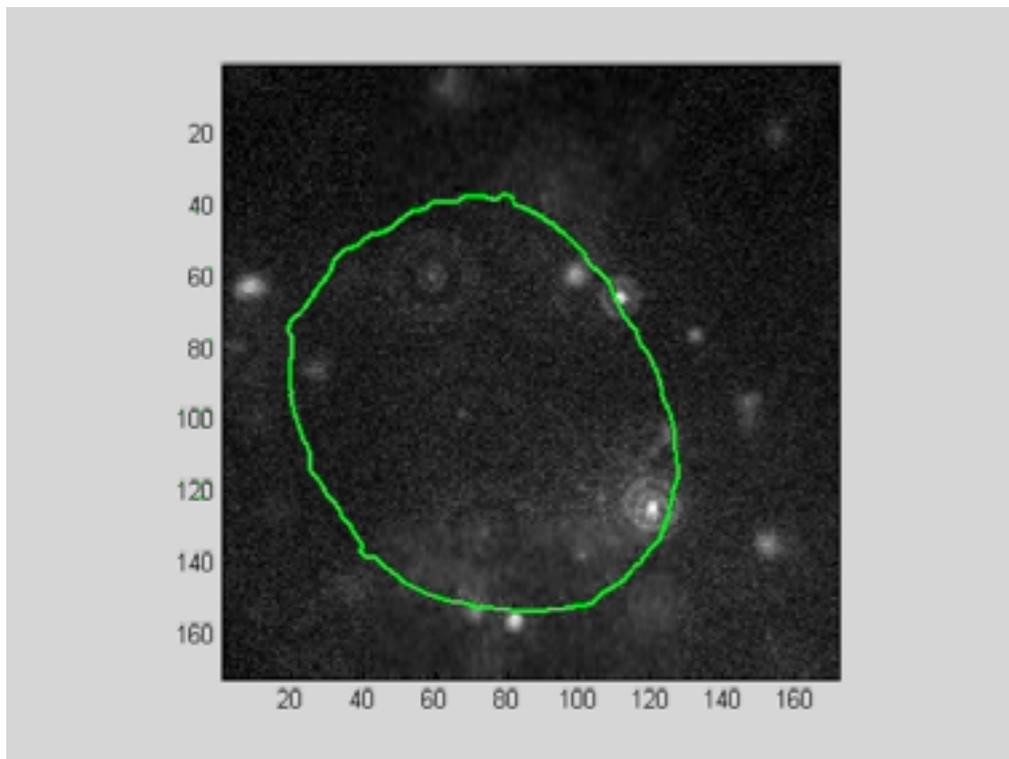
Diffusion and stochastic self-assembly
Membrane curvature as energy minimum

Receptors throughout membrane

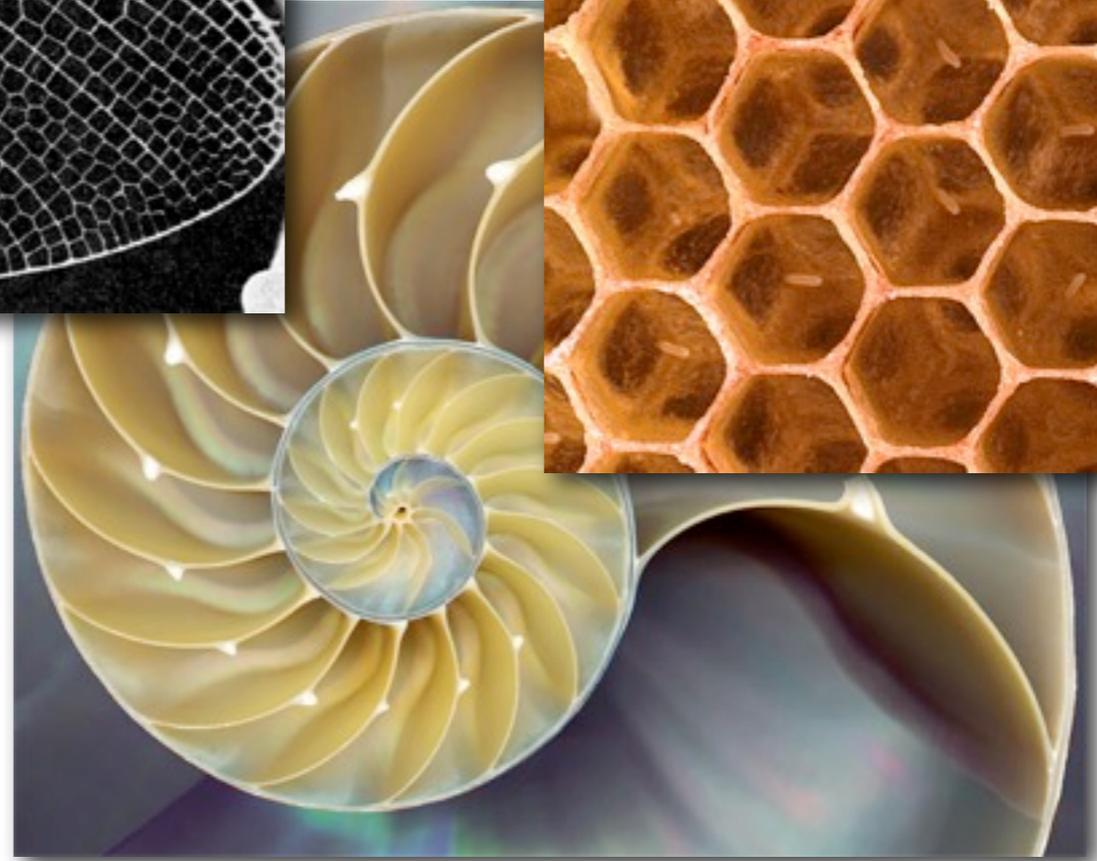
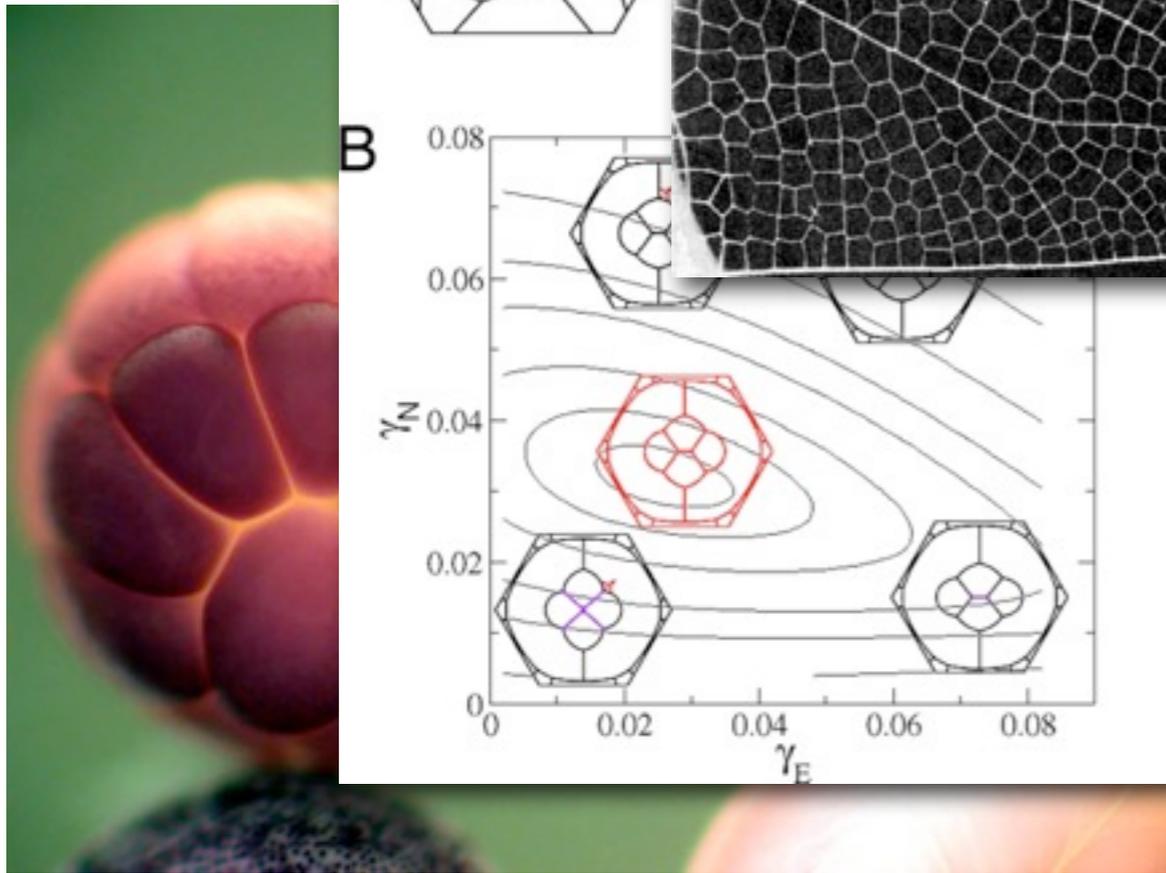
Purely passive mechanism for generating and maintaining periodic protein distributions in membranes



Self-organization of the E. coli Chemotaxis System Imaged with Super-resolution Light Microscopy
 D. Greenfield, A. McEvoy, H. Shroff, G. Crooks, N. Wingreen, E. Betzig, and J. Liphardt
 PLoS Biology (2009)



Efficiency and selectivity via multi-step mechanism, consisting of a cascade of reversible pre-filters and a final irreversible exit step



Problem I: Genes \Leftrightarrow Cell Types

Problem II: Robust Development of an Animal from one Cell

Problem III: Integrating Genes with Mechanical Forces

On Growth and Form,
D'Arcy Wentworth Thompson, 1917

Acknowledgements

Berkeley

Ann McEvoy

Derek Greenfield

Will Draper

Adam Politzer

Alan Lowe

Phil Jess

Jamie Yassif

Rebecca Schulman

Jessica Walter

Brett Schofield

Jake Siegel

Tom Yuzvinsky

Prof. Karsten Weis

Janelia Farm / HHMI

Eric Betzig

Harald Hess

Hari Shroff

Theory

Ned Wingreen (Princeton)

Gavin Crooks (LBNL)

$$\frac{dN}{dt} = \pi\gamma \left[\frac{1}{2} \frac{(R^2 - a^2)}{\ln\left(\frac{R}{a}\right)} - a^2 \right]$$

Support

NIH, NCI, DOE

Bay Area Physical Sciences-Oncology Center

Mohr Davidow Ventures

Searle, Sloan, Hellman, Wit and Will Foundations