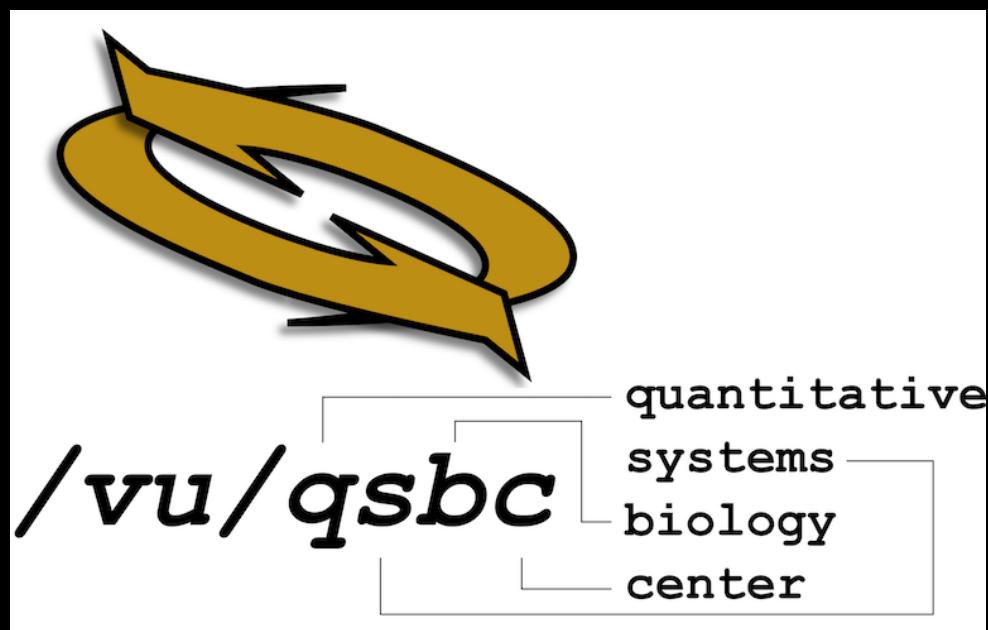
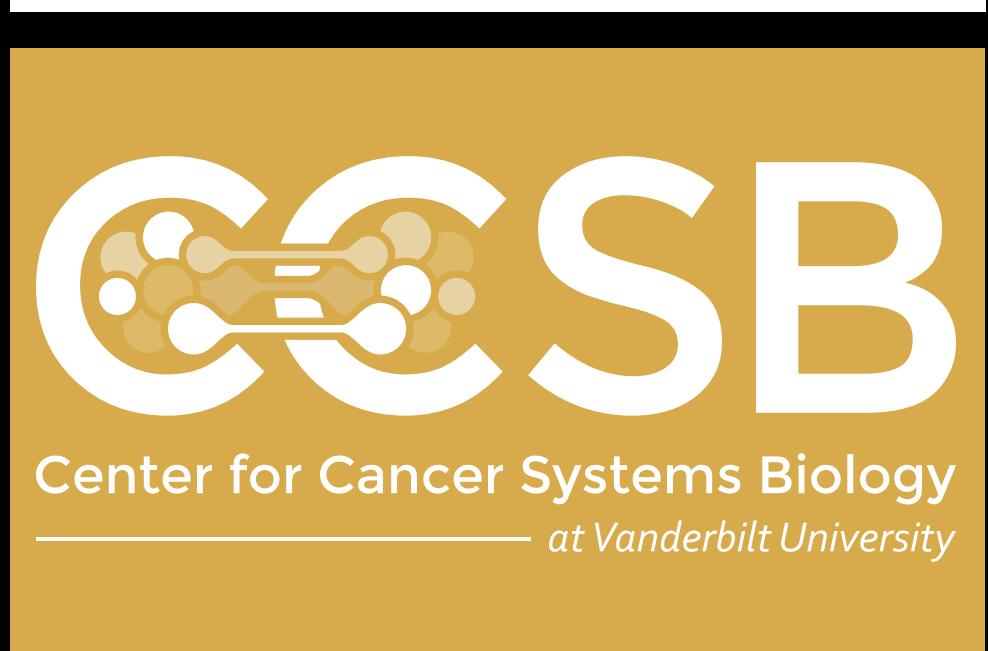


Cancer Systems Biology & Pharmacology



Darren Tyson, Research Associate Professor, Department of Pharmacology



Today's Lecture

- Background on systems biology, cancer systems biology and targeted anticancer therapeutics
- Overview of a data generation and analytical pipeline for understanding the concentration-dependent action of anticancer drugs on cultured human cancer cells
- Introduction to Git/GitHub/GitHub Classroom
- Goals for remaining lectures in Quantitative Science
- Summary of resources available for you to aid your learning

What is Systems Biology?

Systems Biology is a subdiscipline of the study of *Complex Systems*

Non-living and engineered Complex Systems are diverse but with common properties, primarily nonlinear behavior, meaning they may respond in different ways to the same input depending on state or context.

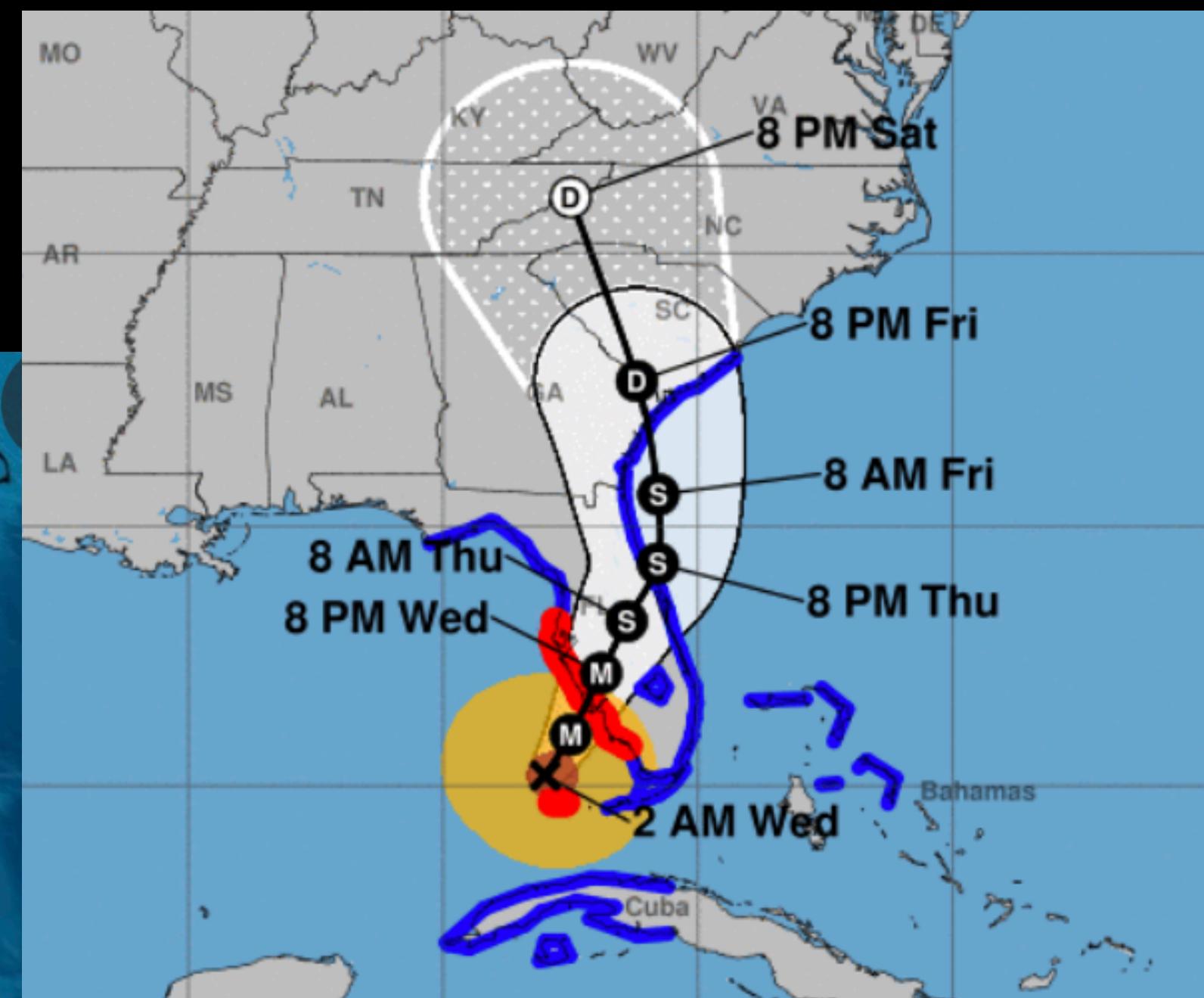
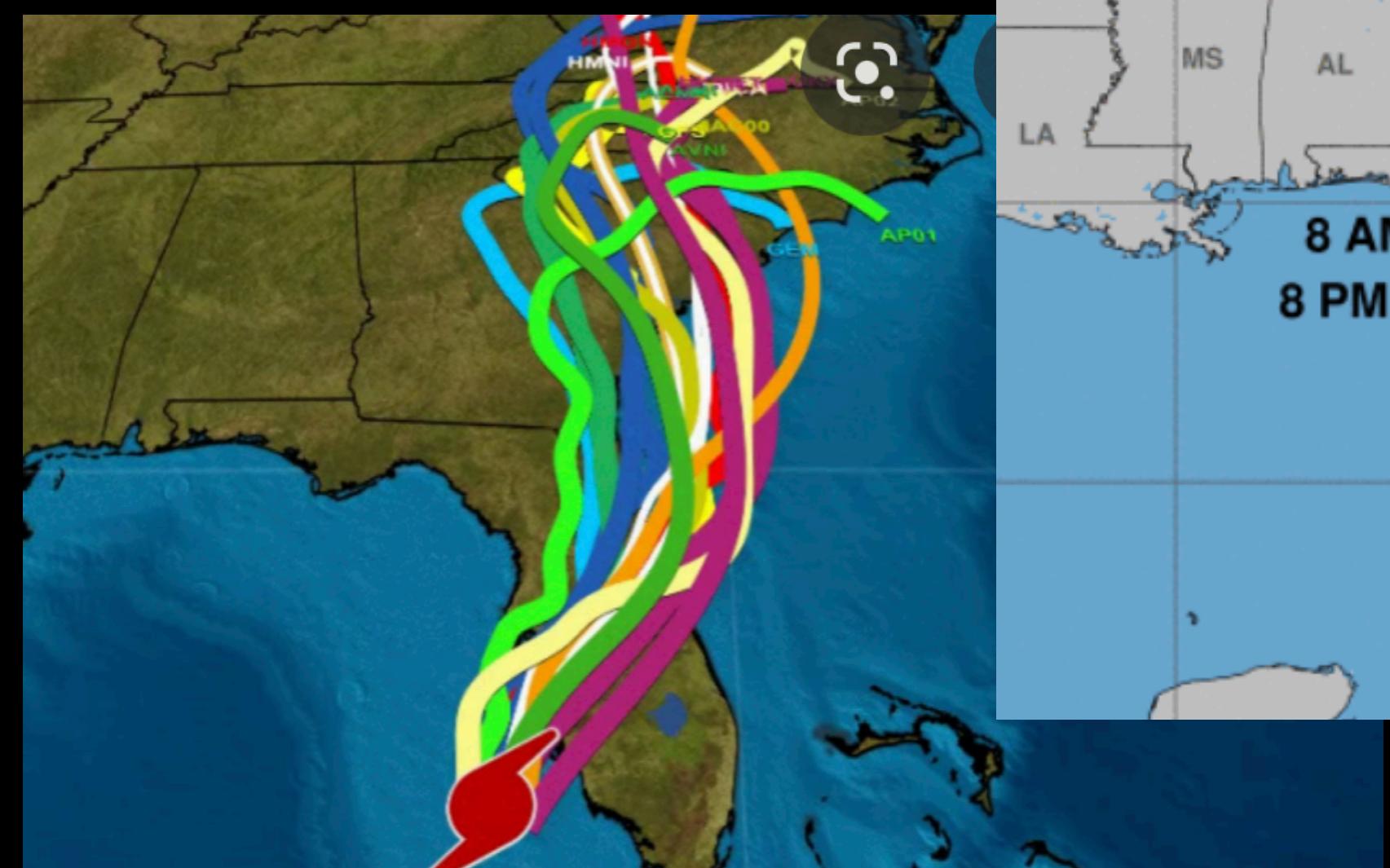
https://en.wikipedia.org/wiki/Complex_system

examples

weather
vehicular traffic
financial markets
climate
world wide web
power grid

Many computational and mathematical tools developed for nonliving systems can be directly applied to biological systems

Hurricane Ian



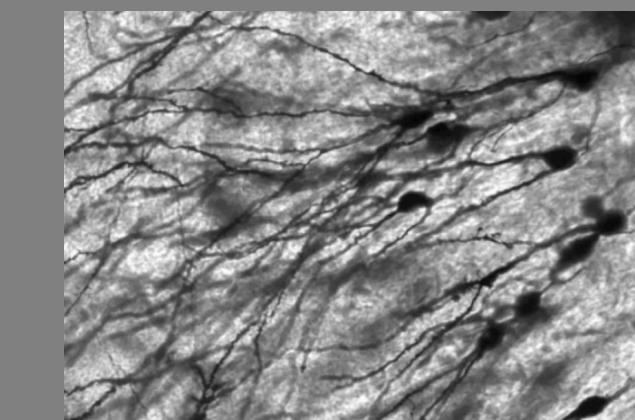
Biological Systems

Properties of Complex Systems (both nonliving and living):

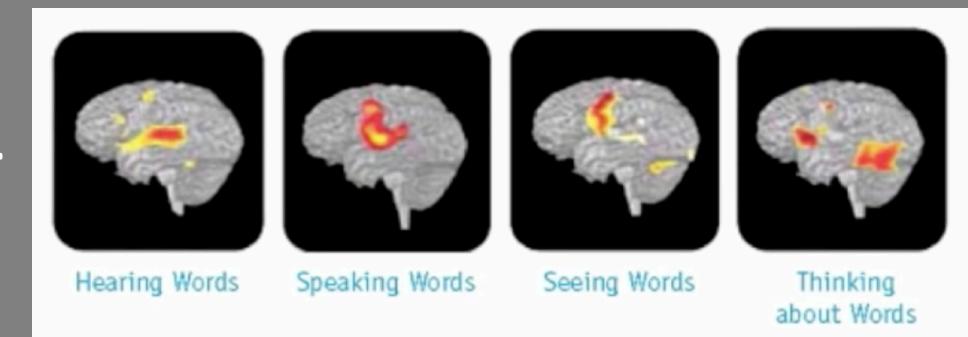
- Simple components or agents
 - *Simple relative to the whole system*
- Emergent behavior
 - *The whole is more than the sum of its parts*
- Nonlinear interactions among components
- Self-organization
 - *No central control*
- Resilience/Robustness
- Hierarchy of scales



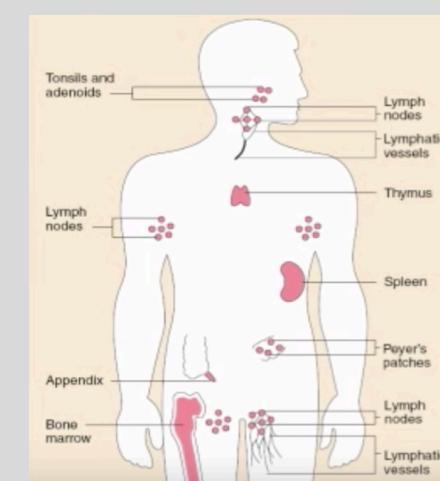
Termites—
Mounds



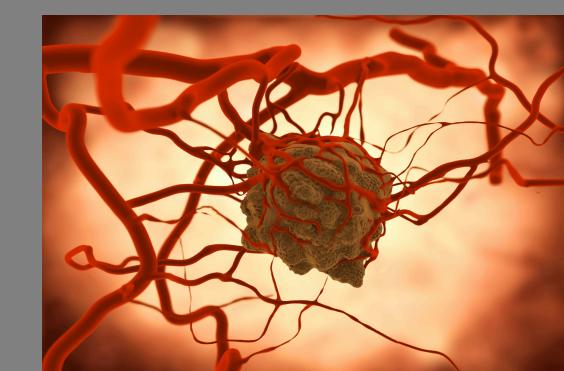
Neurons—
Brain



Lymphocytes



Tumors

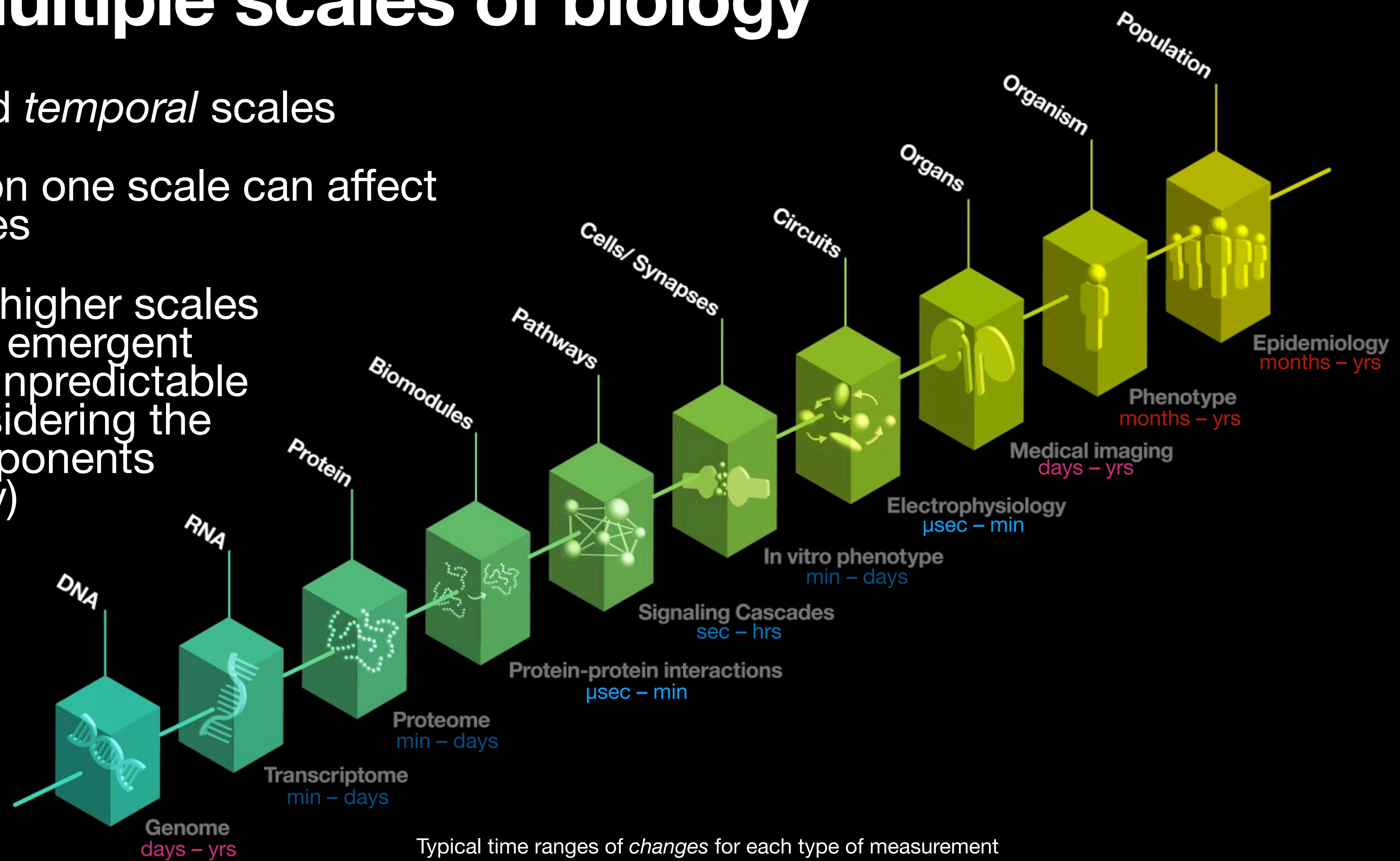


The multiple scales of biology

Spatial and temporal scales

Changes on one scale can affect other scales

Effects on higher scales may be an emergent property (unpredictable when considering the lower components individually)



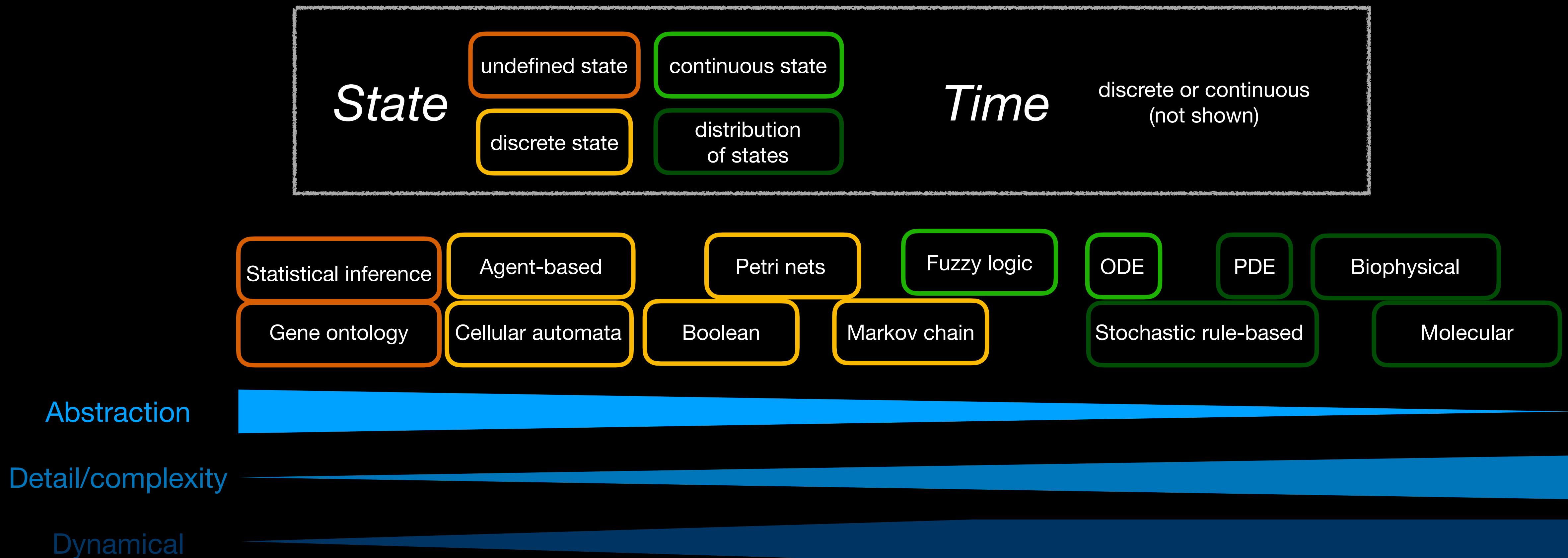
Typical time ranges of *changes* for each type of measurement

What do we mean by “model?”

- A model is a logical machine for deducing conclusions from assumptions
- A model is a precise description of our current (limited) understanding
- Numerous computational and mathematical models/tools developed for nonliving systems can be directly applied to biological systems

Modeling Approaches

(non exhaustive list)

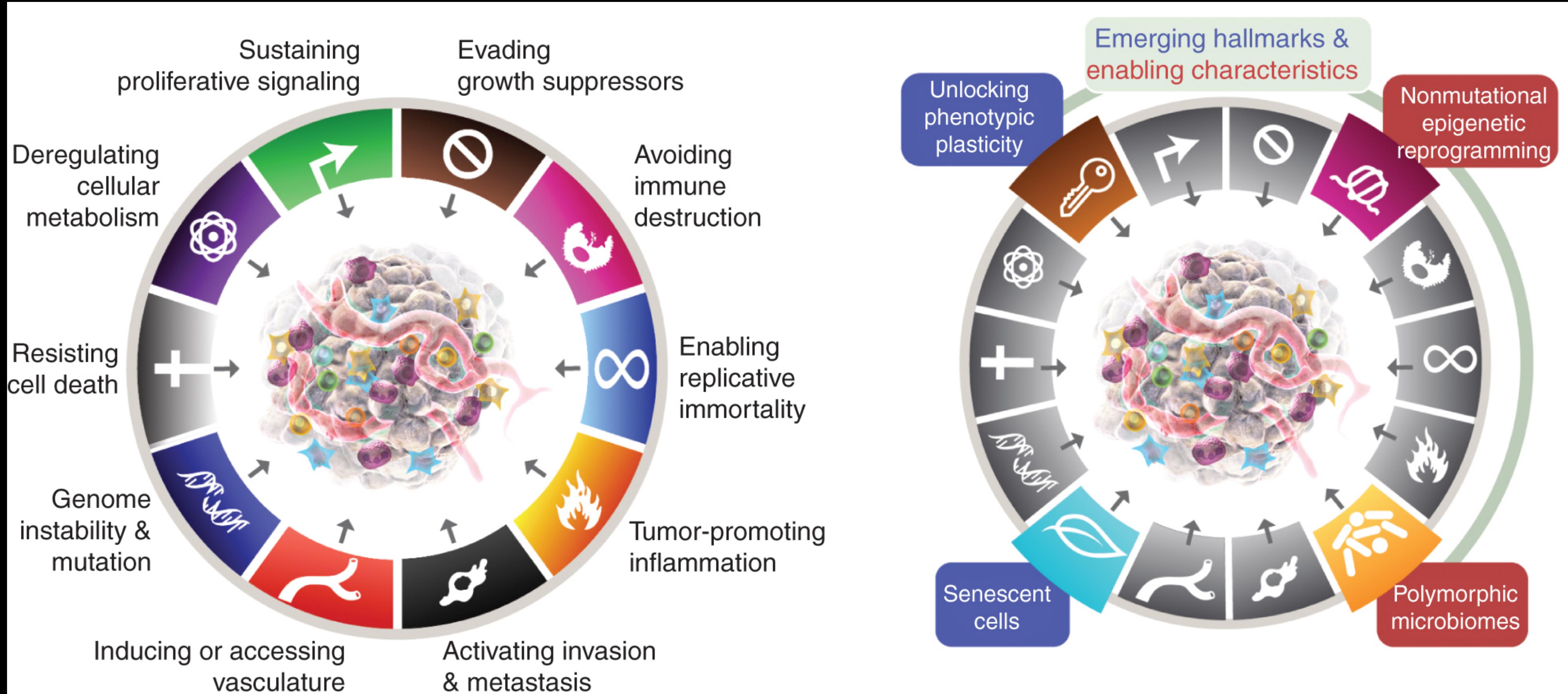


**Each has specific technical foibles and a vast associated technical literature,
but each has the same logical structure**

$$f(\text{inputs}) = \text{output}$$

Hallmarks of Cancer: New Dimensions

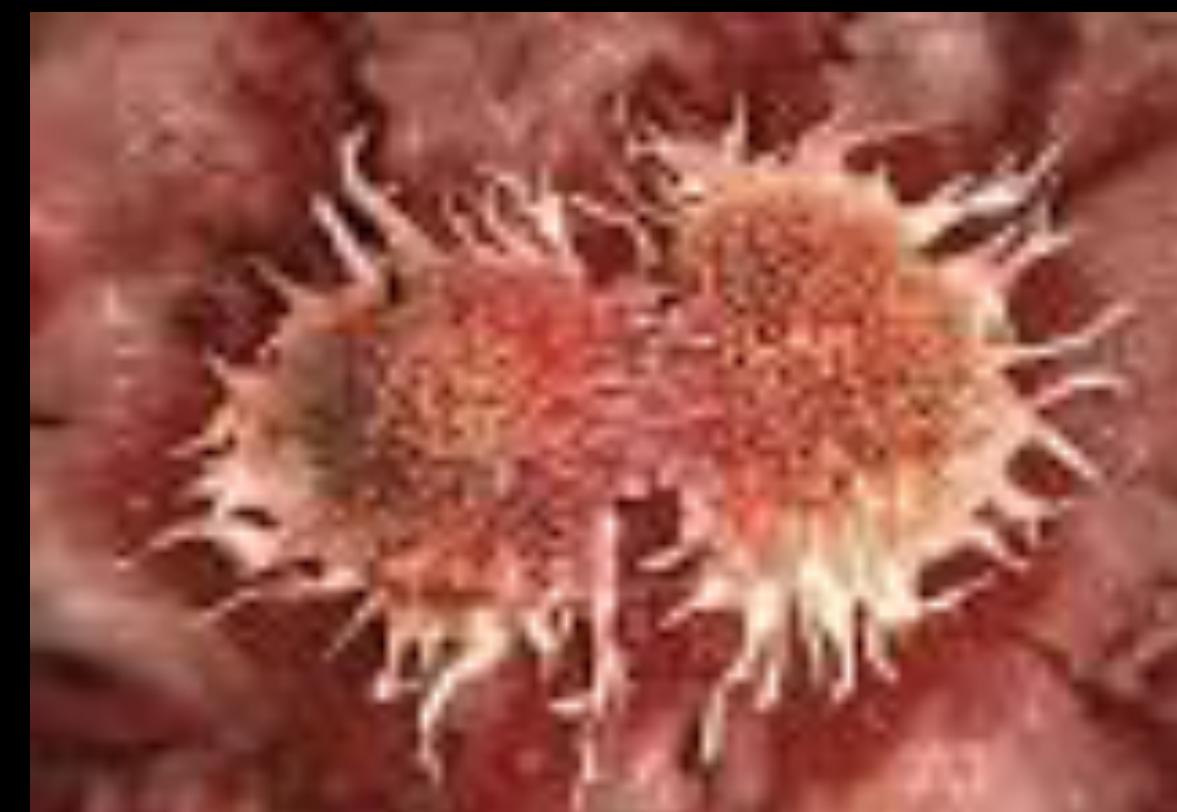
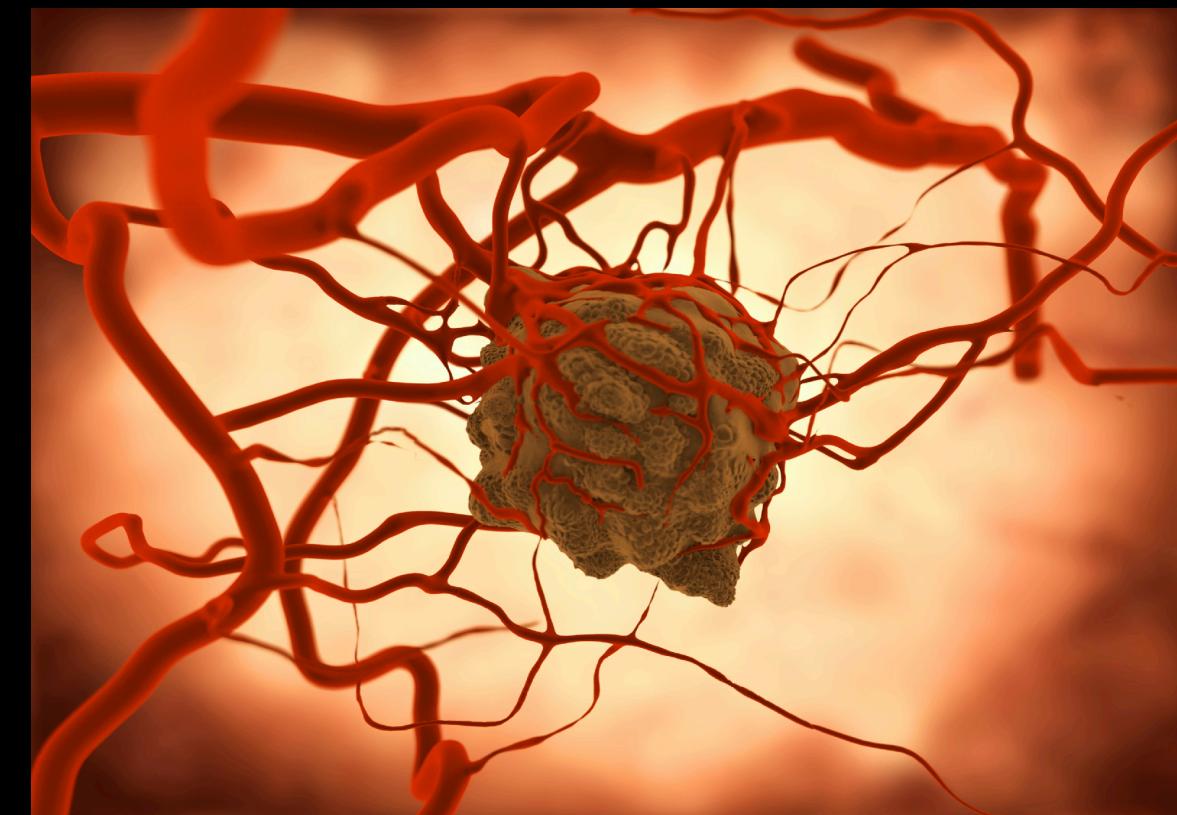
D. Hanahan - *Cancer Discovery*, 2022 (Update of 2000 and 2011 Hallmarks)



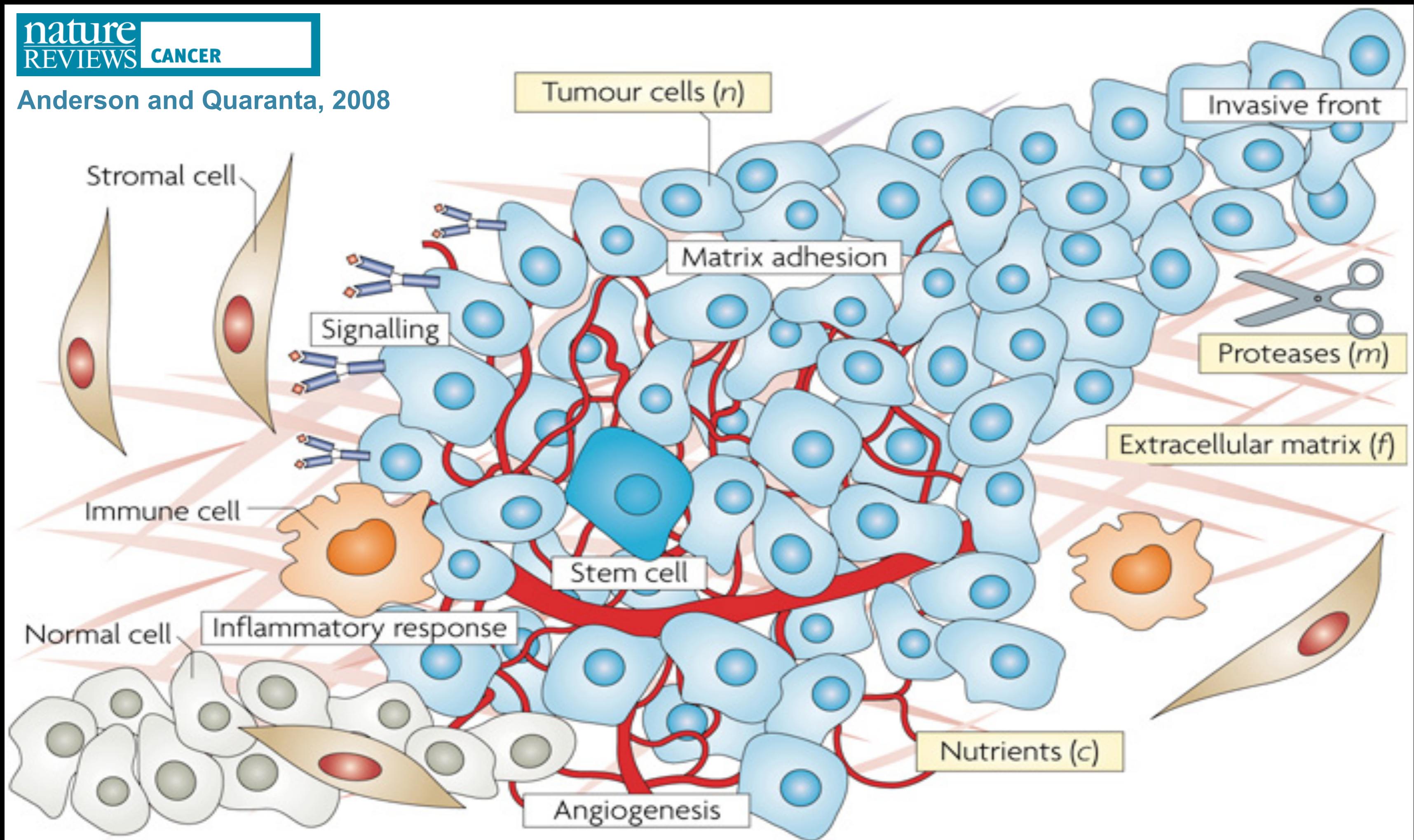
Cancer Systems Biology

Consider tumors as Complex Adaptive Systems and simplify the “Hallmarks” of Cancer to the same “hallmarks” of complex systems.

- Simple components or agents
 - *Simple relative to the whole system*
- Emergent behavior
 - *The whole is more than the sum of its parts*
- Nonlinear interactions among components
- Self-organization
 - *No central control*
- Resilience/Robustness
- Hierarchy of scales



Cancer Progression is a Complex Dynamic Process



- Tissue damage and/or chemical mutagenesis lead to inflammatory responses and tissue remodeling
- Genetic alterations in cells with stem-like features disrupt tissue homeostasis
- Cancer cell phenotypes emerge that respond inappropriately to developmental cues
- Cancer cells co-evolve with their microenvironments

*Some forms of cancer are driven by
oncogenes with genetic alterations that
make them good therapeutic targets*

Targeted anticancer therapeutics

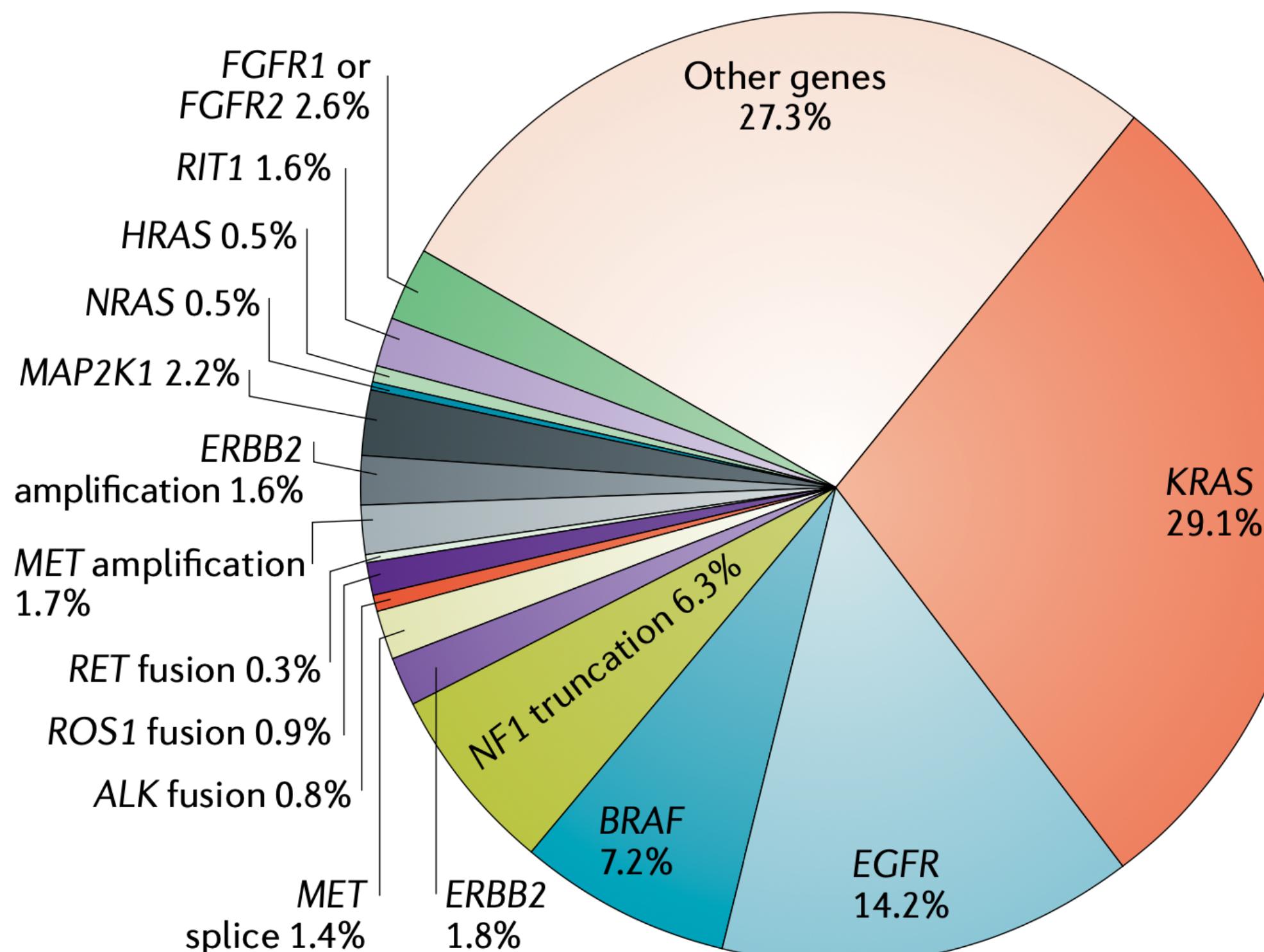
(Oncogene addiction; non-oncogene addiction; synthetic lethality)

- Bcr-Abl fusion in CML and ALL – ABLi (imatinib)
- EGFR mutations in lung cancer – EGFRi (osimertinib)
- BRAF mutations in melanoma – BRAFi/MEKi (vemurafenib/trametinib)
- HER2-amplified BrCa – HER2i (lapatinib) and HER2-specific Ab (trastuzumab)
- KRas mutations in lung/pancreatic cancer – KRASi (sotorasib)
- BRCA1/2- and p53-deficient cancers – PARPi (olaparib)
- Many more

Targetable alterations in lung adenocarcinoma

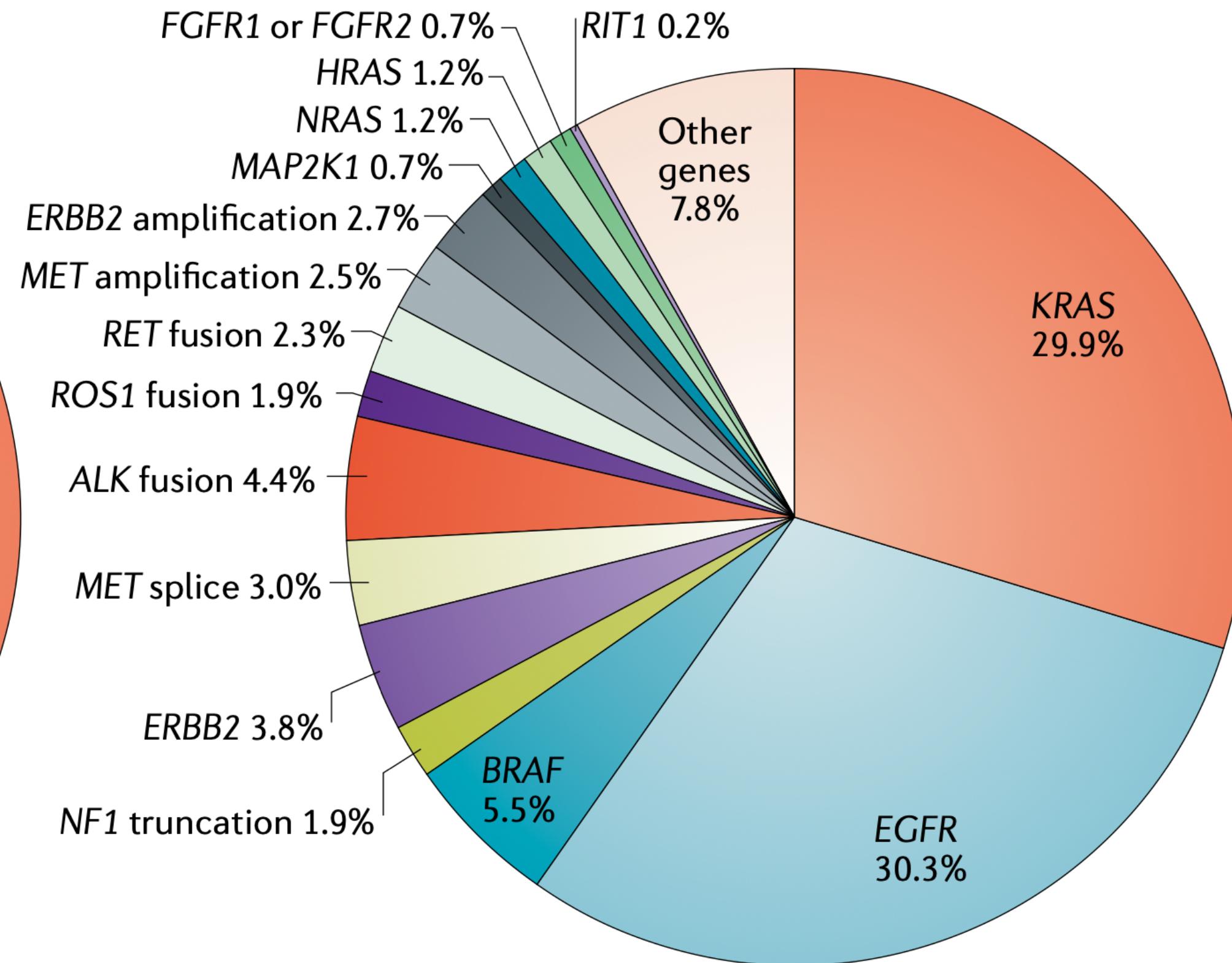
~50–66% of newly diagnosed patients with LuAd have alterations for which targeted therapeutics exist

a Early stage



Data from TCGA (Sanchez-Vega et al.¹⁷⁸, Ellrott et al.¹⁷⁹ and Hoadley et al.¹⁸⁰), Imielinski et al.⁶² and Kadara et al.¹³³ ($n = 741$)

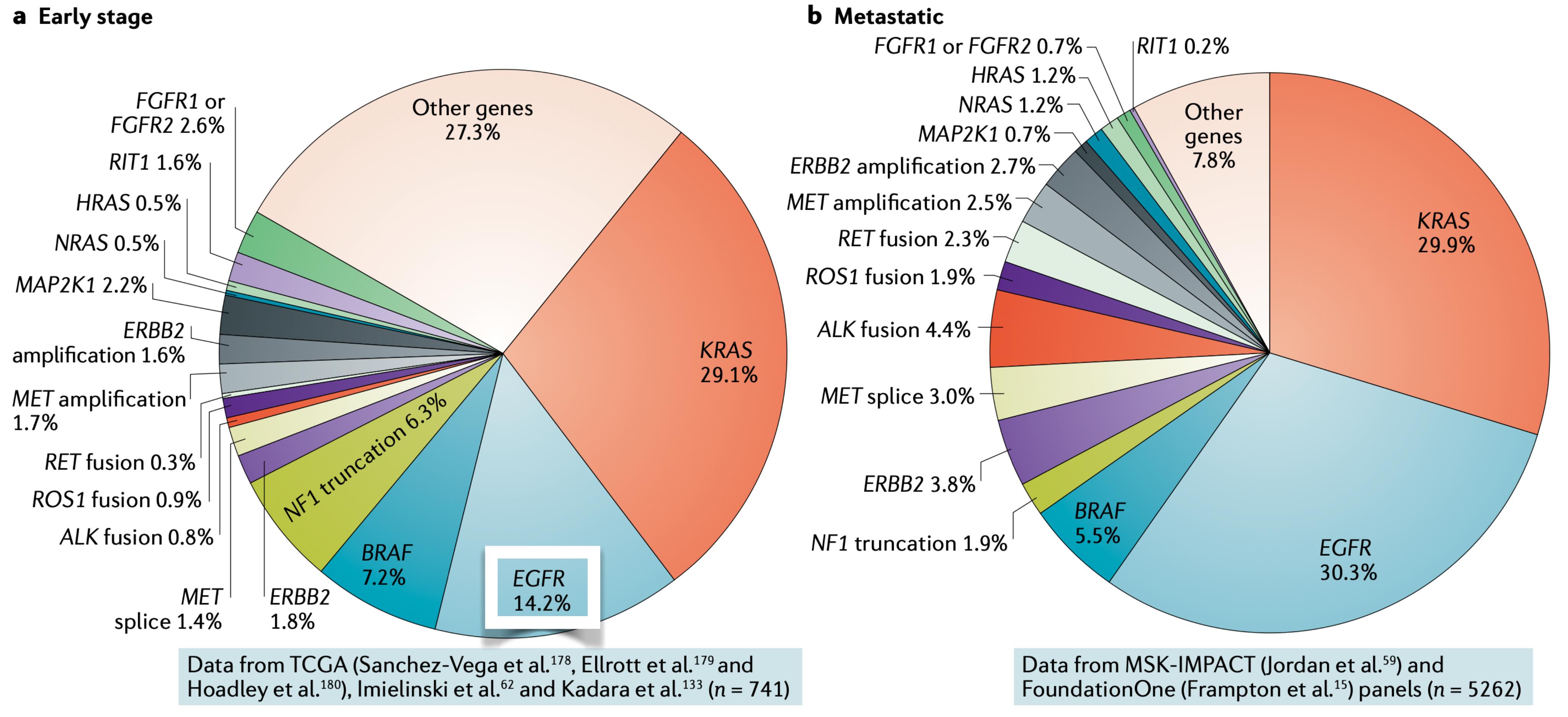
b Metastatic



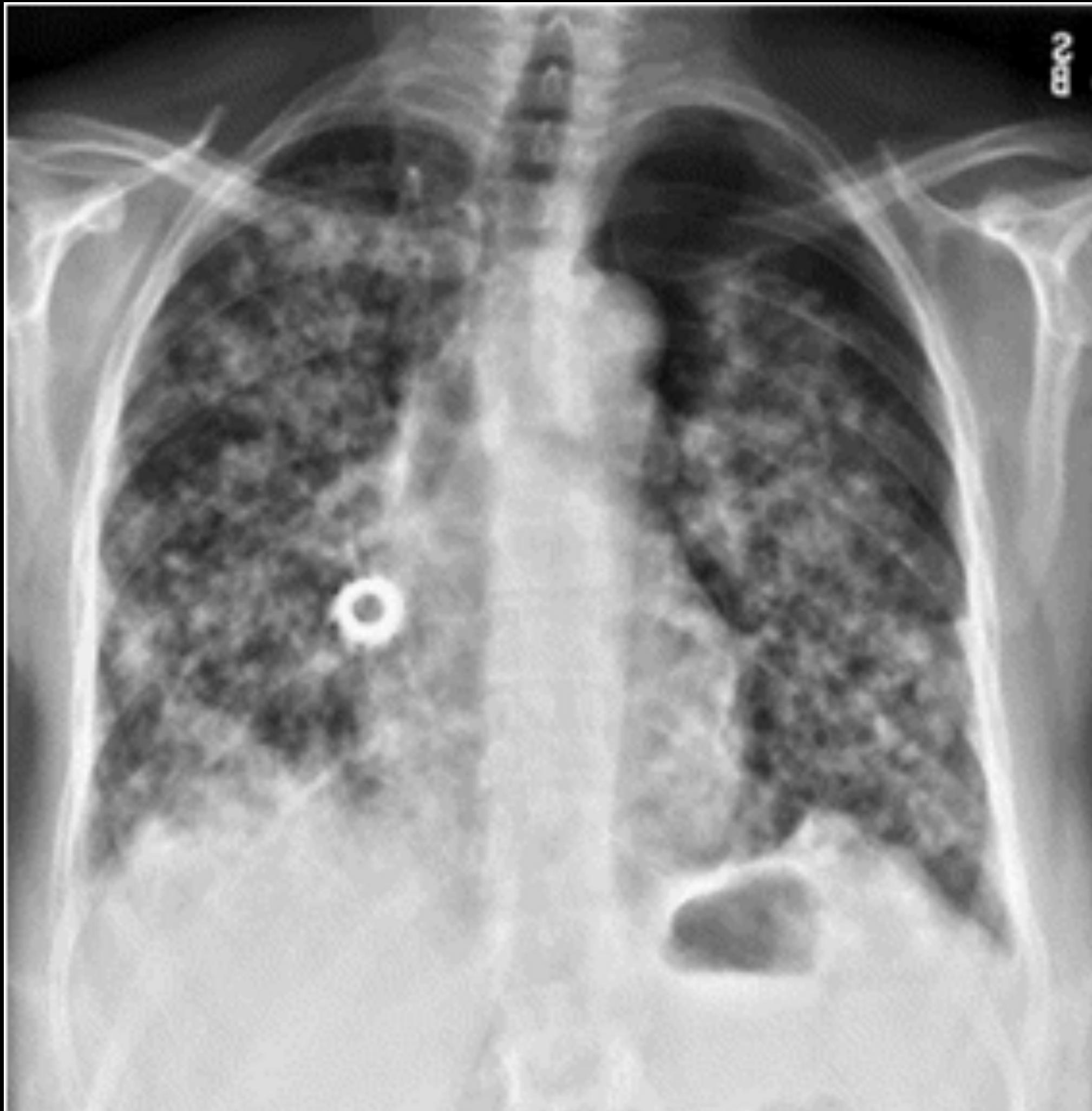
Data from MSK-IMPACT (Jordan et al.⁵⁹) and FoundationOne (Frampton et al.¹⁵) panels ($n = 5262$)

Targetable alterations in lung adenocarcinoma

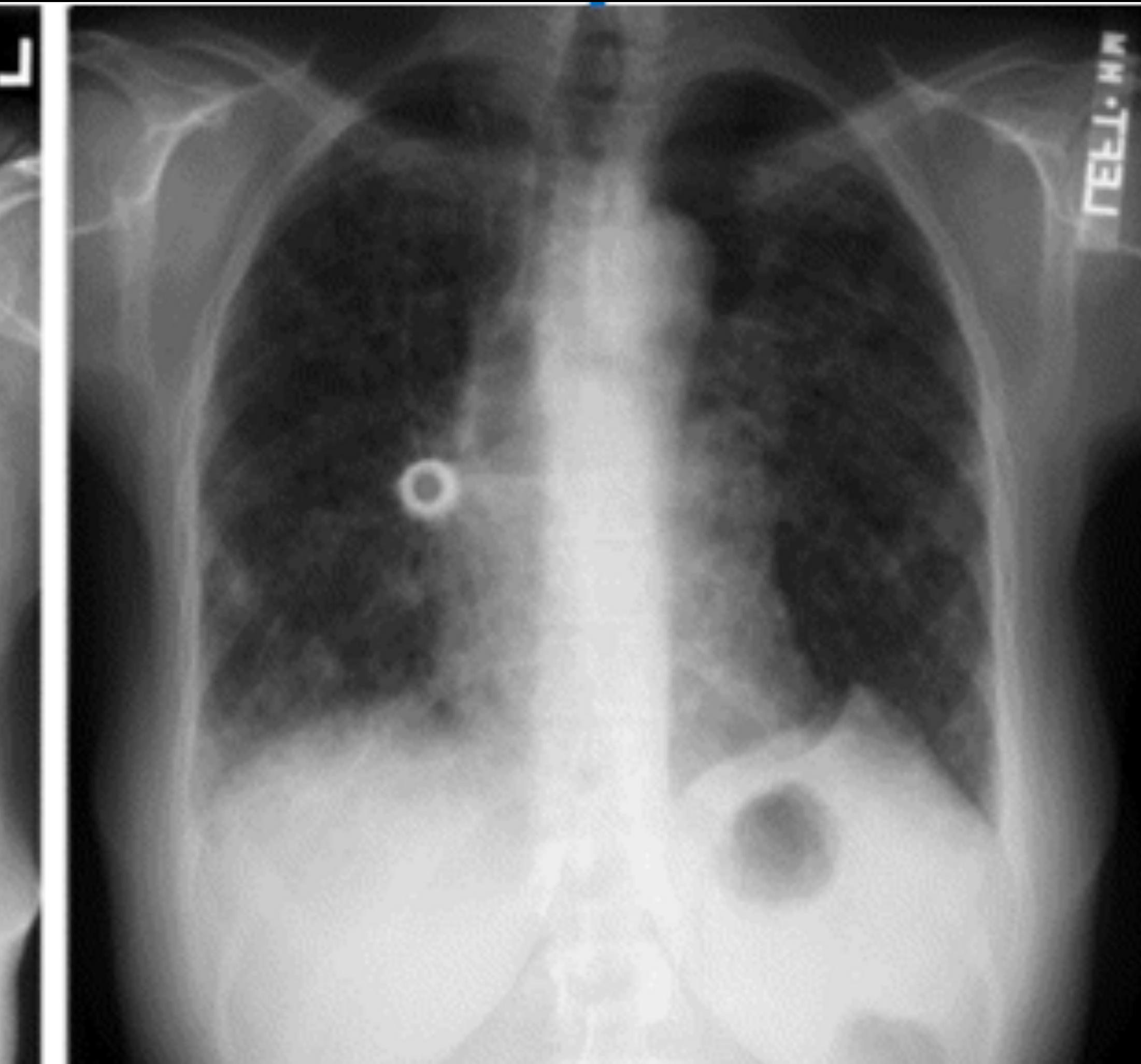
~50–66% of newly diagnosed patients with LuAd have alterations for which targeted therapeutics exist



Therapeutic Progress

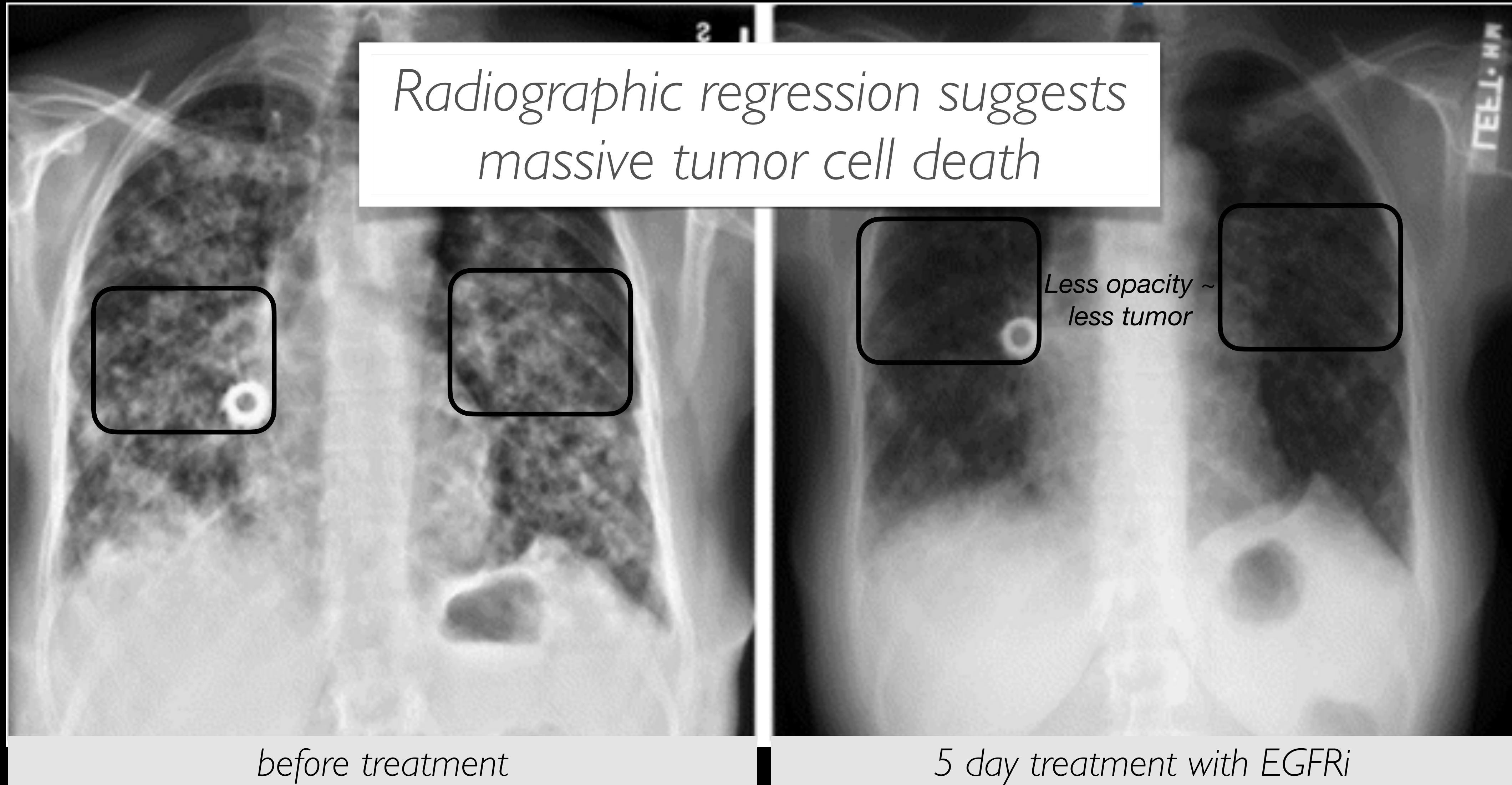


before treatment

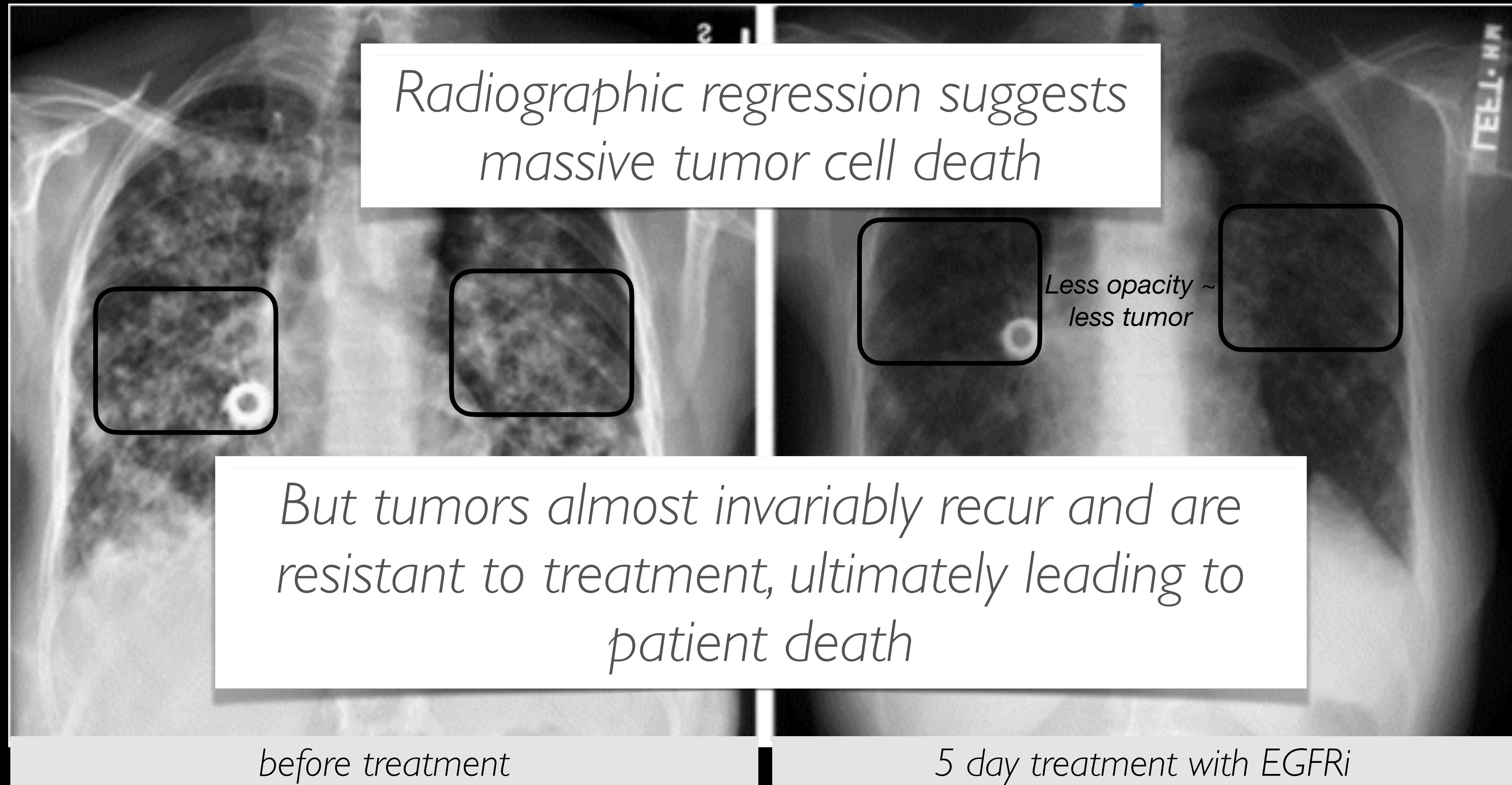


5 day treatment with EGFRi

Therapeutic Progress



Therapeutic Progress



CANCER IS A COMPLEX DYNAMIC PROCESS
AND CELLULAR HETEROGENEITY
IS THOUGHT TO DRIVE RESISTANCE

CANCER IS A COMPLEX DYNAMIC PROCESS
AND CELLULAR HETEROGENEITY
IS THOUGHT TO DRIVE RESISTANCE

But, *how?*

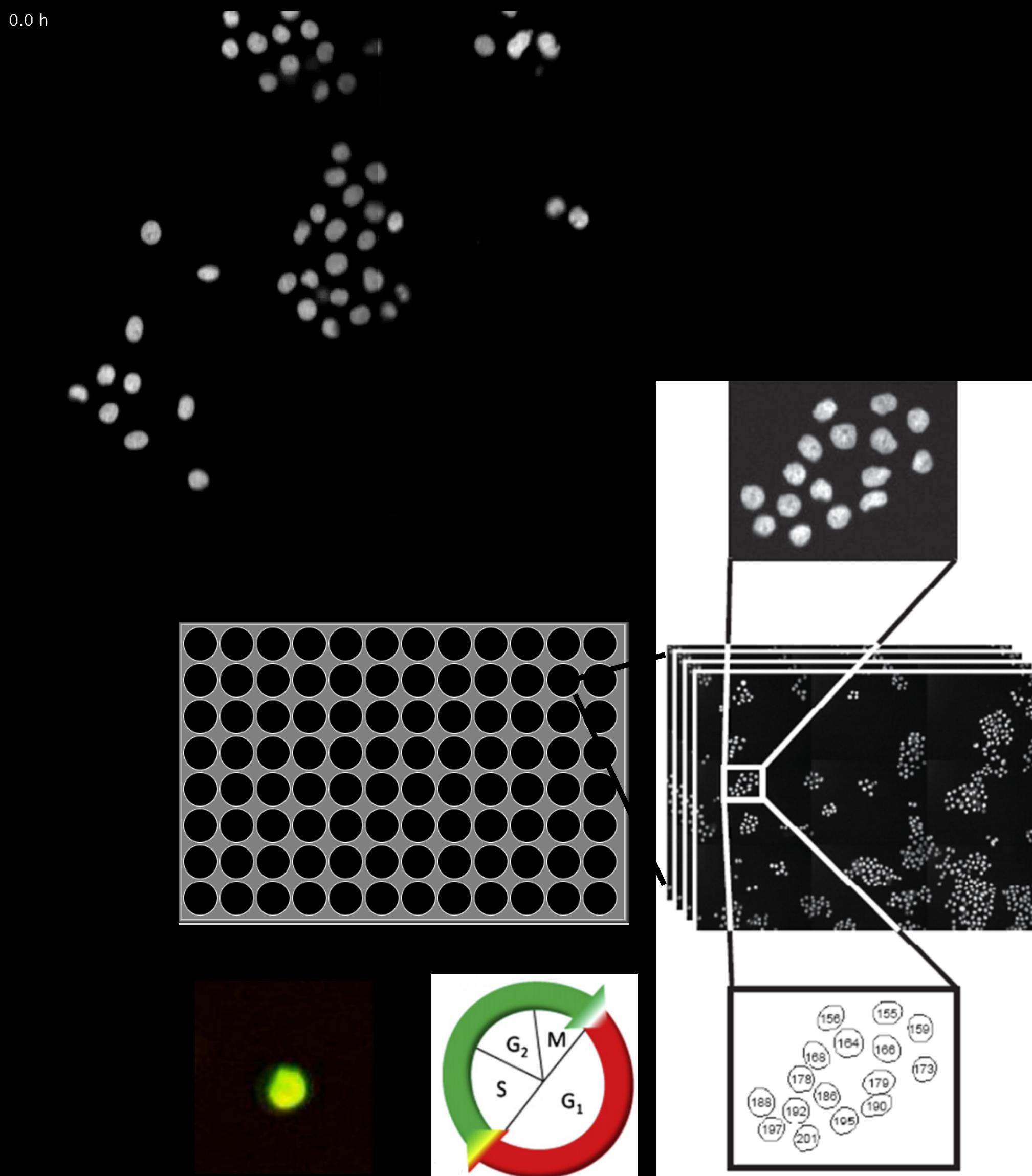
CANCER IS A COMPLEX DYNAMIC PROCESS
AND CELLULAR HETEROGENEITY
IS THOUGHT TO DRIVE RESISTANCE

But, *how?*

Our understanding of how *individual* cells
respond to treatment has been lacking

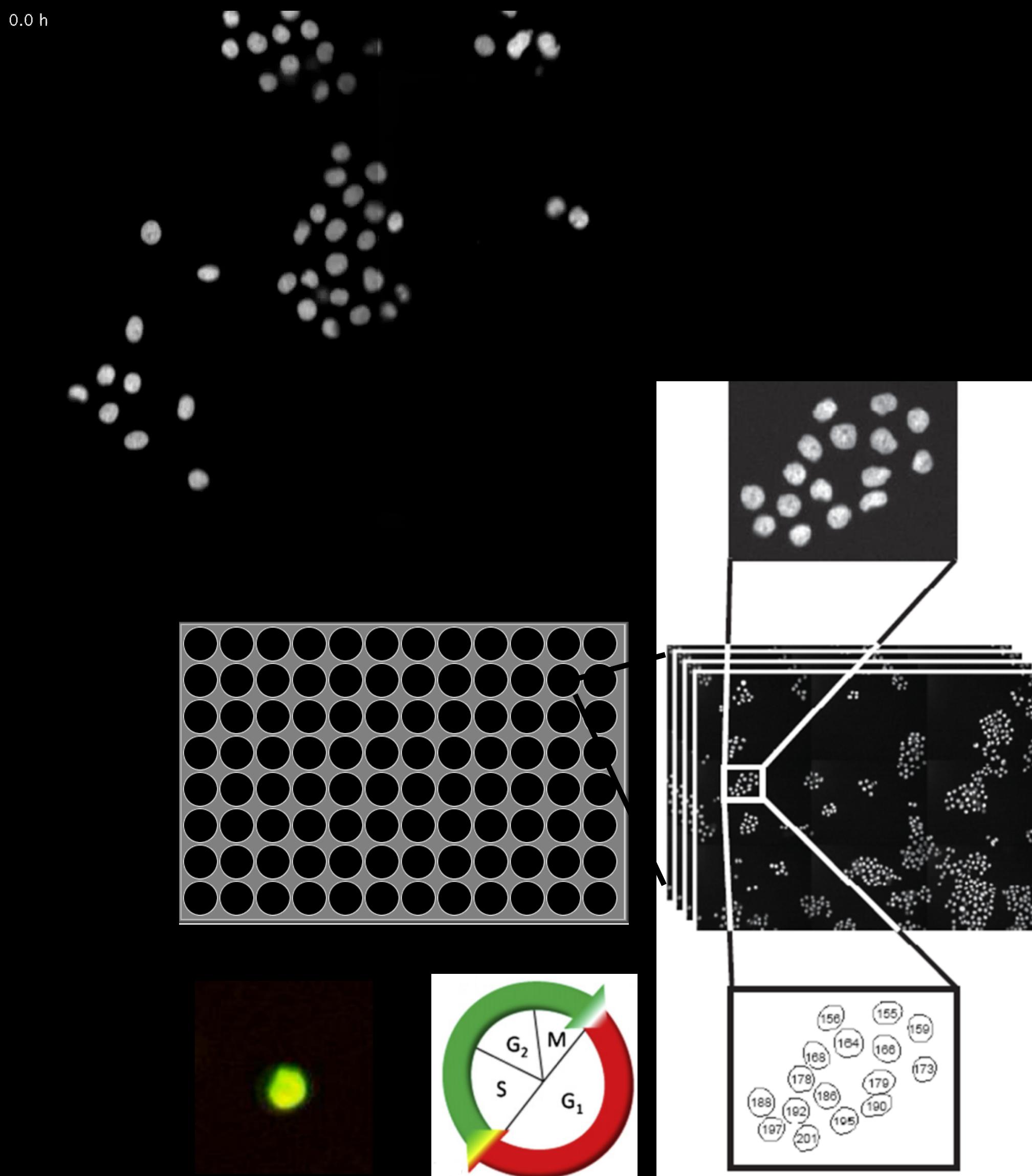
Quantifying Cellular Response Dynamics

- Use relevant cell lines, e.g., EGFR-mut lung cancer
- Genetically encode fluorescent nuclear label (may also include reporter(s) of cell function)
- Perform high-throughput, automated fluorescence microscopy under various treatment conditions
- Count nuclei and single-cell behaviors (birth-death events)
- Model population dynamics



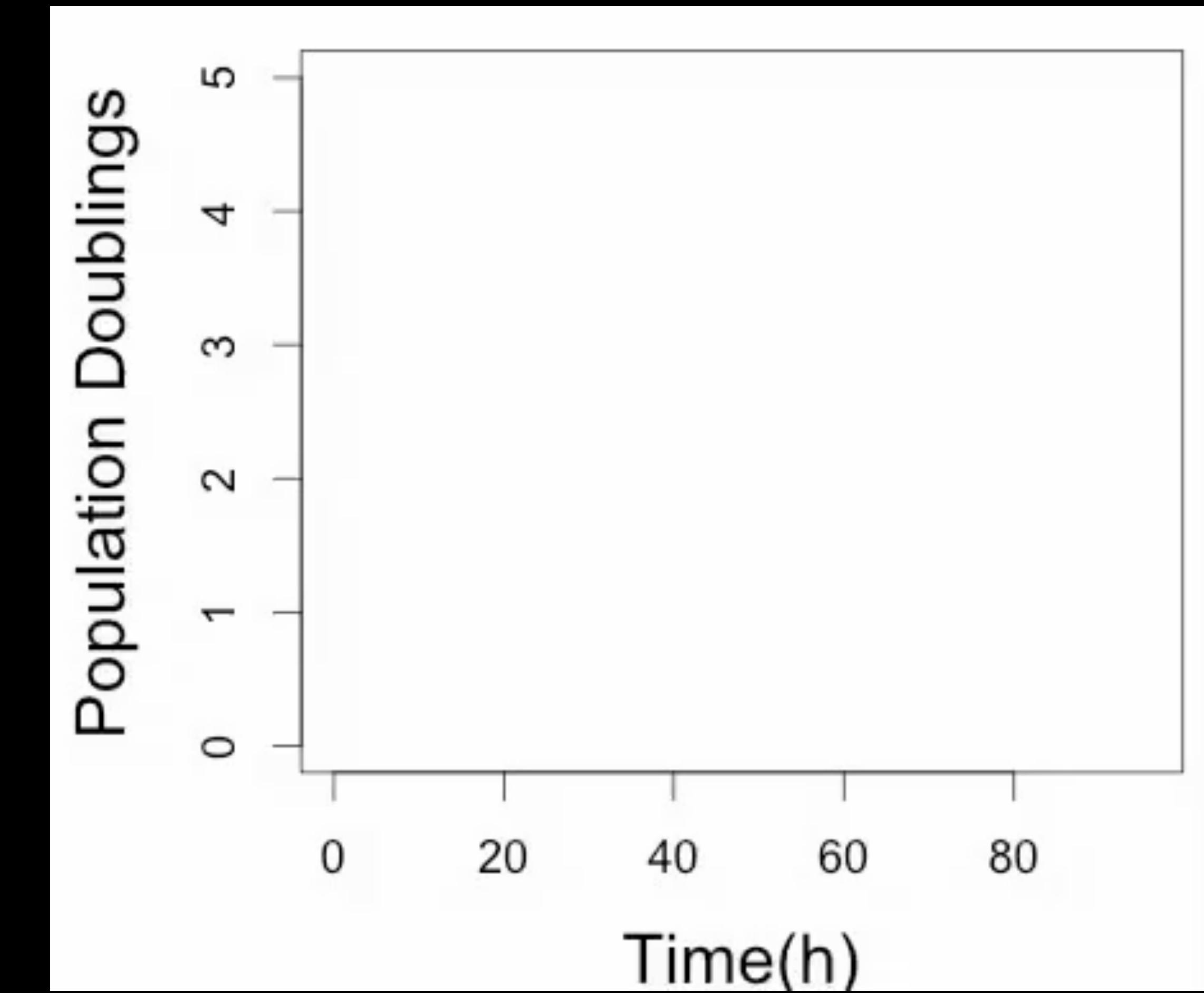
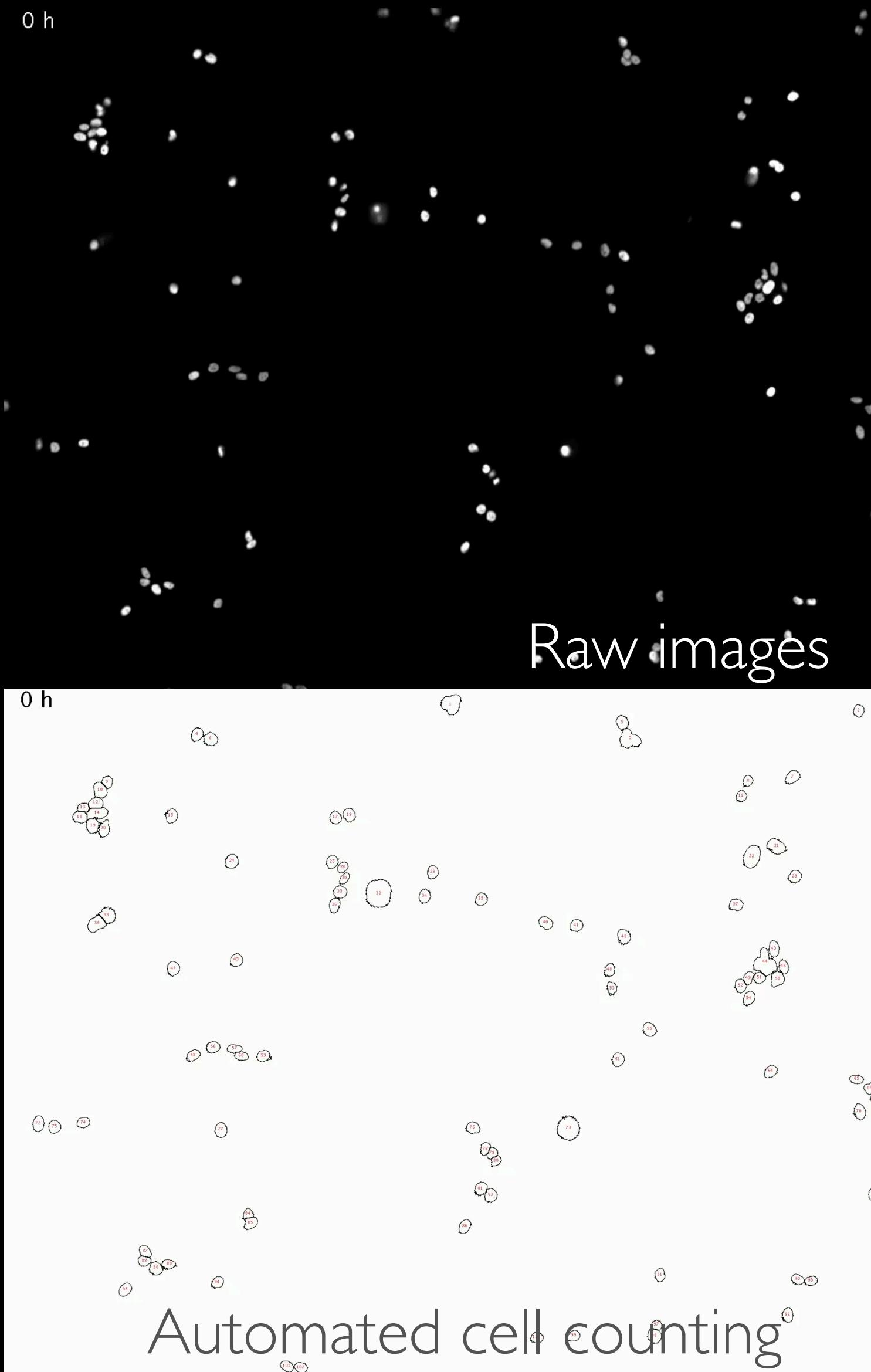
Quantifying Cellular Response Dynamics

- Use relevant cell lines, e.g., EGFR-mut lung cancer
- Genetically encode fluorescent nuclear label (may also include reporter(s) of cell function)
- Perform high-throughput, automated fluorescence microscopy under various treatment conditions
- Count nuclei and single-cell behaviors (birth-death events)
- Model population dynamics



POPULATION GROWTH DYNAMICS

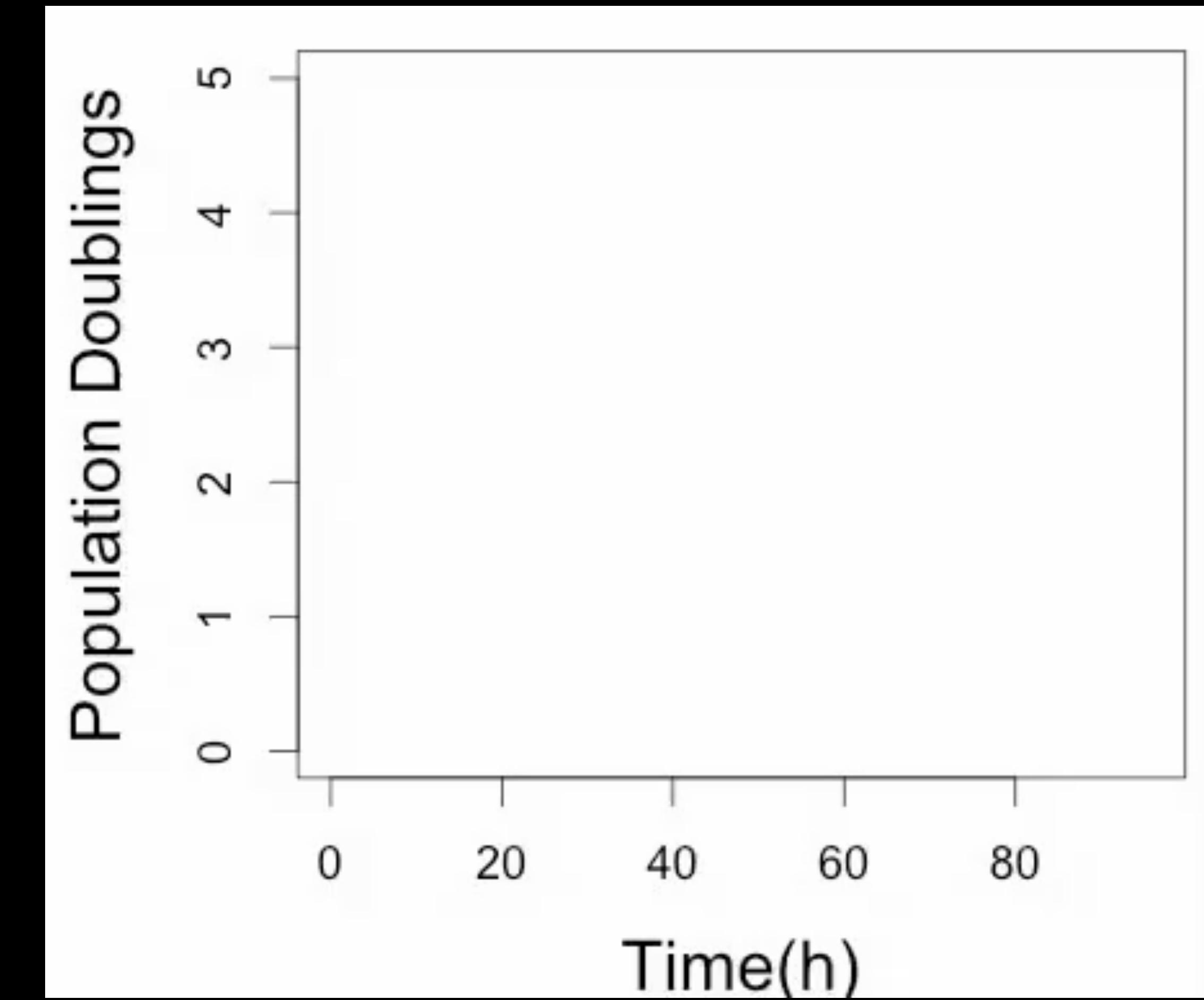
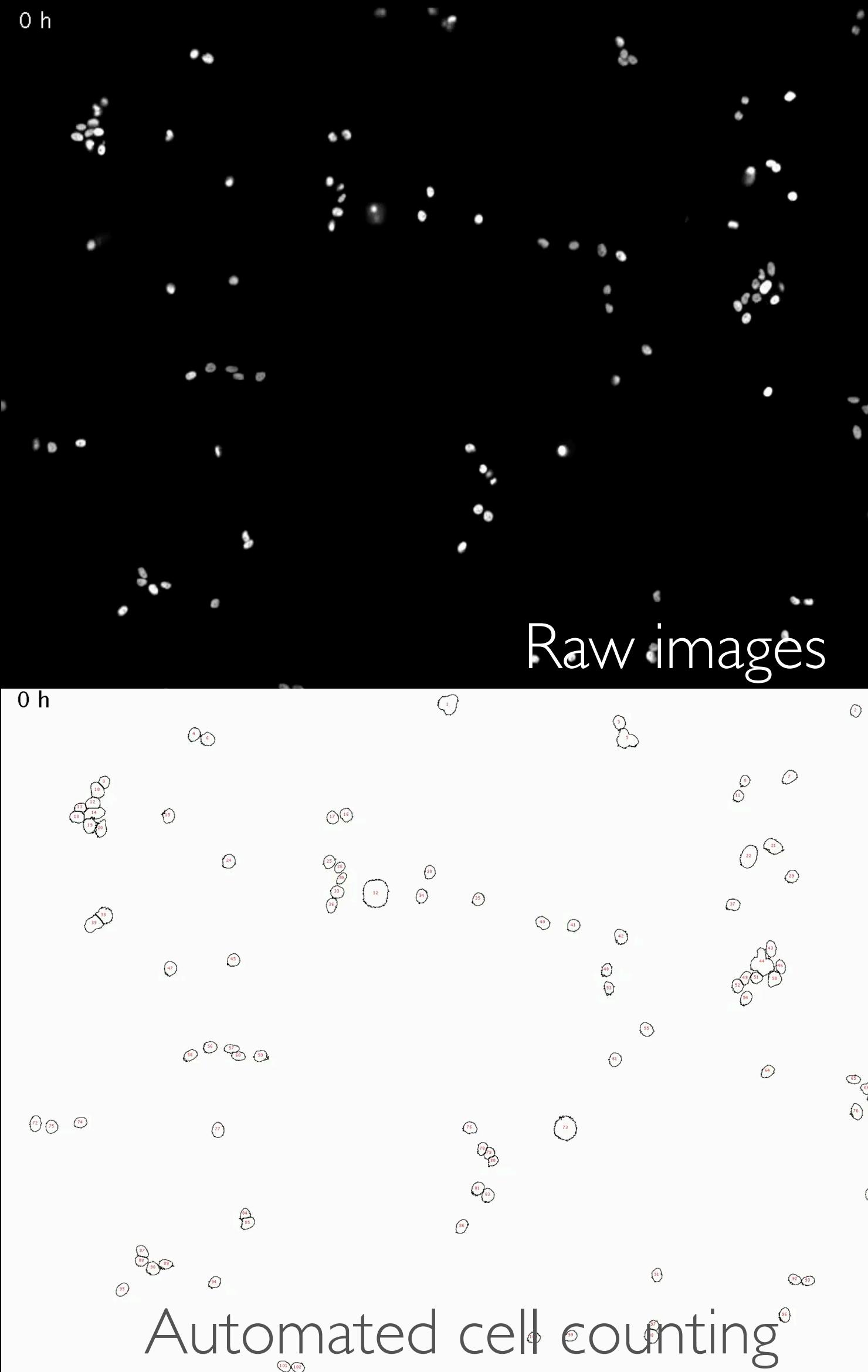
V



What is happening after ~60 h?

POPULATION GROWTH DYNAMICS

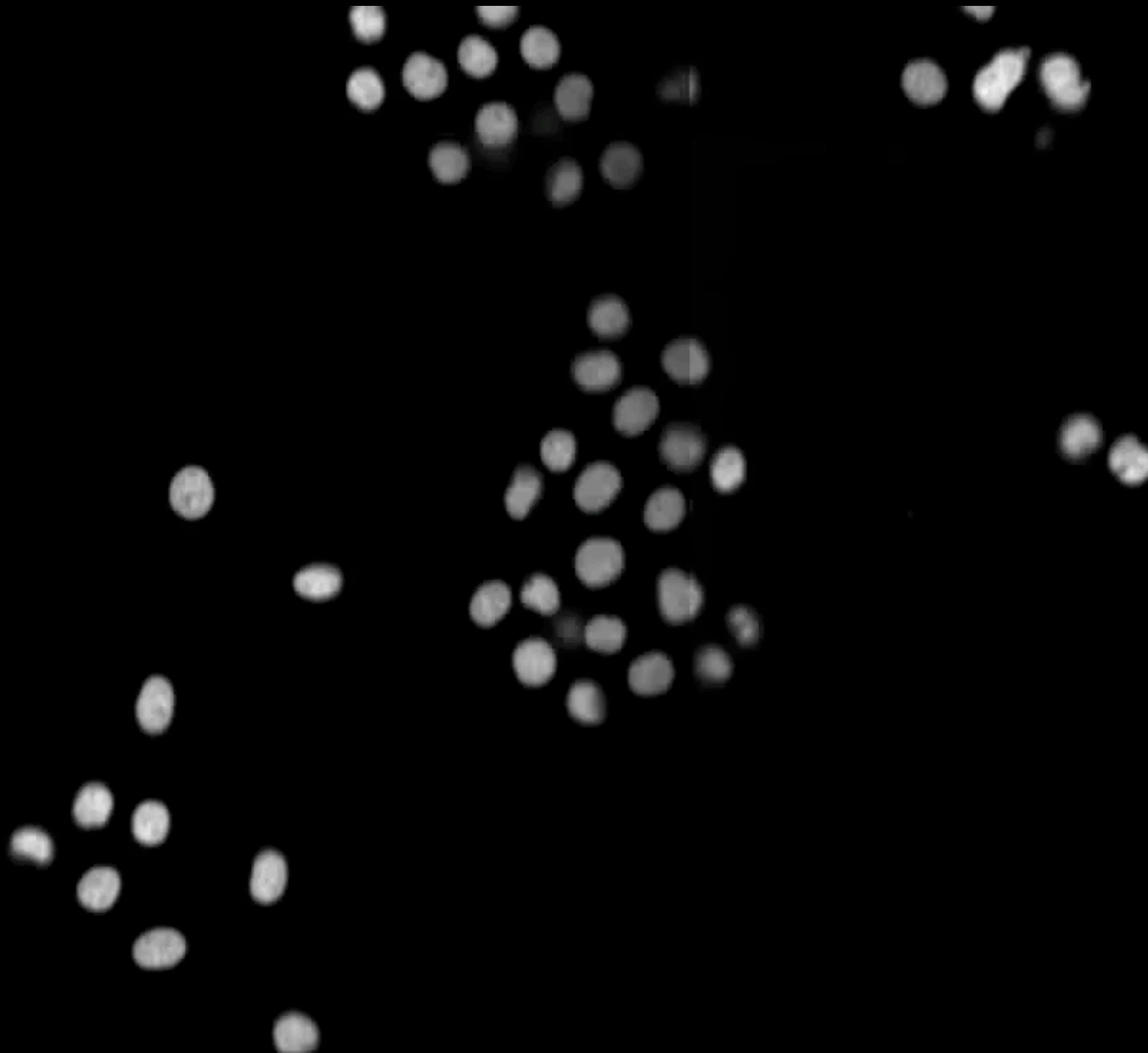
V



What is happening after ~60 h?

Single-cell fate decisions

0.0 h



Expectation

- EGFRi treatment will result in cell death

Result

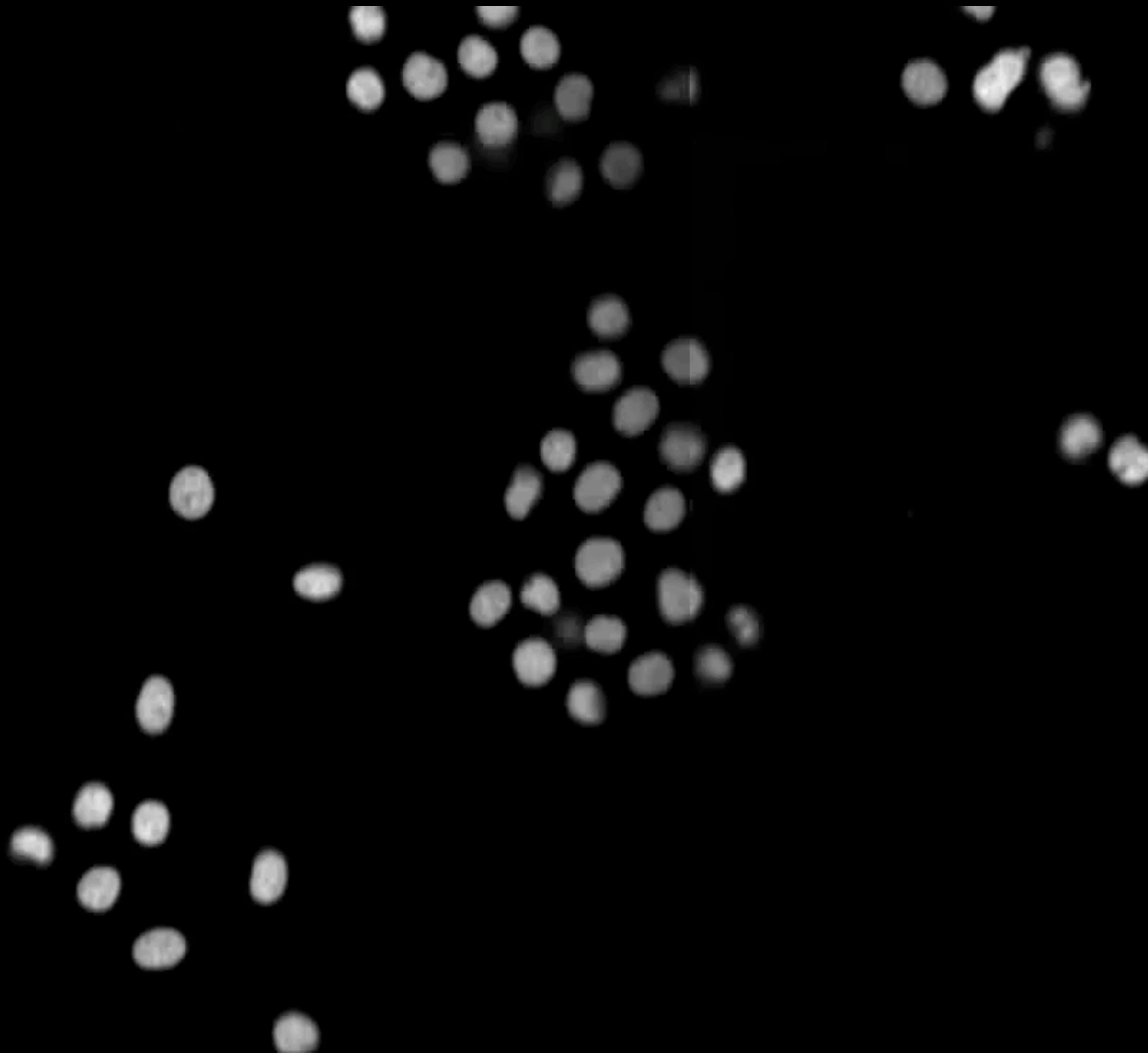
- EGFRi treatment caused different cell fates (death, division, cytostasis) even within the same condition
- Response dynamics are complex, due to time- and possibly cell cycle position-dependent effects

Ovals indicate cell divisions

Rectangles indicate cell death

Single-cell fate decisions

0.0 h



Expectation

- EGFRi treatment will result in cell death

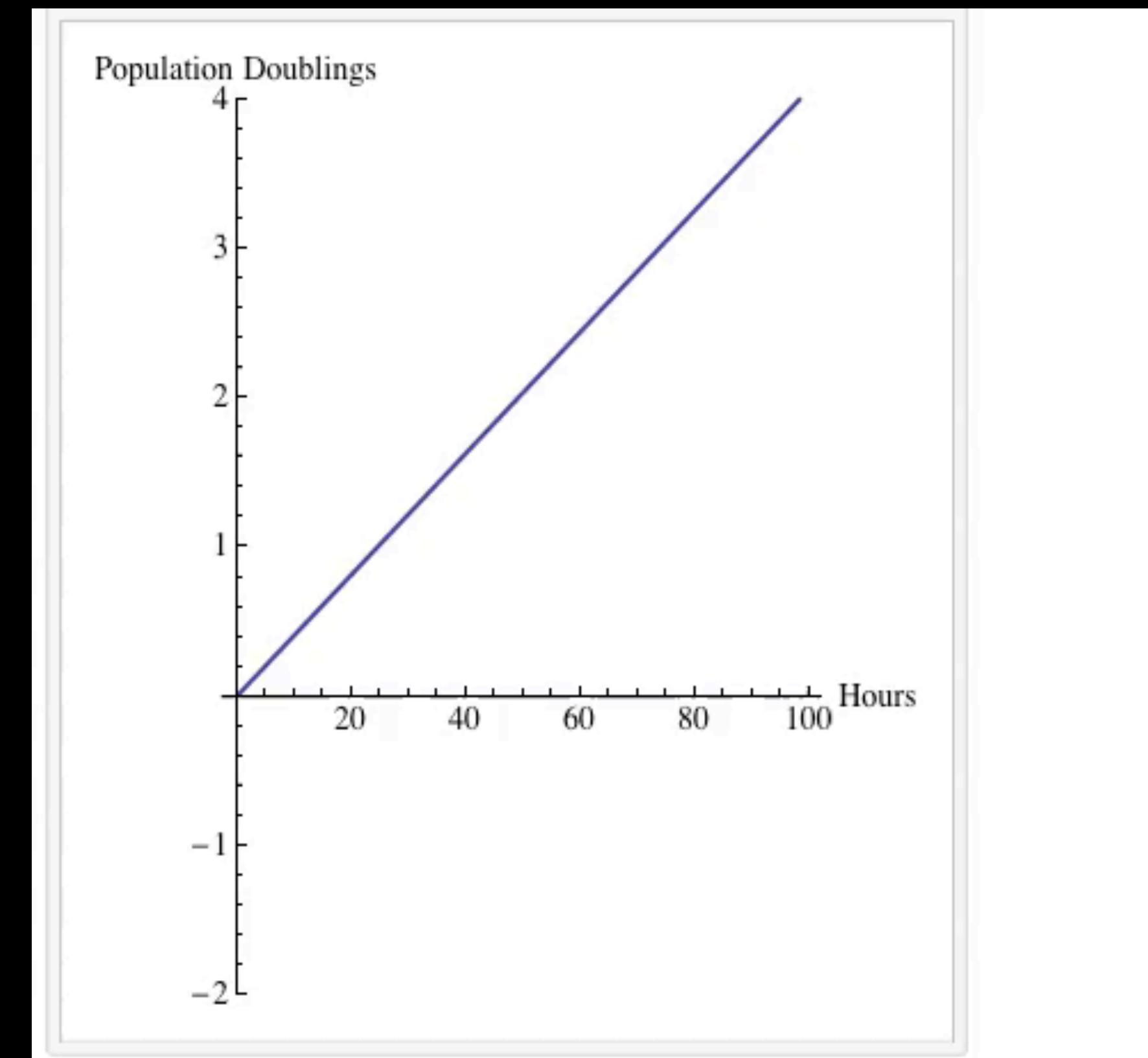
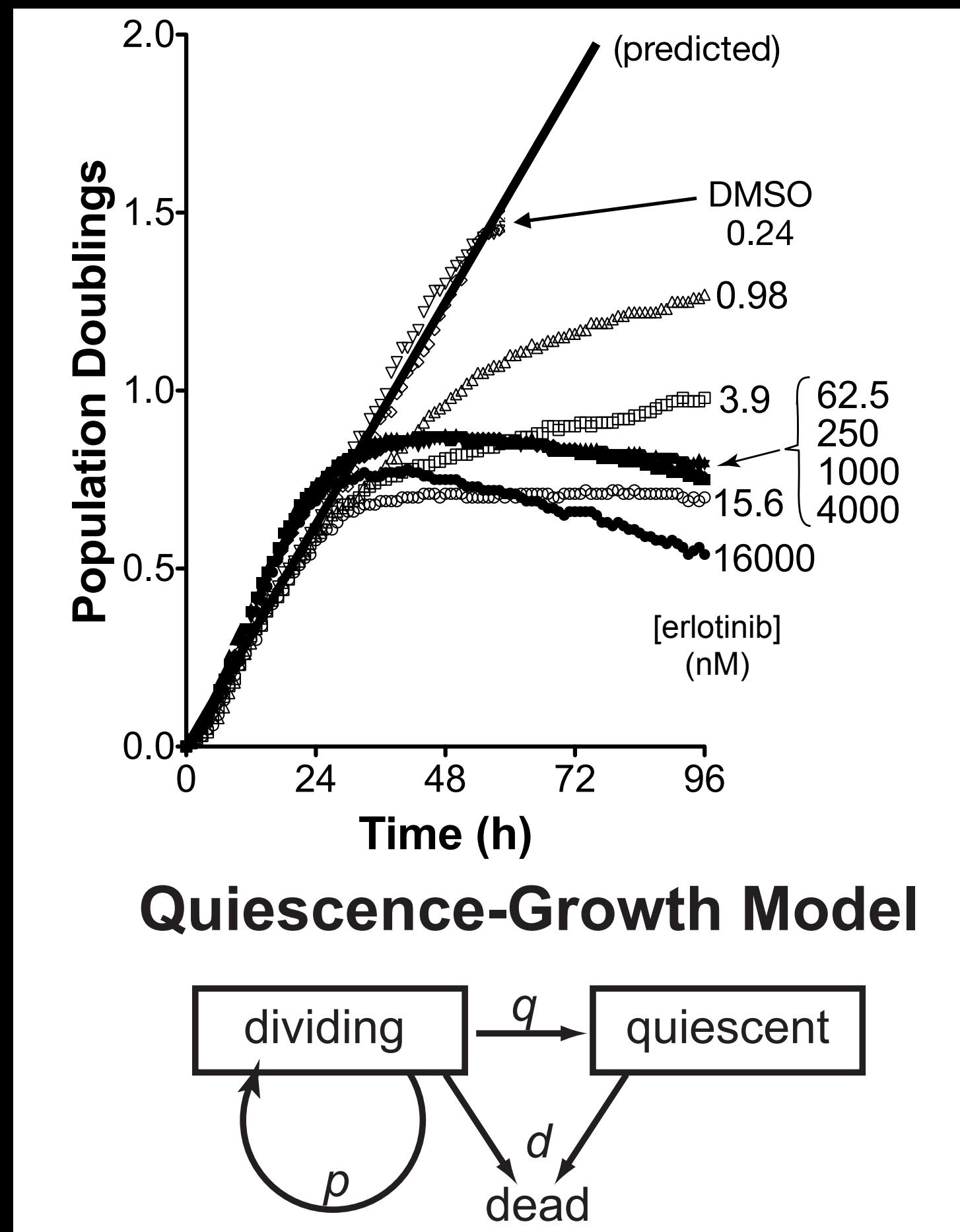
Result

- EGFRi treatment caused different cell fates (death, division, cytostasis) even within the same condition
- Response dynamics are complex, due to time- and possibly cell cycle position-dependent effects

Ovals indicate cell divisions

Rectangles indicate cell death

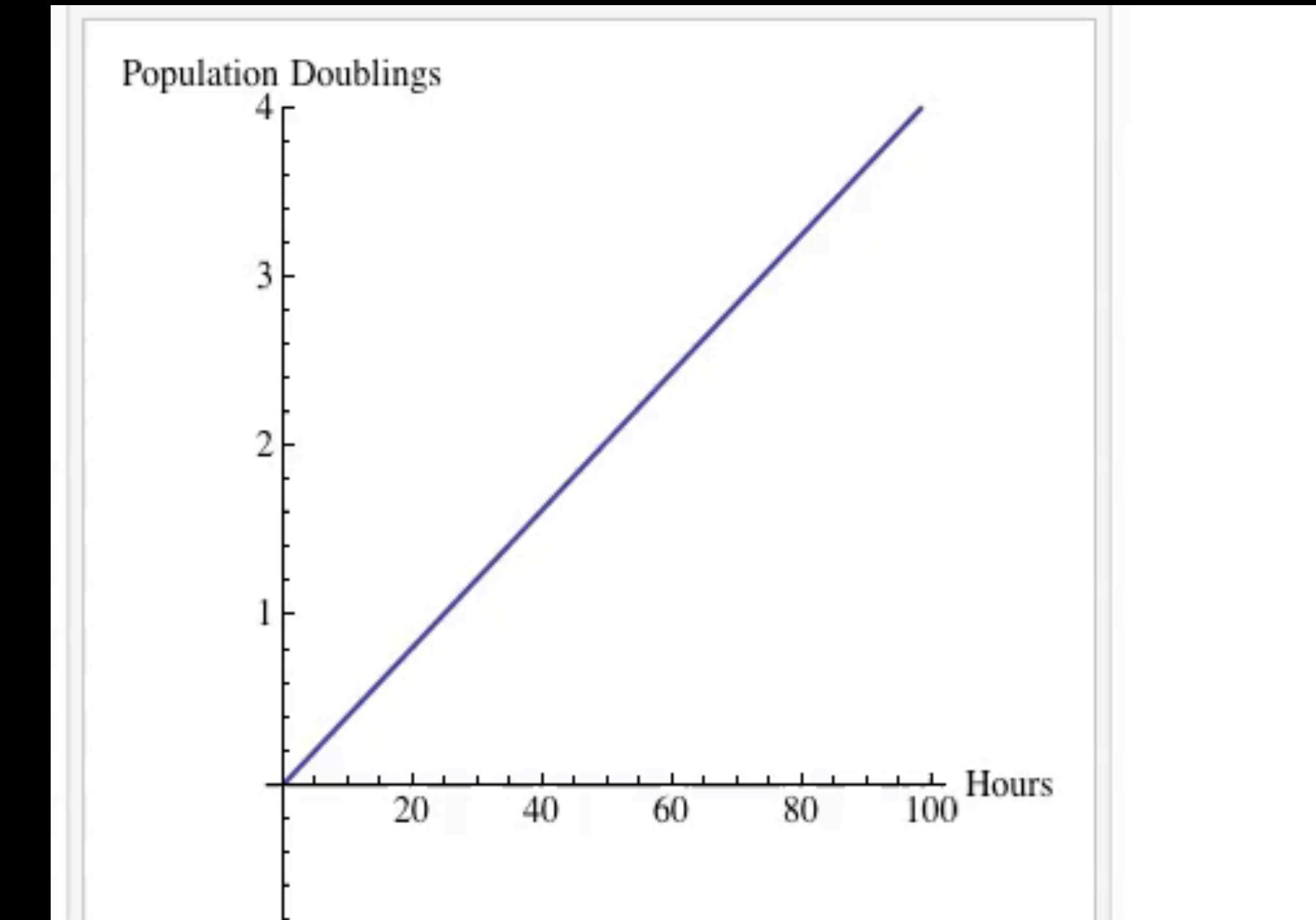
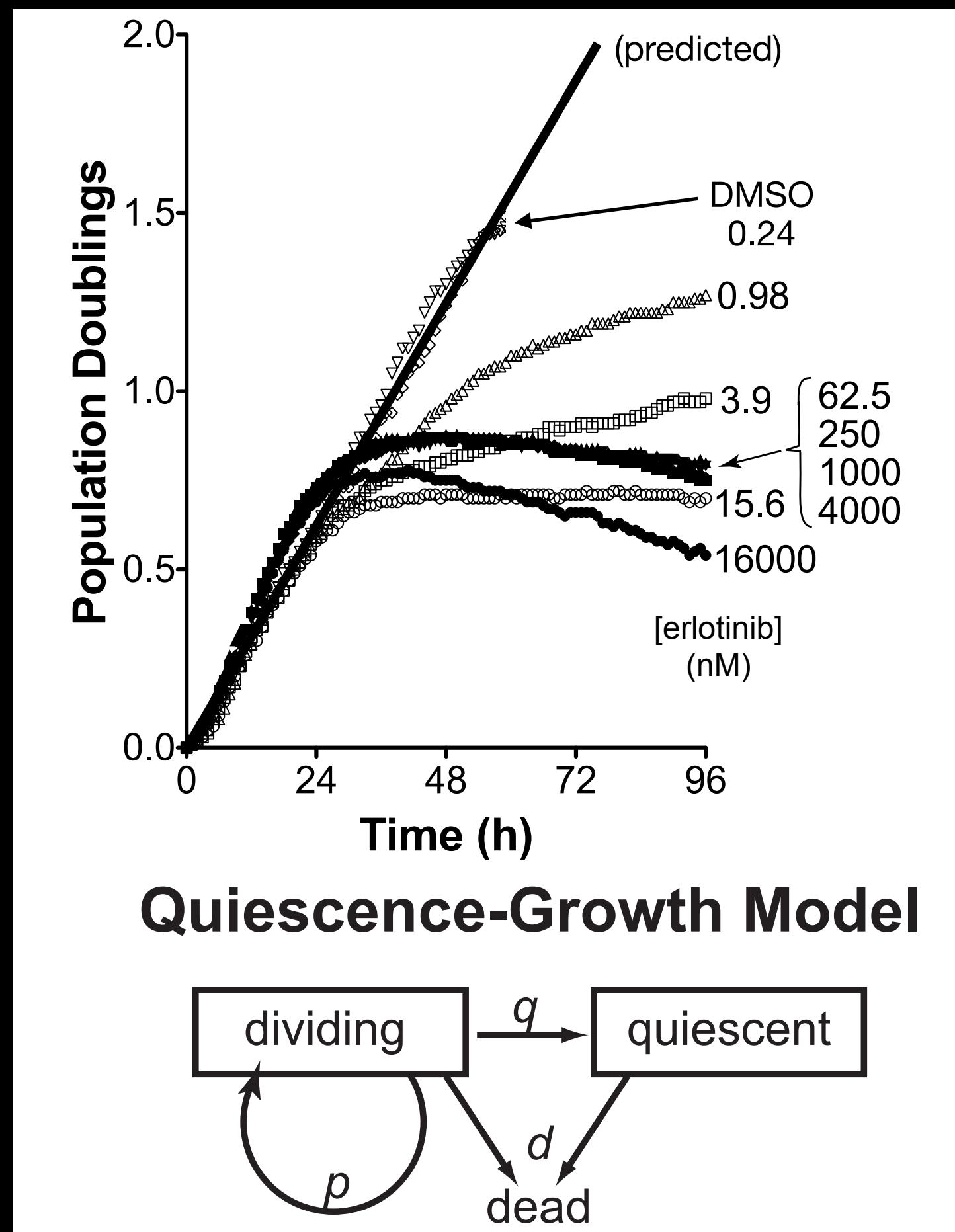
Our model approximates cell population growth dynamics



$$y(t|p, q, d) = y(0)[e^{(p-q-d)t} + \frac{q}{q-p} e^{-dt}(1 - e^{(p-q)t})]$$

dividing
 quiescent

Our model approximates cell population growth dynamics

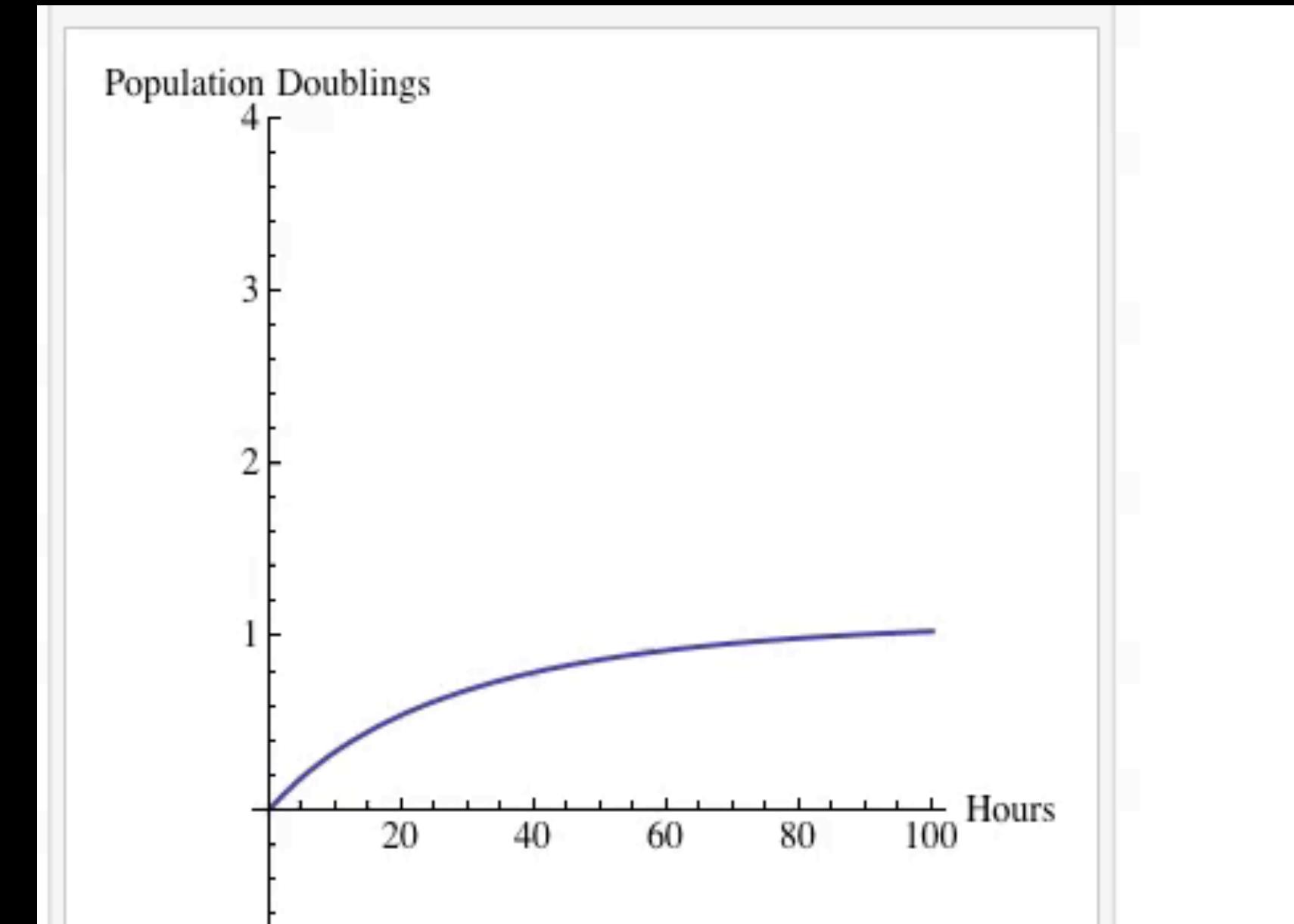
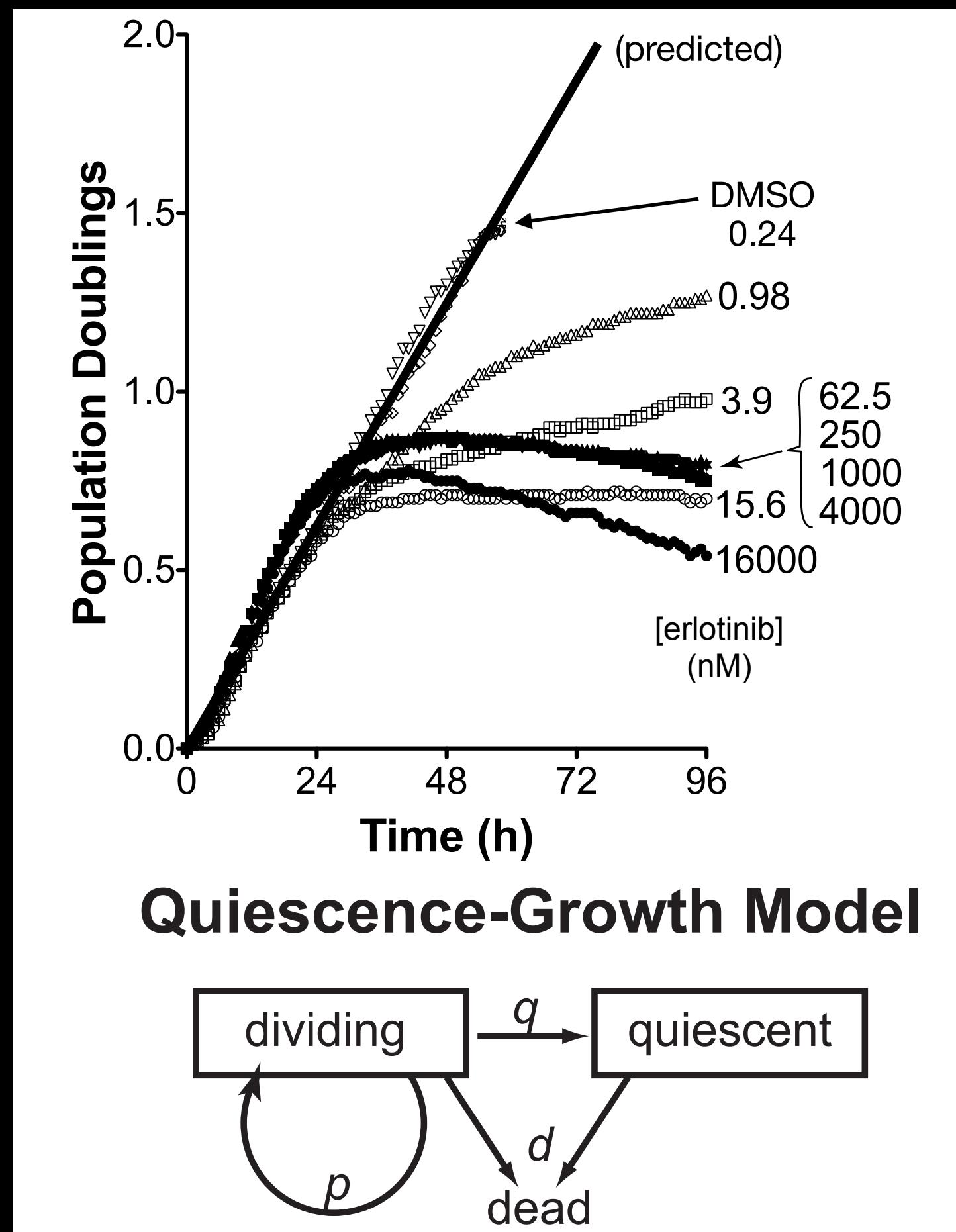


division or death rates only affect slope of line

$$y(t|p, q, d) = y(0)[e^{(p-q-d)t} + \frac{q}{q-p} e^{-dt}(1 - e^{(p-q)t})]$$

\downarrow *dividing* \downarrow *quiescent*

Our model approximates cell population growth dynamics

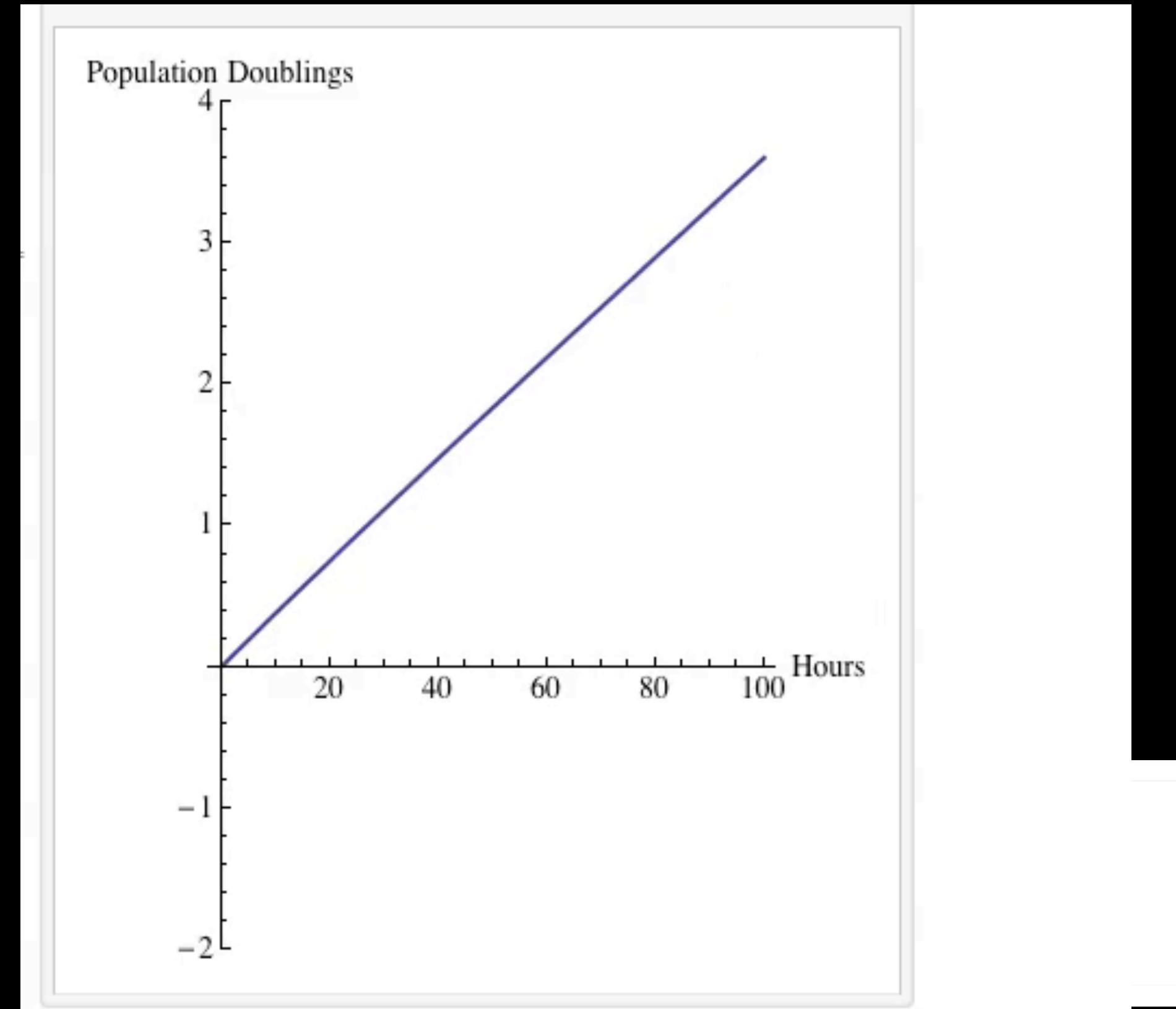
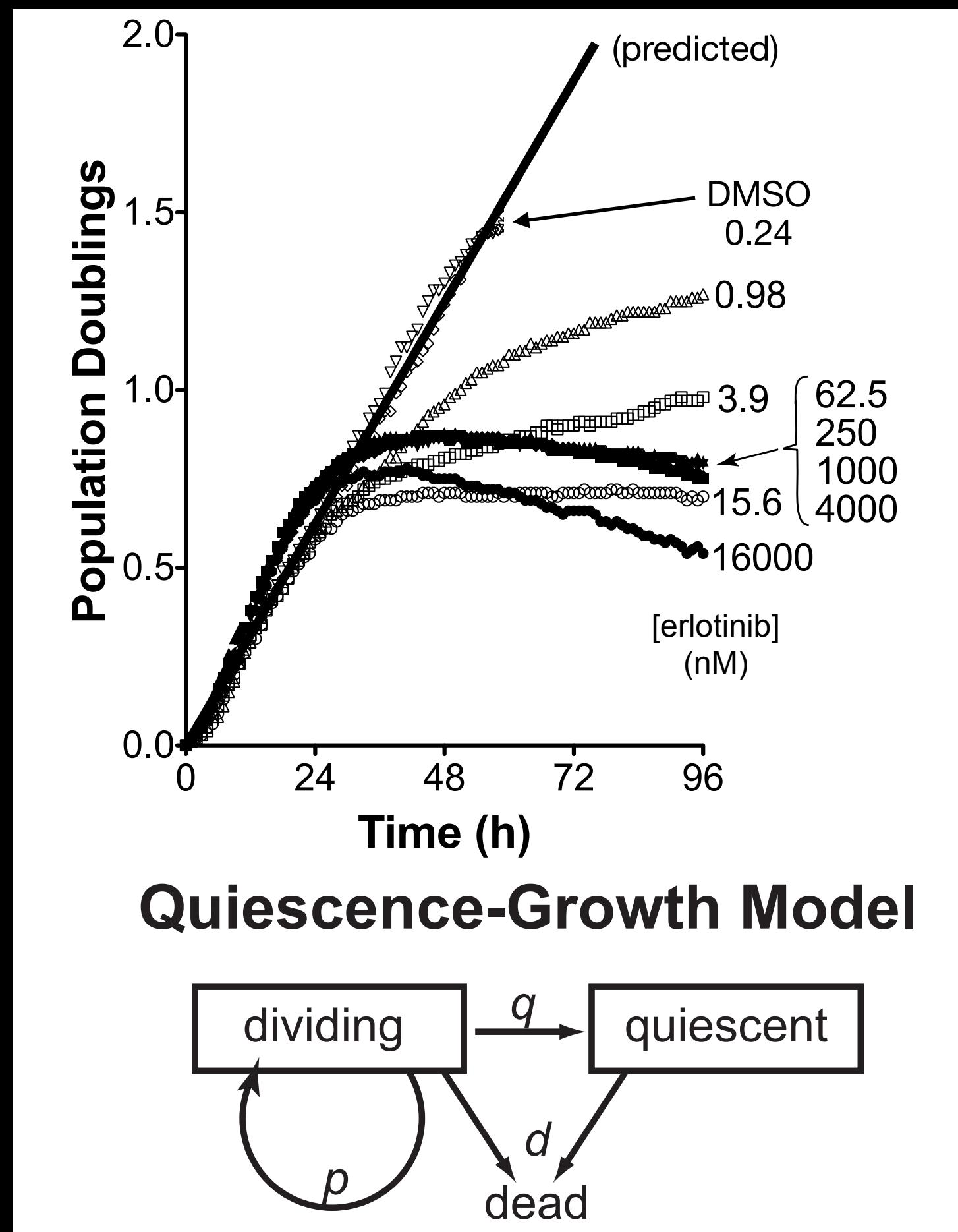


quiescence rate affects
nonlinearity

$$y(t|p, q, d) = y(0)[e^{(p-q-d)t} + \frac{q}{q-p} e^{-dt}(1 - e^{(p-q)t})]$$

dividing
 quiescent

Our model approximates cell population growth dynamics



$$y(t|p, q, d) = y(0)[e^{(p-q-d)t} + \frac{q}{q-p} e^{-dt}(1 - e^{(p-q)t})]$$

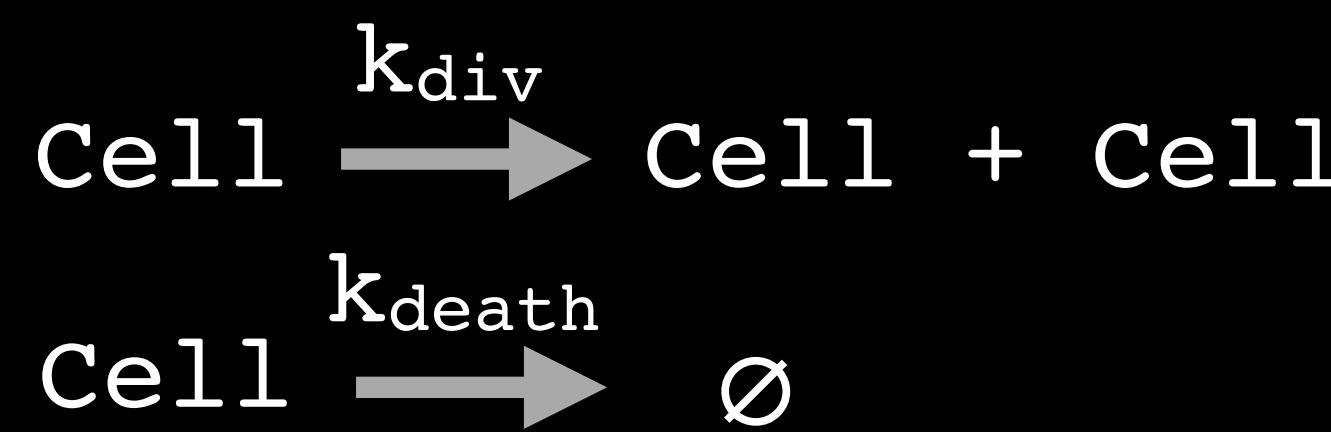
dividing
 quiescent

Anticancer drugs frequently
have a time-dependent effect:
*delayed transition to new stable
rate of population change*

Drug-Induced Proliferation (DIP) Rate

Steady-state rate of proliferation

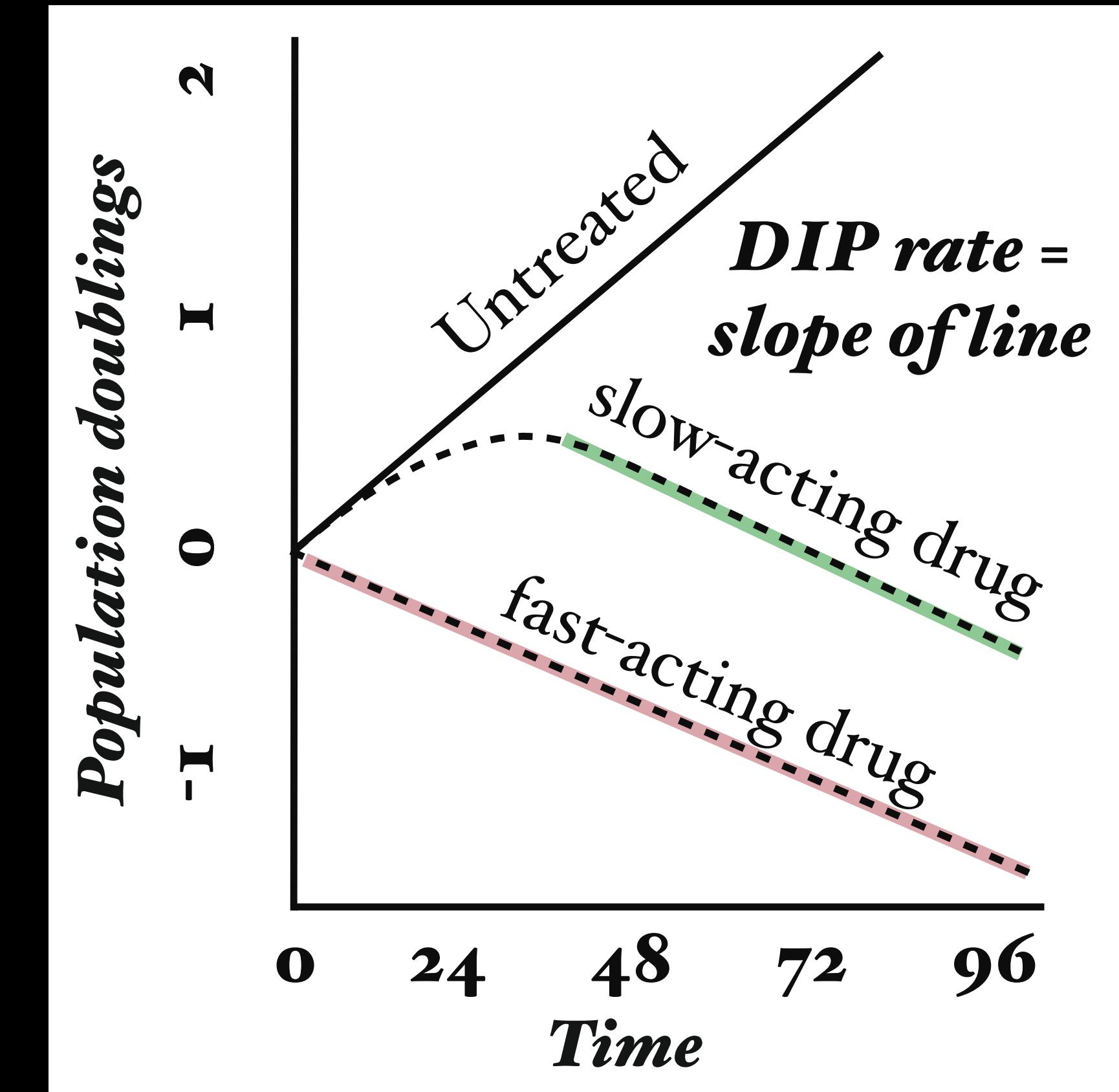
Simple mathematical model of
a birth-death process



$$\frac{d\ln\text{Cell}(t)}{dt} = k_{\text{div}} - k_{\text{death}}$$

(slope of line)

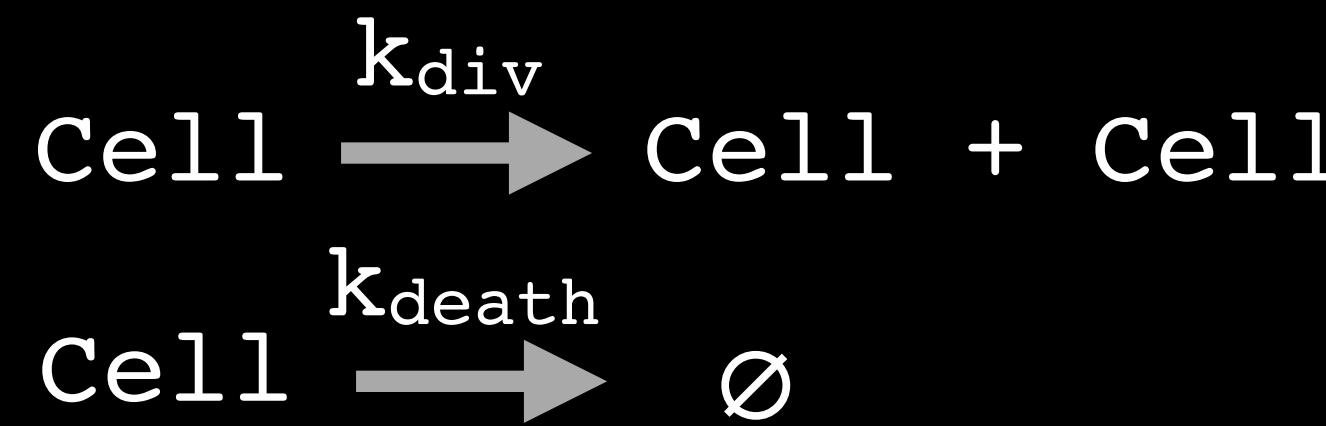
DIP rate



Drug-Induced Proliferation (DIP) Rate

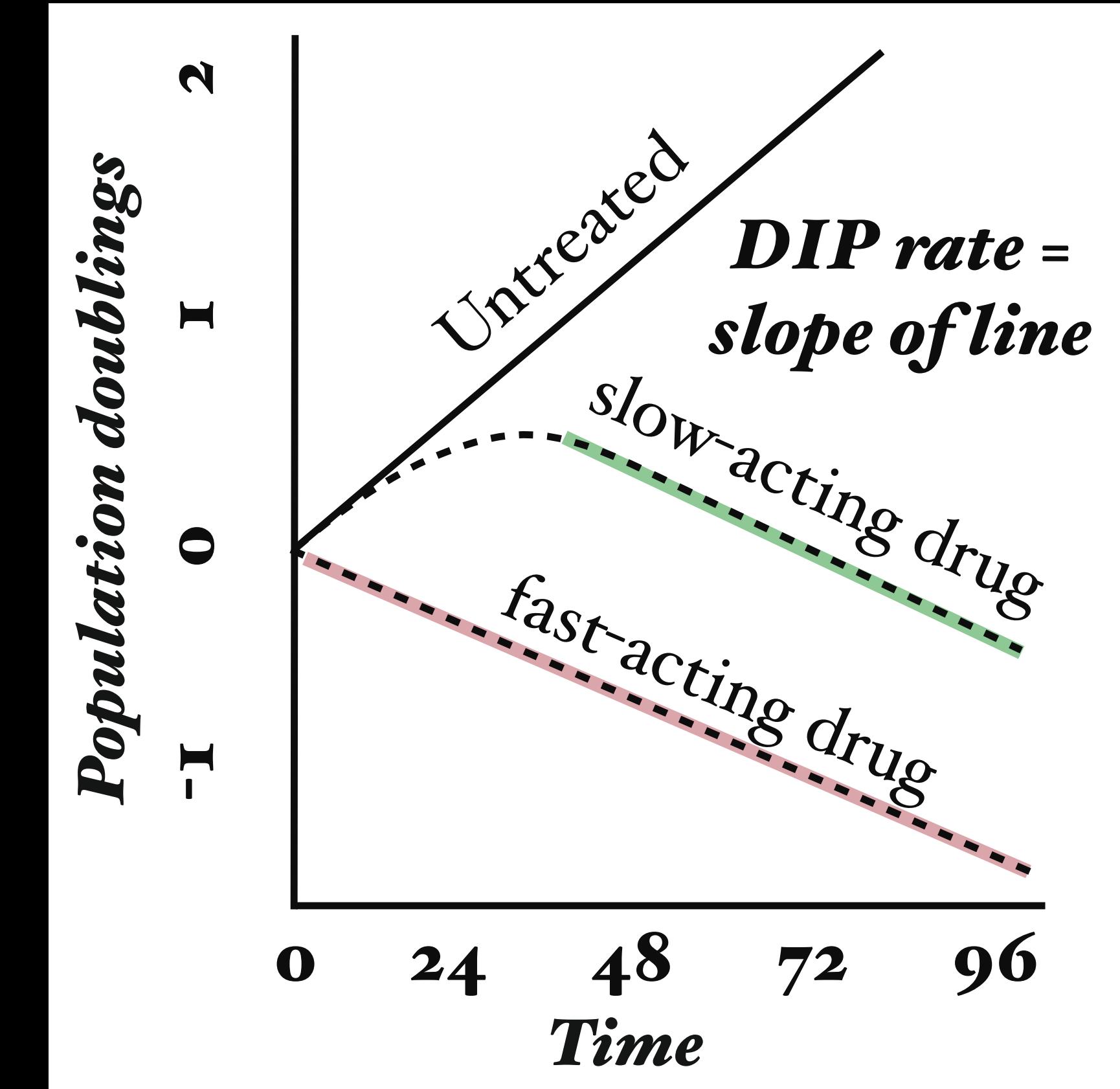
Steady-state rate of proliferation

Simple mathematical model of
a birth-death process



$$\frac{d\ln\text{Cell}(t)}{dt} = k_{\text{div}} - k_{\text{death}}$$

(slope of line) DIP rate

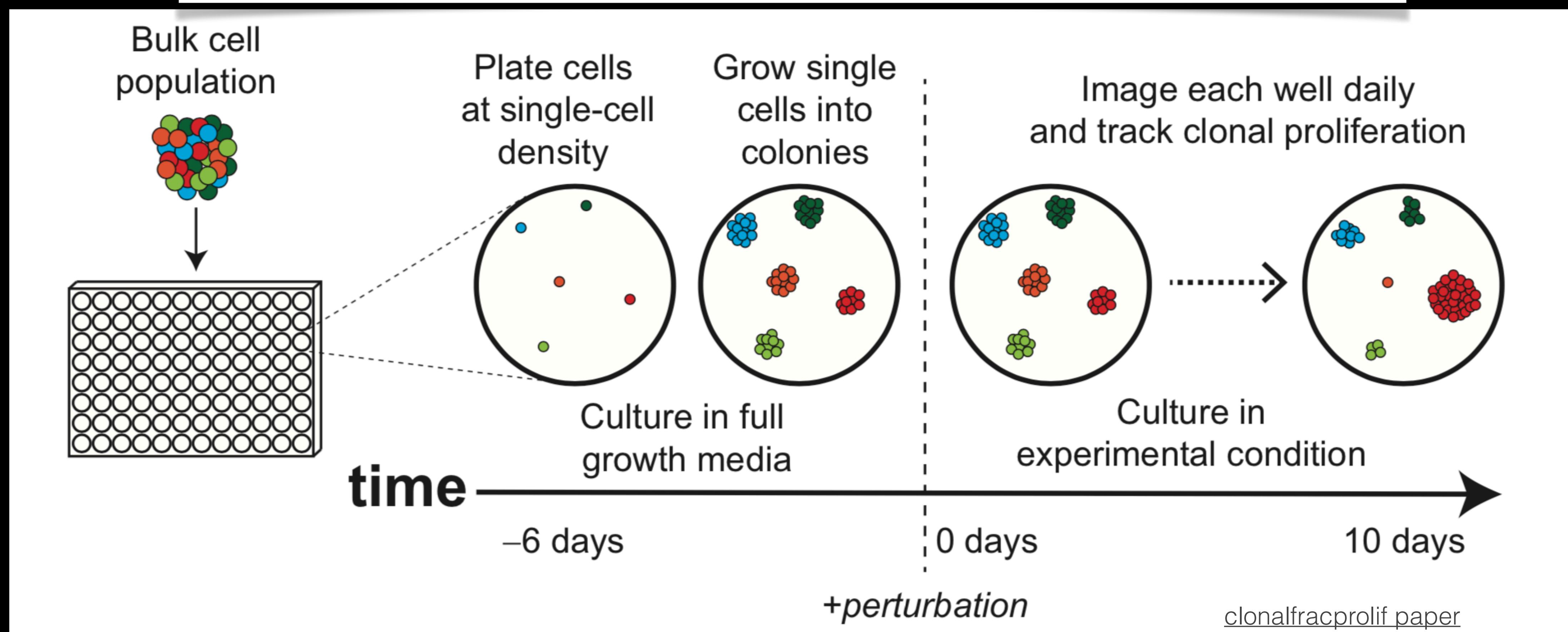


*DIP rate ~ fitness of the observed cell population
within a specific environment*

DIP Rate as a Metric of Clonal Fitness

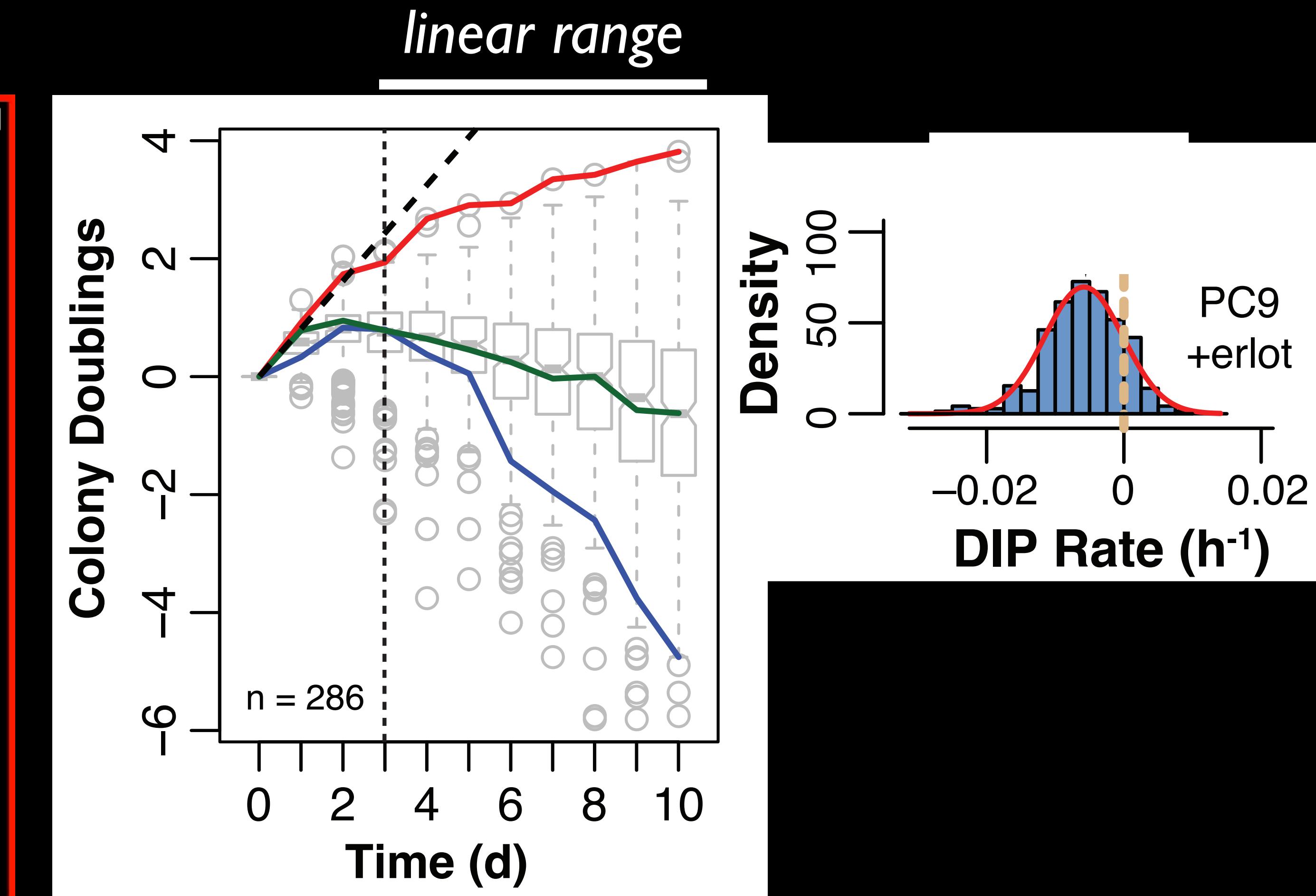
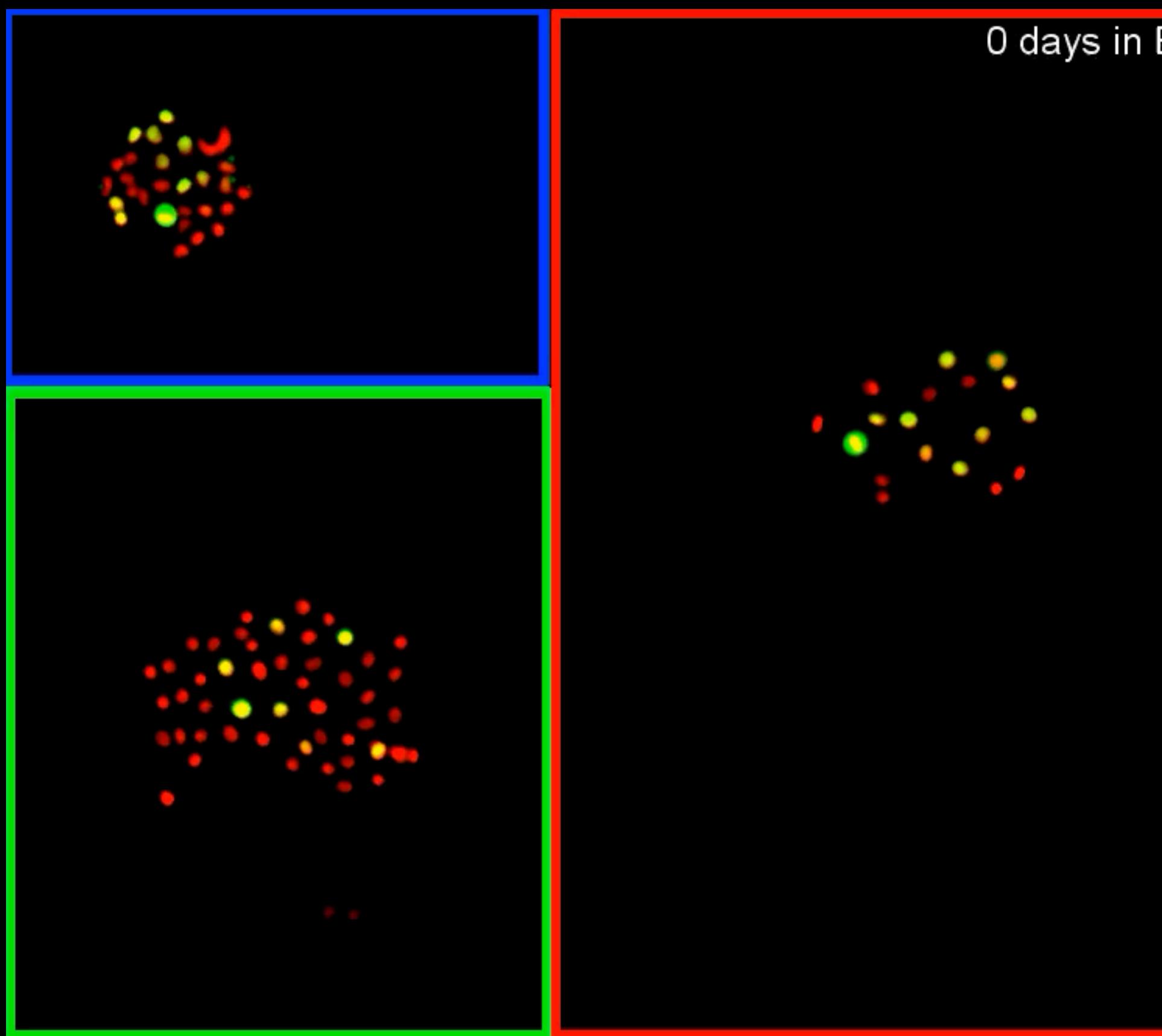
Rather than infer from a change in relative cell abundance, measure directly (proliferation rate)

Spatially segregated single-cell derived cell populations



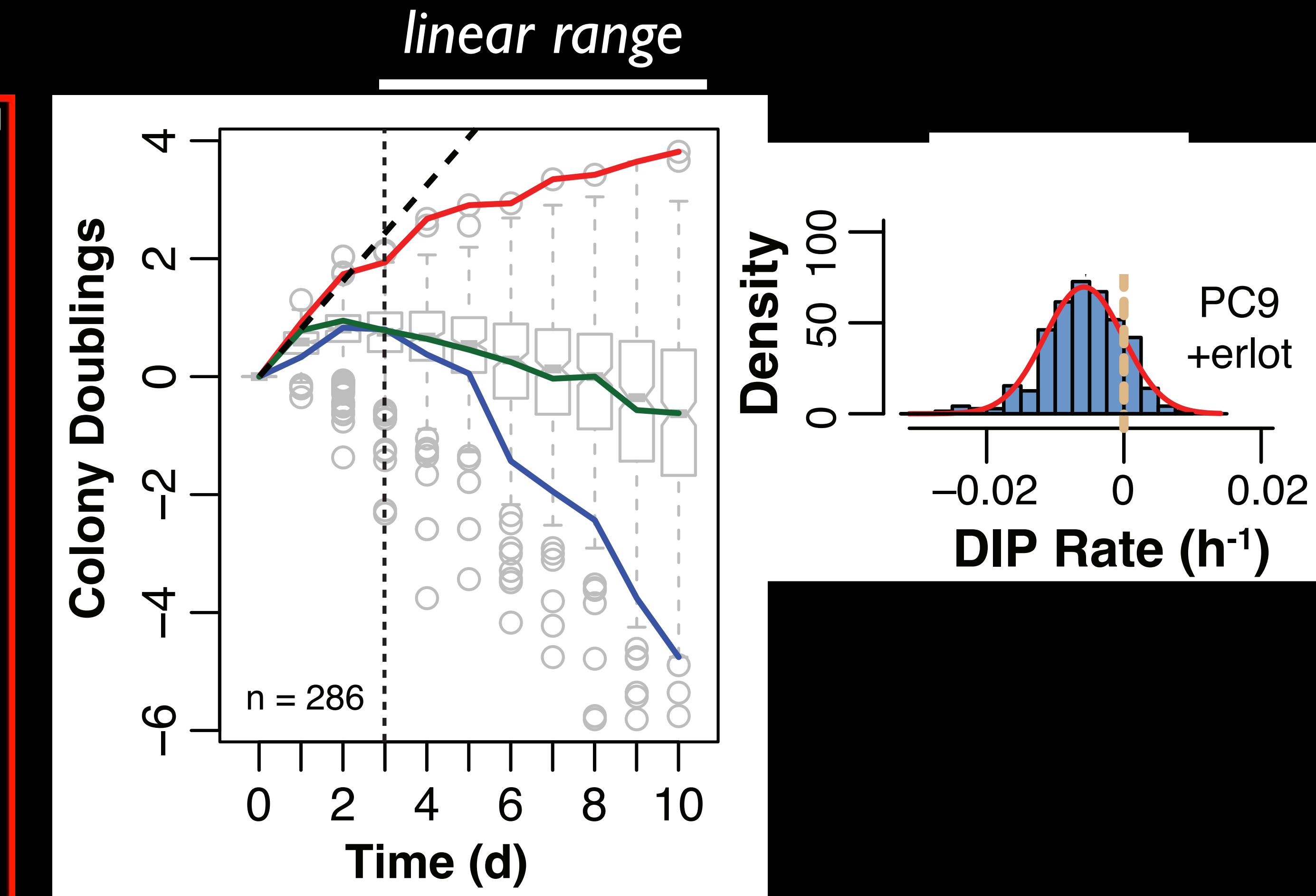
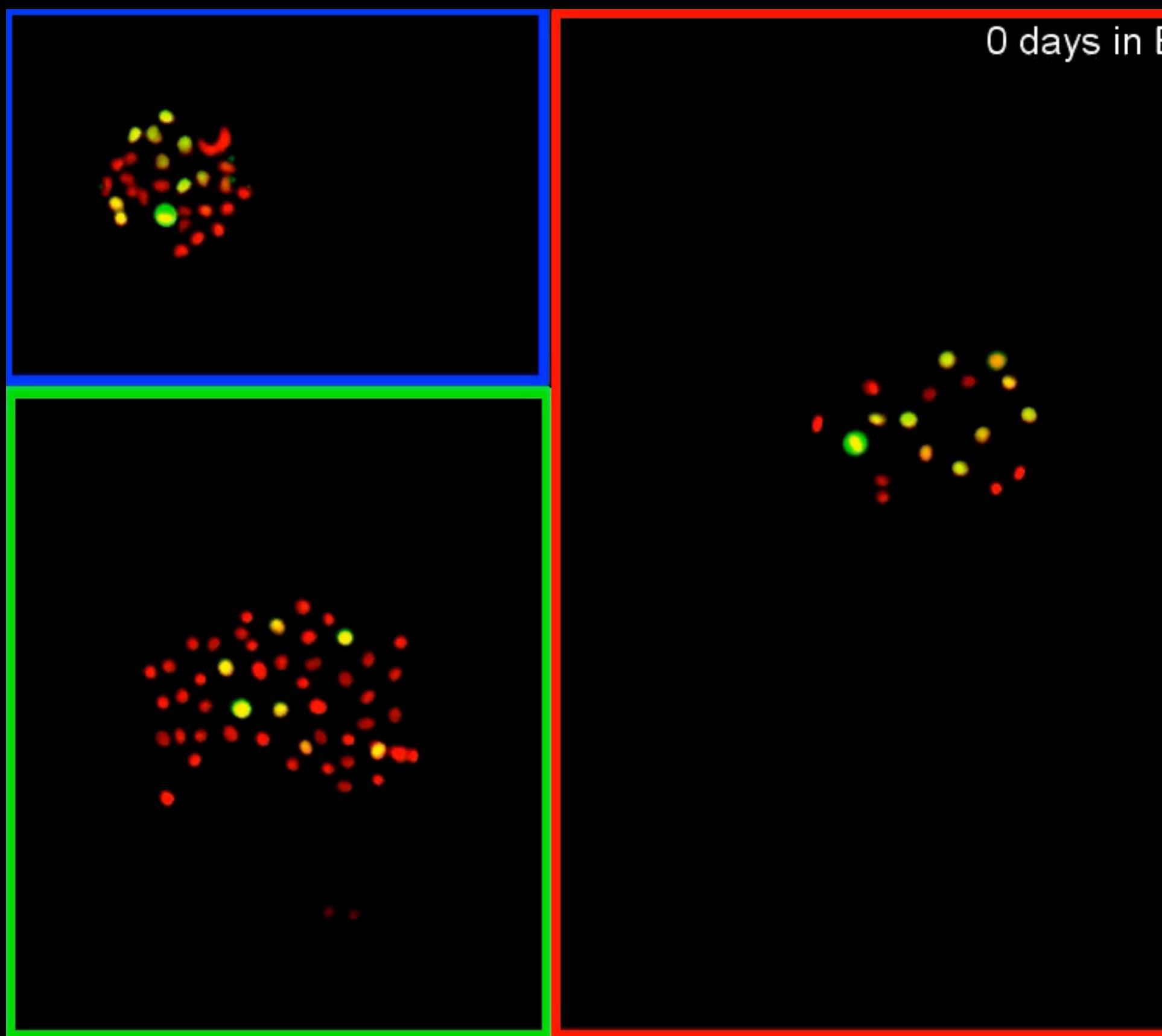
EGFR-mut Lung Cancer Response is Variable

PC9 Colonies Have Distribution of DIP rates in EGFRi



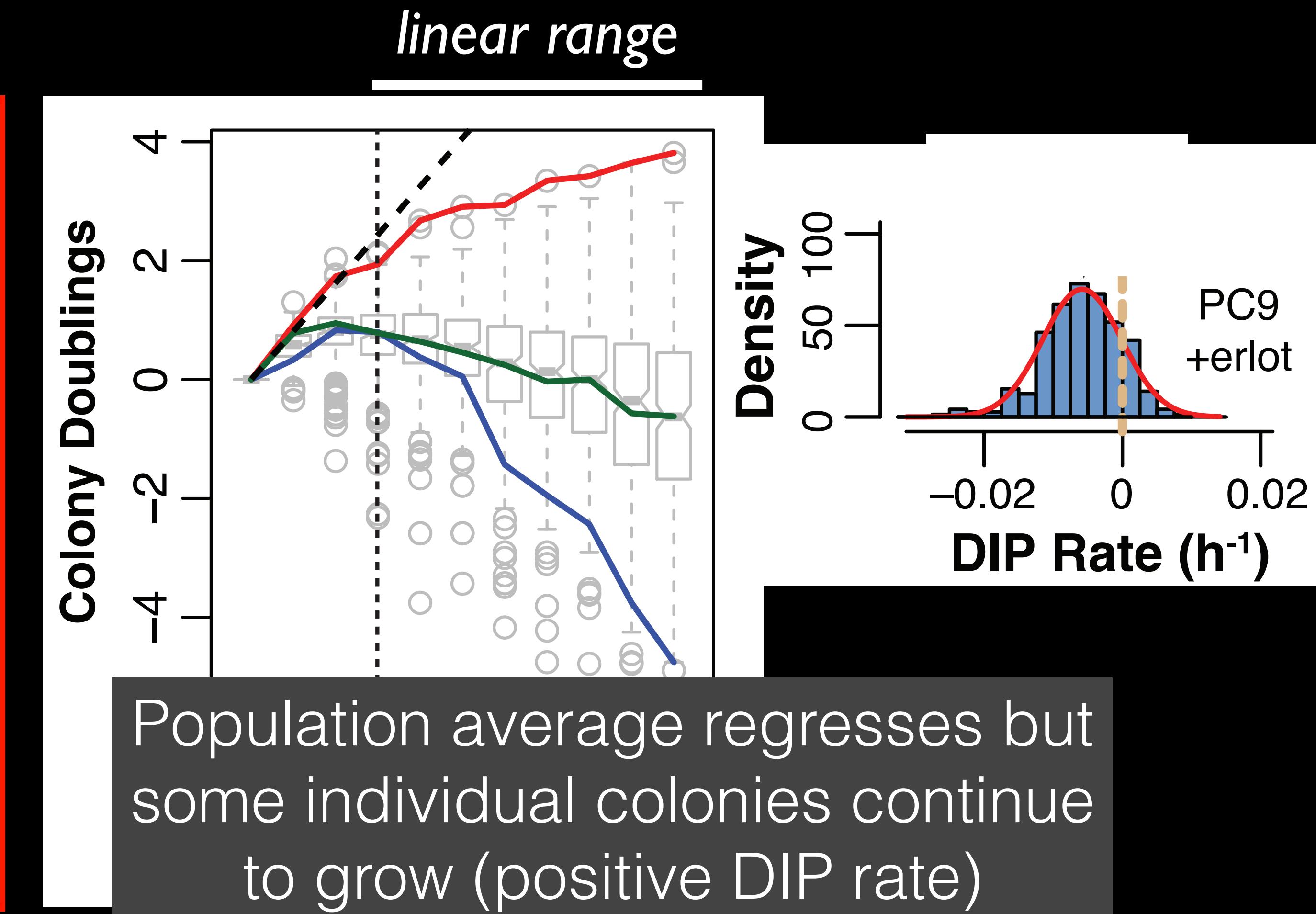
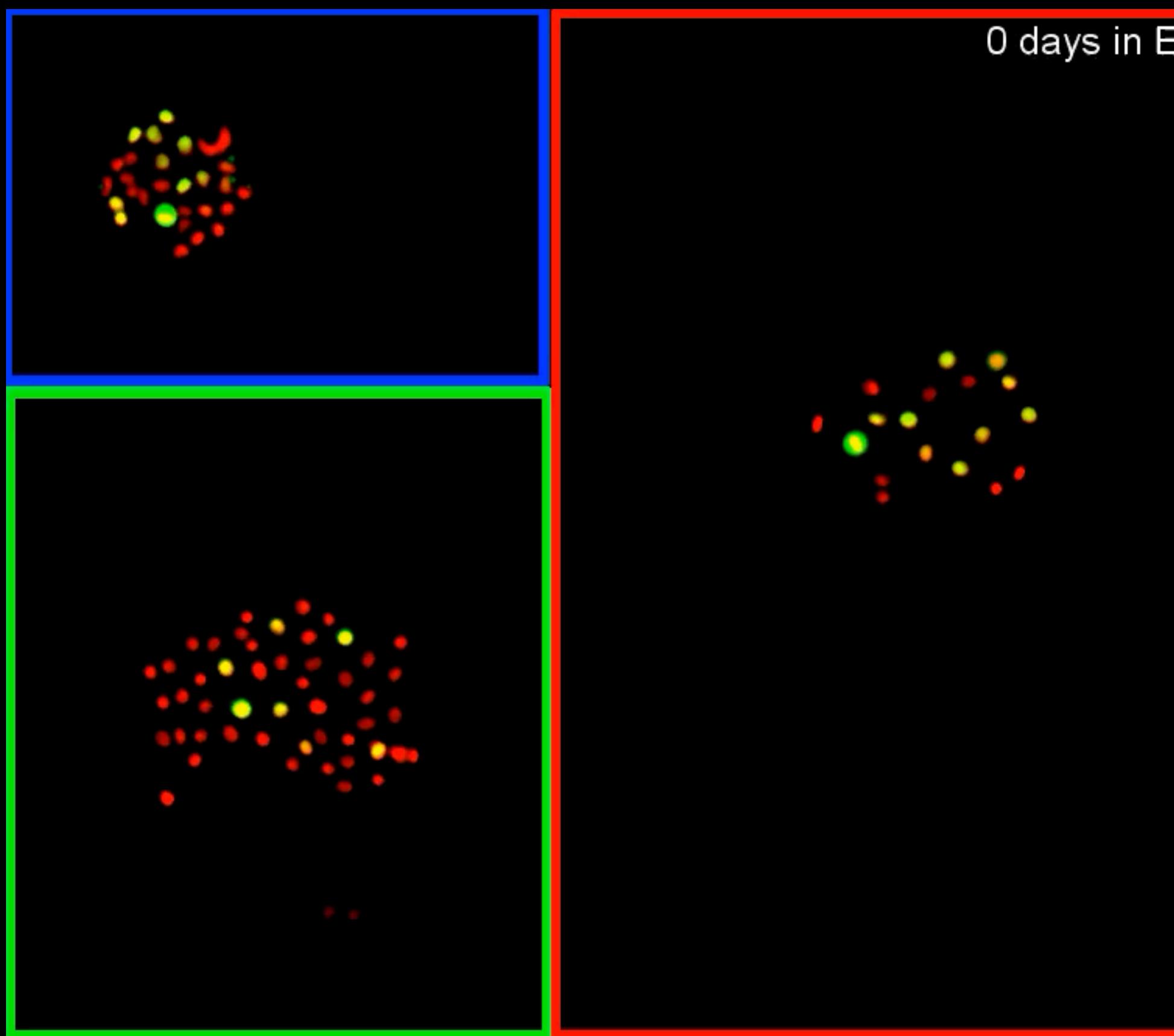
EGFR-mut Lung Cancer Response is Variable

PC9 Colonies Have Distribution of DIP rates in EGFRi



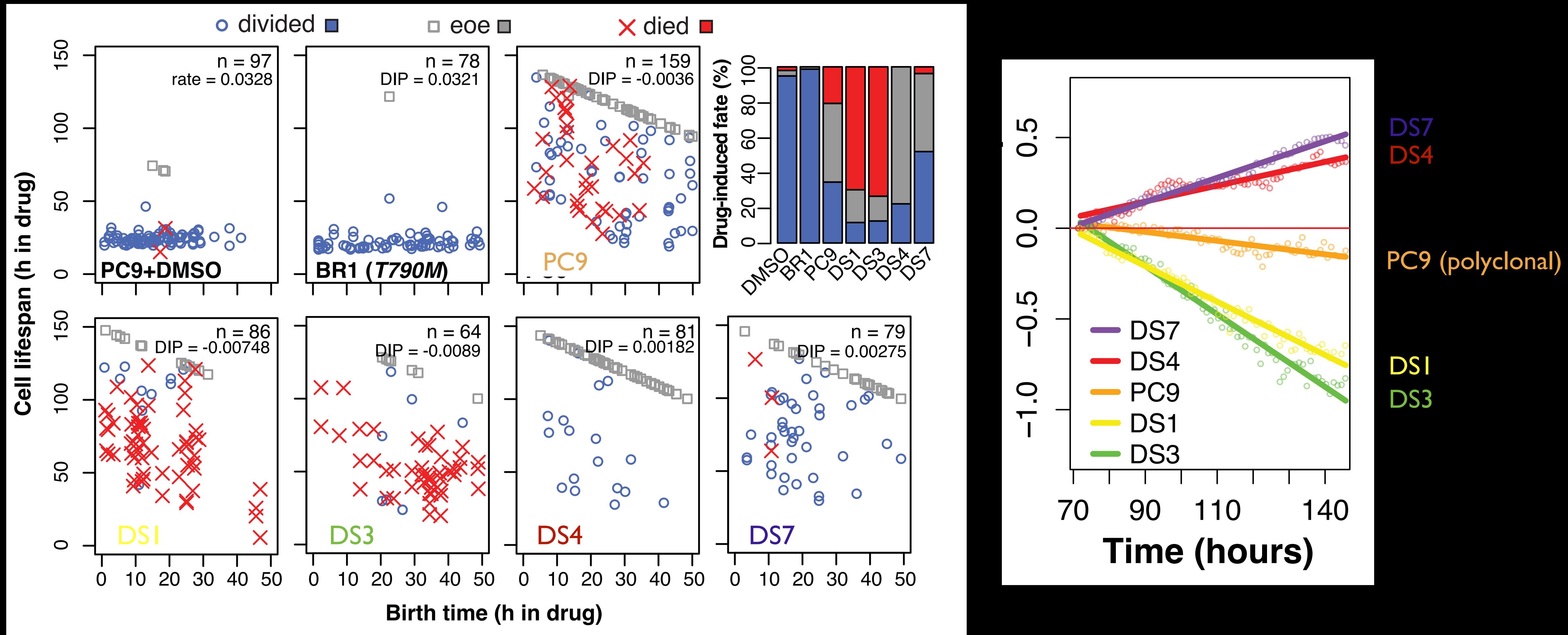
EGFR-mut Lung Cancer Response is Variable

PC9 Colonies Have Distribution of DIP rates in EGFRi

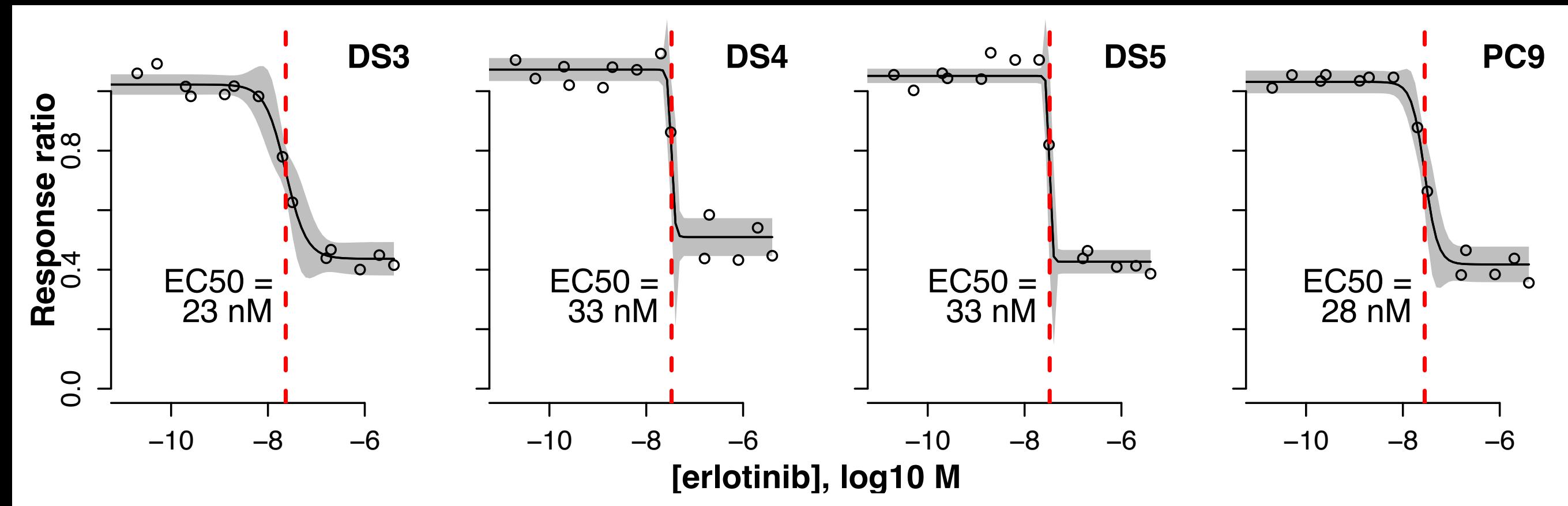


Stochastic Single-Cell Behaviors Correspond to DIP Rate

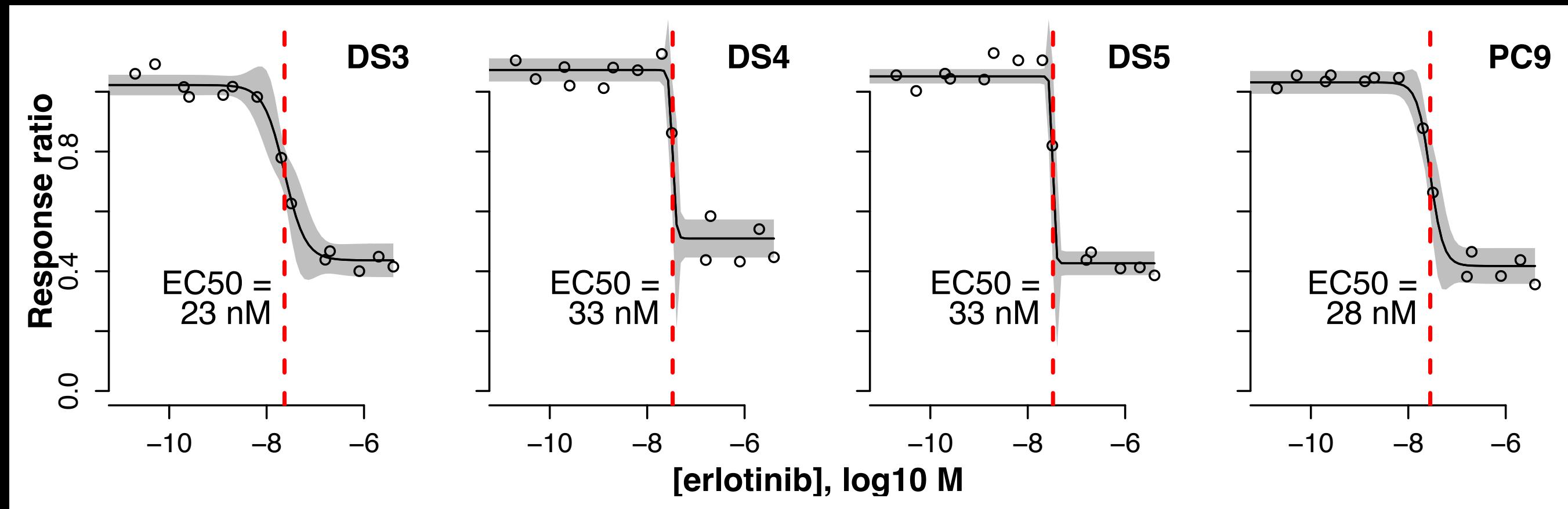
Single-cell fate decisions of erlotinib-treated sublines



Differences Exposed by Analysis of Cell Population Dynamics

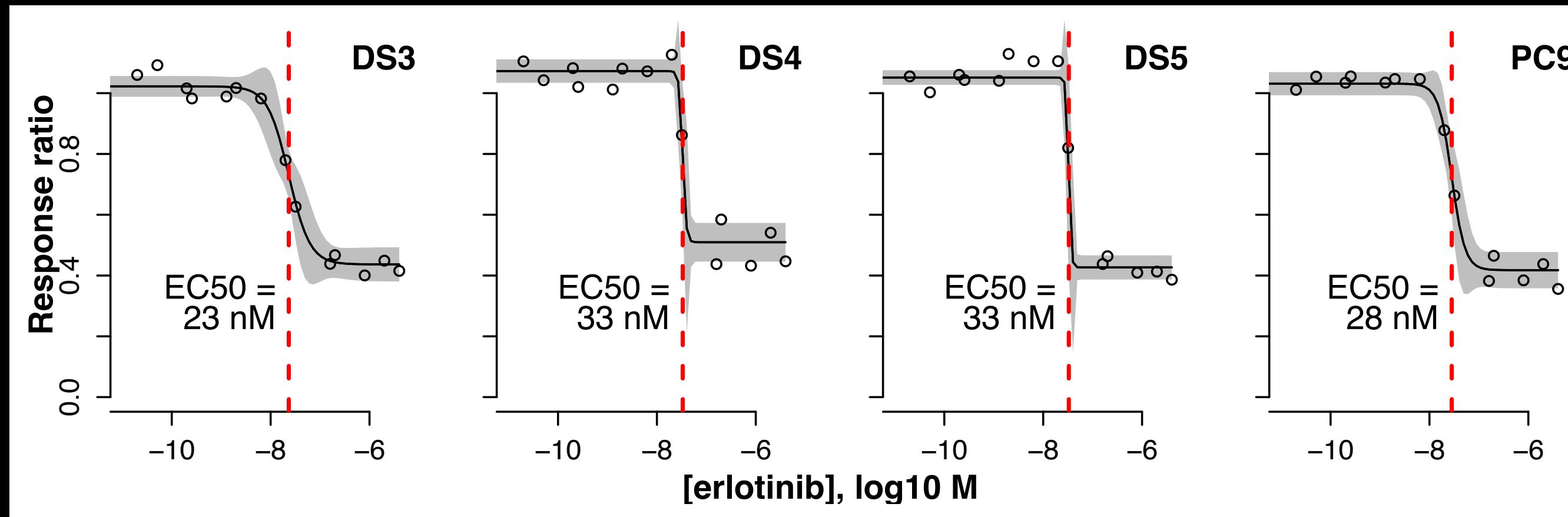


Differences Exposed by Analysis of Cell Population Dynamics

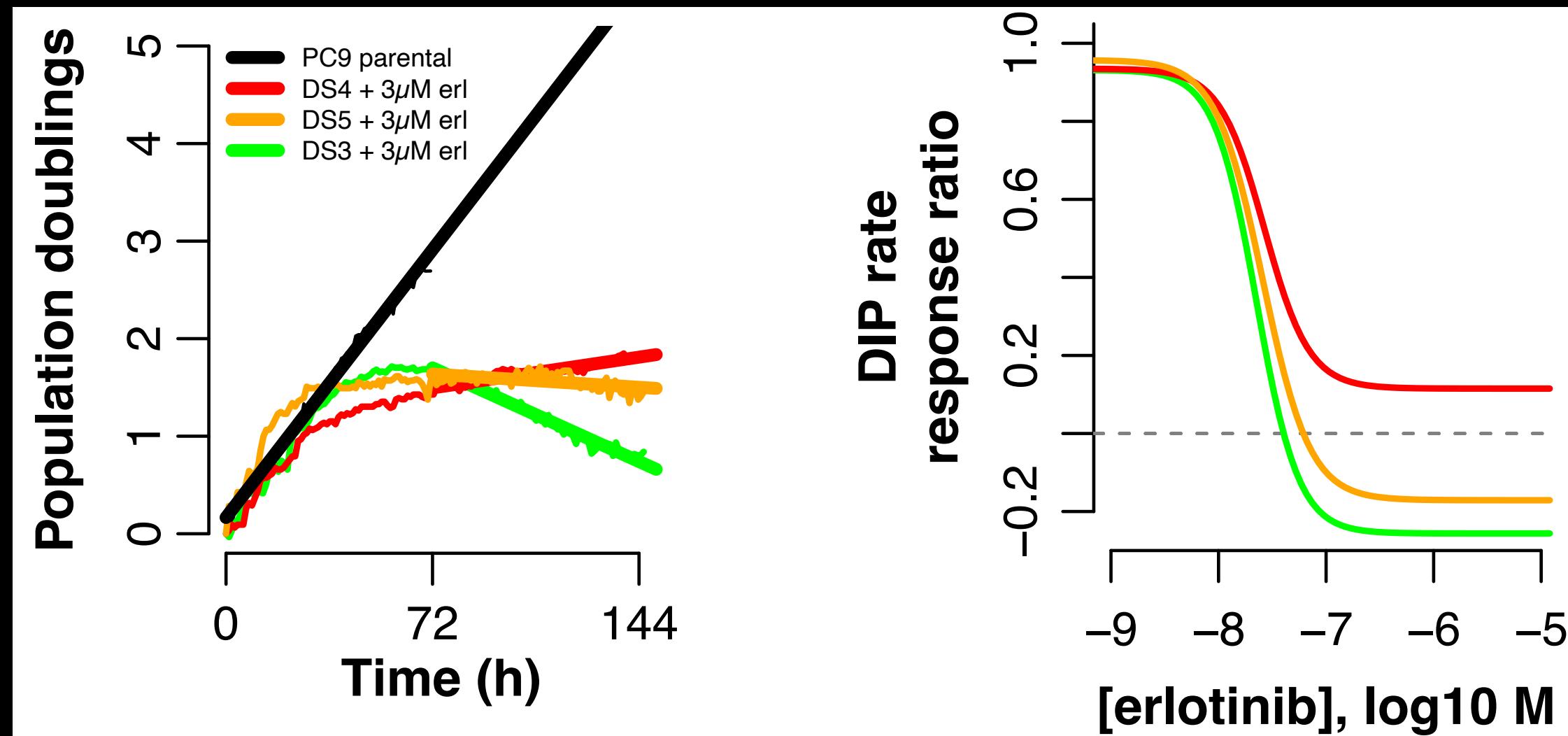


Standard endpoint assay of drug response cannot distinguish cell lines

Differences Exposed by Analysis of Cell Population Dynamics

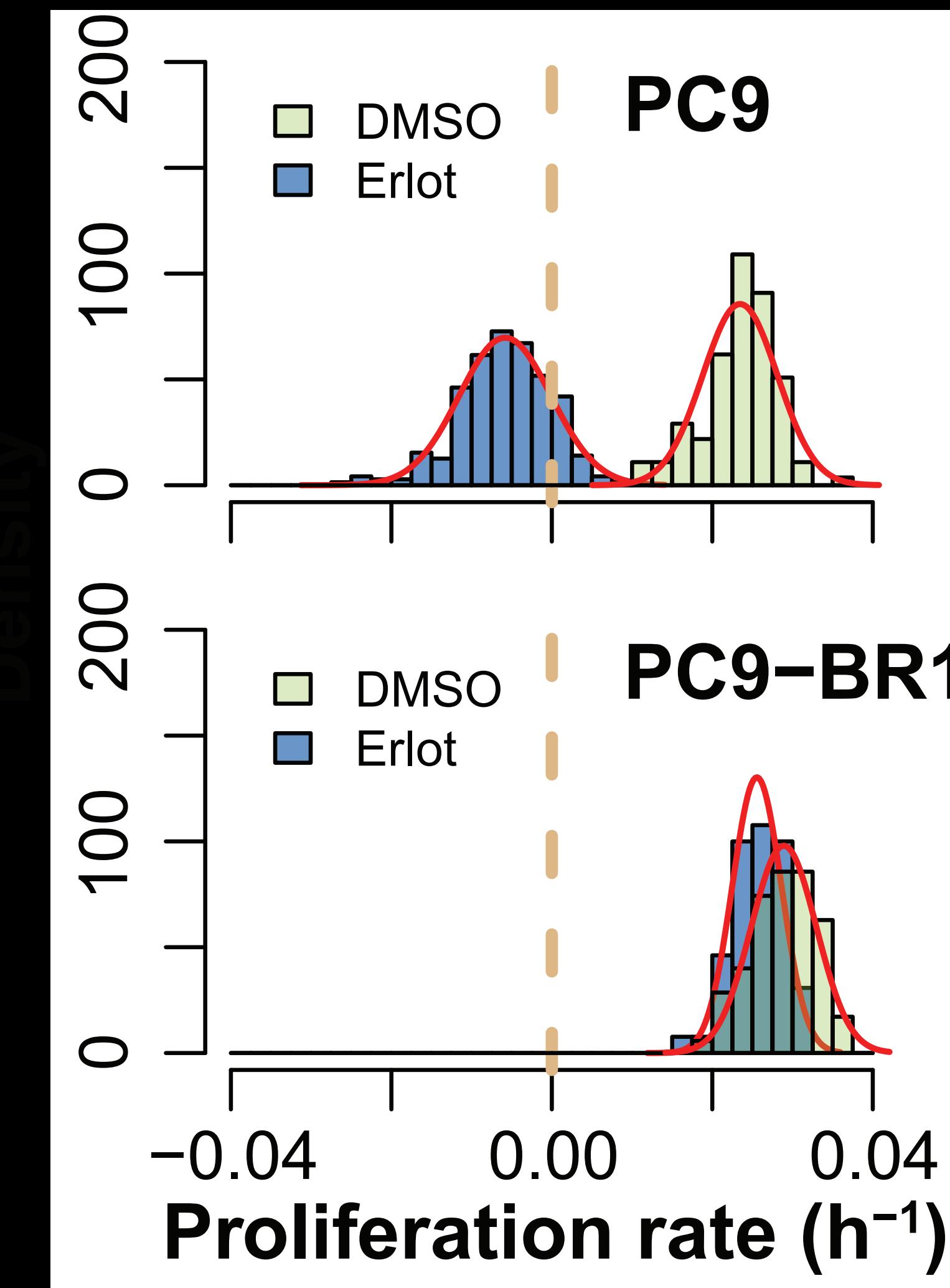


Standard endpoint assay of drug response cannot distinguish cell lines

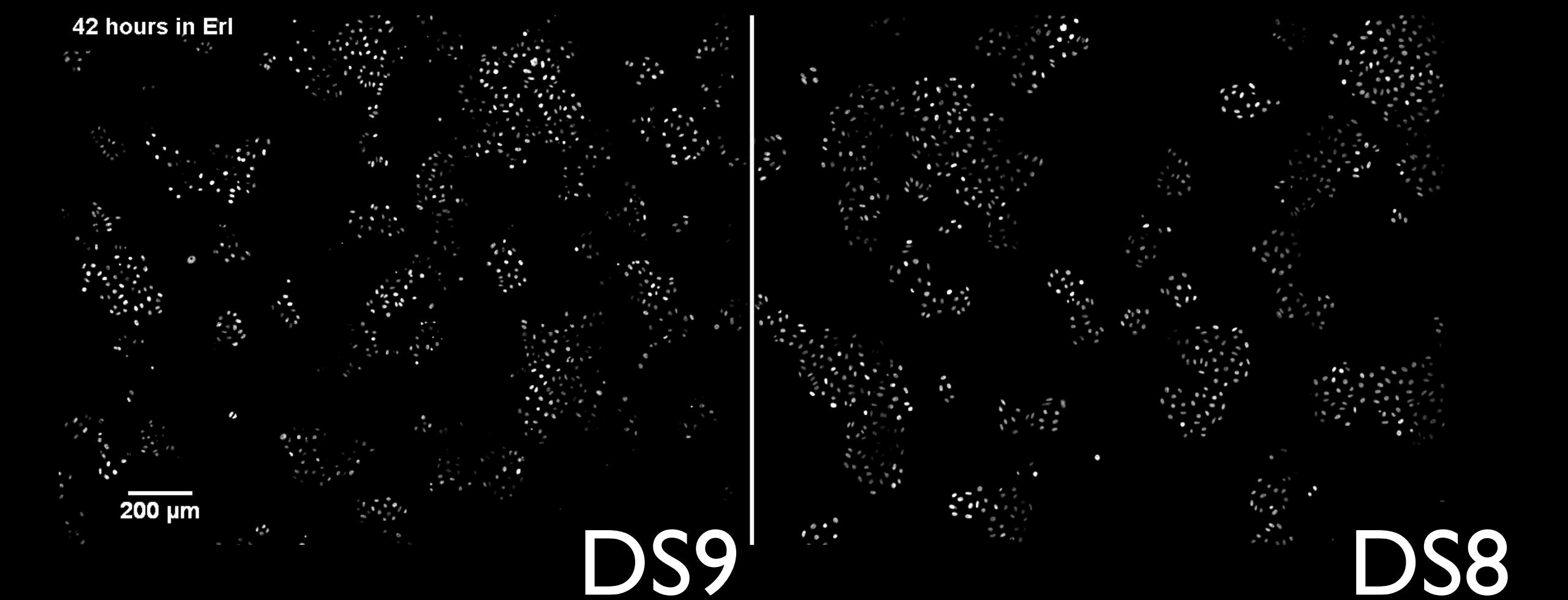
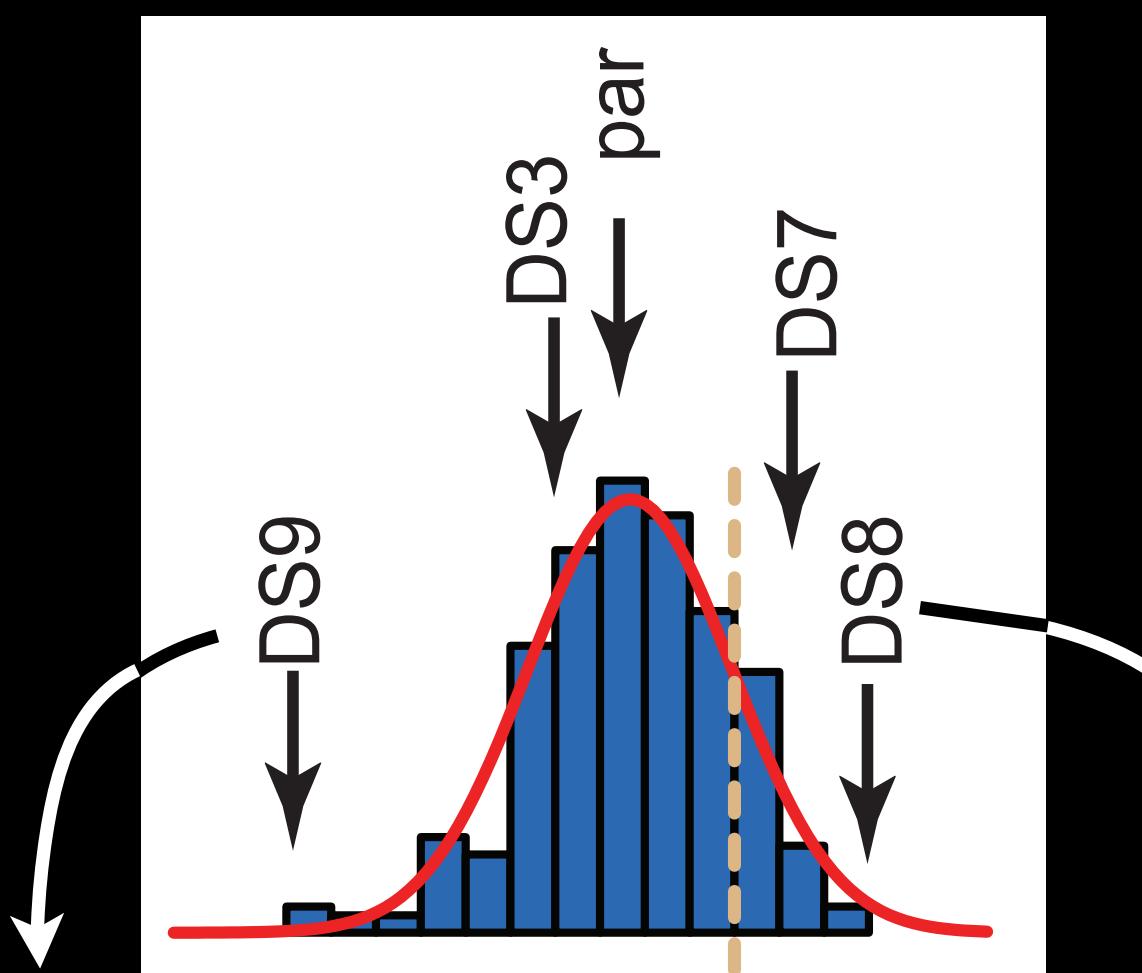


When the time-dependence of the response is considered, the dose-dependent effects are easily distinguished

DIP rates quantify variation in drug sensitivity as distinct V



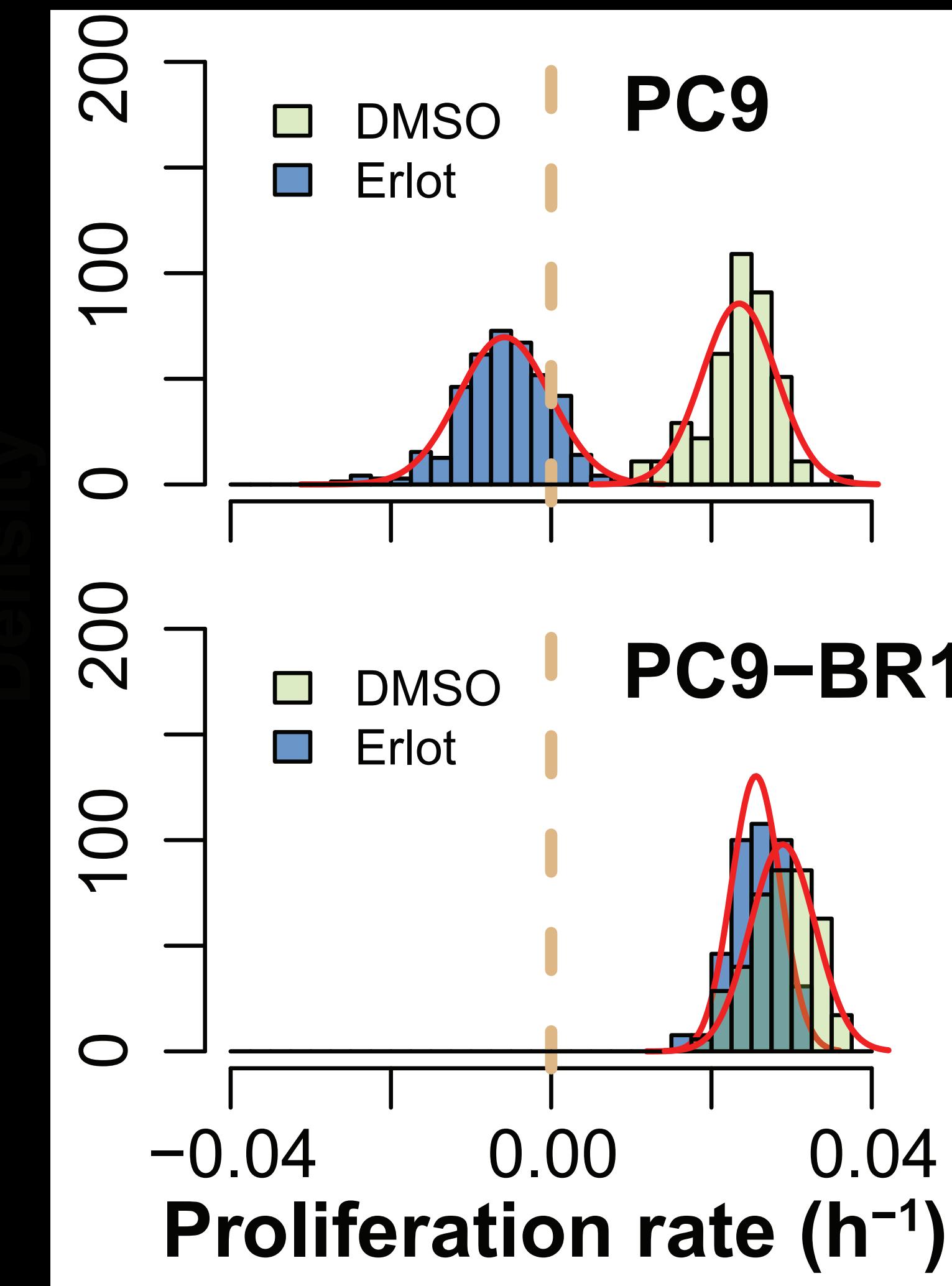
95 single-cell-derived discrete sublines



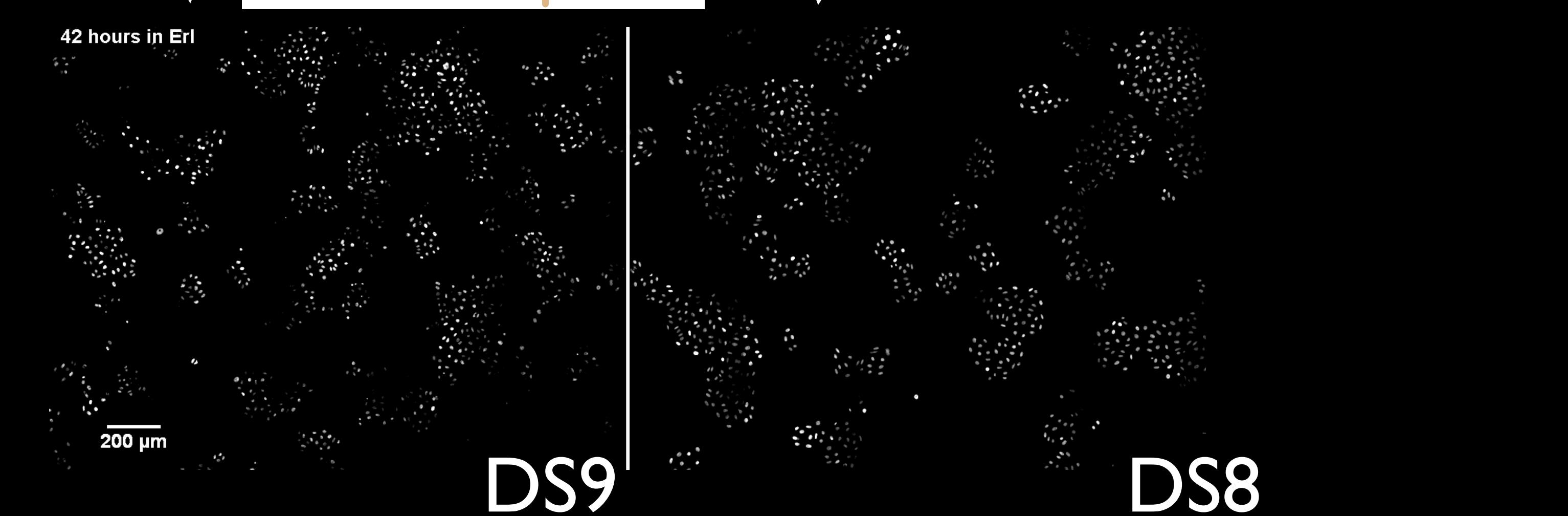
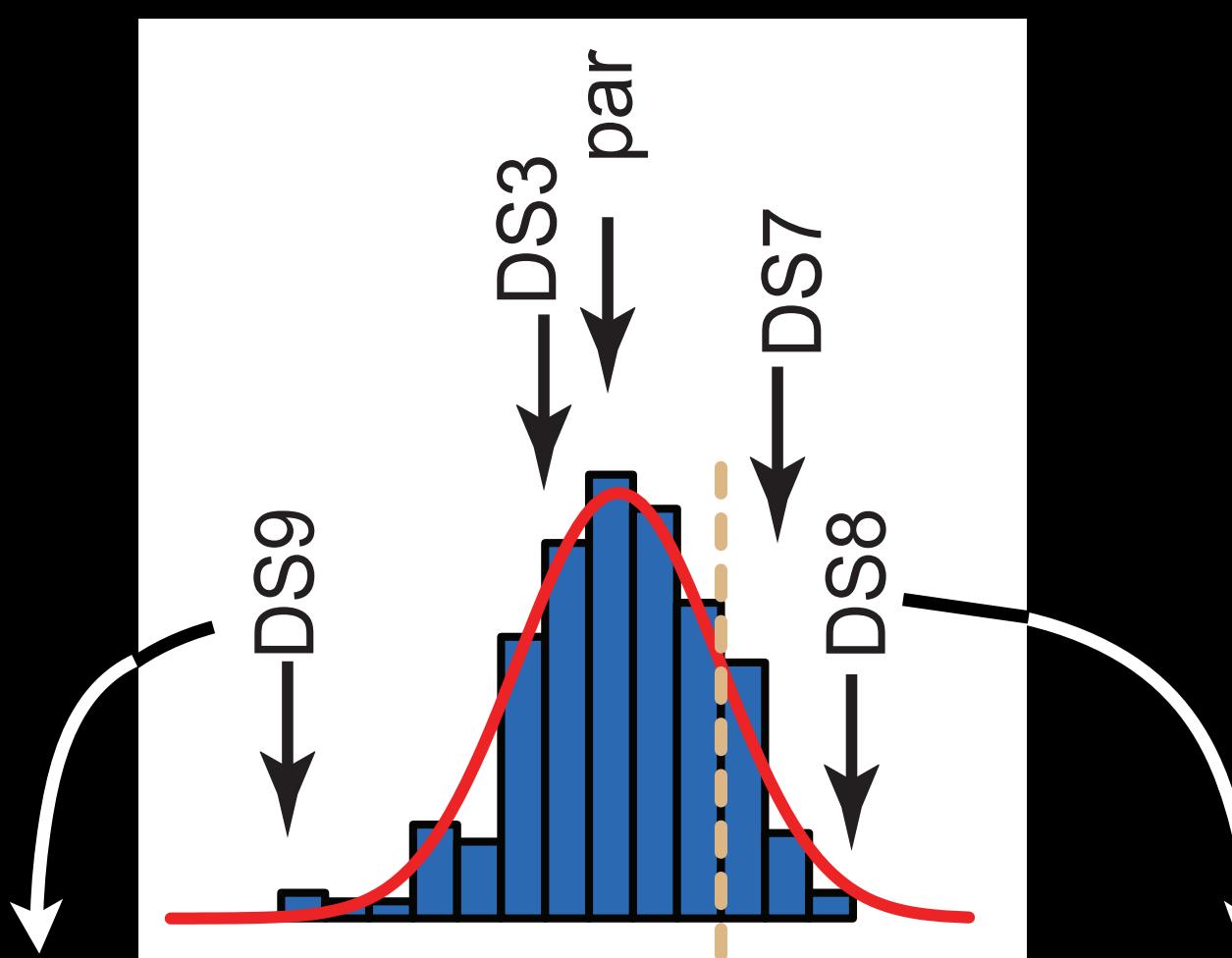
Discrete sublines from extremes of the distribution exhibit dramatically different behaviors

Long-term culture does not result in change of DIP rate but DS8 << BR1

DIP rates quantify variation in drug sensitivity as distinct V



95 single-cell-derived discrete sublines



Discrete sublines from extremes of the distribution exhibit dramatically different behaviors

Long-term culture does not result in change of DIP rate but DS8 << BR1

Assessing concentration
dependence of drugs using DIP rate

Image Acquisition

- Performed in Vanderbilt HTS Core with help of Josh Bauer
- Up to 42 384-well plates in incubator
- Each plate pulled from incubator by robotic arm, its barcode read, and placed into imager (MD imageXpress)
- Each well (inner 308; viewed thru 10X objective) gets imaged once per channel (usually red/green)
- ~10 min per plate; ~8–9 h before plate is imaged again
- Each experiment duration ~5 days (imaging only)
- ~2.5–3.5 TB of image data (308 wells X 42 plates X 2 channels X 13 time points = 336,336 images per experiment)



*This represents
1/12,936 of the total
wells imaged in the
experiment (and only
1 channel)*

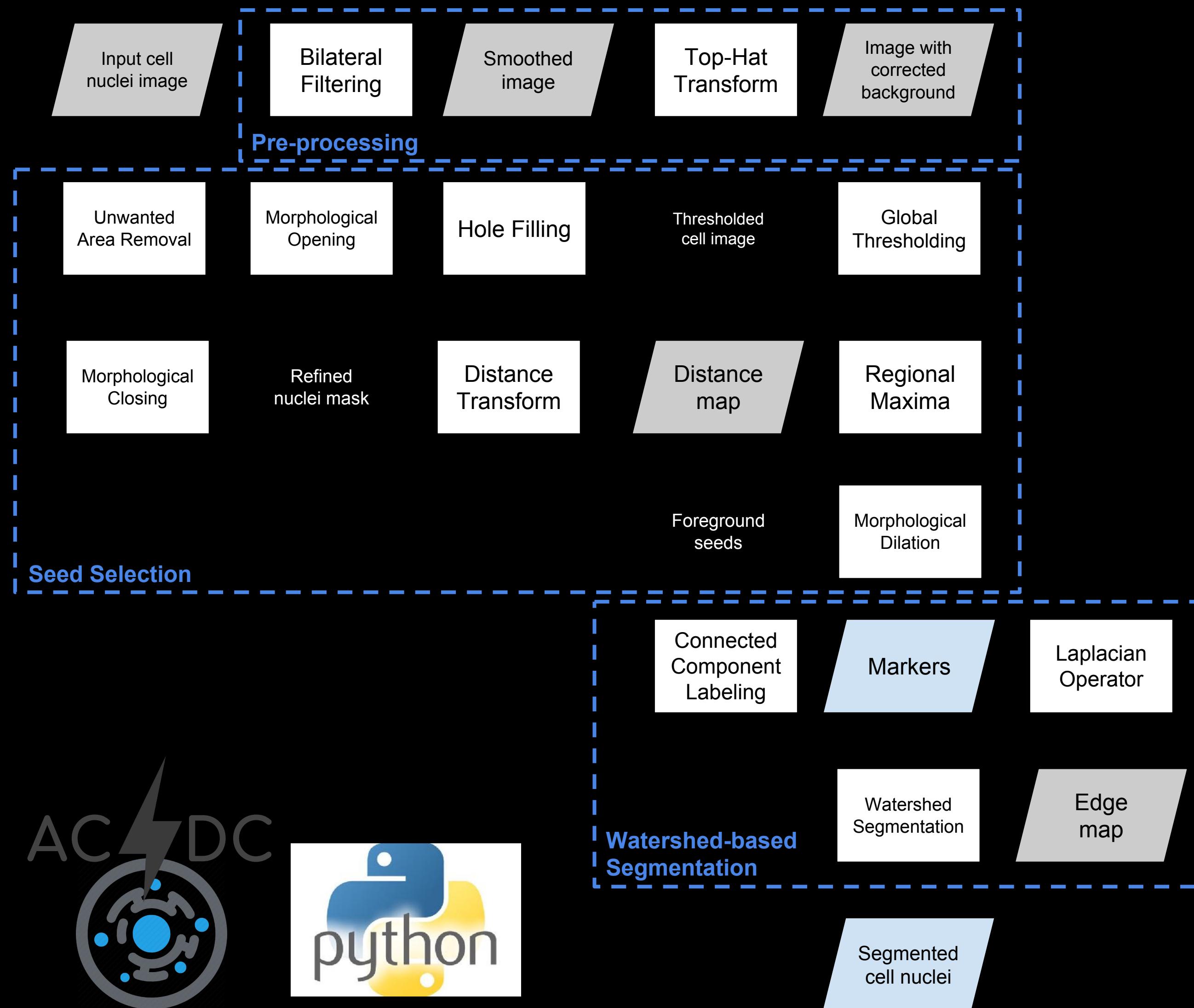
Image Acquisition

- Performed in Vanderbilt HTS Core with help of Josh Bauer
- Up to 42 384-well plates in incubator
- Each plate pulled from incubator by robotic arm, its barcode read, and placed into imager (MD imageXpress)
- Each well (inner 308; viewed thru 10X objective) gets imaged once per channel (usually red/green)
- ~10 min per plate; ~8–9 h before plate is imaged again
- Each experiment duration ~5 days (imaging only)
- ~2.5–3.5 TB of image data (308 wells X 42 plates X 2 channels X 13 time points = 336,336 images per experiment)

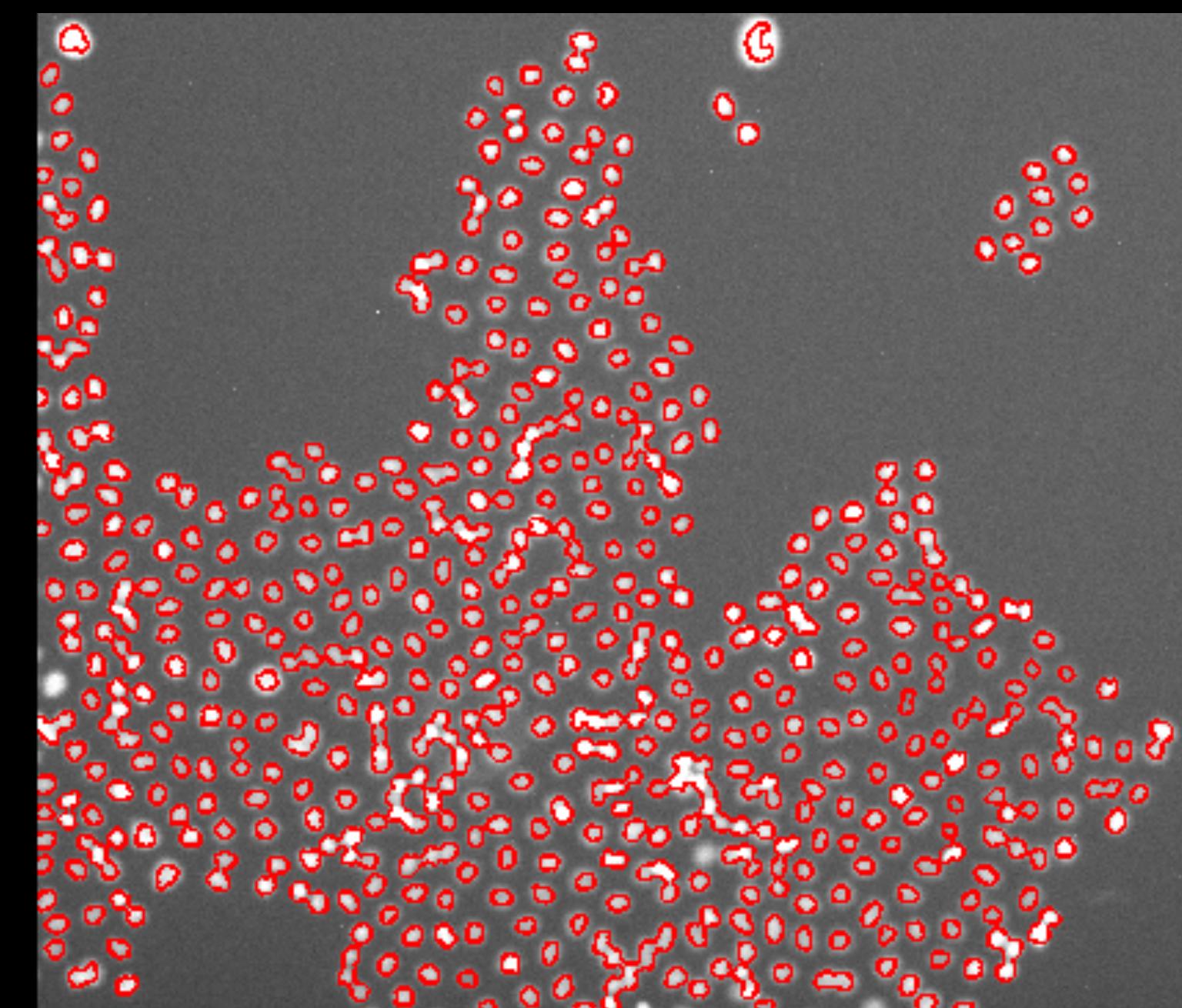


*This represents
1/12,936 of the total
wells imaged in the
experiment (and only
1 channel)*

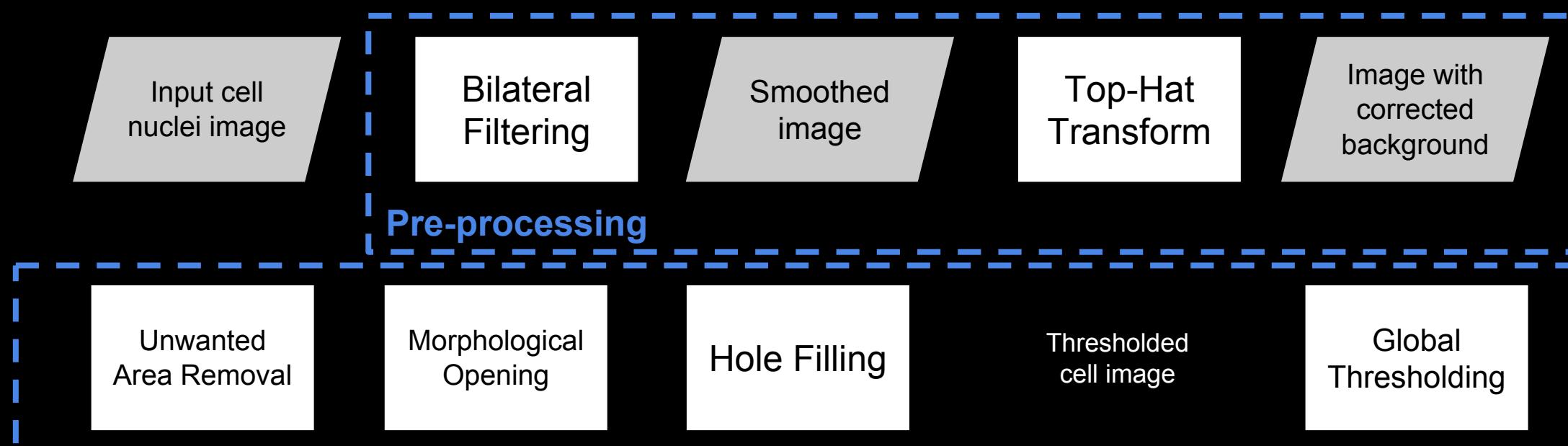
AC/DC (Automatic Cell Detection and Counting)



Python-based image segmentation code optimized to perform background correction, identify nuclei, and quantify pixel intensities in other channel for a pair of images

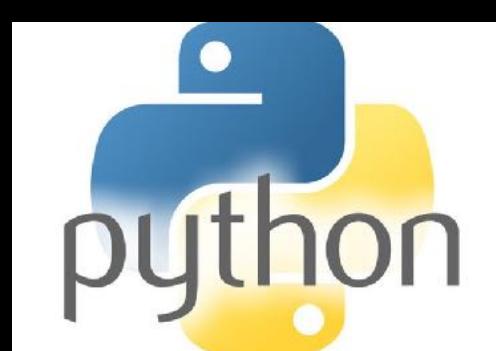
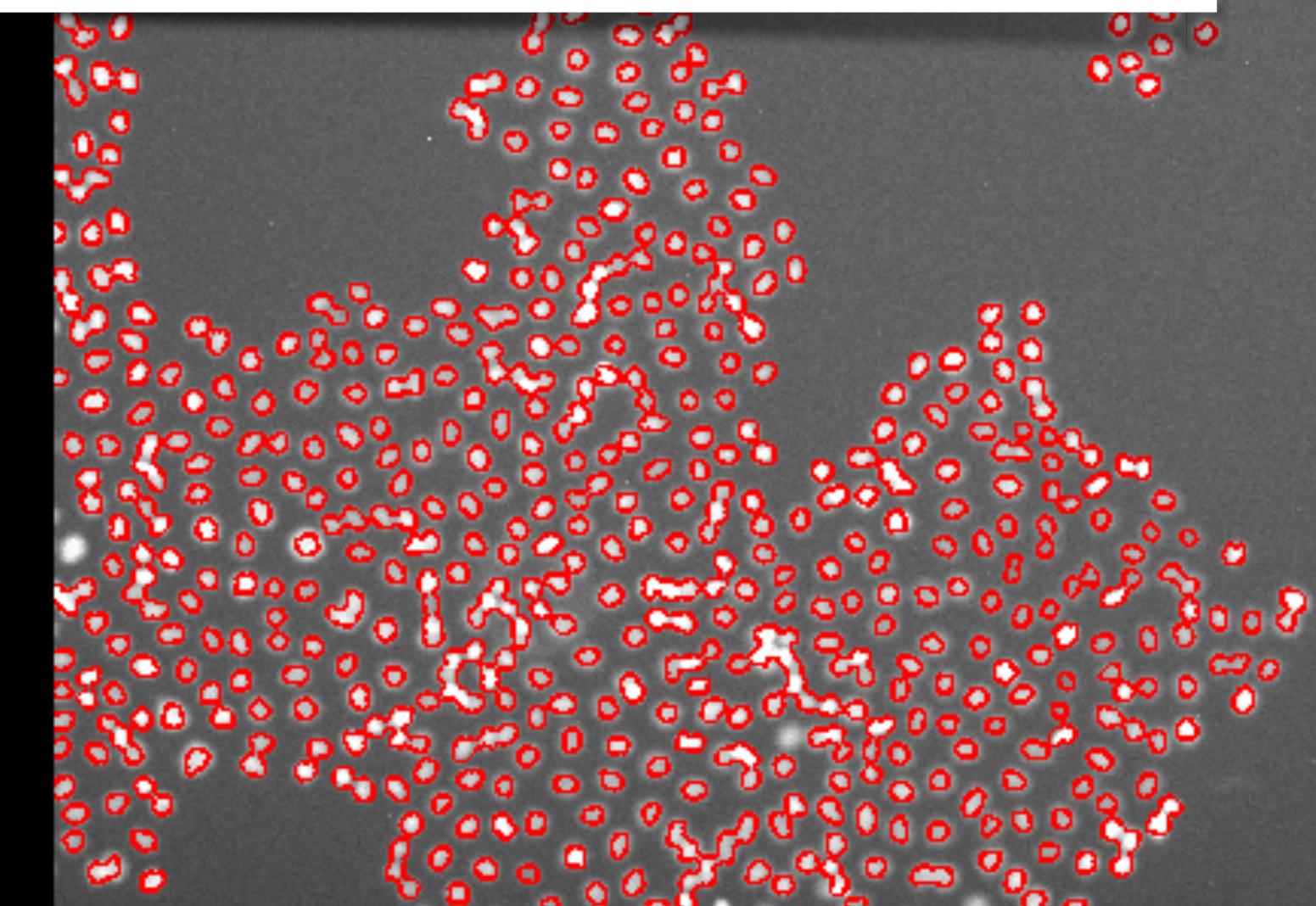
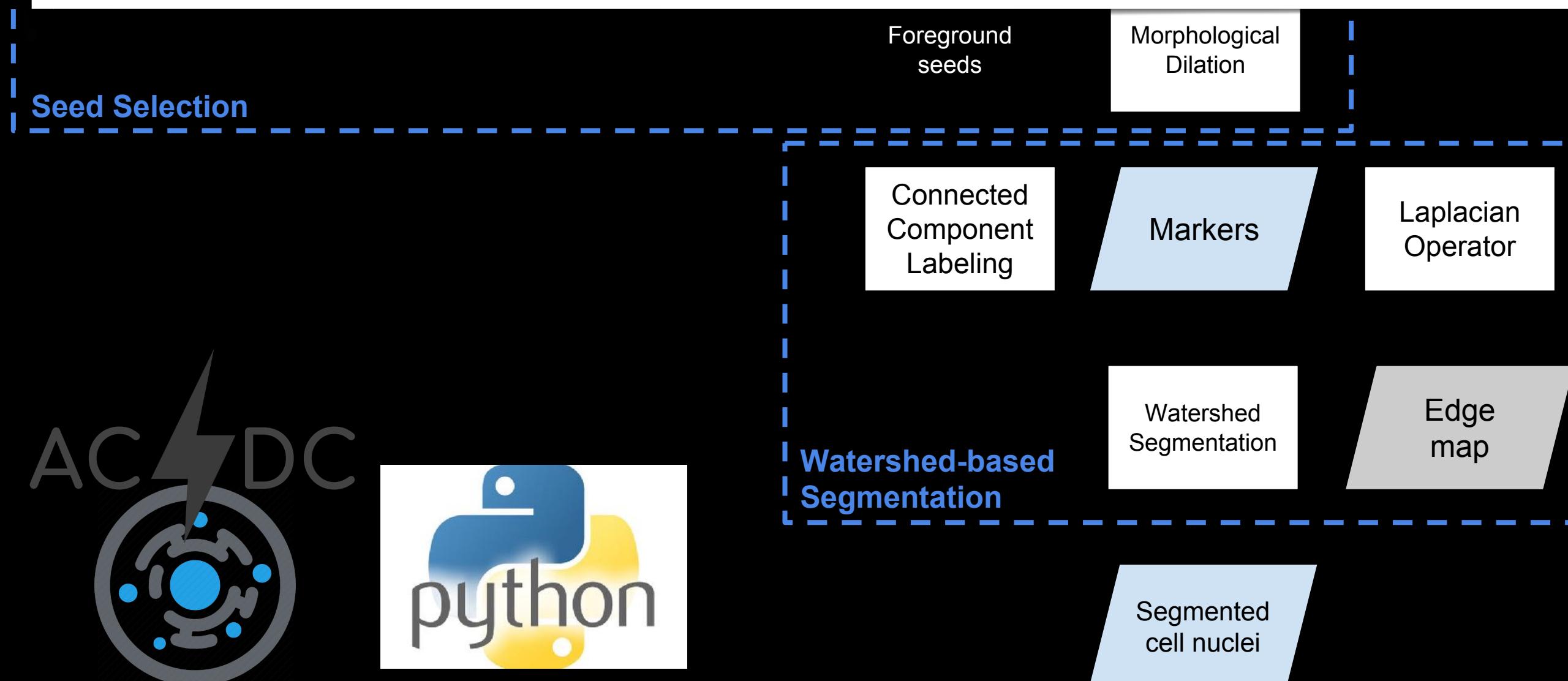


AC/DC (Automatic Cell Detection and Counting)



Python-based image segmentation code optimized to perform background correction, identify nuclei, and quantify pixel intensities in other channel for a

Used to quantify cells within image, but single-cell segmentation also allows analysis of cellular heterogeneity

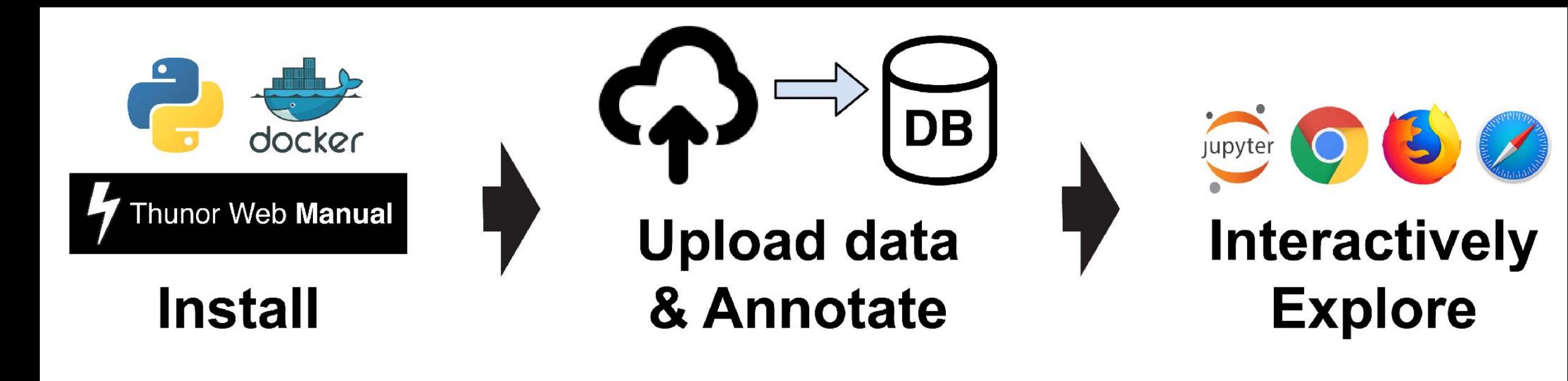


Analyzing big DIP rate datasets



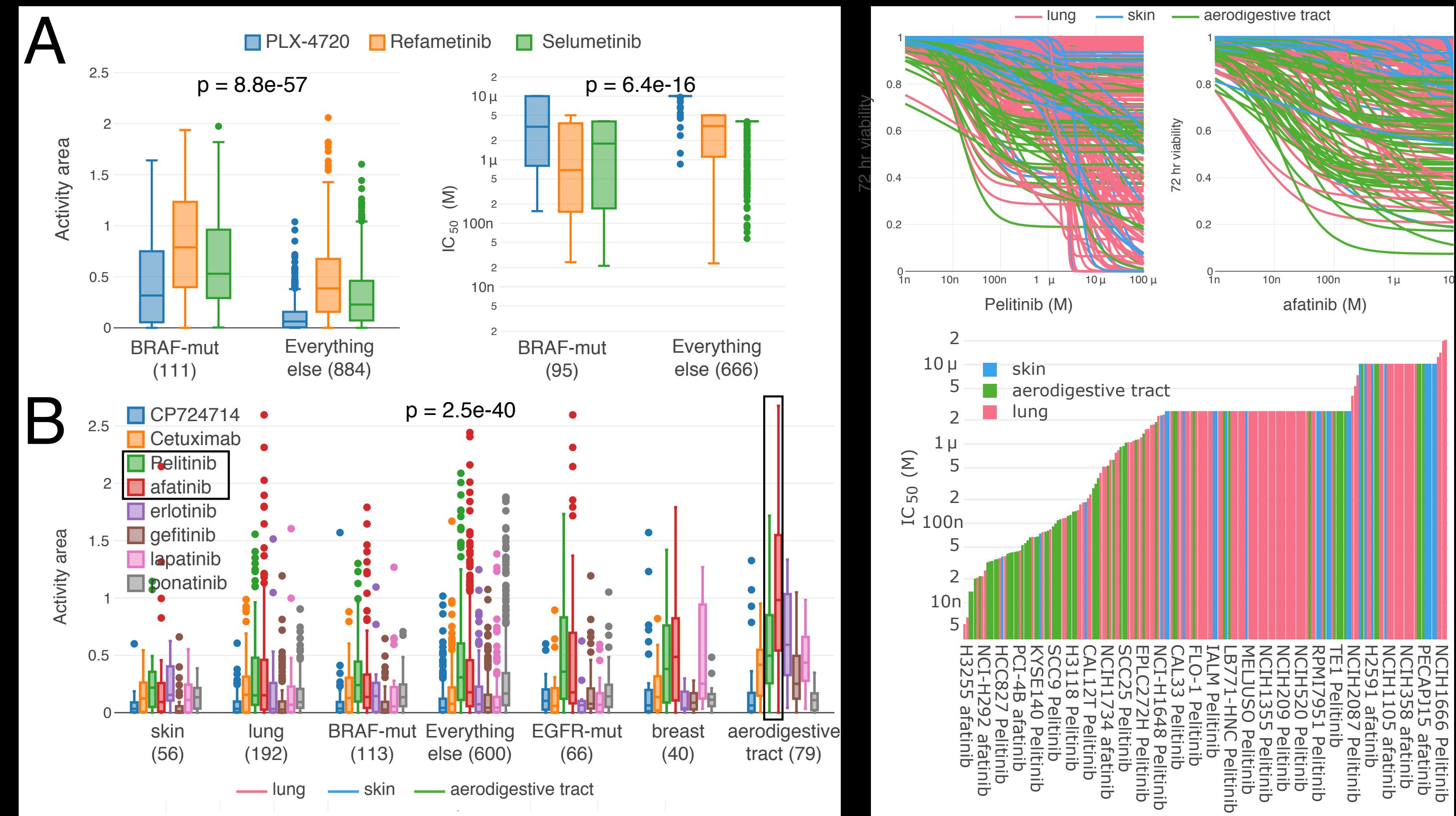
<https://www.thunor.net>

Open source software solution for high-throughput cell proliferation data management/analysis



Designed for cell count over time data but also accommodates traditional end-point (viability) measurements

Provides import capability for large pre-existing datasets (GDSC, CTRP, NCI DTP)



Potential Drug Synergy

Assess using MuSyC (Multidimensional Synergy of Combinations)

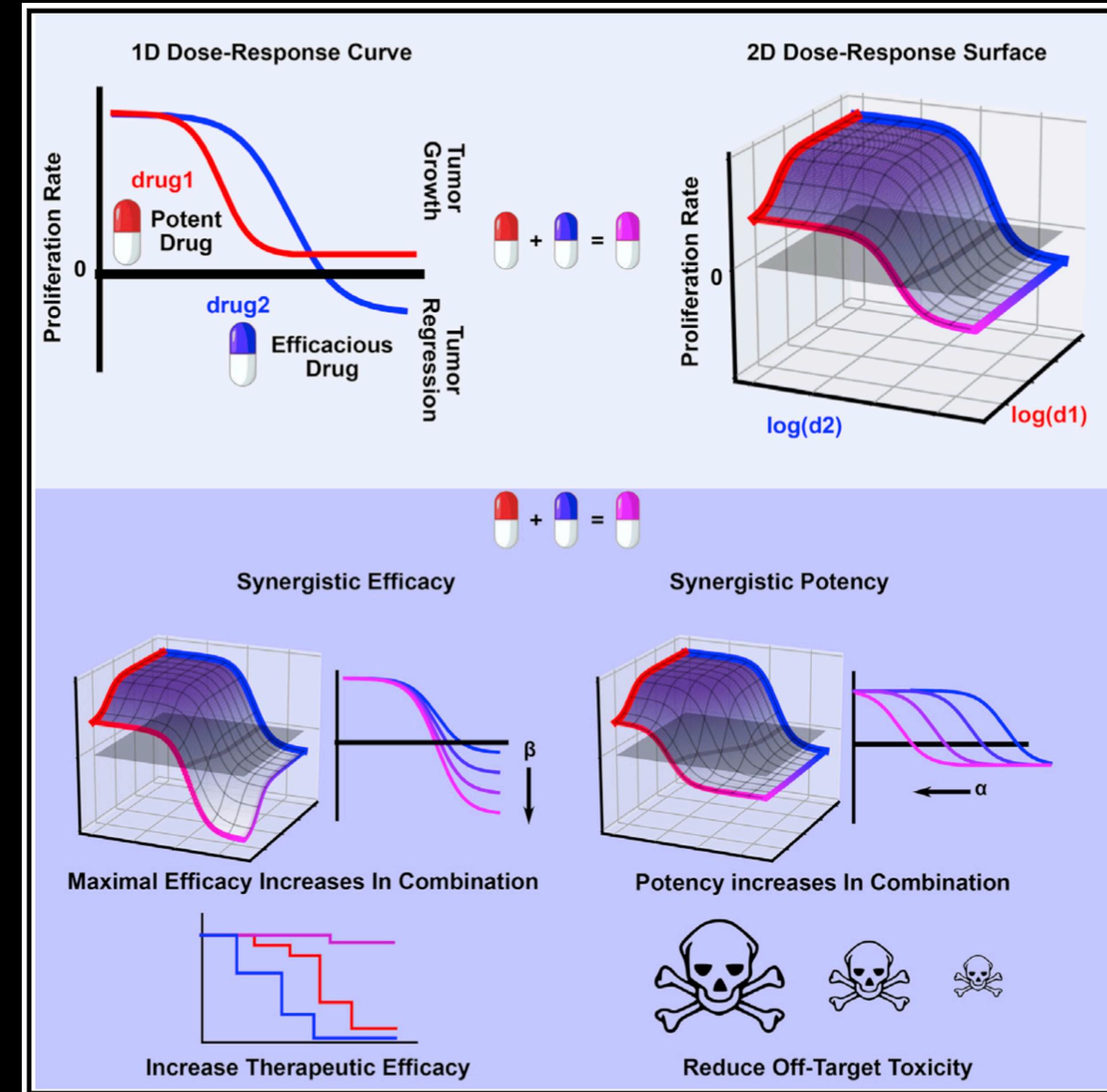
Cell Systems

Quantifying Drug Combination Synergy along Potency and Efficacy Axes

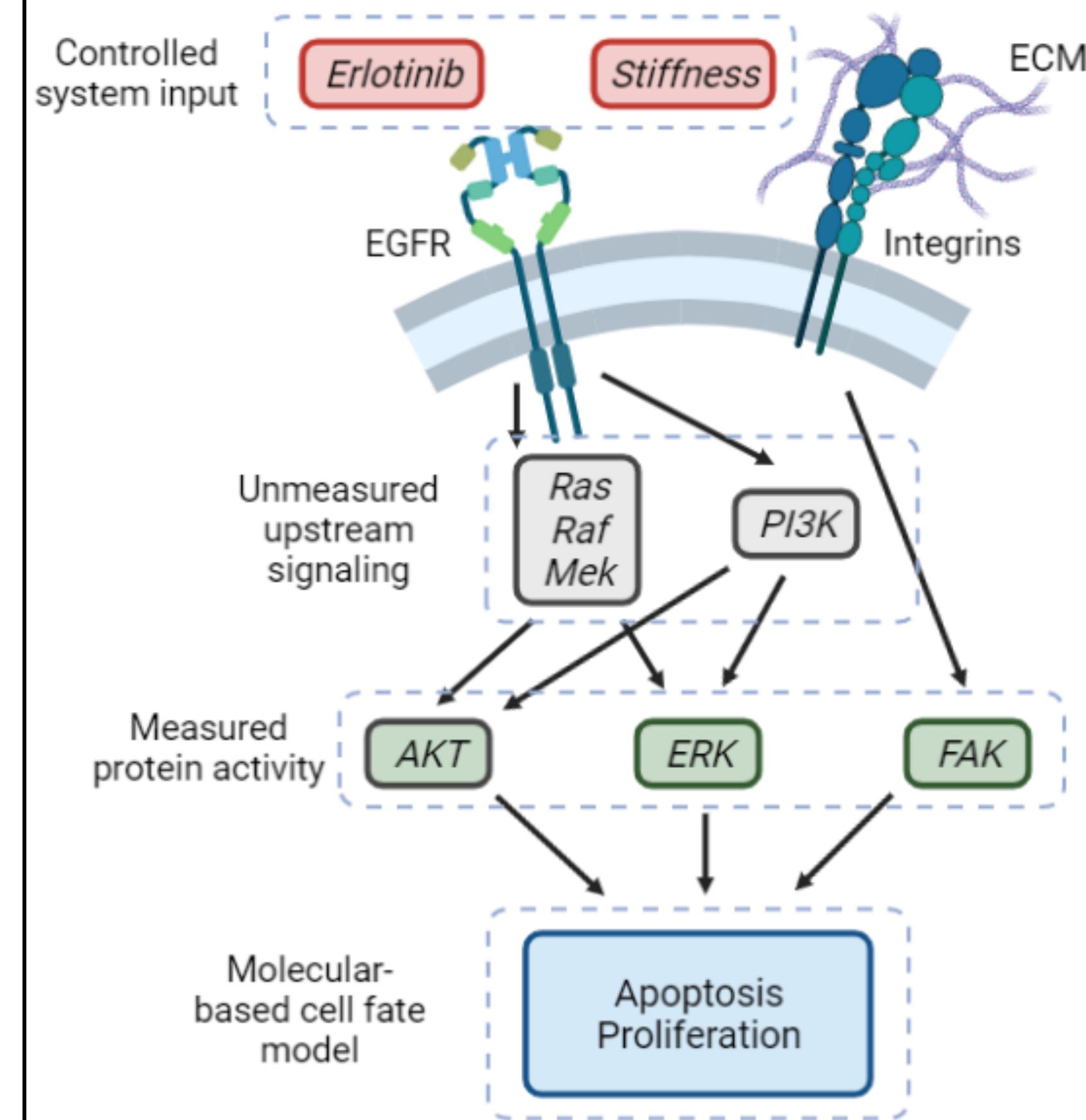
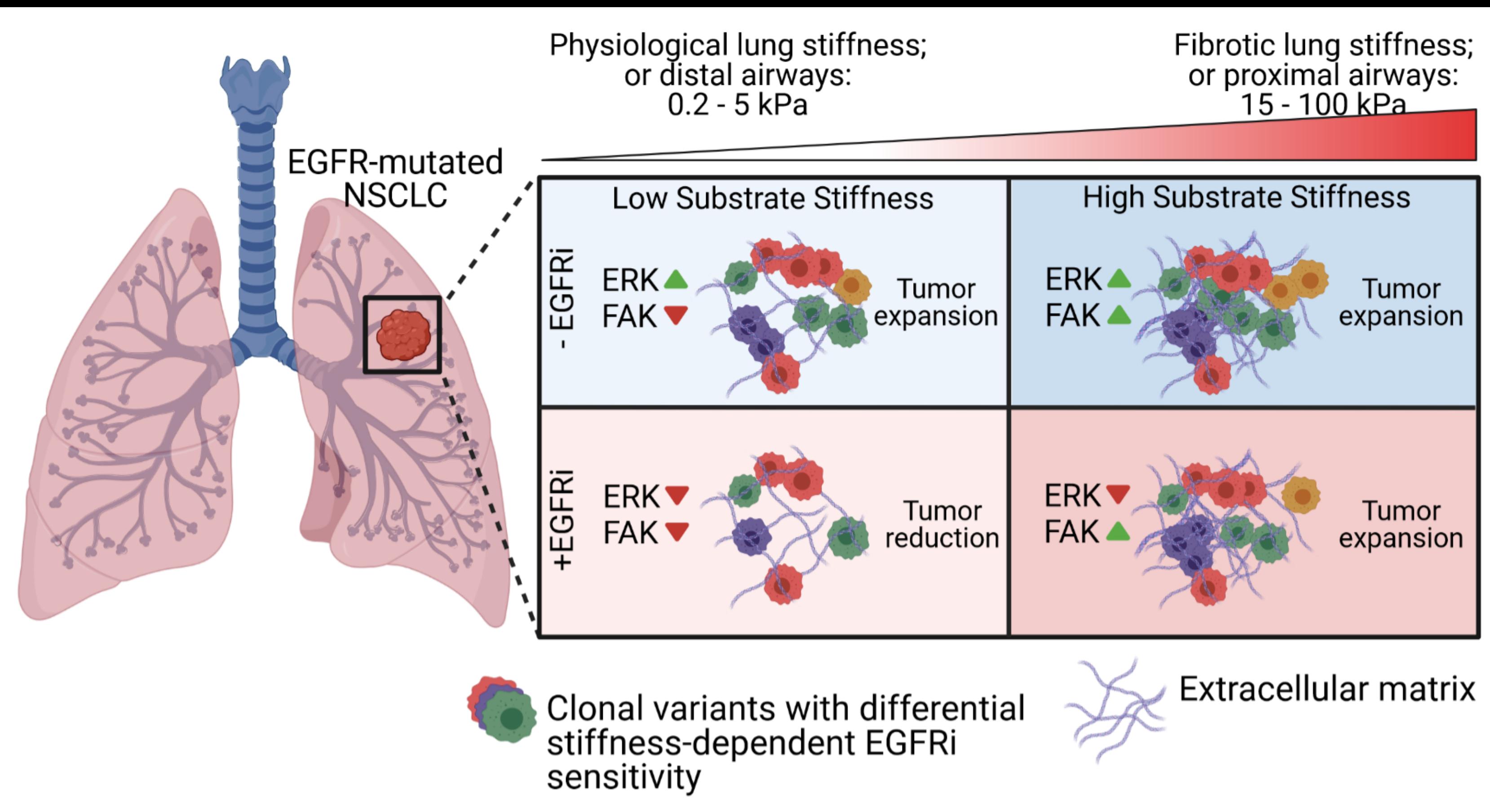
Formalism extending Hill equation to 2D

$$\frac{E_C}{E_0} = \frac{E_0 C_{0.5}^h + E_{\max} C^h}{E_0 (C_{0.5}^h + C^h)}$$

$$\frac{E_{C_{i,j}}}{E_0} = \frac{C_{0.5_i}^h C_{0.5_j}^h E_0 + C_i^h C_{0.5_j}^h E_i + C_{0.5_i}^h C_j^h E_j + (\alpha_j C_i)^h C_j^h (\min(E_i, E_j) - \beta \cdot (E_0 - \min(E_i, E_j)))}{E_0 (C_{0.5_i}^h C_{0.5_j}^h + C_i^h C_{0.5_j}^h + C_{0.5_i}^h C_j^h + (\alpha_j C_i)^h C_j^h)}$$



Linking molecular events to cell fates and population dynamics



We will use some of the code and data from this manuscript in class

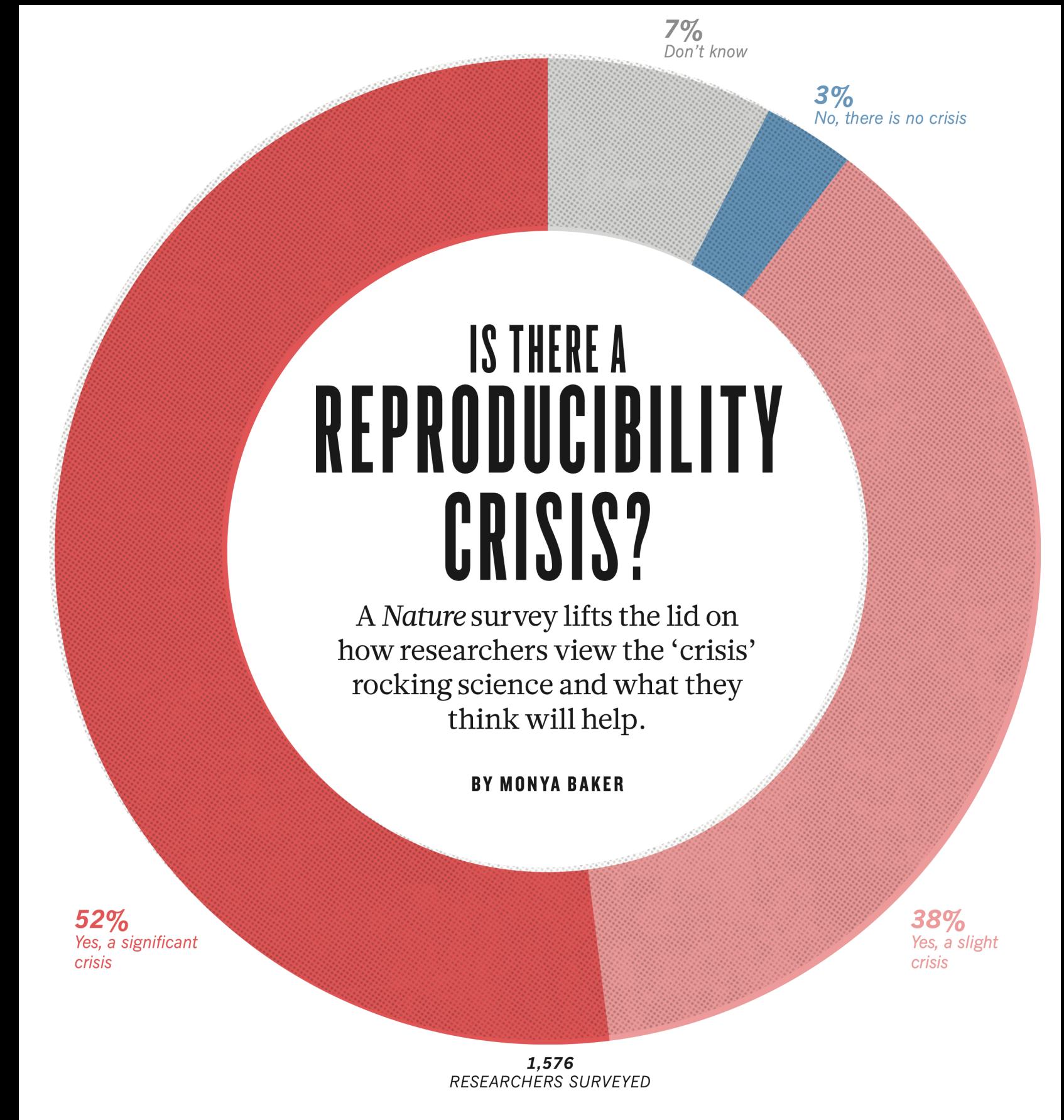
Manage digital data

Computation tools and Open Science

- Good digital data practices can counteract the "reproducibility crisis"
- Your digital notebooks can keep a record of everything you do with your data
- Some basic principles:
 - * *Do not modify raw data directly* (consider it read-only!); save processed data to a new file, if needed
 - * Use digital notebooks to keep track of all changes to the raw data and to graphically summarize your results
 - * Use version control to save snapshots of code, data and visualizations as they change over time

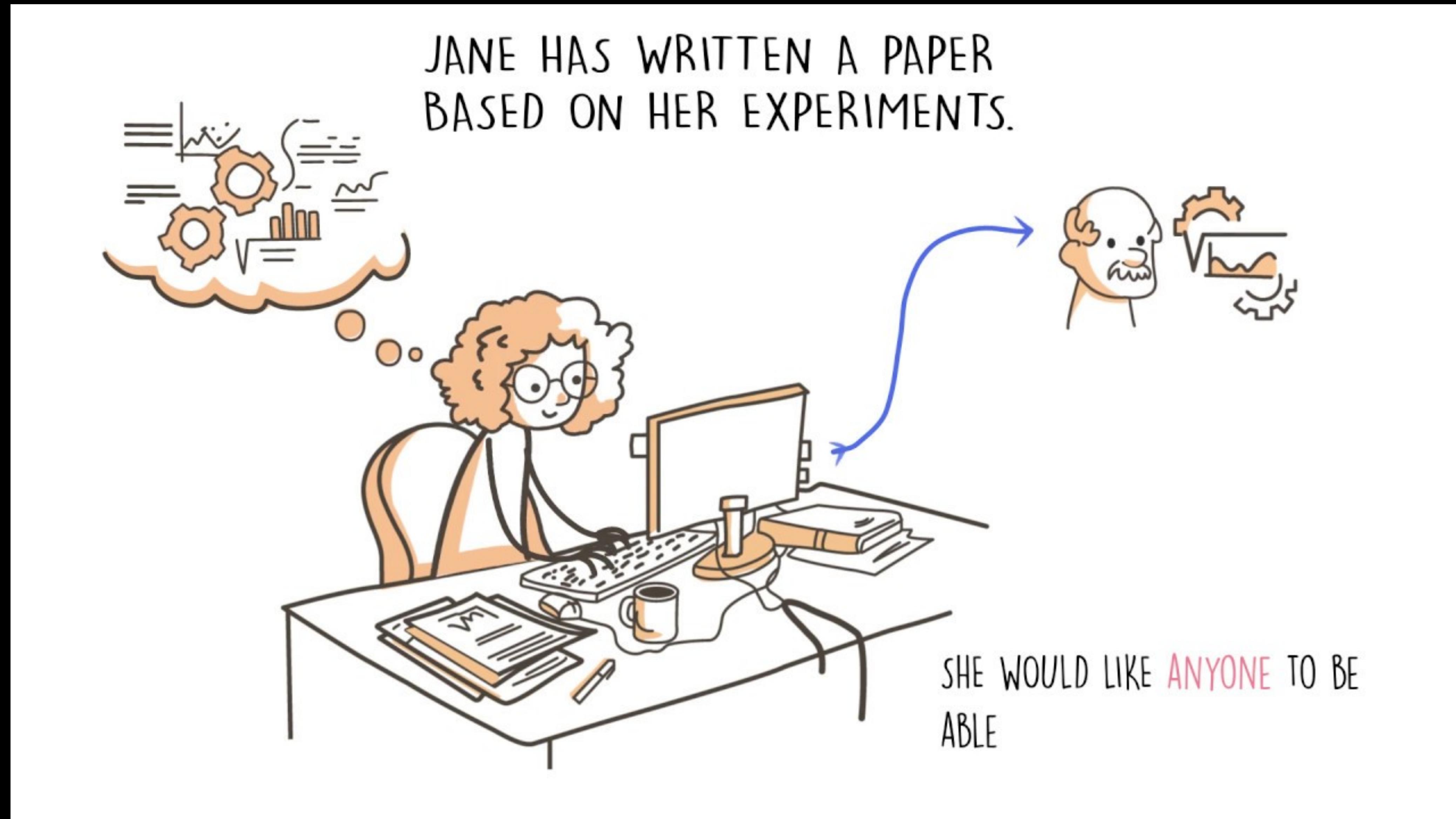
Address the science “reproducibility crisis”

- ~70% of cell biologists reported not being able to reproduce a published experiment (go.nature.com/kbzs2b)
- Many more examples of irreproducible results
- A Reproducibility Initiative is directly assessing this by attempting to reproduce results from dozens of high-impact cancer biology papers <http://www.reproducibilityinitiative.org>
- Open Science can counteract this



<https://www.nature.com/articles/533452a>

Data sharing example



Data sharing, the FAIR way

Box 2 | The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

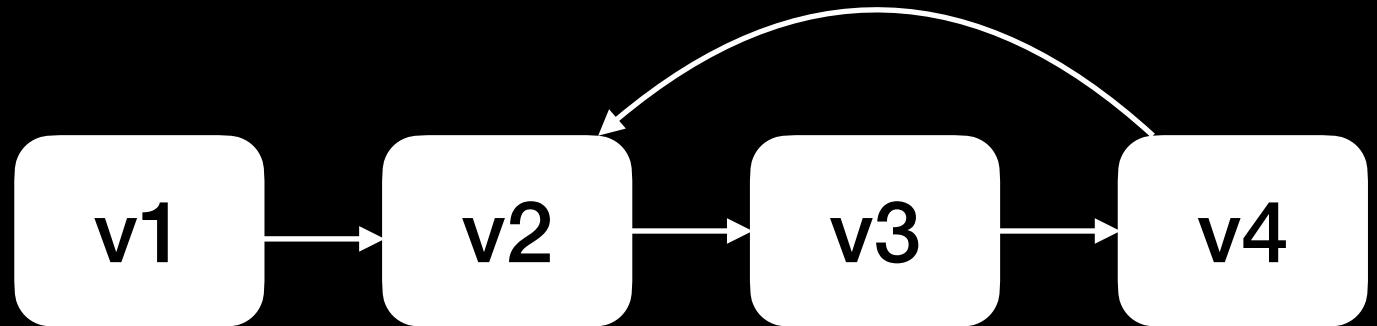
To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

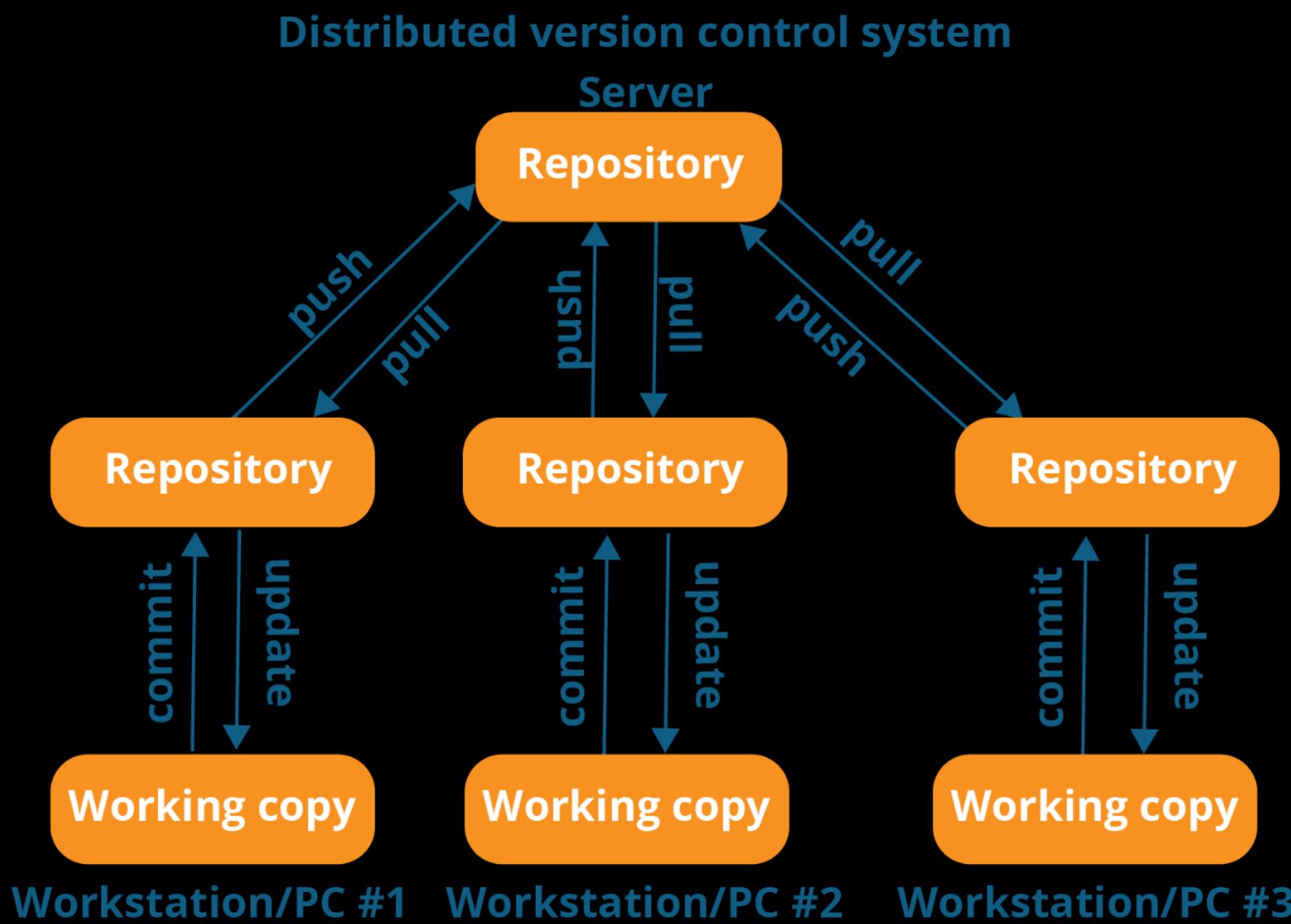
- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

Version Control Systems



- A system for keeping track of the files (and their locations) on your computer and any changes that are made to them; essentially a backup system for your files
- Many applications now have embedded version control systems that allow you to “revert” to an earlier version of your file(s)
- General version control systems can do this for any file type and provide additional capabilities, e.g., when working with other people and their files
- VCS provides a complete long-term change history of every file and traceability (when a change occurs, what was changed and by whom)
- Use of a VCS is in keeping with the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles to promote open science

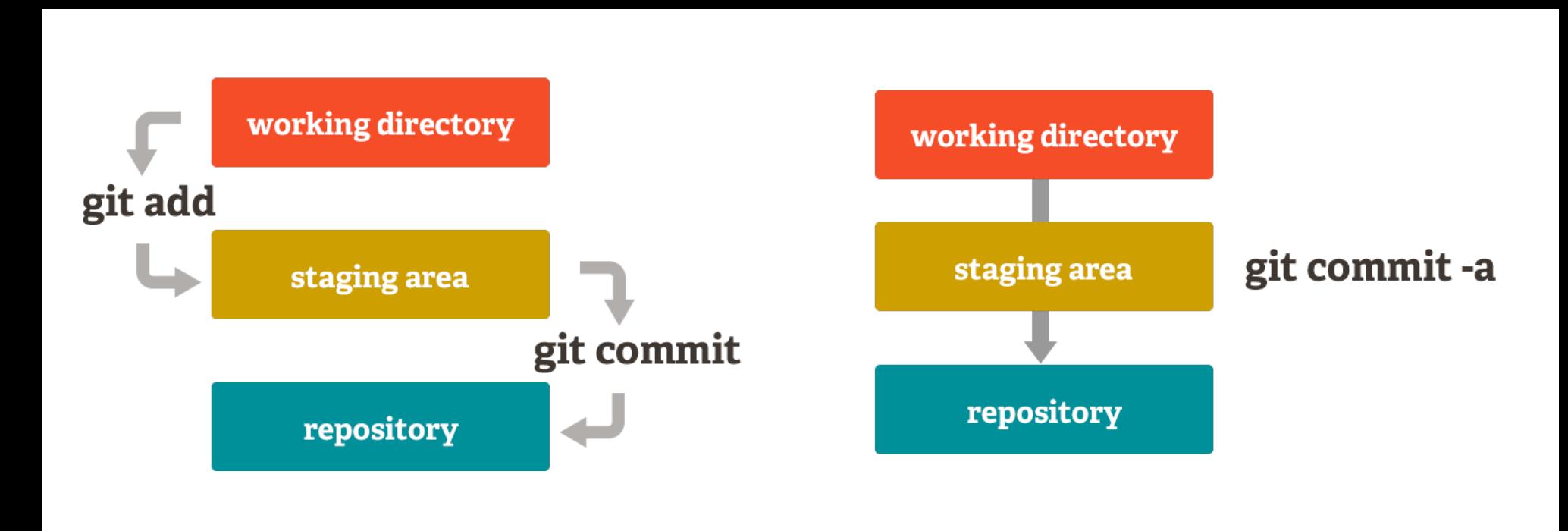
Git is a Distributed VCS



- Git can be used on a single computer to manage files/changes
- Its strength is in collaborative environments
- Branching and merging capabilities allow user-specific changes to be integrated with others'
- Git also has a “staging area” that allows you to precisely control how changes are managed but can also be easily bypassed

Distributed VCS

- Multiple copies of the repository
- No network connection needed (except for merging)
- Local changes to repositories merged with others
- Control of how local changes are merged
- No need for a central repository (can merge with any copy of the repository)



GitHub



- Can be used as a central server to manage Git repositories (repos)
- Provides other tools
 - ◆ User access management
 - ◆ Direct visualization
 - ◆ Modification without Git client
 - ◆ Handles many Git repos
 - ◆ Organization- and individual-level management
 - ◆ Integration with other tools
- Now owned by Microsoft

Git, GitHub & GitHub Classroom

Overview of differences

- Git is a distributed version control system that tracks file changes with *repositories* and facilitates collaboration
- GitHub is a web-based server that can act as a central location for coordinating Git repositories
- GitHub Classroom helps teachers manage how students interact with GitHub
- *You will not need to learn or use Git on your own computers for this class;* you will use Git by your interactions with GitHub and GitHub Classroom
- Your first assignment will help you learn/understand how Git and GitHub work



Learning objectives

- Basics of Git/GitHub (useful resource; only minor use in class)
- Digital image analysis: segmentation and object counting (ImageJ/FIJI)
- Digital image analysis: basics with Python
- Basics of data analysis: linear and nonlinear regression
- Several population dynamics (growth) models
- Analysis of drug dose–response relationships
- Some general data management (loading, manipulating, plotting)

Upcoming lectures

- 2024-02-27: Digital image analysis (ImageJ/FIJI); NOTE class in LH 512 instead of LH 208
- 2024-03-05: Digital image analysis using Python (Hossein Jashnsaz)
- 2024-03-12: Population growth models and single-cell fate analysis
- 2024-03-19: Dose-response relationships
- 2024-03-26: Data analysis exercises in class

Resources

- Class website: <https://vu-csp.github.io/QuantBio> (NOTE: caps in QuantBio required!)
- Thunor @ VU: <https://thunor.app.vanderbilt.edu>
- Assignments posted on class website and in Slack (different GitHub Classroom link for each assignment)

Git & GitHub References

- Learn Git on GitHub: <https://docs.github.com/en/get-started/using-git/about-git>
- Git basics: <https://www.atlassian.com/git>
- Connect to GitHub using SSH keys: <https://docs.github.com/en/authentication/connecting-to-github-with-ssh/about-ssh>
- GitHub-Flavored Markdown: <https://github.github.com/gfm/>

First assignment

- Complete Git/GitHub Classroom assignment: [https://classroom.github.com/a/
zJxKK9A5](https://classroom.github.com/a/zJxKK9A5)