Using simulations to explore treatment strategies in high throughput

Paul Macklin, Ph.D.

Intelligent Systems Engineering
Indiana University

November 16, 2018

Cancer is a systems problem

Interconnected systems and processes:

- Single-cell behaviors
- Cell-cell communication
- Physics-imposed constraints (e.g., diffusion)
- Systems of systems (e.g., immune system)

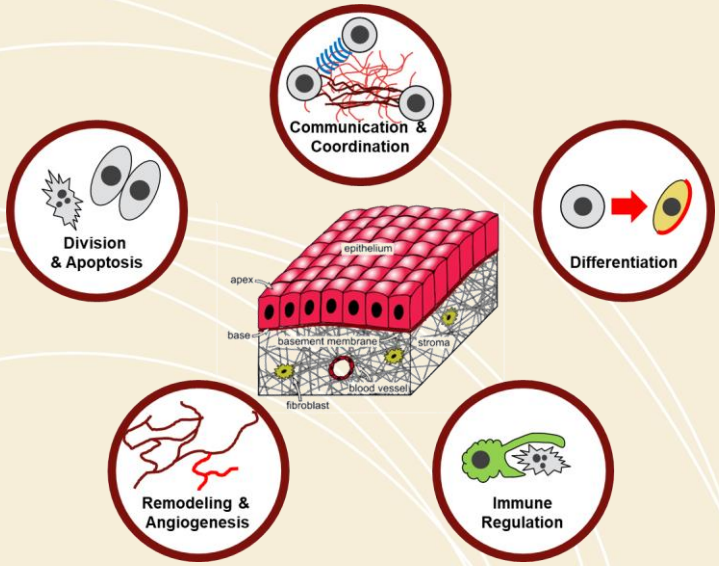
In cancer (and diseases), these systems become dysregulated.

Treatments target parts of these systems.

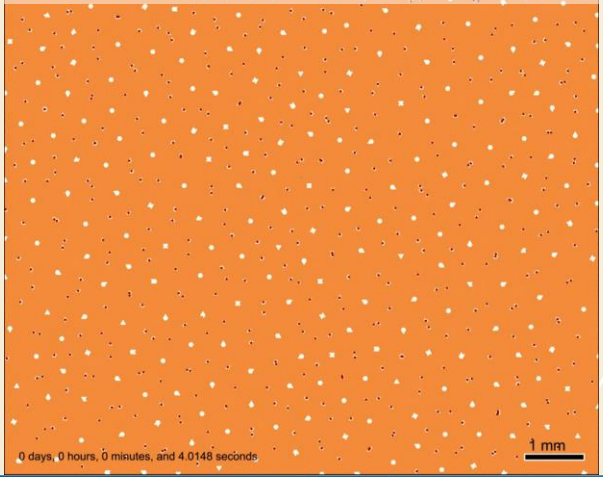
As with any complex system, changing one part can have surprising effects!

Modeling can help **understand** this system. This is **multicellular systems biology.**

If we can **control** these systems, we've arrived at **multicellular systems engineering.**



Metastatic seeding in 1 cm^2 of liver parenchyma



Analogy: multicellular biology as a play

- The **microenvironment** is the **stage**.
- The **cells** are the **actors**.
- The **cells** follow their own **scripts**.

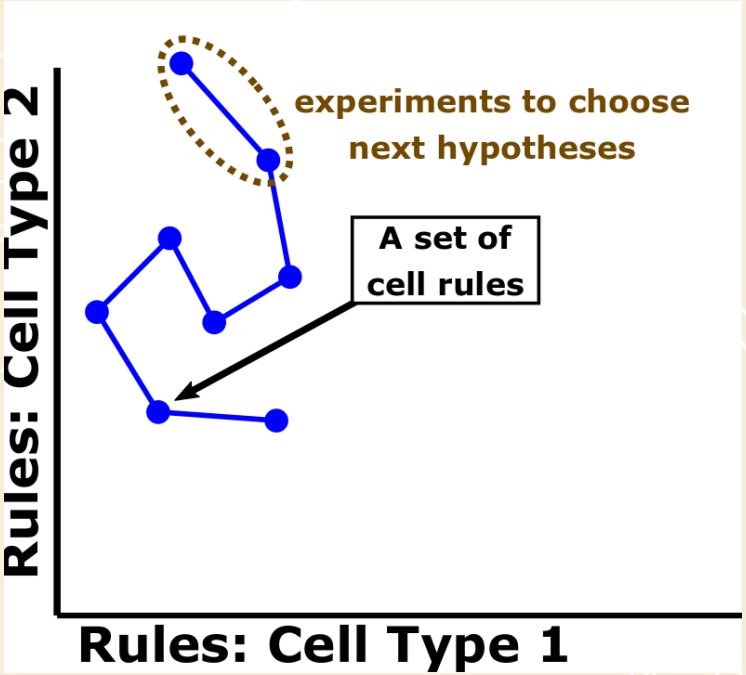
- ***BUT*:**
  - The scripts change based on the stage. (ME-dependent phenotype)
  - The actors' dialog is critical. (cell-cell communication)
  - The actors can tear up and remodel the stage. (tissue remodeling)
  - The actors can ignore their scripts and ad lib. (Mutations, evolution)

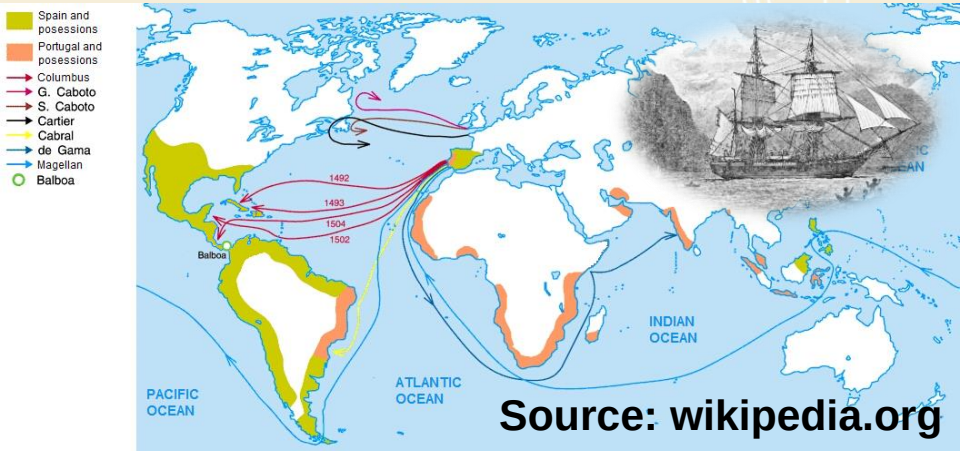
It's our job as scientists to figure out each actor's script by watching the play.

Clinicians and engineers want to rewrite the script.

Exploring the Space of Ideas

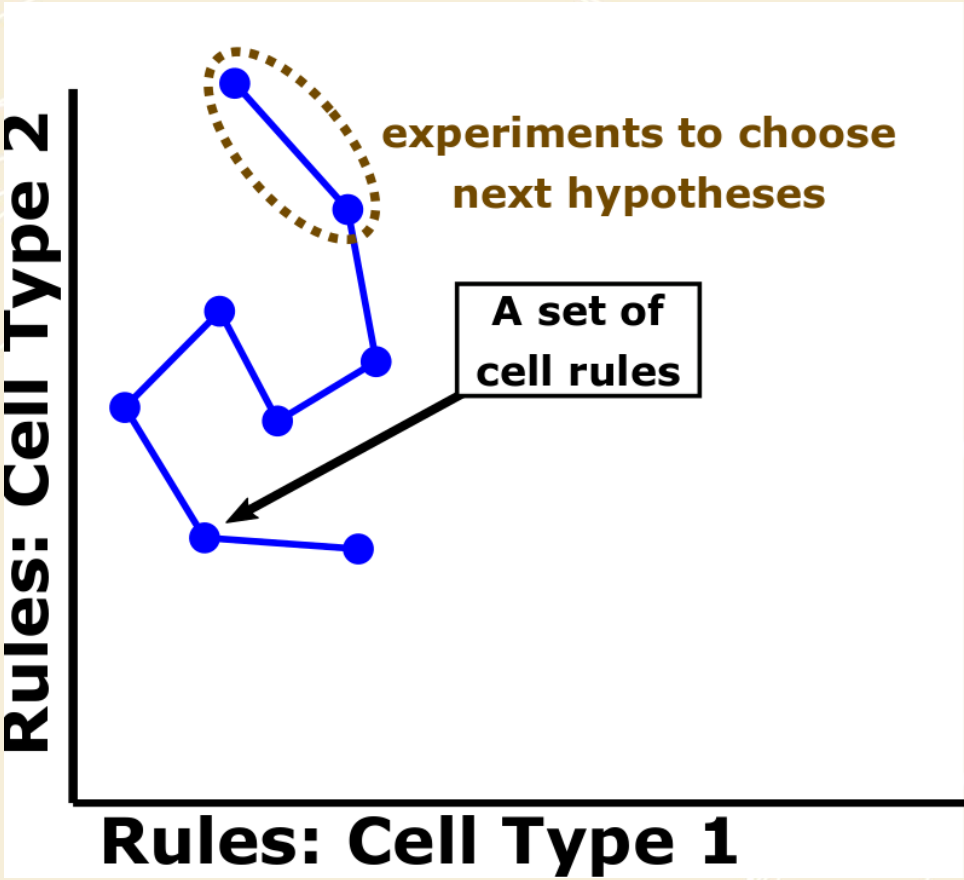
- Observations and intuition drive hypotheses:
  - What are the rules that drive the system's behavior?
  - What therapeutic strategies can disrupt this?
- A **hypothesis set** is a point in a high-dimensional **hypothesis space**
- Experiments help us **trace a path** through hypothesis space.
  - If cells don't behave the way we expect, adjust our rules.
  - If the therapy didn't work as expected, adjust our strategy.
- This exploration leads to **discovery**





Simulations can explore space faster

- Experiments can be expensive:
  - Fixed costs:
    - Cutting-edge instruments
    - Lab facilities
  - Expensive marginal costs:
    - Days or weeks to run
    - Lab personnel
    - Lab supplies
- Simulations have different economics
  - Fixed costs:
    - Software and model development
    - Computing hardware
  - Cheaper marginal costs:
    - Seconds, hours, or days to run
    - Single person can run 100s of experiments
    - Supplies = disk space, electricity



Our simulation toolbox

BioFVM: Simulating 3-D biotransport

Design goal: Simulate multiple diffusing substrates in 3D with desktops or single HTC/HPC nodes

Typical use: pO_2, glucose, metabolic waste, signaling factors, and a drug, on 10 mm^3 at 20 μm resolution

Features:

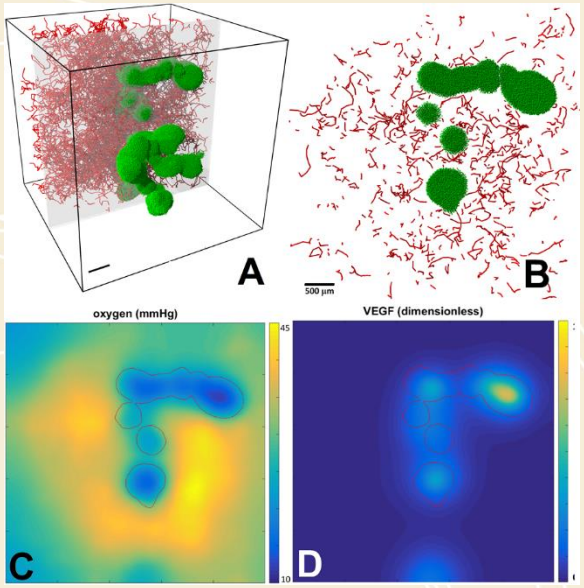
- Off-lattice cell secretion and uptake
- 2^{nd}-order accurate (space), 1^{st}-order accurate (time), numerically stable

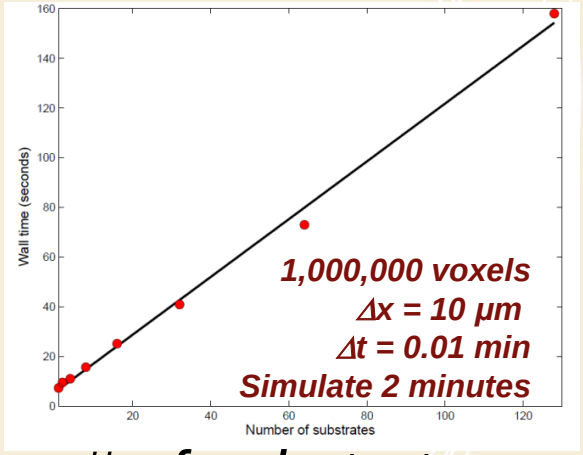
Method:

- Operator splitting, LOD, customized Thomas solvers, etc.
- Standard C++11, cross-platform
- OpenMP parallelization
- O(*n*) cost scaling in # substrates, # voxels
- Easy to simulate 5-10 substrates on 10^6 voxels

Reference: Ghaffarizadeh et al., *Bioinformatics* (2016)

DOI: 10.1093/bioinformatics/btv730





Simulation time

# of substrates

PhysiCell: A multicellular simulator

Design goal: Simulate 10^6 or more cells in 3D on desktops or single HTC/HPC nodes

Features:

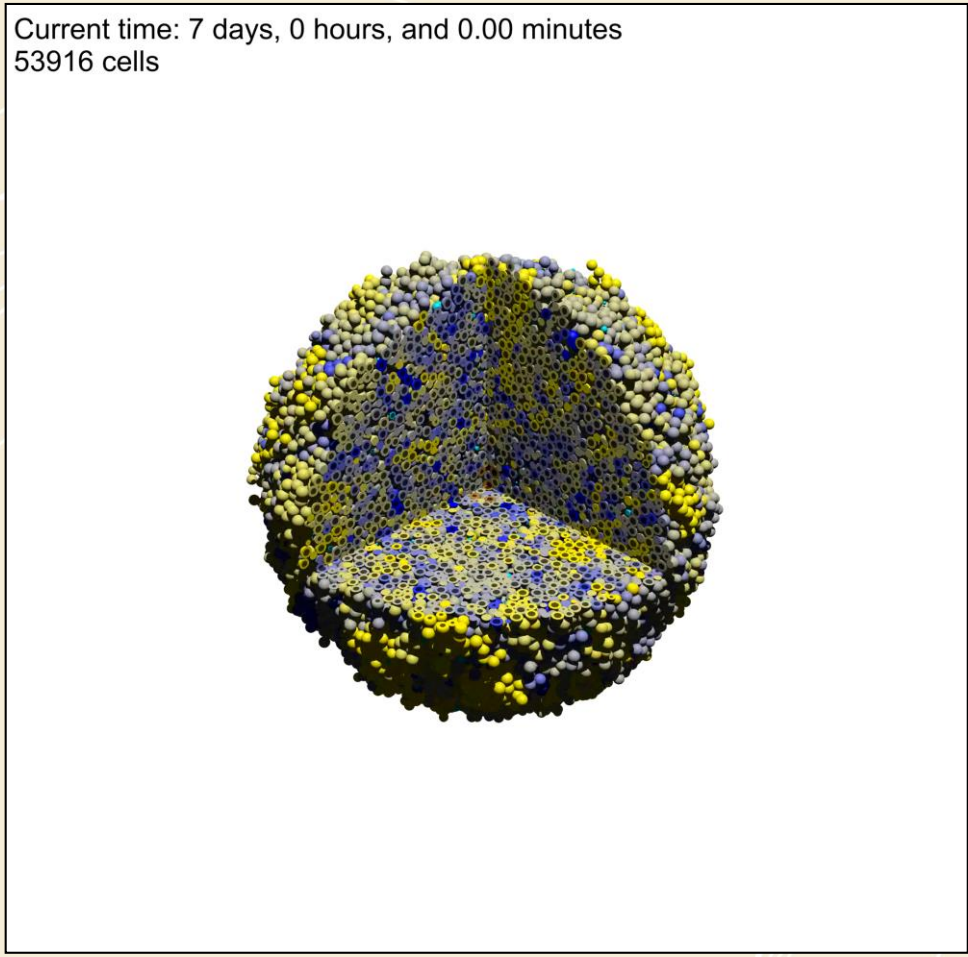
- Off-lattice cell positions
- Mechanics-based cell movement
- Cell processes (cycling, motility, …)
- Signal-dependent phenotype
- Can dynamically attach custom data and functions on a cell-by-cell basis
- **Deployed from Raspberry Pi to Crays**

Method:

- Standard C++11, cross-platform
- OpenMP parallelization
- O(*n*) cost scaling in # cells

Reference: Ghaffarizadeh et al., *PLoS Comput. Biol.* (2018)

DOI: 10.1371/journal.pcbi.1005991



Competition in a 3-D tumor

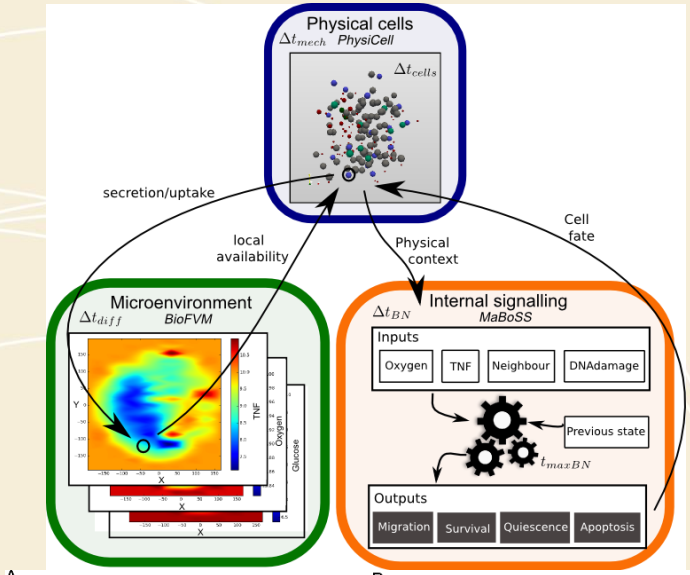
[View on YouTube (8K)]

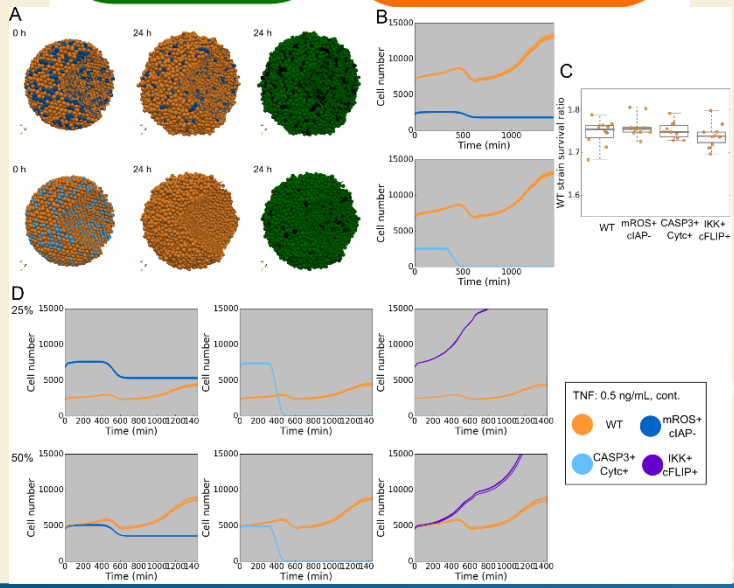
Adding cell signaling …

- First major PhysiCell functionality contributed by an outside group!
- Work lead by Gaelle Letort, with Montagud, Stoll, Barillot, Zinovyev, Calzone (Institut Curie)

  PhysiCell (multicellular simulation framework)
  + MaBoSS (Boolean signaling network framework)
  = PhysiBoSS [**DOI:** 10.1093/bioinformatics/bty766]

- Add a MaBoSS signaling network (with independent parameters and state) to each PhysiCell agent
- **This is a key strength of open source!**
  - Other groups can freely adapt and extend the work, then share the improvements with all.
- **This is also a win for preprints!**
  - Letort found preprint ~1 year before publication
  - PhysiBoSS preprint online before PhysiCell paper
  - Preprint also kicked off dialog with *GigaScience*





Example 1:

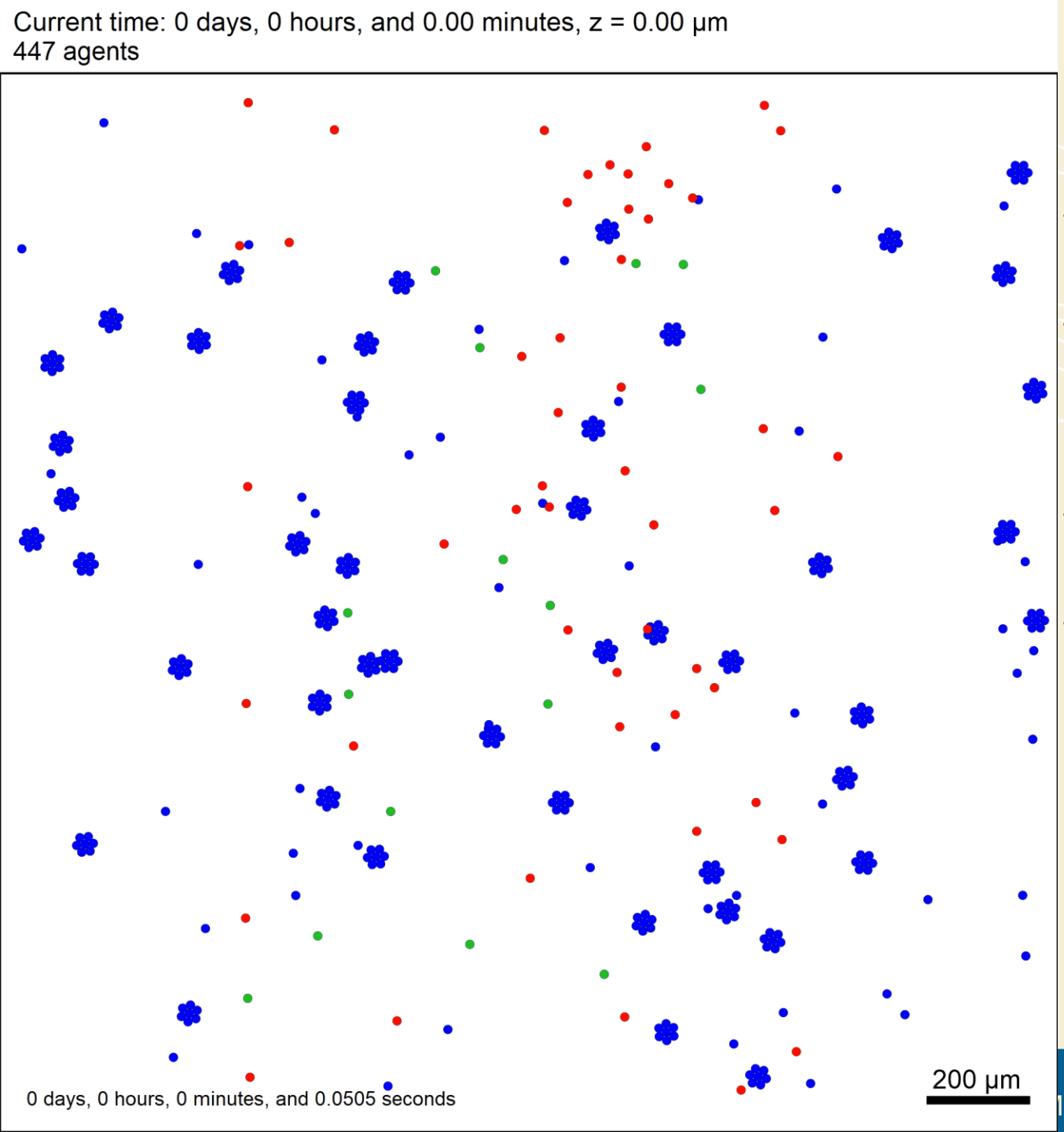
Design rules for synthetic cell colonies

Synthetic multicellular systems engineering

- Suppose we have a handful of phenotypic “programs” we can implement:
  - Secretion of a chemical signal
  - Chemotaxis towards a chemical signal
  - Switching adhesion on or off
  - Switching secretion on or off
  - Switching between directed and random motility

  …
- What happens if we “program” our cells with these rules?
  (e.g., CRISPR-Cas9, opto-genetic switches, mRNAs, etc.)
- Can we use these to deliver a cargo? Could we delivery a therapeutic?
- We can use PhysiCell to test our design choices!

Our goal: Use PhysiCell for high-throughput testing of biorobot designs!



Director cells:

- *Secrete chemoattractant 1 (c1)*

Cargo cells:

undocked:

- *Secrete chemoattractant 2 (c2)*
- *Express receptor (R)*

docked:

- *Stop secreting c2*
- *Stop expressing R*

Worker cells:

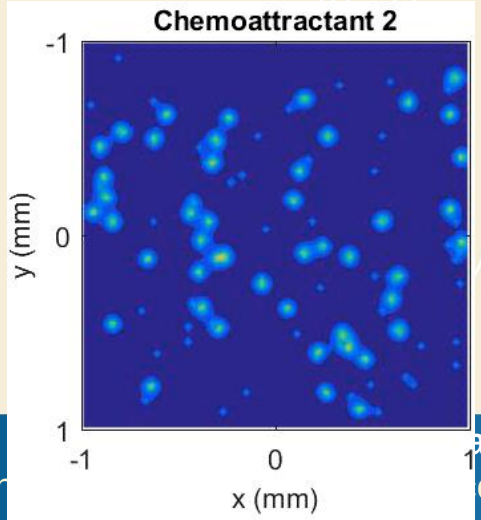
undocked:

- *Chemotaxis towards c2*
- *Attempt docking with cells with R*

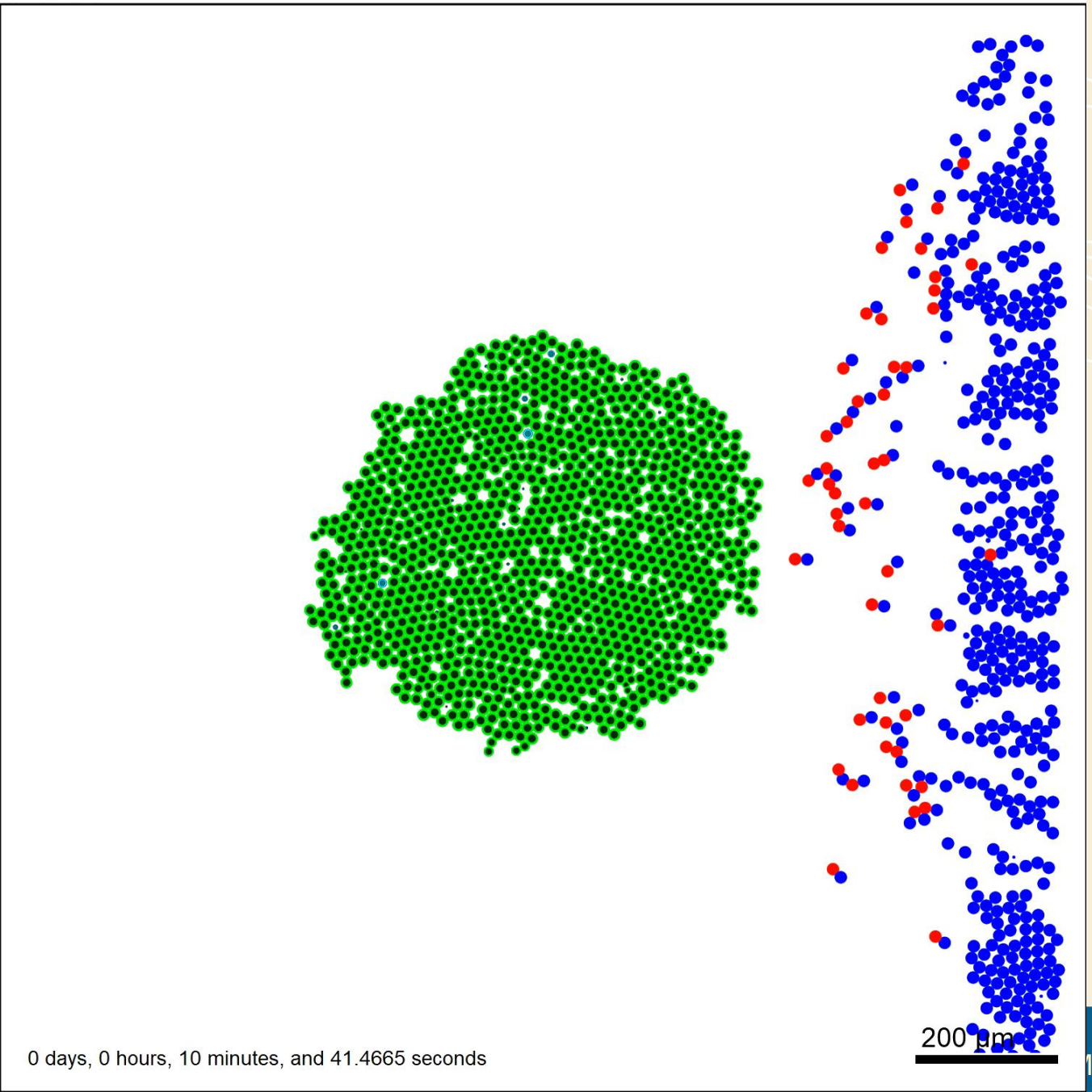
docked:

- *Chemotaxis towards c1*
- *Release cargo if c1 > min*

YouTube: [link]



If we could "program" cells this way, we could envision very Sci-Fi therapies …



Cancer cells:

- Consume oxygen
- Are damaged by the therapeutic (darker = more damaged)
- Can repair their damage
- Apoptose proportionally to their current damage

Cargo cells:

undocked:

- Secrete chemoattractant (c)
- Express receptor (R)

docked:

- Stop secreting c
- Stop expressing R

delivered:

- Secrete the drug

Worker cells:

undocked:

- Chemotaxis towards c
- Attempt docking with cells with R

docked:

- Chemotaxis towards hypoxic zones (along - ∇pO_2)

YouTube: [link]

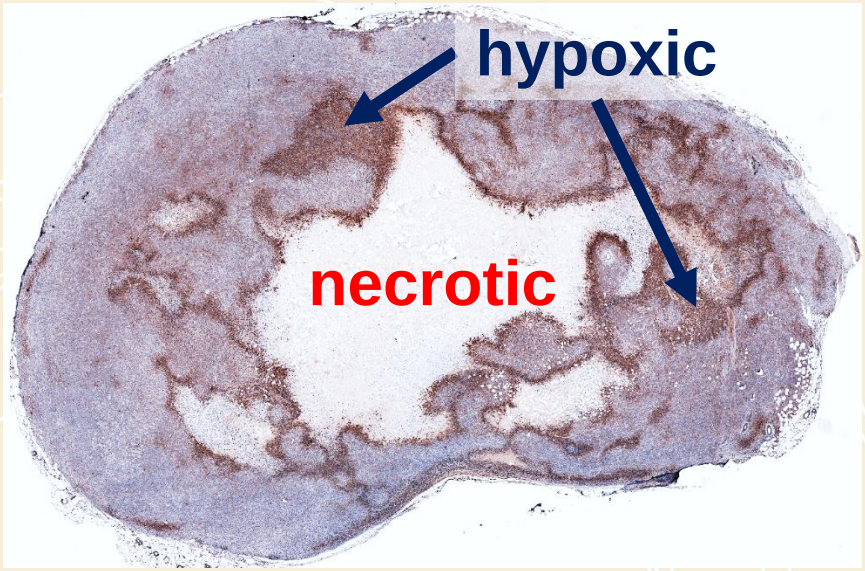
Example 2:

Cancer cell response to hypoxic stress (low pO_2)

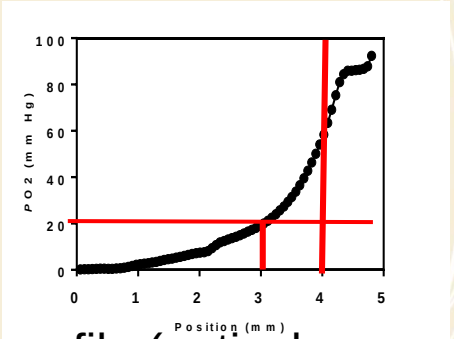
Hypoxia in breast cancer

- **Most breast cancers are hypoxic**
  - normal breast: pO_2 ~ 65 mmHg
  - breast cancer: pO_2 ~ 10 mmHg
  - (Tatum *et al. Int. J. Radiat. Biol.* 2006)

- **Hypoxia drives phenotype changes**
  - Hypoxic responses at ~8-10 mmHg
  - transformation into stem-like cells
  - Increased **motility**
  - Increased **ECM remodeling**
  - Increased glycolysis
  - Increased acidosis
  - (maybe) decreased adhesion
  - VEGF secretion (+angiogenesis)



Hypoxic breast tumor (via hypoxyprobe)
Source: Gilkes lab, Johns Hopkins



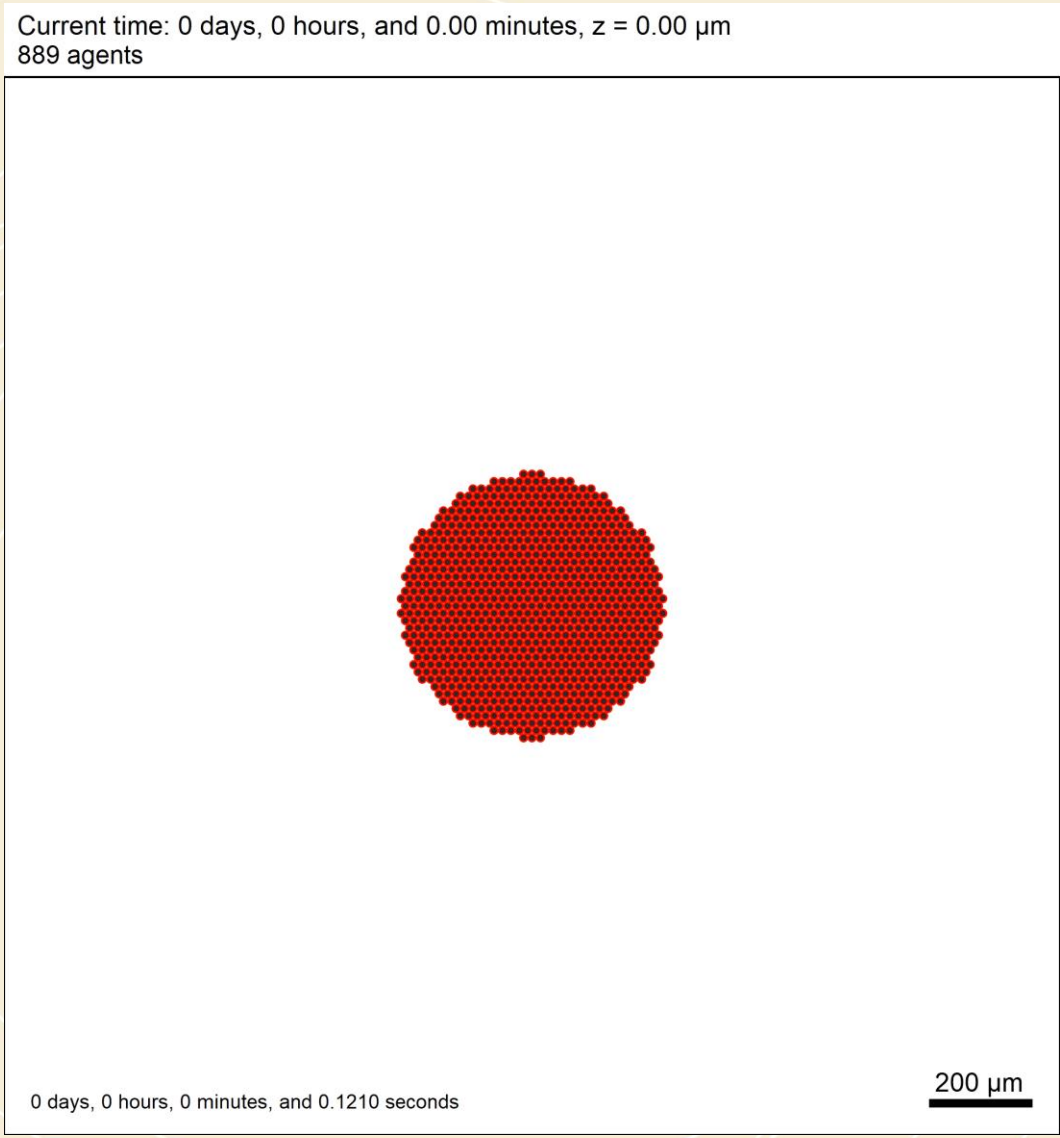
Radial pO_2 profile (optical measurement)
Source: Gilkes lab, Johns Hopkins

What are the rules of hypoxic cell motility?

How persistent is their response to hypoxic stress?

No phenotypic persistence

- GFP+ cells:
  - If pO_2 < 10 mmHg:
    - Same division rates
    - Speed: 0.25 μm / min
    - 50% bias along ∇pO_2
  - If pO_2 > 10 mmHg
    - Set speed = 0.0

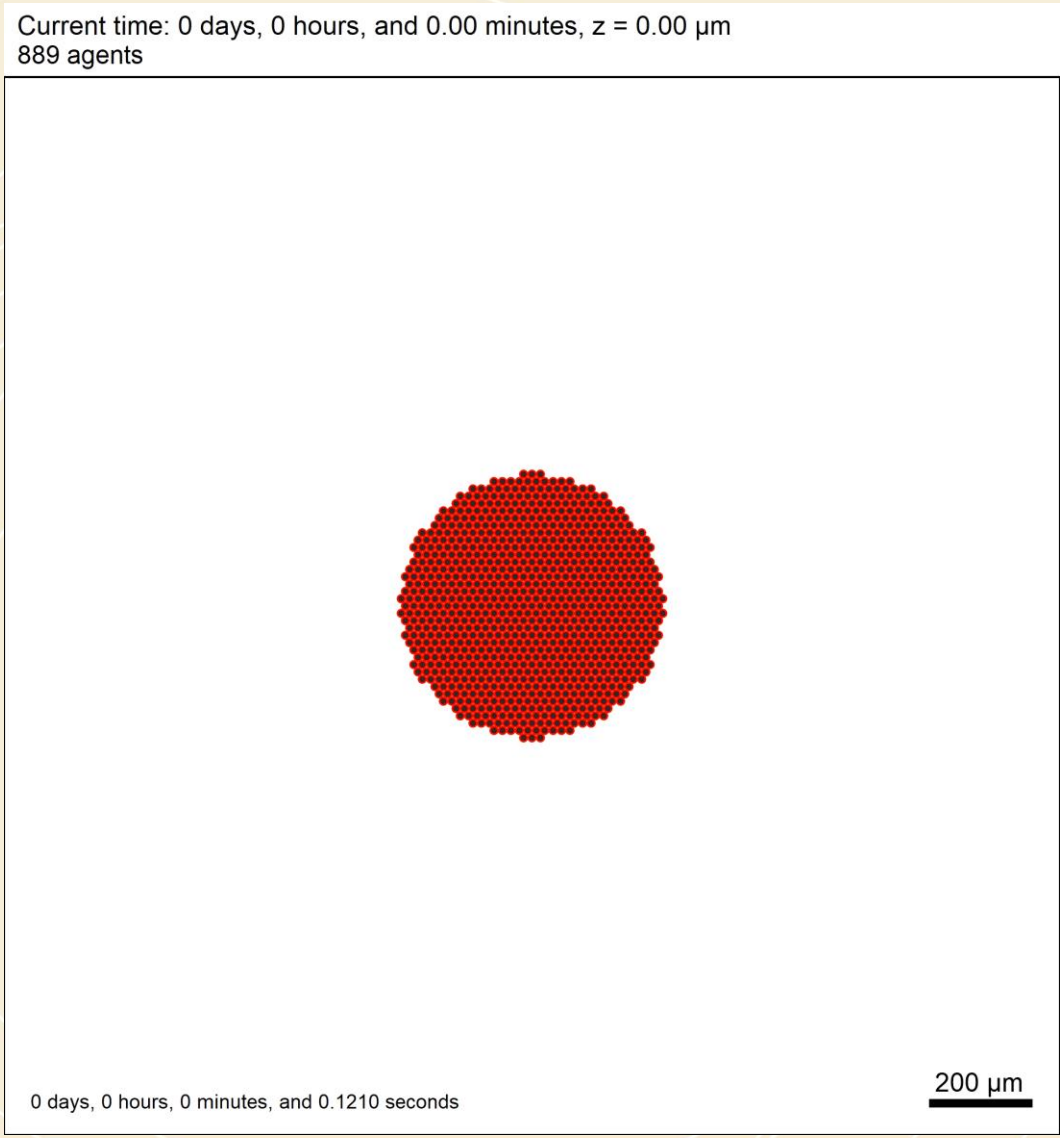


Matching observations:

[ ] GFP+ cells reach edge
[x] Necrotic core (a bit)
[ ] GFP+ microcolonies

Phenotypic permanence

- GFP+ cells:
  - If $pO_2 < 10$ mmHg:
    - Same division rates
    - Speed: 0.25 μm / min
    - 50% bias along ∇pO_2
  - If $pO_2 > 10$ mmHg
    - No change



Matching observations:

[x] GFP+ cells reach edge

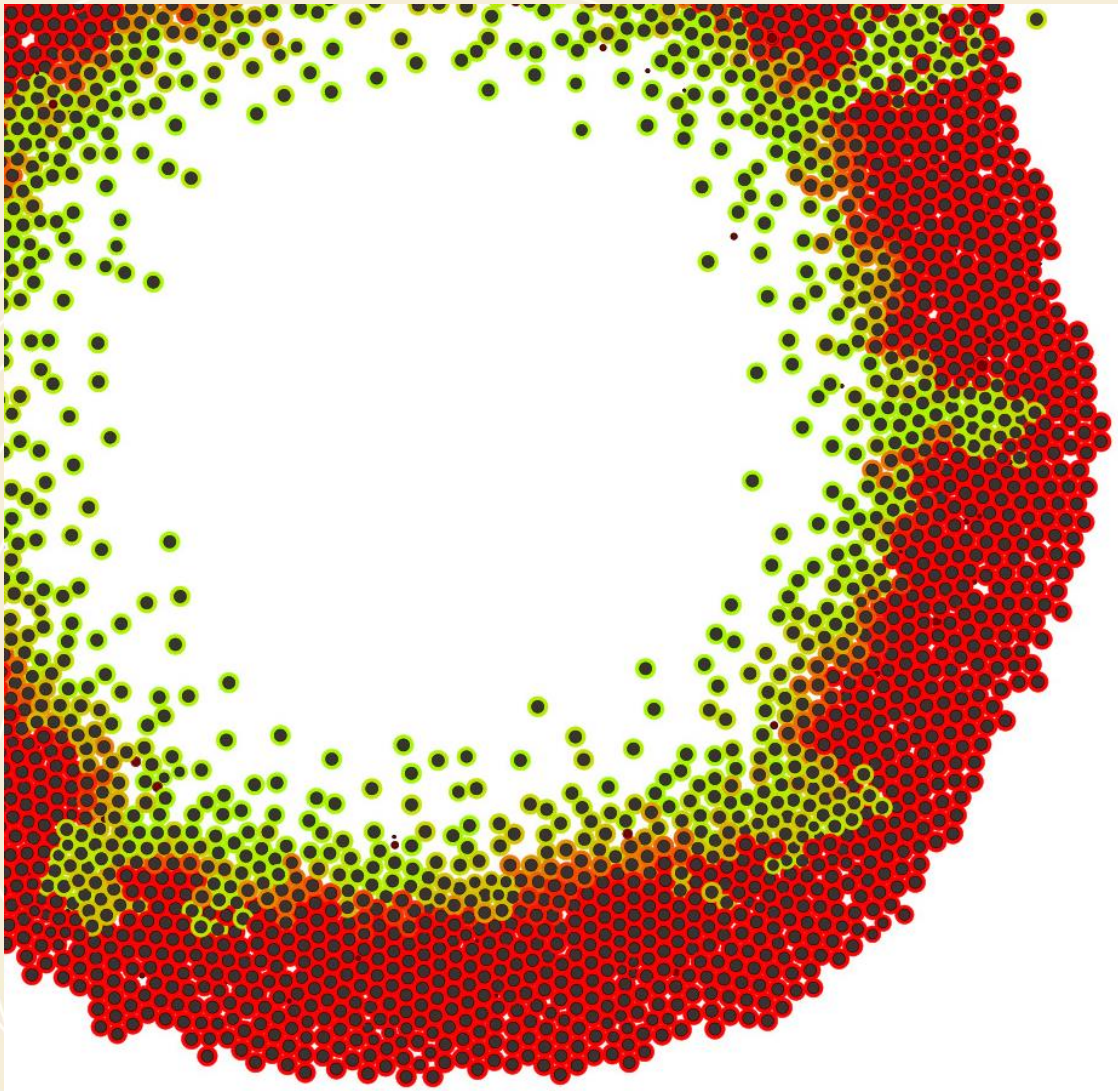
[ ] Necrotic core

[ ] GFP+ microcolonies

Novel prediction: hypoxic plumes

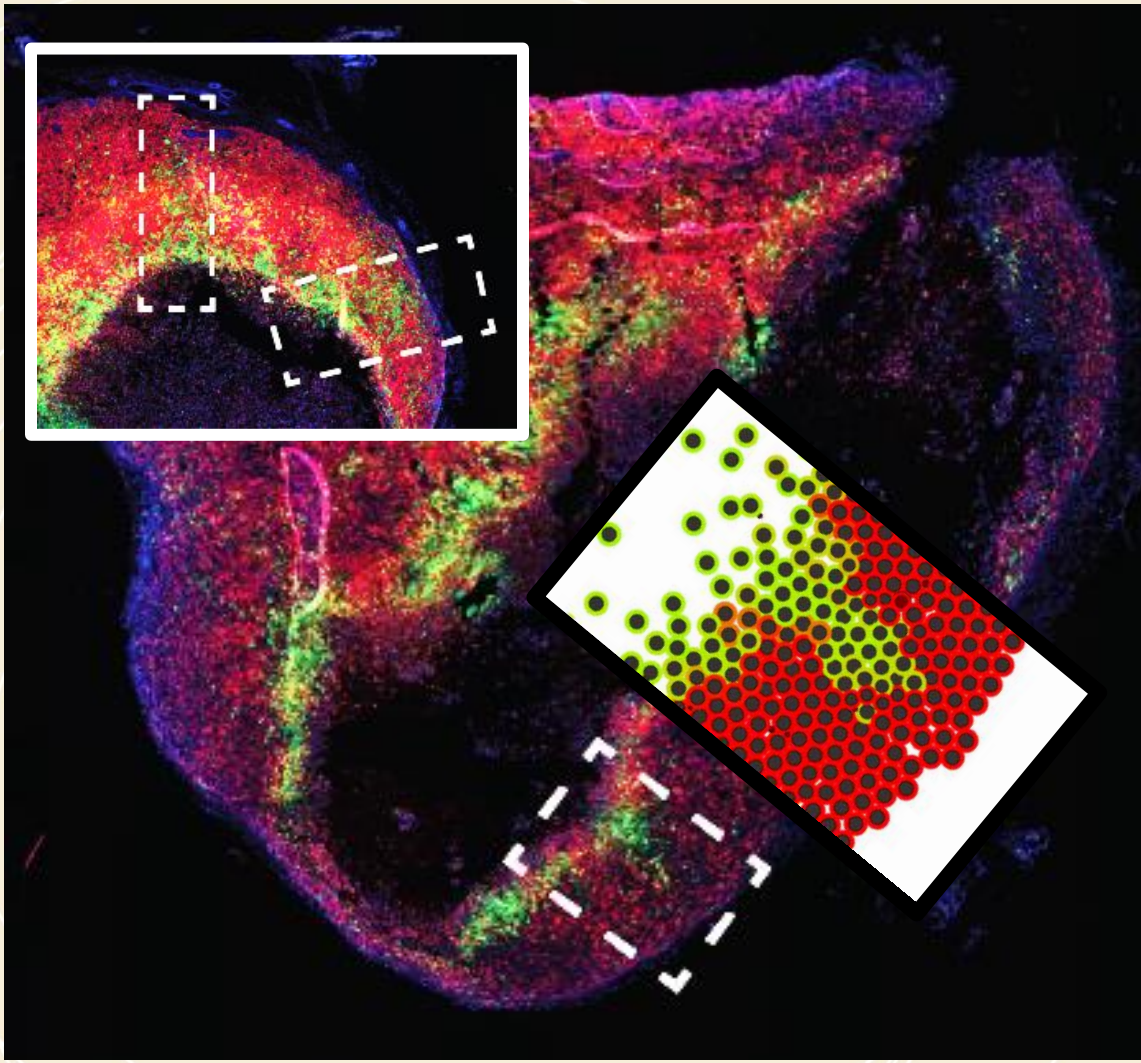
- It ***looks like*** collective motion, but it's purely mechanics
  - Cells are motile
  - If one motile cell finds a gap, it's easier for others to exploit it
  - A "plume" of hypoxic cells grows

- Model suggests a **therapeutic strategy**:
  - Make hypoxic response less persistent to reduce escape.



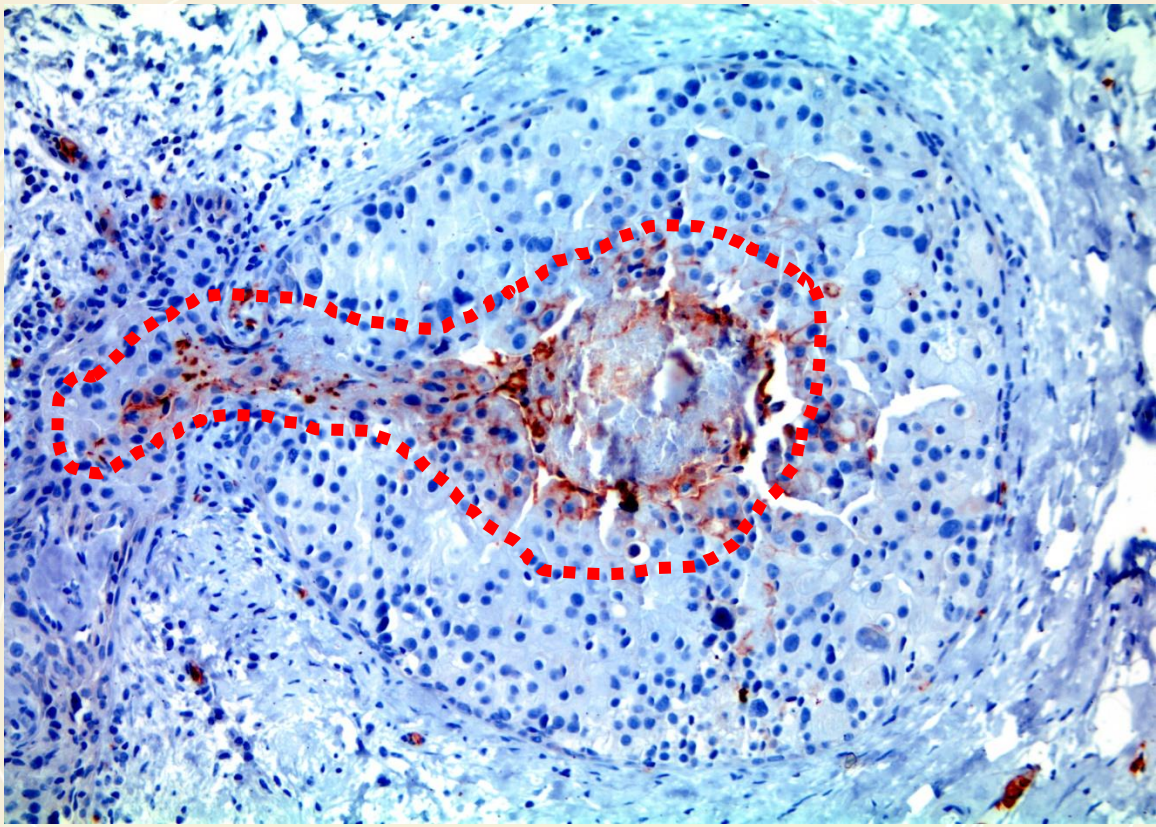
Novel prediction: hypoxic plumes

- It ***looks like*** collective motion, but it's purely mechanics
  - Cells are motile
  - If one motile cell finds a gap, it's easier for others to exploit it
  - A "plume" of hypoxic cells grows
- Model suggests a **therapeutic strategy**:
  - Make hypoxic response less persistent to reduce escape.
- **They're observed *in vivo***
  - MDA-MB-231 in mice at ~20 days
  - **Source:** Gilkes lab (JHU)
- Also observed clinically
  - DCIS pathology (GLUT-1)
  - Source: Bob Gatenby



Novel prediction: hypoxic plumes

- It ***looks like*** collective motion, but it's purely mechanics
  - Cells are motile
  - If one motile cell finds a gap, it's easier for others to exploit it
  - A "plume" of hypoxic cells grows
- Model suggests a **therapeutic strategy**:
  - Make hypoxic response less persistent to reduce escape.
- **They're observed *in vivo***
  - MDA-MB-231 in mice at ~20 days
  - **Source:** Gilkes lab (JHU)
- **Also observed clinically!**
  - DCIS pathology (GLUT1)
  - **Source:** Bob Gatenby



Example 3:

Immunosurveillance

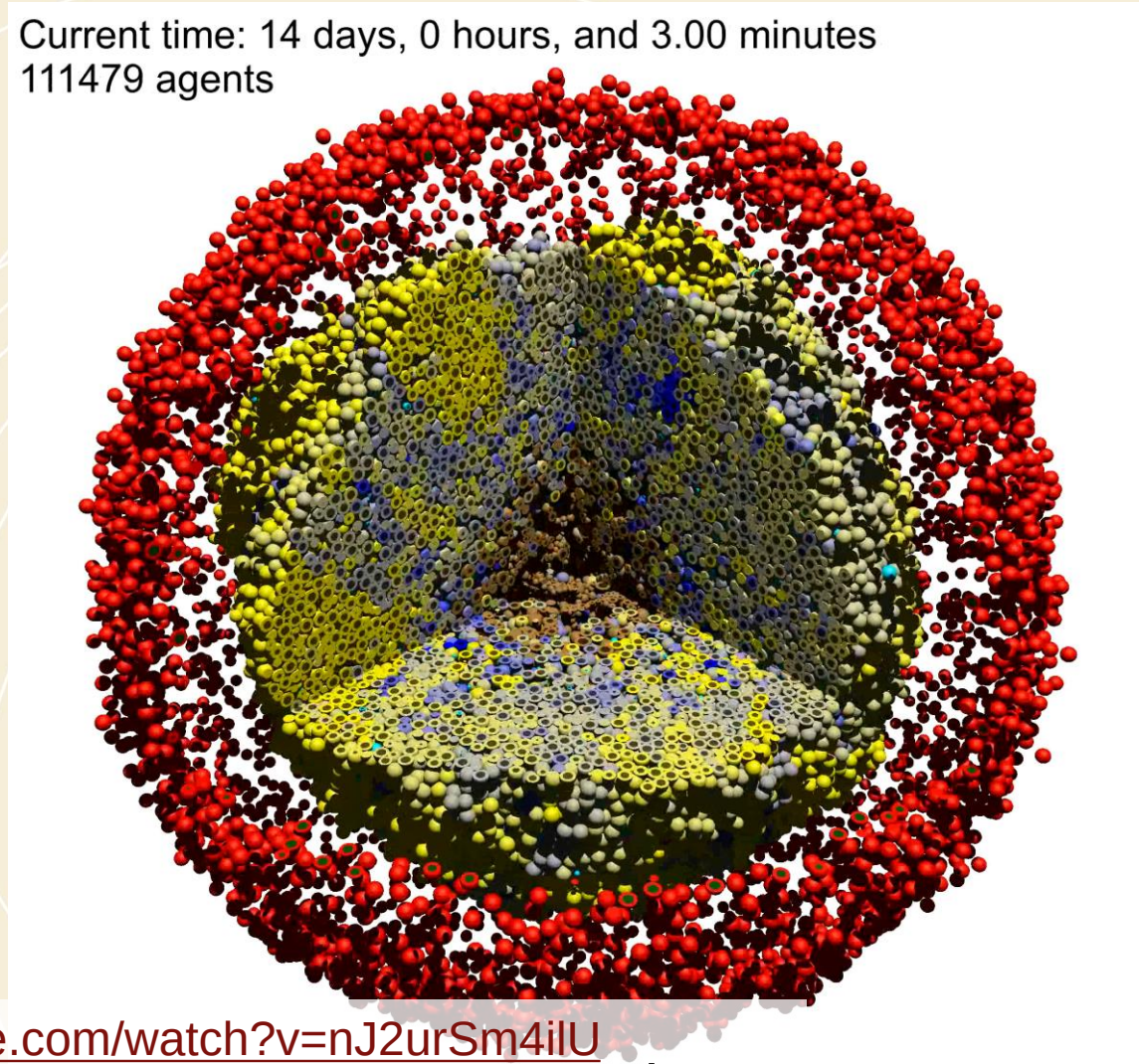
Simple model of cancer immune response

Heterogeneous tumor cells:

- Cycle entry rate scales with O_2
- Cells necrose in very low O_2
- Yellow cells are most proliferative; blue are least
- Yellow cells are most immunogenic (simplified model of MHC)

Immune cells (red):

- Biased random walk towards tumor
- Test for contact with cells
- Form adhesion
- Attempt to induce apoptosis (e.g., via FAS receptor), with rate dependent on immunogenicity
- Eventually detach from cell, continue search



Immune attack on a 3-D tumor

YouTube (4K): https://www.youtube.com/watch?v=nJ2urSm4ilU

Paper: https://doi.org/10.1101/088773

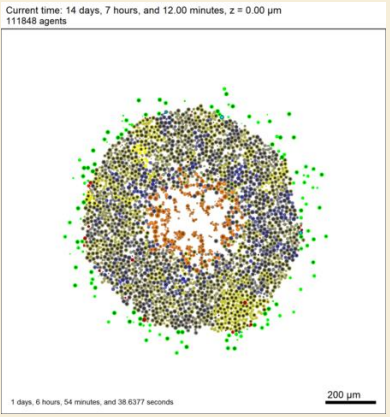
Scaling up

Scaling up from demo to science …

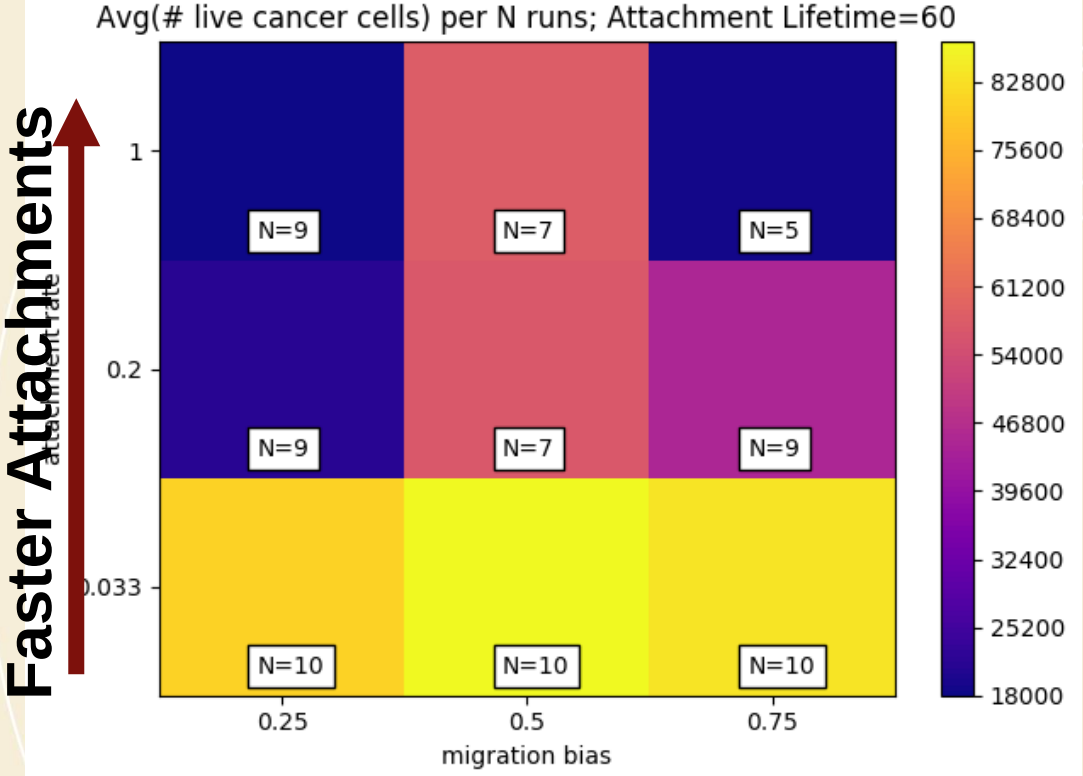
- Early **insight**: immune cell homing is ***non-intuitive***
- Key immune cell parameters:
  - Random motility bias (biased random walk):
    - How much randomness to we allow in motility?
  - Immune cell attachment rate:
    - How quickly do immune cells form new adhesions, instead of wandering?
  - Immune cell attachment lifetime:
    - How long do immune cells try to kill before giving up?
- **Combinatorics**:
  - 3 parameters, 3 levels per parameter
  - $3^3 = 27$ simulations
- Simulations are **stochastic**! Need at least 10x replicates for each condition!
  - $3^3 \times 10 = 270$ simulations
  - 2 days per simulation → **1.5 years** of computing!!

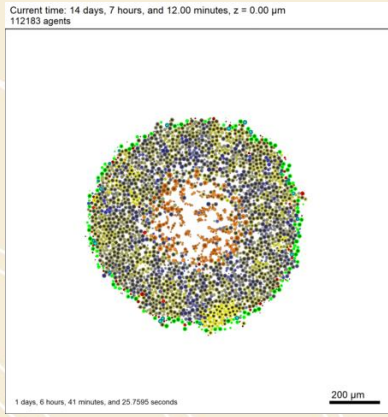
We need high-throughput computing to do the science!

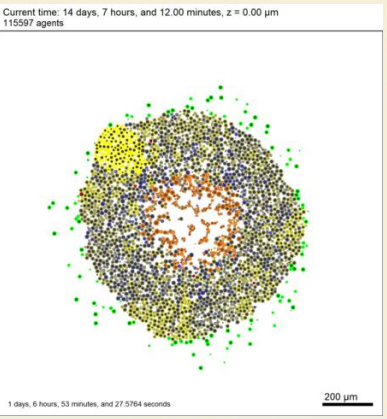
Varied: migration bias & attachment rate



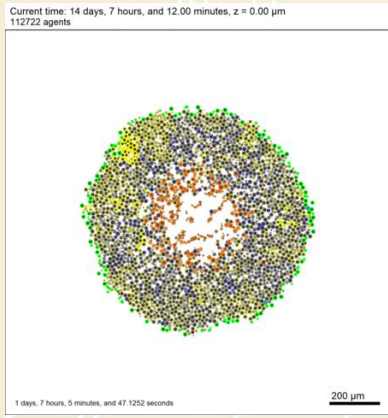
Blue = better





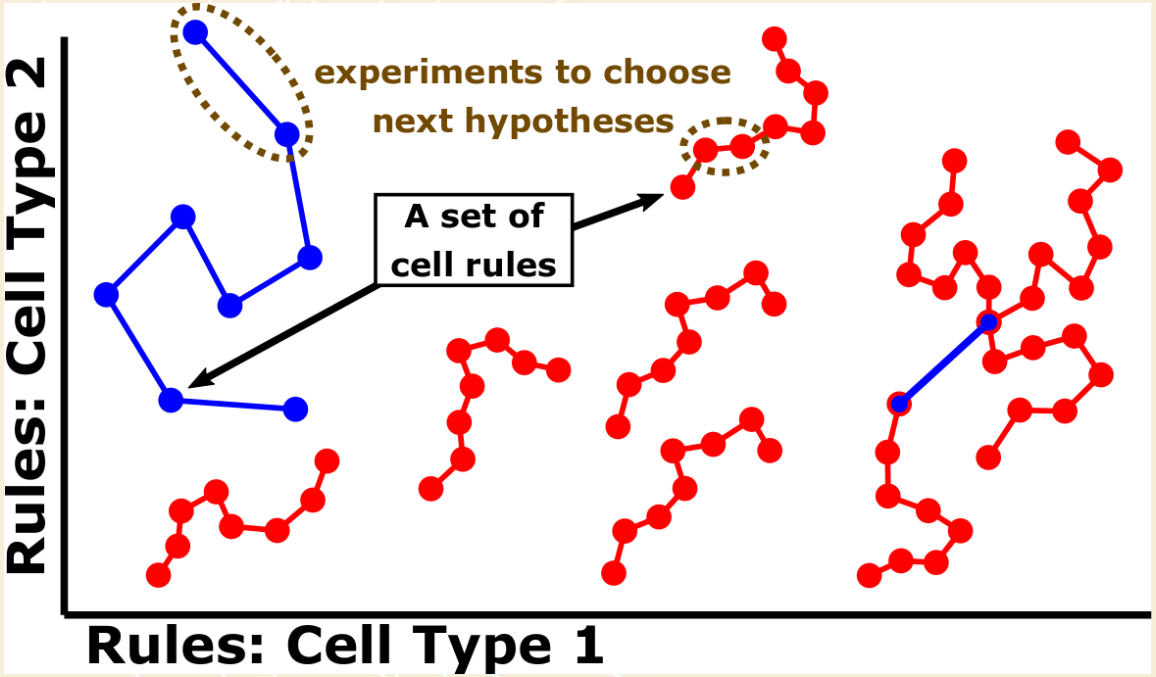


Less Random



High-throughput exploration

- Run many copies of the model at once with high-throughput computing (HTC)
  - Explore more of the space of treatment ideas at once
  - More likely to discover a winning strategy.
- **Next:** accelerate models runs with AI
- **Next:** Use reinforcement learning to guide our exploration

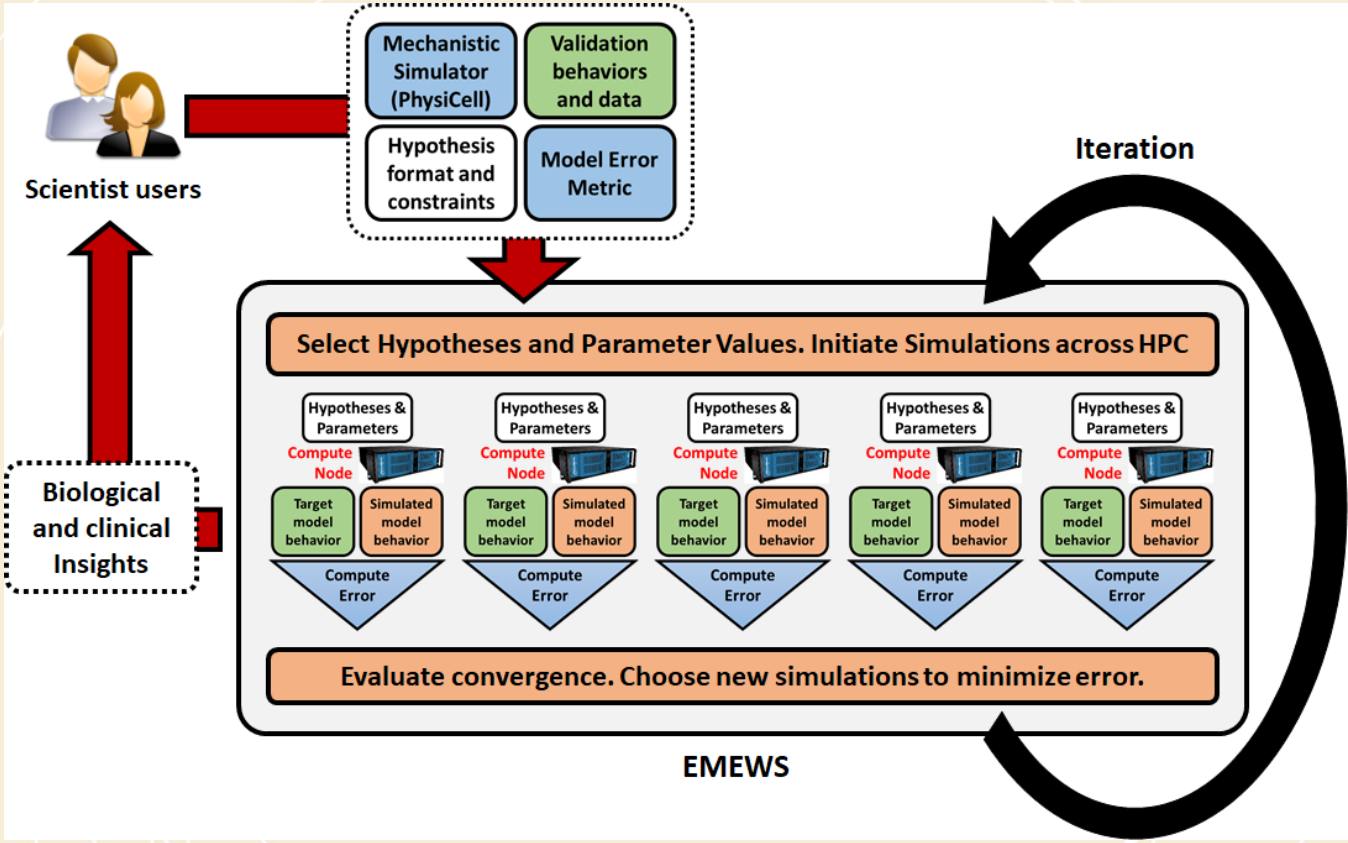


Discovery as optimization

How do we reframe discovery as an optimization problem?

Key Components:

- [X] Simulate multiple substrates (BioFVM)
- [X] Simulate many cells (PhysiCell)
- [X] A mechanistic model of the problem
- [X] High-throughput run capacity (EMEWS)
- [X] A model error metric



We use the EMEWS platform from Argonne National Lab to manage this adaptive workflow. *BMC Bioinformatics* (2018, accepted), *bioRxiv* 196709

Challenges I

- **Scientific:**
  - How do we compare simulation behavior to experiments?
    - Next generation: machine vision to "see" similarity
  - How do we connect single-cell behaviors with molecular biology and drugs?
    - Simulation models explore the *dynamics* of therapy strategies
    - Machine learning finds the *connections* between molecules and phenotype
  - How do we "translate" from short-ish simulations to clinical endpoints?
    - Need surrogate, mid-term markers that correlate with long-term clinical outcome
  - Combine strengths of simulations (test *dynamics* of strategies) with machine learning (find the connections between molecular biology and phenotype)
- **Computational / Technical:**
  - How do we improve speed of simulations?
    - GPU computing, Hybrid OpenMP+MPI, deep neural networks to approximate models

Challenges II

- **Supporting exploration**
  - Funding agencies favor hypothesis-driven projects
  - Exploration is viewed as a "fishing expedition" (pejoratively)
  - Key discoveries are serendipitous, found in exploration or "failed" experiments.
  - Would the NIH or NSF have funded Darwin's "fishing expedition"?
    - Situation is improving today for consortia, but less so for investigator-driven work.

- **Partnering with industry**
  - How do we sort out licensing and IP?
    - Community should "own" the public goods – the core libraries
    - Industry funders should "own" the specific applications – the cancer simulators they pay for.
    - Pharma shouldn't insist on "owning" the IP to the entire software stack. Just the IP they create with the software.
      - A playwright gets copyright for her plays, not the office suite she used to write it.
  - How can we incentivize industry ….
    - to take advantage of "free" software resources?
    - to share data for mutual benefit?
    - to "pay" for free by contributing to the software?

Challenges III

- **Sustainable software:**
  - Creating and maintaining *polished and tested* software
  - Writing documentation
  - Training and supporting new users
  - Grant agencies tend to fund *software applications* but not the software
    - They fund cancer projects that use PhysiCell, but it takes multiple grants to support developers. (20% on NIH grant 1, 15% on NSF grant 2, ….)
  - University bureaucracy makes it difficult, too!
    - Can users buy $500 of support? (No! We have to negotiate a subcontract! Ugh.)
    - Can users donate to the lab? (Not easily!)
    - Can I use crowdsourcing like Patreon? (Not sure! Each office says to ask another office.)

Acknowledgements: Partners

- **Breast cancer hypoxia:** Daniele Gilkes lab (Johns Hopkins)
- **Breast cancer invasion:** Andrew Ewald lab (Johns Hopkins)
- **Colon cancer metabolism:** Stacey Finley (USC)
- **Colon cancer organoids:** Shannon Mumenthaler (USC)

- **PhysiCell:** Randy Heiland (IU)
  - **alumni:** Samuel H. Friedman (OKSI), Ahmadreza Ghaffarizadeh (USC)

- **IU PhD students:** John Metzcar (hypoxia, invasion), Yafei Wang (liver metastases, nanotherapy), Furkan Kurtoglu (multicellular metabolism), Aneequa Sundus (cyanobacteria, synthetic multicellular systems)

- **IU Undergraduates:**
  - **Metastasis:** B. Fischer, D. Murphy, K. Konstantinopoulos, B. Duggan
  - **Nanotherapy:** T. Mahjan
  - **Jupyter GUIs:** E. Bower, D. Mishler, T. Zhang
  - **PhysiCell tech:** E. Freeman, G. Lahman

- **HTC via EMEWS:** Jonathan Ozik, Nicholson Collier, Justin Wozniak, Charles Macal (Argonne National Lab), Chase Cockrell, Gary An (U. Chicago)

Acknowledgements: Funders

- **NIH:**
  - NIH CSBC U01 (1U01CA232137), **PIs** Finley / Macklin / Mumenthaler
  - Provocative Questions grant (1R01CA180149), **PIs** Agus / Atala / Soker
  - NIH PS-OC center grant (5U54CA143907), **PIs** Agus / Hillis
  - High-End instrumentation grant (1S10OD018495-01), **PI** Foster
  - HuBMAP CCF contract (OT2 OD026671), **PI** Borner

- **NSF:**
  - Engineered nanoBIO Hub (1720625), **PI** Fox. **Co-PIs** Douglas, Glazier, Macklin, Jadhao
  - Cyanobacteria / Synthetic Biology (1818187), **PI** Kehoe, **Co-PI** Macklin

- **Breast Cancer Research Foundation & JKTGF**, **PI** Macklin
  - projects with **PIs** Agus, Gilkes, Peyton, Ewald, Newton, Bader

BCRF BREAST CANCER RESEARCH FOUNDATION

Thank you!