Virtual-Tissue Modeling as a Tool in Biology

James A. Glazier
Biocomplexity Institute
Indiana University
Bloomington, IN 47405
USA

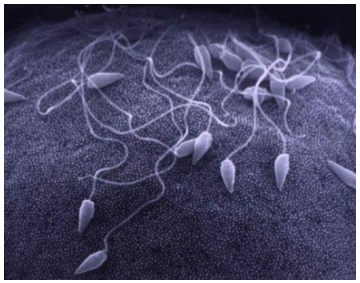
CompuCell3D User Training Workshop
Indiana University, Bloomington, Indiana
Sunday, August 09, 2015

IU Bloomington: [Dr. Dragos Amarie], [Dr. Julio Belmonte], Dr. Stephen Burns, Dr. Sherry Clendenon, [Alin Comanescu], [Clayton Davis], [Dr. Rita de Almeida*], Xiao Fu, [Aaron Dy], [Garth Gast], Dr. Thomas Gast, Dr. J. Scott Gens, [**Randy Heiland**], [Dr. Susan Hester], [**Dr. Mitja Hmeljak**], [Dr. Srividhya Jayaraman], Dr. James Klaunig, [Dr. Roeland Merks*], Nazanin Mosavian, Guilherme Oliveira, [Dr. Nikodem Poplawski], [Ryan Roper], [Dr. Abbas Shirinifard*], Dr. James Sluka, **Dr. Endre Somogyi**, **Dr. Maciej Swat**, [Dr. Gilberto Thomas*], [Ruei Wu], [Benjamin Zaitlen], [Dr. Ying Zhang]. Indiana University, School of Medicine: Dr. Robert Bacallao, Dr. Kenn Dunn. [University of Houston: Dr. Maria Bondesson, Dr. Jan-Ake Gustafsson, Dr. Catharine McCollum]. EPA: Dr. Thomas Knudsen, Dr. Imran Shah, [Dr. Nicole Kleinstreuer]. [University of Michigan: Dr. Santiago Schnell]. [KUMC: Dr. Charles Little]. University College London: Ana S. Dias, Irene de Almeida, Dr. Claudio Stern, University of Dundee: Dr. Mark Chaplain. Moffitt Cancer Center: Dr. Heiko Enderling. CRG Barcelona: Dr. James Sharpe. [Cambridge University: Dr. Octavian Voicelescu]. University of Paris 6: **Dr. Francois Graner**. University of Wisconsin Milwaukee: Dr. Roshan D'Souza. [UCSF: Dr. Tony Hunt], Emory University: Dr. Fereydoon Family, Dr. Hans Grossniklaus. Georgia State University: Dr. Yi Jiang.

Support: NIH NIGMS-R01GM076692, NIGMS-R01GM077138, NIH U01-GM111243, Falk Medical Research Trust, IUCRG, NSF

Key Biological Questions

Development: How does Fertilized Egg Self-Organize into an Organism **without** a road map or plan?



→

http://www.stanford.edu/group/Urchin/LP/
[Lauren Palumbi]

http://www.kvarkadabra.net/images/articles/Regeneracija-organov_1_original.jpg

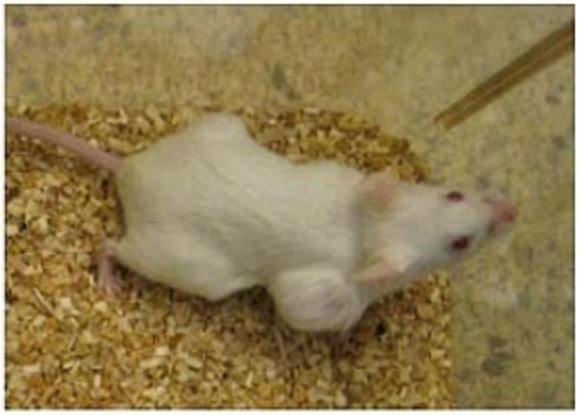
Homeostasis: How does an Organism Maintain itself without an absolute standard of reference?

→

Key Biological Questions

Developmental Diseases: How does Failure of Homeostasis Lead to Redeployment of Developmental Mechanisms in Pathological Ways?

→



e.g., liver cirrhosis, cancer, diabetic retinopathy, polycystic kidney disease, osteoporosis,..

Focus will be on how cell-level behaviors and subcellular control organize into tissue- and organism-level outcomes

Need for Mechanistic Understanding of Developmental Mechanisms

- Improved treatment regimes for cancer (ranging from more accurate surgery to more effective and less toxic therapies).
- Control of stem and other human-derived cells for engineering of tissue replacements both *in vivo* and *in vitro*
- Treatments of developmental diseases
- Prediction of chemical toxicities
- …

Nomenclature

Models—embody the underlying (biological) hypotheses to be tested (elements of a model are **components**)

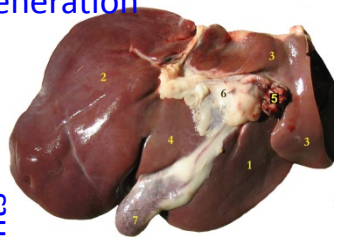
Simulations—an executable program containing the specific **methods** needed to numerically evaluate a model's components in a specific instance

Many layers of abstraction between Model and Simulation

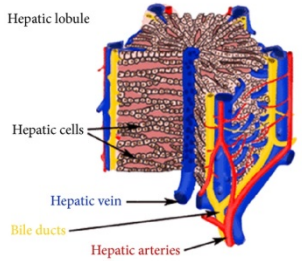
Nomenclature

Workflow—The path from Experiment to Model to Simulation and Back

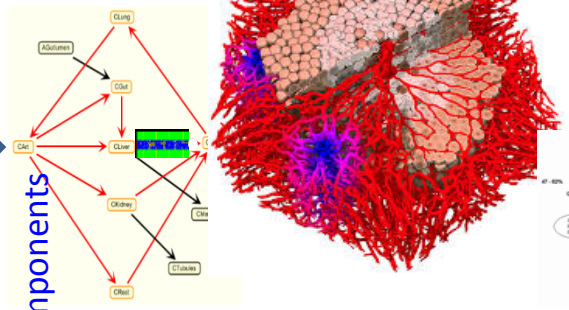
Experiments, Biological Observation Extraction and Hypothesis Generation

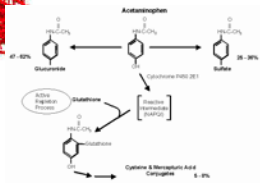


Biological Model Specification



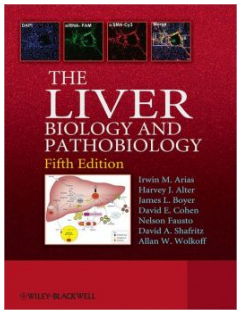
Mathematical Model Specification





Design of New Experiments

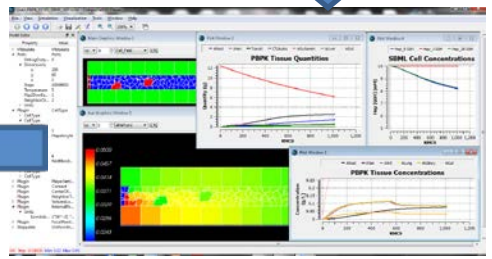
Knowledge Capture and Presentation

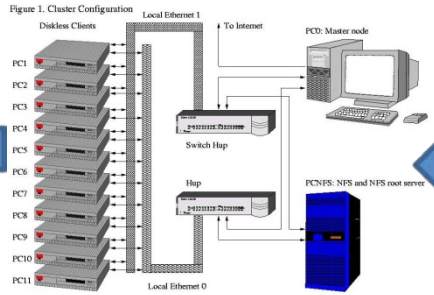


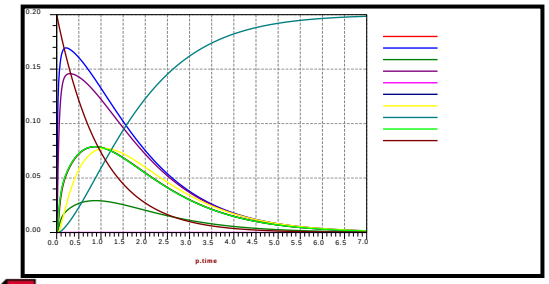
Assembly of Model Components

Integrated Model

Instantiation as simulation







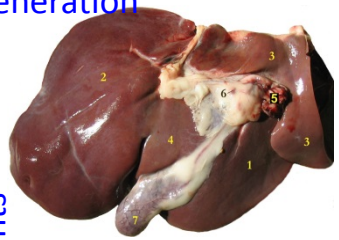
Validation of Simulation Against Experiment

Simulation Processing, Output Organization, Model Archiving,

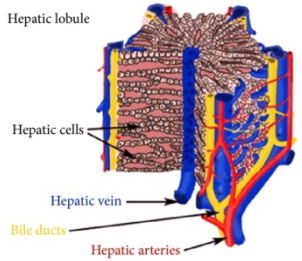
Knowledge Generation Workflow: Not Just a Simulation Engine

Computational infrastructure (models, software and hardware) that allows simulating tissue development, homeostasis and disease, interpreting and comparing simulation results to experimental outcomes and **GENERATING, ENCAPSULATING and CAPTURING KNOWLEDGE**

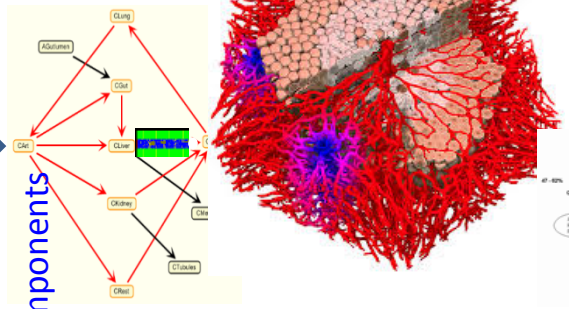
Experiments, Biological Observation Extraction and Hypothesis Generation

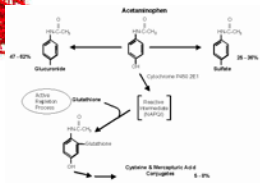


Biological Model Specification



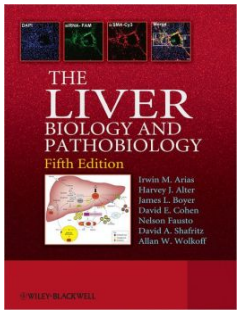
Mathematical Model Specification





Design of New Experiments

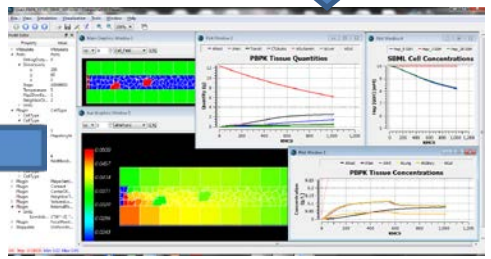
Knowledge Capture and Presentation

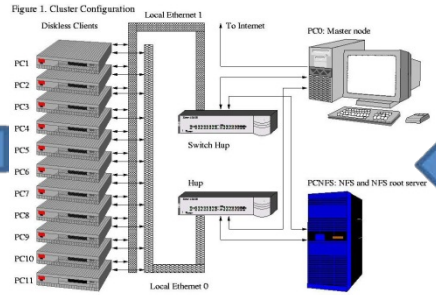


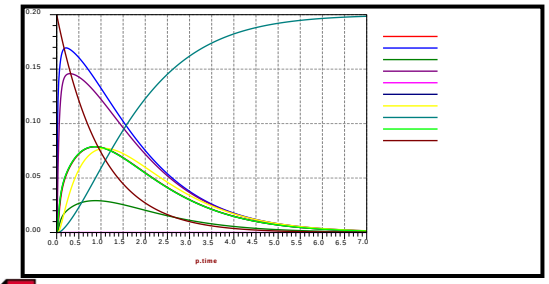
Assembly of Model Components

Integrated Model

Instantiation as simulation







Validation of Simulation Against Experiment

Simulation Processing, Output Organization, Model Archiving,

Why is Bioinformatics Software so Successful?

Tool USERS Don't care about methods, they care about results

Tools provide information that is useful, even essential to modern biology

Tools easily integrate with experimental workflows

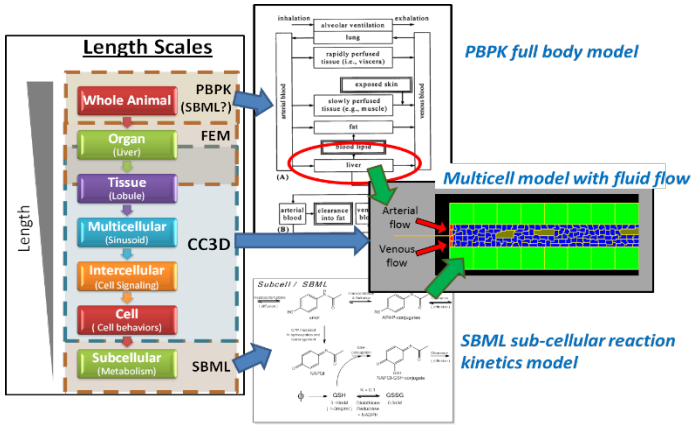
Supports Scientific Workflows—Tools support every step from literature review to hypothesis development, experiment design, experiment execution, analysis, data archiving and knowledge capture and archiving

Standards based—a few basic standards for data, model and software interchange

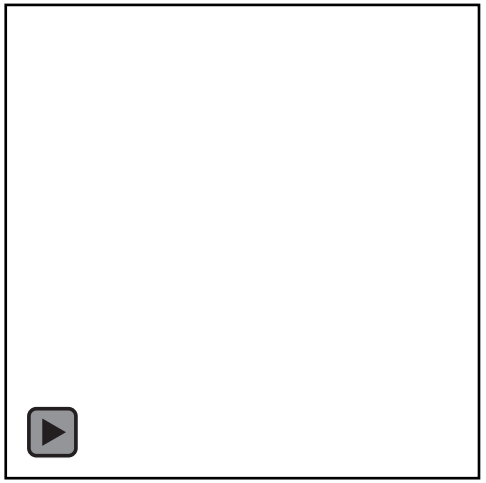
Software and methods are modular—tools interconnect and replacing or updating a tool or method doesn't require changes elsewhere in the ecosystem

Tools use biologically natural representations of data and workflows

Virtual-Tissue Modeling Should be Transformative to Modern Biology



Virtual Tissue Schematic



*"The **user community remains small**, with so far **limited impact** in terms of publications, and those almost **exclusively in the modeling literature**; hardly any published work in the primary biomedical literature has yet arisen."*

Recent NIH report on virtual tissue approaches

To be accepted as a normal part of biological science, we need to move well beyond the development of methodologies

Key Requirements for Productive Science

Verifiability/Falsifiability
Reproducibility
Extensibility

- **Need to make Mathematical and Computational Biology more Scientific**
- **Most modeling results cannot be reproduced with reasonable effort**
- **Most published simulation code is not extensible without undue effort**
- **Most methodologies are not easily combinable**
- **Most models are not designed as reconfigurable extractable and extensible components**
- **Most model components are not designed to allow independent validation**

There are standards of best practice in all these areas, however, we don't always follow them or teach them to our students. This workshop tries to teach some of them

Replicability and Reproducibility

Replicability: The ability to replicate original results using identical tools and methods as the original

Replicability shows that the methods presented produce the results stated (*i.e.* they provide technical validation of the **precision** of results)

Reproducibility: The ability to confirm original results using independent methods (*i.e.* they confirm the **accuracy** of results)

Reproducibility is the cornerstone of scientific method—Ordinarily don't accept a scientific finding until it has been reproduced

Replication primarily addresses simulations, reproducibility addresses models

Significance of Reproducibilty

Reproducibility:

- Provides confidence in validity of results: reproduction with independent methods helps validate underlying hypotheses and confirm that results are free of technical errors and method-dependent artifacts
- Allows reuse, adaptation and extension of tools and scientific knowledge
- Allows science to accumulate reliable knowledge cumulatively

Lack of Reproducibility:

- **Insecurity**
  - Can't validate simulations (is the simulation performing as reported) or models (are the hypotheses the simulation is supposed to encode correct)
  - Prevents model use in critical applications (clinical, regulatory)
- **Inefficiency**
  - The knowledge (model) encoded in simulations is lost and must be reinvented by each researcher for each problem
  - Results in wasteful duplication of effort
  - Slows scientific progress

Key Requirements for Models to be Useful (I)

Models must be Reproducible (not just replicable)

Most models are hard coded in computational form and as a result are not reproducible

Solution—Use standard tools and build models using appropriate work flows

Key Requirements for Models to be Useful (II)

Models must be intelligible

Can't in general map computer code and mathematics back to the biology they represent

- Design models as much as possible in terms of natural biological concepts
- Be consistent in your terminology, units,…
- Document your decisions at all levels.
- Keep specifications as high-level (biological) as possible.
- Maintain connection between more biological and more computational levels of description

Key Requirements for Models to be Useful (III)

Models must be extensible and their components extractable and reusable

Monolithic code mixing model definition and implementation makes model extension and submodel extraction nearly impossible

- Separate model specification from implementation specification.
- Design models so that components are modular and capable of running independently
- Maintaining connection between more biological and more computational levels of description helps in model extension and reuse

Key Requirements for Models to be Useful (IIIb)

Can't predict how people will combine model components, so methods must be flexible and reconfigurable

E.g. if you needed to use a different word processor to write a letter and an article (pace LaTeX fans), you would use a typewriter

Most simulations mix methods together creating hard-to-parse monolithic code, without considering the consequences

Costs:

- **Can't know if the method does what it says**
- **Can't replace one method with another that has a similar function**
- **Adapting a method for another use requires disentangling it from its context, rewrapping it and possibly recoding it outright**
- **Can't reuse or extend methods**
- **Can't build libraries of methods**

As if we had to spend three months on C++ recoding each time we wanted to use a new app on our cell phone!

Key Requirements for Models to be Useful (IV)

Building models must use abstractions and workflows that are natural to those who will use them. Understanding biology will always be hard, implementing it as a model or simulation should be easy

Most modeling approaches are framed in terms of mathematical and computational rather than biological abstractions—odd since the results of a given biological model should be independent of the solution method

***E.g.* if you had to enter your English-language text in Japanese characters to write a letter on a computer, you would use a typewriter**

- **Mathematical and Computational Biology should be useful to the people who matter—Biologists!**

- **Specify models in terms of natural biological concepts**

Biologically Natural and Modular Model Specification

Biology full of modular concepts

- **Objects: Individuals, Organs, Tissues, Cells, Molecules…**
- **Behaviors: Division, Growth, Circulation,…**
- **Interactions: Chemotaxis…**
- **Initial Conditions: Epithelium, Tubule,…**

Without Modular Model specification:

- **Can't identify model hypotheses clearly**
- **Can't reuse Objects, Behaviors,… in a different context**
- **Can't build libraries of Objects, Behaviors**
- **Can't connect Objects, Behaviors, Interactions, Initial Conditions**

As if we organized all DNA sequence data as 1000 bp blocks with no concept of genes, pathways,…

Key Requirements for Models to be Useful (IVb)

Models must separate biological structure from implementation

Most modeling methods mix description of the model biology with the solution methodology (limiting portability and cross-method validation)

Problems
Need to reverse engineer the hypotheses from the simulation
Can't easily substitute one simulation method with an alternative method to reproduce a model's results
Can't extract model components for reuse to build other models
Can't extract methods from simulations for reuse to apply to other models

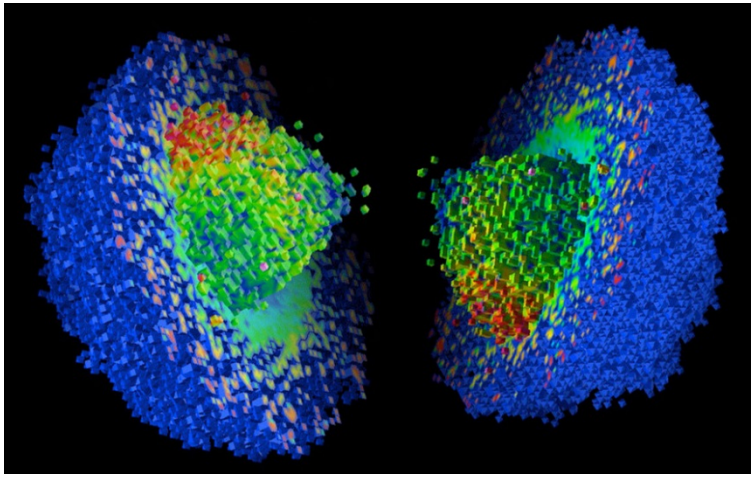
As if each DNA sequencing method only worked on one specific gene in one specific organism, and we had to redevelop each method each time we wanted to work on a new gene or organism or you had to rewrite MSWord each time you wanted to write a new letter

E.g., Need to Separate Biology from Implementation

Lots of Ways to Represent Cells

Cell Representations: Cellular Automata

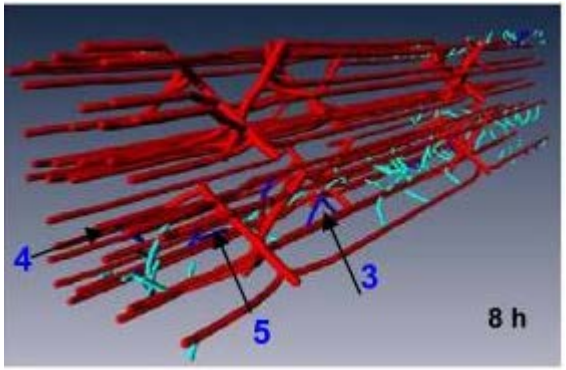
- **Lattice-based**, **probabilistic model**
- Cells **represented as lattice points**
- Fast, allow simulations of fairly **large tissue volumes**
- **Rule-based**
- **Cells have states** and **can move** to different lattice sites. Transition probabilities define cell properties/behaviors
- Detailed representation of single cell behaviors or interactions difficult
- It may be hard to match experimentally measured quantities to
- model parameters
- Implementation is simple. Most researchers write their own code but also many open source packages available e.g. Cellumat3D, CelLab etc…



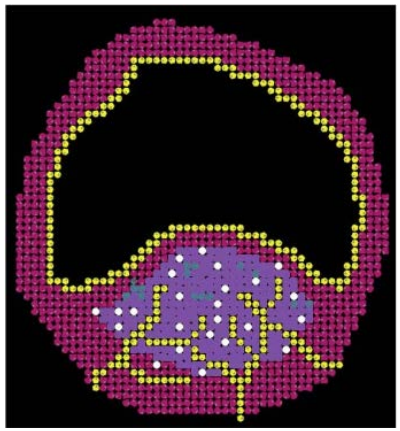
3D model of *in-silico* tumor, Alexander Anderson

Cell Representations: Rule-based Models

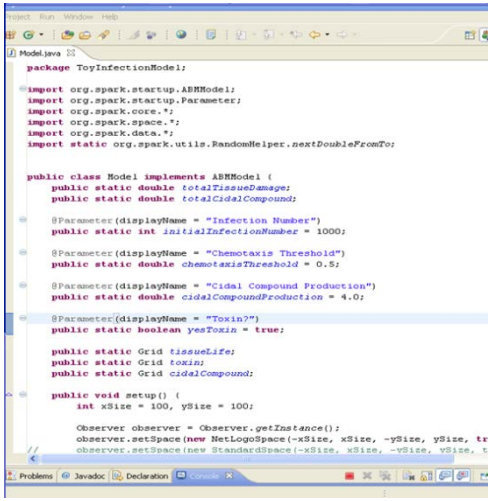
- Agents are **independent entities with pre-programmed set of rules** which interact leading to emergent phenomena
- Agents can **represent cells**, **parts of cell**, **ECM** components etc…
- Model description often consists of **many heuristic rules**.
- Models usually **represented as programing language classes/functions**
- It may be hard to translate lab measured quantities into Agent Based Model parameters/rules.
- Like neural networks require training/tuning
- When tuned/trained properly, have predictive power
- Many software packages: SWARM, NetLogo, Mason, RePast, AnyLogic, SPARK, FLAME, GAMA, plus research code developed by various labs.



Agent-Based Model angiogenesis, Gang Liu, Amina A Qutub , Alexander Popel



Agent-Based Model of Atherosclerotic Plaque Progression, Shayne Peirce-Cottler *et al*



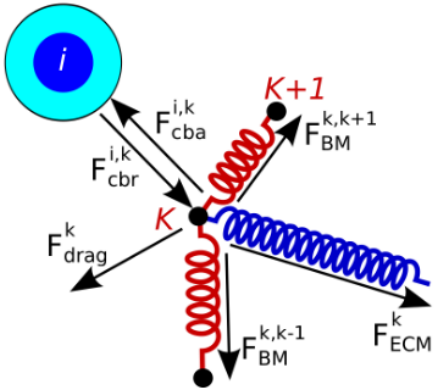
ABM code for SPARK simulation, Alexey Solovyev, Gary An

Cell Representations: Physics-based models

- Represent dynamics of cells and extracellular matrix by assuming

$$\vec{F} \sim \vec{v}$$

- Use auxiliary equations such as reaction-diffusion, Navier Stokes, elasticity relations etc…

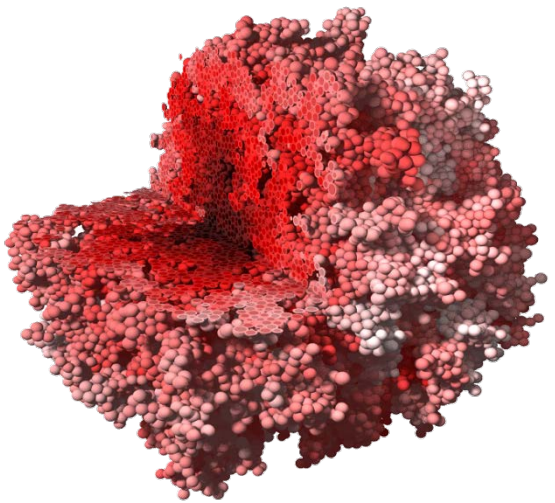


$$\mathbf{v}_{\mathrm{cell},i} = \frac{1}{\nu}\left[\sum_{k=1}^{N_p}\left(\mathbf{F}^{ik}_{\mathrm{cba}} + \mathbf{F}^{ik}_{\mathrm{cbr}}\right) + \sum_{\substack{j=1 \\ j\neq i}}^{N_c}\left(\mathbf{F}^{ij}_{\mathrm{cca}} + \mathbf{F}^{ij}_{\mathrm{ccr}}\right)\right]$$

Paul Macklin: An agent-based model for elasto-plastic mechanical interactions between cells, basement membrane and extracellular matrix

Center Models

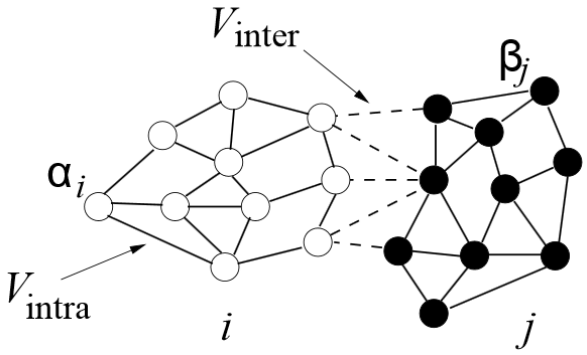
- Represent **cells as spheres or ellipses**
- **Force-based formalism** to describe cell dynamics
- Inspired by **molecular dynamics** but computationally simpler because **first order ODEs**
- **Drawback**: limited cell shape representation.
- **Advantage**: fast and allow easy mapping of lab-measured quantities to model parameters.
- Representation of time is **explicit** (i.e. does not have to be deduced)
- **Software**: CellSys – free binaries closed-source, CHASTE – open source



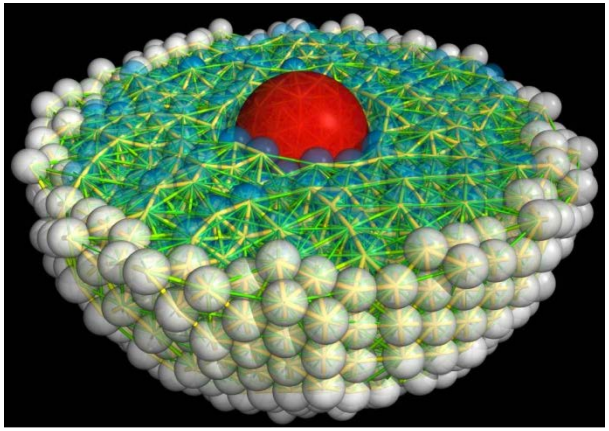
Stefan Hohme, Dirk Drasdo: Growing multi-cellular population

Cell Representations: Subcellular Element Model

- Introduced by Tim Newman (U. Dundee, Scotland).
- Represents single **cell as a collection of interacting subcellular elements – points/spheres**. Interaction forces between elements of the same cell are much stronger than interactions between elements of difference cells
- **Advantages:** force-based formalism, allows for fairly **detailed cell shape representation**. Representation of time is **explicit**
- **Drawbacks:** computational cost is quite high, number of parameters can be large
- **Software:** no established framework – research code of Tim Newman and others

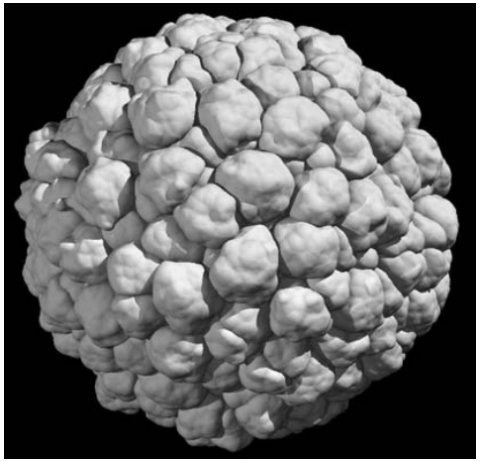


$$\dot{\mathbf{y}}_{\alpha_i} = \eta_{\alpha_i} - \nabla_{\alpha_i} \sum_{\beta_i \neq \alpha_i} V_{\text{intra}}(|\mathbf{y}_{\alpha_i} - \mathbf{y}_{\beta_i}|) - \nabla_{\alpha_i} \sum_{j \neq i} \sum_{\beta_j} V_{\text{inter}}(|\mathbf{y}_{\alpha_i} - \mathbf{y}_{\beta_j}|)$$



10 um

Detailed SEM model of a single cell,
Sebastian Sandersius, Tim Newman

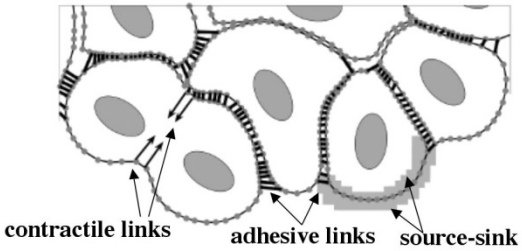


100 um

Modeling cellular aggregate
using SEM, Tim Newman

Cell Representations: Immersed Boundary Method

- Computational Fluid Dynamics approach that models **single cell as a fluid enclosed inside a elastic membrane**. Fluid motion affects membrane and membrane affects fluid motion.
- Deals gracefully with shape changes, can simulate subcellular organelles,
- Excellent spatial resolution for single cell modeling.
- **Drawbacks**: mathematical complexity.
- Computationally expensive (simulating e.g. 1000 cells can be challenging)
- Usually 2D, 3D hard
- **Software**: IBCell (Kasia Rejniak), research codes of Michael King, David Gee



$$\rho\left(\frac{\partial \mathbf{u}(\mathbf{x},t)}{\partial t} + (\mathbf{u}(\mathbf{x},t)\cdot\nabla)\mathbf{u}(\mathbf{x},t)\right) = -\nabla p(\mathbf{x},t) + \mu\Delta\mathbf{u}(\mathbf{x},t) + \frac{\mu}{3\rho}\nabla s(\mathbf{x},t) + \mathbf{f}(\mathbf{x},t), \quad (1)$$

$$\rho\nabla\cdot\mathbf{u} = s(\mathbf{x},t), \quad (2)$$

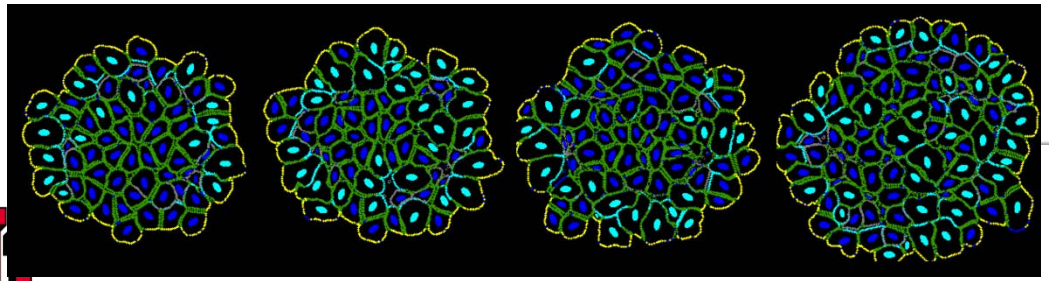
$$\mathbf{f}(\mathbf{x},t) = \sum_i \mathbf{f}_i(\mathbf{x},t) \quad \text{where} \quad \mathbf{f}_i(\mathbf{x},t) = \int_{\Gamma_i} \mathbf{F}_i(l,t)\,\delta(\mathbf{x}-\mathbf{X}_i(l,t))\,dl,$$
$$\text{and} \quad \mathbf{F}_i(l,t) = \sum_\alpha \mathbf{F}_{\alpha(i)}(l,t), \quad (3)$$
$$\text{and} \quad \mathbf{F}_{\alpha(i)}(l,t) = \mathcal{F}_\alpha \frac{\|\mathbf{X}_{i,k}(t)-\mathbf{X}_{i,l}(t)\|-\mathcal{L}_\alpha}{\|\mathbf{X}_{i,k}(t)-\mathbf{X}_{i,l}(t)\|}(\mathbf{X}_{i,k}(t)-\mathbf{X}_{i,l}(t)),$$

$$s(\mathbf{x},t) = \sum_i s_i(\mathbf{x},t) \text{ where } s_i(\mathbf{x},t) = \sum_{k\in\Xi_i^+} S^+(\mathbf{Y}_{i,k}^+,t)\,\delta(\mathbf{x}-\mathbf{Y}_{i,k}^+) + \sum_{k\in\Xi_i^-} S^-(\mathbf{Y}_{i,k}^-,t)\,\delta(\mathbf{x}-\mathbf{Y}_{i,k}^-),$$
$$\text{and} \sum_{k\in\Xi_i^+} S_+(\mathbf{Y}_{i,k}^+,t) + \sum_{m\in\Xi_i^-} S_-(\mathbf{Y}_{i,m}^-,t) = 0. \quad (4)$$
$$\text{and} \quad S_+(\mathbf{Y}_{i,k}^+,t) = \begin{cases} \mathcal{S}_+^0, & \text{if } \sum_i \gamma(\mathbf{X}_i(l,t),t) > \gamma_{thr} \\ 0, & \text{otherwise.} \end{cases}$$

$$\gamma(\mathbf{X}_i(l,t),t) = \int_\Omega \gamma(\mathbf{x},t)\,\delta(\mathbf{x}-\mathbf{X}_i(l,t),t)\,d\mathbf{x}. \quad (5)$$

$$\frac{\partial \mathbf{X}_i(l,t)}{\partial t} = \mathbf{u}(\mathbf{X}_i(l,t),t) = \int_\Omega \mathbf{u}(\mathbf{x},t)\,\delta(\mathbf{x}-\mathbf{X}_i(l,t))\,d\mathbf{x}. \quad (6)$$

$$\frac{\partial \gamma(\mathbf{x},t)}{\partial t} = \mathcal{D}_\gamma\,\Delta\gamma(\mathbf{x},t) - \mathcal{V}_\gamma \frac{\gamma(\mathbf{x},t)}{\kappa_\gamma+\gamma(\mathbf{x},t)}\cdot\chi(\Theta_\Gamma), \quad (7)$$
$$\gamma(\mathbf{x},t_0) = \gamma_0 \quad \text{for} \quad \mathbf{x}\in\Omega,$$
$$\gamma(\mathbf{x},t) = \gamma_0 \quad \text{for} \quad \mathbf{x}\in\partial\Omega \quad \text{and} \quad t\geq t_0.$$

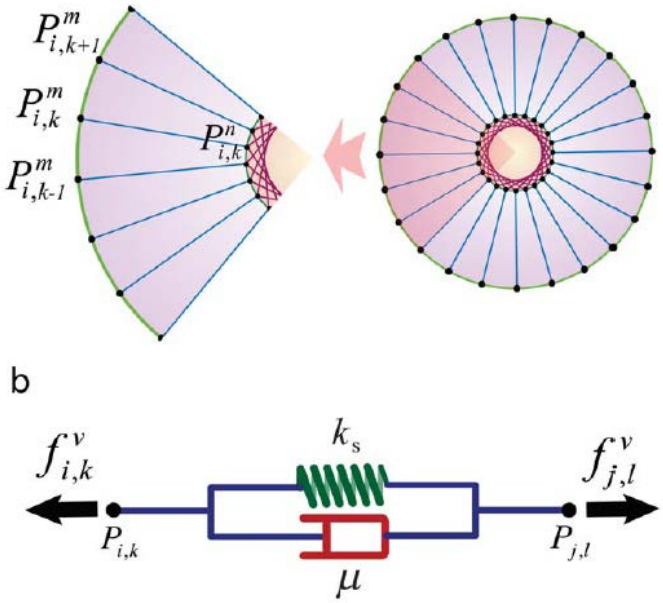


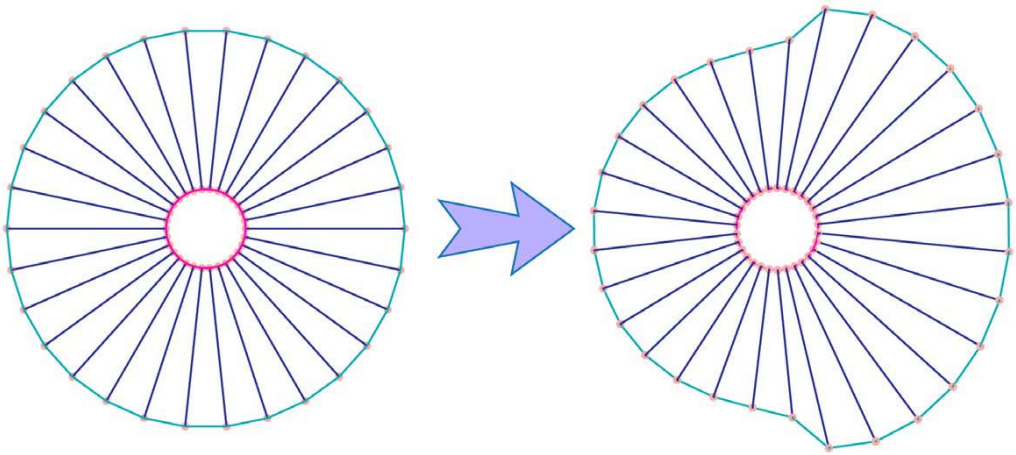
Images courtesy of Kasia Rejniak, Moffitt Center, Tampa

Cell Representations: Sub-Cellular Viscoelastic Model

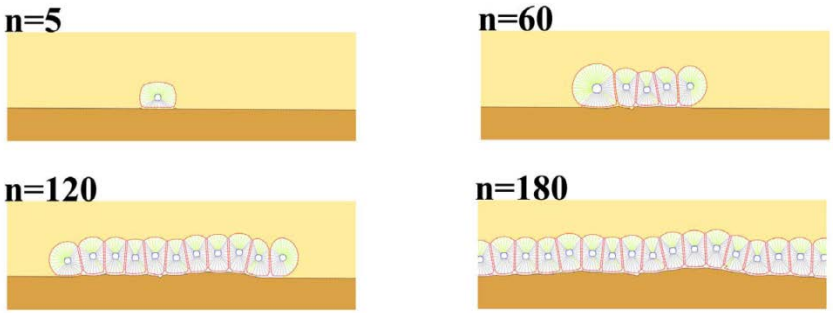
connected by visco-elastic links

- Sits somewhere between Subcellular Element Model and Immersed Boundary Method
- Models in detail intracellular components
- Relatively new
- **Software**: research code
- Reference: Jamali Y, Azimi M, Mofrad MRK (2010) A Sub-Cellular Viscoelastic Model for Cell Population Mechanics. PLoS ONE 5(8): e12097. doi:10.1371/journal.pone.0012097





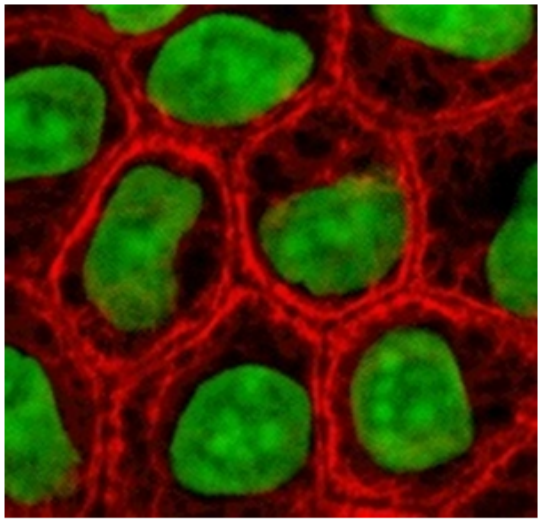
Cell Shape change during movement

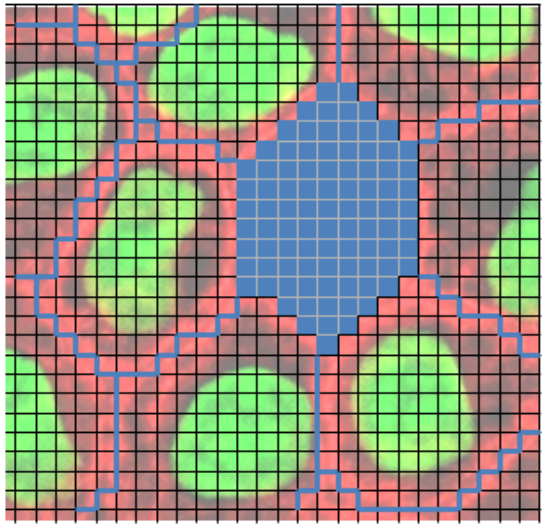


Formation of epithelial layer

Images Courtesy Yousef Jamali*, Mohammad Azimi, Mohammad R. K. Mofrad, UC Berkeley

Cell Representations: CPM/GGH





$$E = \sum_{\substack{\vec{x},\vec{x}' \\ \text{neighbors}}} J\left(\tau(\sigma(\vec{x})),\tau(\sigma(\vec{x}'))\right)\left(1-\delta\left(\sigma(\vec{x}),\sigma(\vec{x}')\right)\right)$$

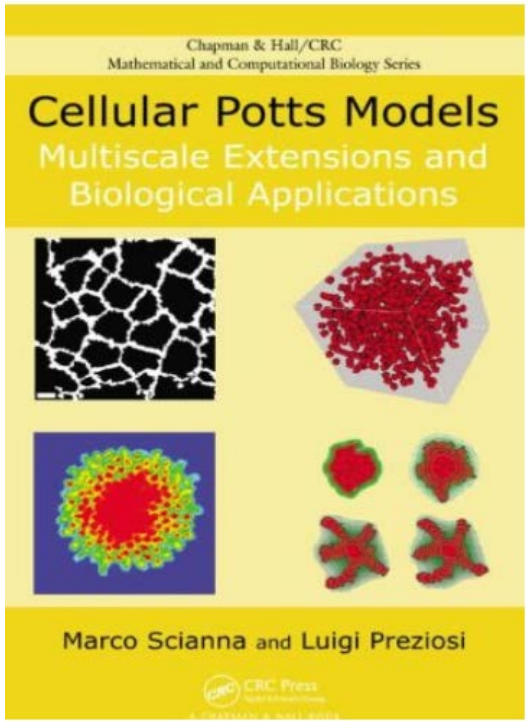
$$+\sum_{\sigma}\lambda_s(\sigma)\left(s(\sigma)-S_{\text{Target}}(\sigma)\right)^2$$

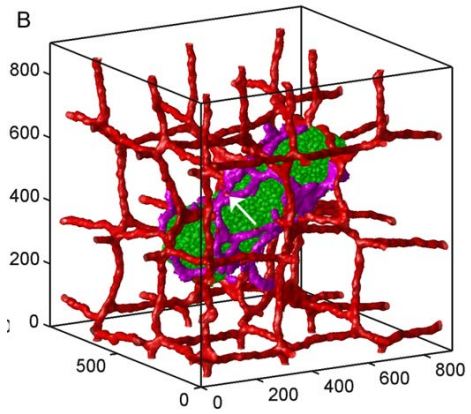
$$+\sum_{\sigma}\lambda_v(\sigma)\left(v(\sigma)-V_{\text{Target}}(\sigma)\right)^2$$

$$+E_{chem}+E_{hapt}+\ldots$$

$$P(\Delta E) = 1,\ \Delta E \le 0$$

$$P(\Delta E) = e^{-\Delta E/kT},\ \Delta E > 0$$





A poor man's method, but can add any conservations and physics by adding constraints

Should All Give the Same Answer for the Same Biological Model

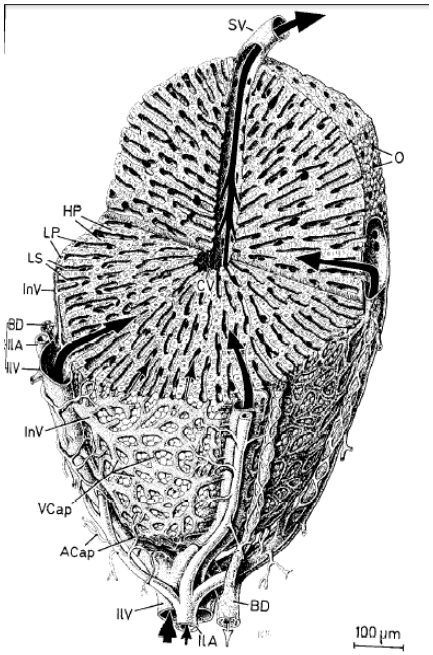
Difference between results using different solvers are artifacts of the methodologies, and don't represent the biology

Often, higher accuracy methods aren't the issue: the biological approximations and uncertainties dominate the numerical ones. Biological noise and robustness help

Why is Scientific, Mathematical and Computational Biology Difficult?

Biology Occurs Heterogeneously Across Spatial and Temporal Scales--Biology is Hard!

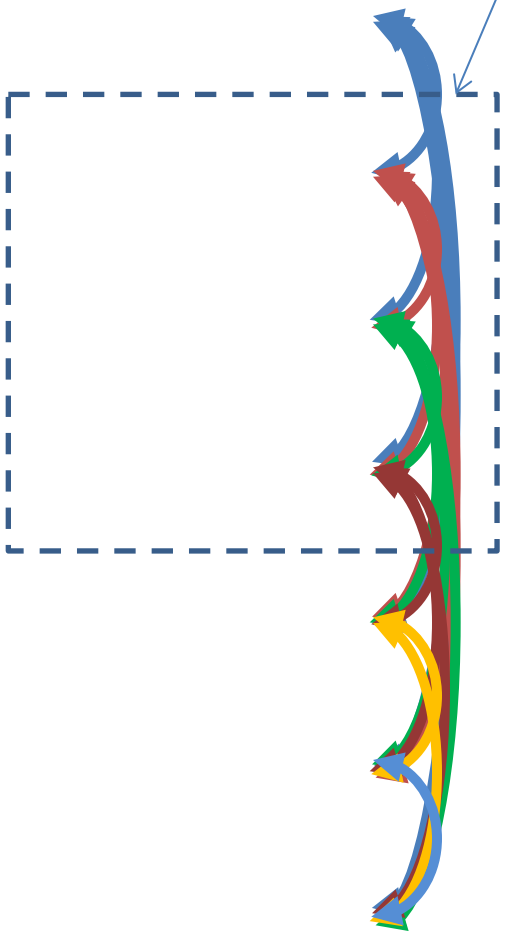
- Distance scales range from sub-nanometer to meters.
- Time scales range from seconds to decades.



1. Human Microscopic Anatomy: An Atlas for Students of Medicine and Biology, R. V. Krstic, Springer-Verlag, 1991 (ISBN 978-3-540-53666-6).

Length

Virtual-Tissues



Coupled ODE, Fluid Dynamics,…

Continuum Mechanics, PDE,…

Continuum Mechanics, PDE,…

Agent-Based, CPM/GGH, Center, Vertex,…

PDE, Coupled ODE,…

Continuum Mechanics, PDE, Coarse-Grained Molecular Dynamics,…

Coupled ODE, Stochastic, MD,…

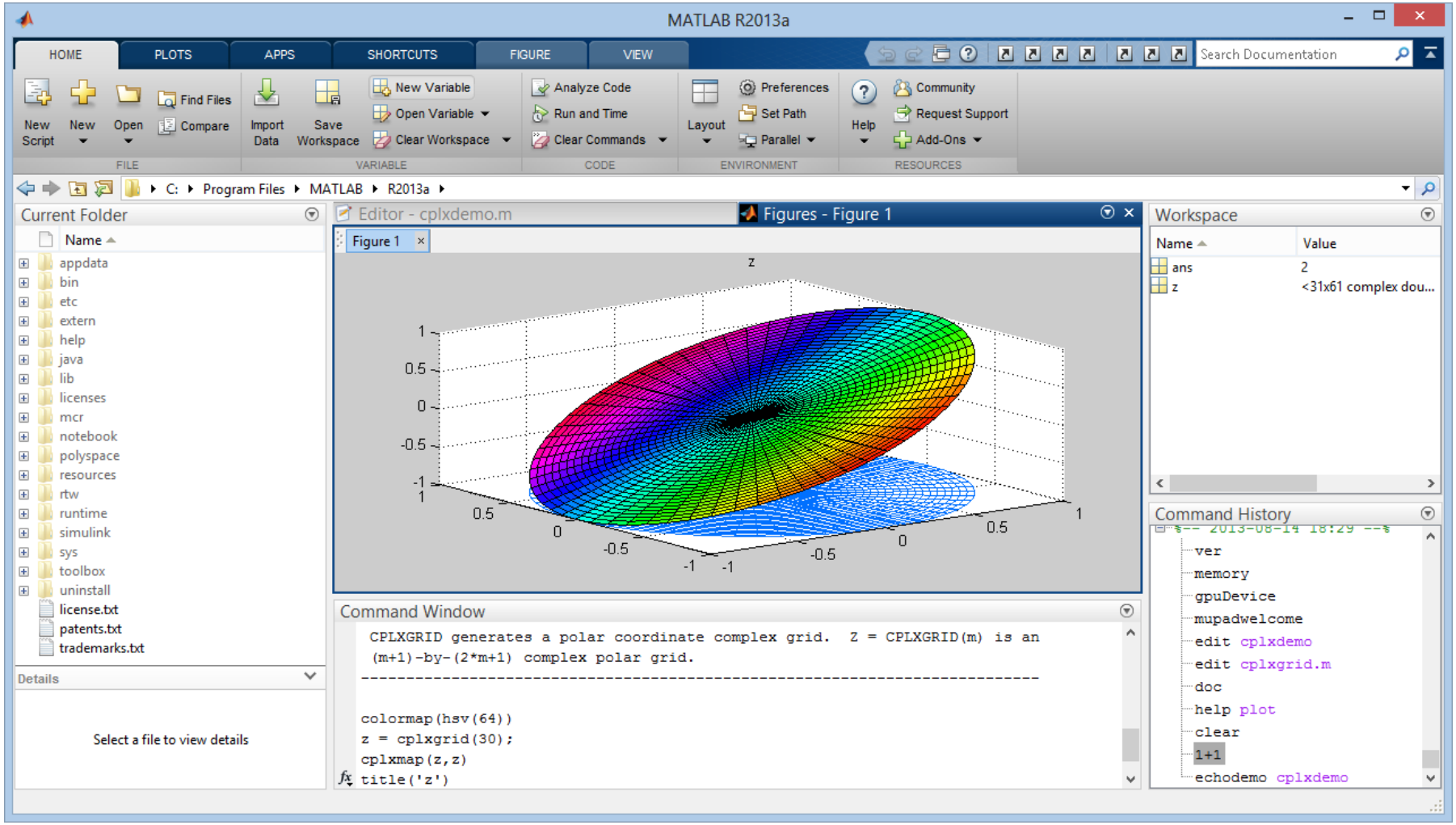
Challenges (1)

- **Biology is messy**
  - Terms not consistent (*e.g.*, species, cell division, cell growth,…)
  - Boundaries not sharp (*e.g.*, gap junctions between cells)
  - Structures can overlap in multiple ways, not always spatially or temporally hierarchical
  - Concepts fuzzy (*e.g.* two cells of the same type may have different behaviors)
  - **DON'T HAVE UNAMBIGUOUS DEFINITIONS OF BIOLOGICAL CONCEPTS**
- **Biology is diverse**
  - Many components
- **Biology is Heterogeneous**
  - Can't always scale simply
  - Have to couple 1-D, 2D and 3D objects, with complex long-range and inhomogeneous mesh structures

Challenges (II)

- **Some of the Science is still not well developed**
  - Mesh dynamics is poorly understood
  - Complex Viscoelasticity
  - Complex and poorly understood transport and advection
- **Cross-Scale Connections are Complicated**
  - Hard to define a limited set of parameters describing cell, tissue and organ behaviors and infer them from elements of other scales
  - *E.g.* very hard to say how a level of a given chemical affects a specific cell behavior (*e.g.* motility) or how a gene mutation affects a specific function of an organ
- **Natural biological model components may not correspond to natural simulation methods**
  - *E.g*. a cell may secrete a chemical which then diffuses, but we may want to implement secretion and absorption in a methods that solves diffusion of chemicals rather than in a method that specifies cell movement
- Developing Workflow tools, standards and user interfaces is **unglamorous, even boring**.
- Model Sharing is virtually non-existent so **bootstrapping costs are large**
- **Software sustainability** is tough

Importance of Integration Platforms: A Commercial Success Story—Matlab



Importance of Integration Platforms: A Commercial Success Story

Is Matlab really “one-size-fits-all” solution? In many cases it certainly appears to be the case.

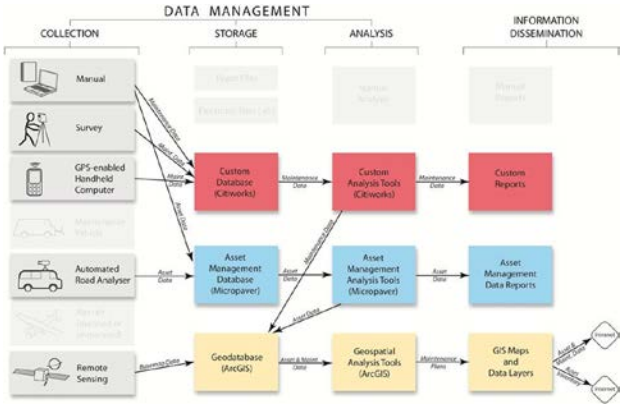
Features which make Matlab sell:

1) Easy to use scripting language – makes Matlab extensible, and the syntax does not change every 2 years
2) Plethora of computational tools under one roof
3) User interface – visualization, code editor, well defined data formats etc…
4) Matlab models are shareable and exchangeable (within Matlab community)
5) Can be used in classroom setting to teach students modeling
6) Documentation and books teaching how to effectively use Matlab
7) Outreach efforts – training
8) Professionally maintained

Potential of Reproducible Methods

E.g. Your cell phone allow you to click on an address and pull up a navigator that finds where you are then takes you to your destination (on foot, by train or in a car), including the time and traffic alternatives to your destination

Depends on a host of independently developed and interchangeable computational components working together using heterogeneous data at multiple length and time scales

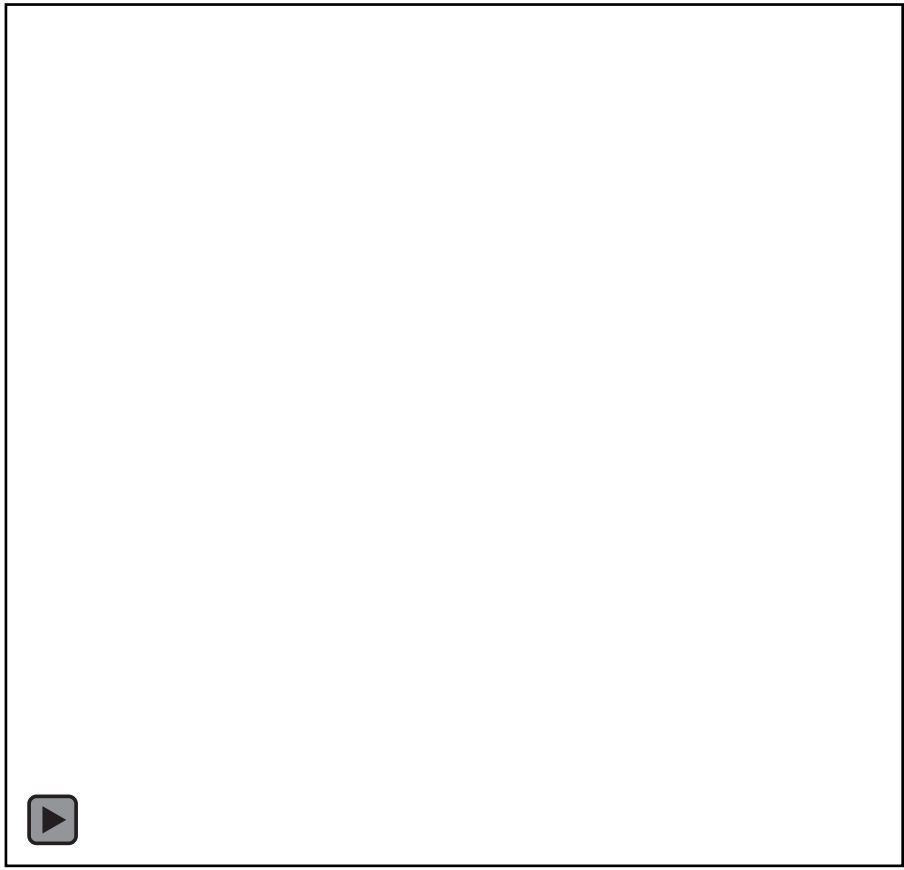


http://www.gis.fhwa.dot.gov/documents/images/AssetMgmt_fig14.gif Models

E.g. Fuse Experiment and Simulation

A Virtual Tissue Environment:

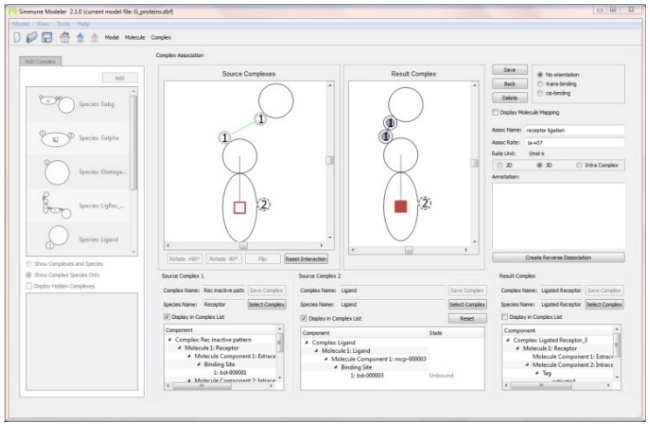
- Reads an Annotated Image to Identify the Locations and Identity of Components
- Builds the Model by Populating the Representation of the Image with Components from a Cell Type Repository and Other Repositories
- Translates the model into a Simulation using Standardized Specifications of Components, Properties, Behaviors and Interactions Outputs the Simulation Results as Annotated Simulation Images for Analysis and Comparison with Experiment
- Functions as a Variable Power Microscope, Handling Refinement/Coarse Graining Automatically
- Simulates all Cells in Embryo, Tissue,…

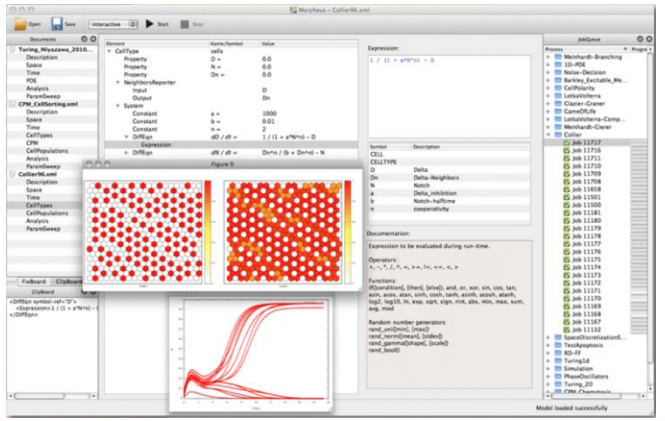


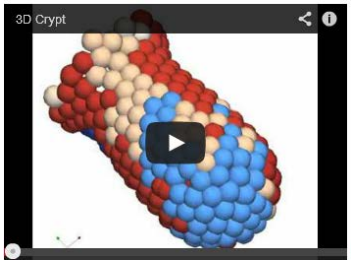
Reconstructed zebrafish embryonic development from P. J. Keller, *et al.,* "Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy," *Science* **322**, 1065 (2008).

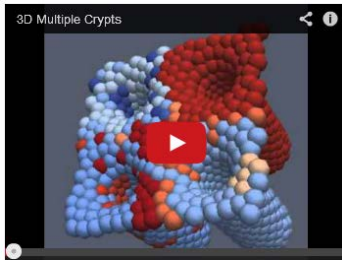
Virtual-Tissue Environments

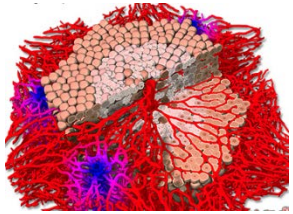
- Many tools for generic agent based modeling but few for physics-based simulations
- CompuCell3D (www.compucell3d.org) – implements CPM/GGH. Powerful and easy to use. Cross-platform and open source. Uses Python and XML model specification. Many users
- Morpheus (http://imc.zih.tu-dresden.de/wiki/morpheus/doku.php) – CPM/GGH. XML model specification. Free but closed-source
- Simmune (http://www.niaid.nih.gov/labsandresources/labs/aboutlabs/lsb/Pages/simmuneproject.aspx) – CPM/GGH. Free but closed-source
- CHASTE (http://www.cs.ox.ac.uk/chaste/) – multi-method (CPM/GGH, Center Model, Vertex Model, FEM continuum). Open source. C++ library. Linux only. Hard to learn. C++ model specification
- CellSys (Dirk Drasdo and Stefan Hohme)– Center model. Free but closed-source. GUI-based model specification is limiting
- LBIBCell (Dagmar Iber)…

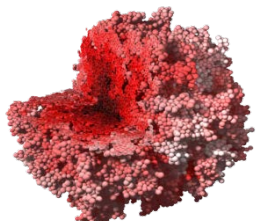






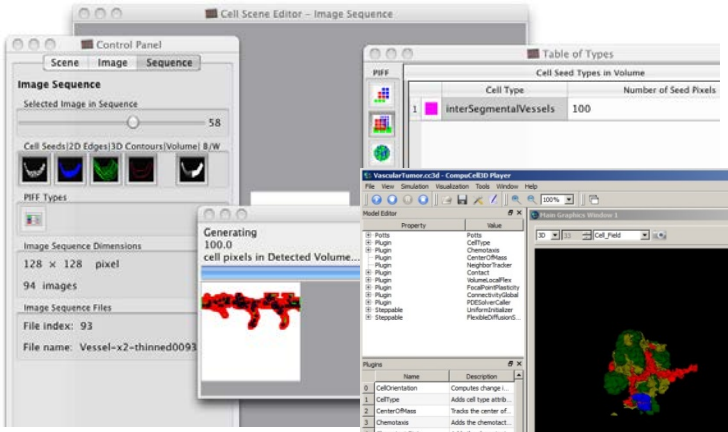


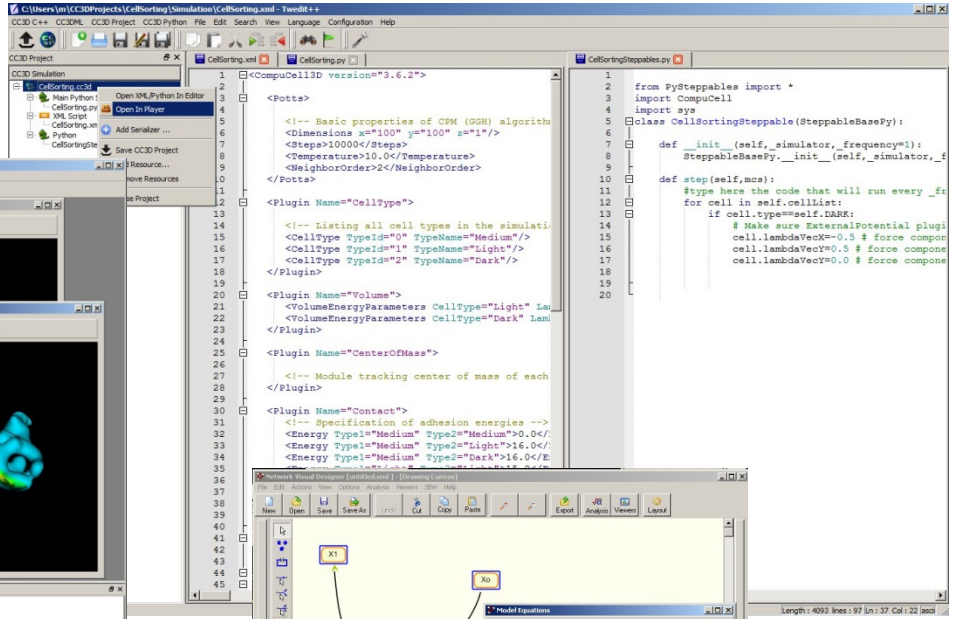


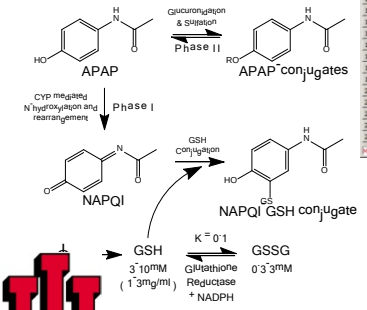


CompuCell3D

- **Cross-platform, open-source environment for building and running Virtual-Tissue simulations**
- Written in **C++** and **Python**
- Describes models using combination of **XML** (CC3DML) and/or **Python**
- **Parallel implementation** on GPU's and multi-core machines
- Simulation-wizards, simulation editors, graphical interfaces, visualization tools, SBML Dynamic Network Solvers, and FE solvers (experimental support)

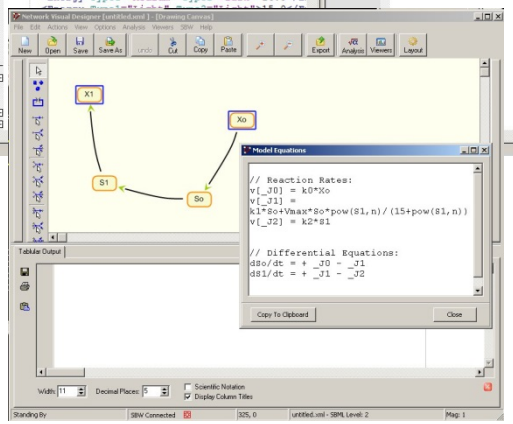






SBML models can run inside each cell

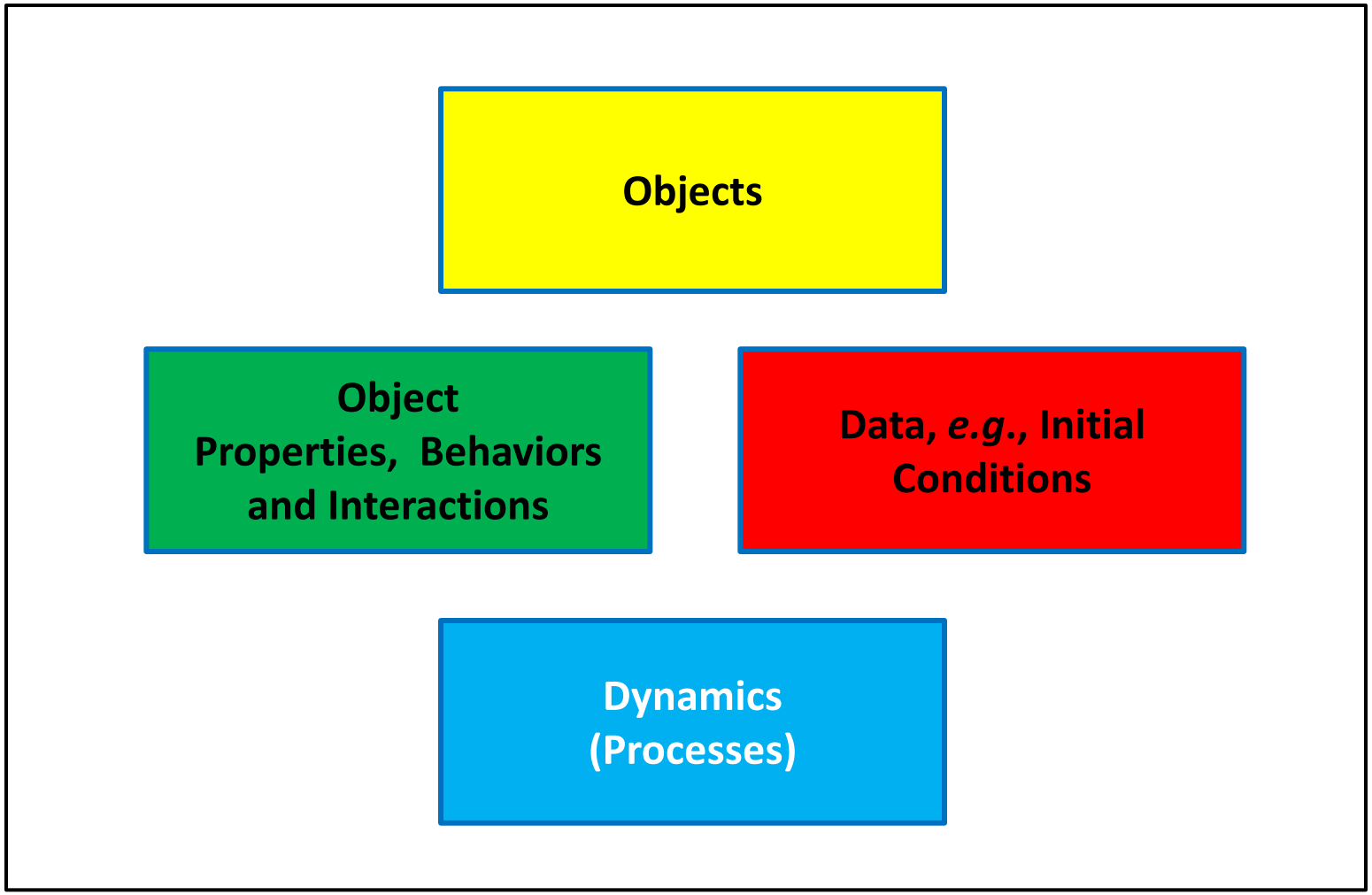
Model Generation, *e.g.* using H. Sauro's JDesigner



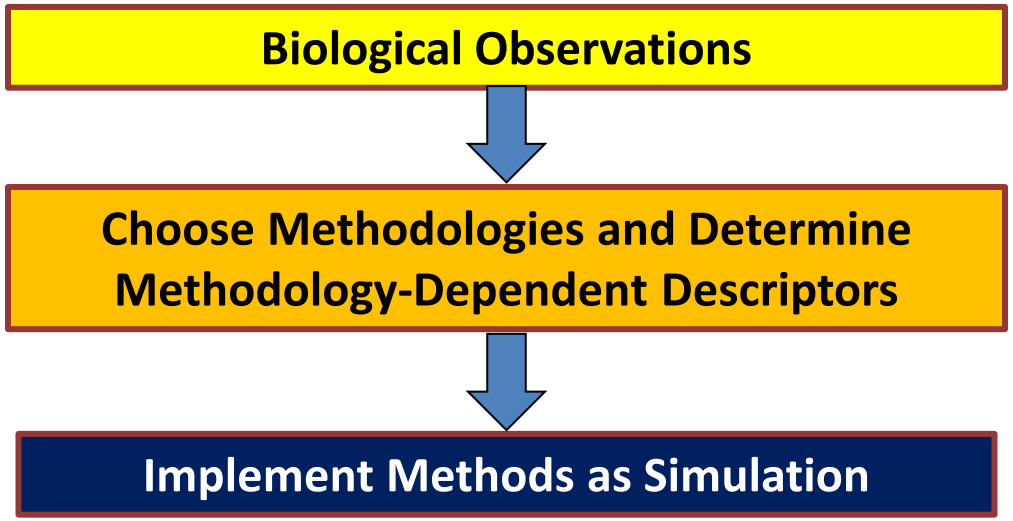
How to Start Building a VT Model?

- What are the specific questions you are trying to answer (hypotheses you are trying to test)?
- **Models can only show sufficiency, not necessity!**
- What are the scales you need to include in your model to test these hypotheses?
- What do you see as the most important components (Objects) and Mechanisms in the biology you are trying to understand?
- **A model cannot predict the importance or role of a Mechanism not included in the Model!**
- Do you have enough information to describe the Objects and Mechanisms at the scale you have chosen?
- **You cannot include a Mechanism unless you can describe it quantitatively!**
- How will you compare simulation results with experiment to test your model?
- VT models are useful for understanding emergent properties of tissues, when the behaviors of their components are relatively well understood.

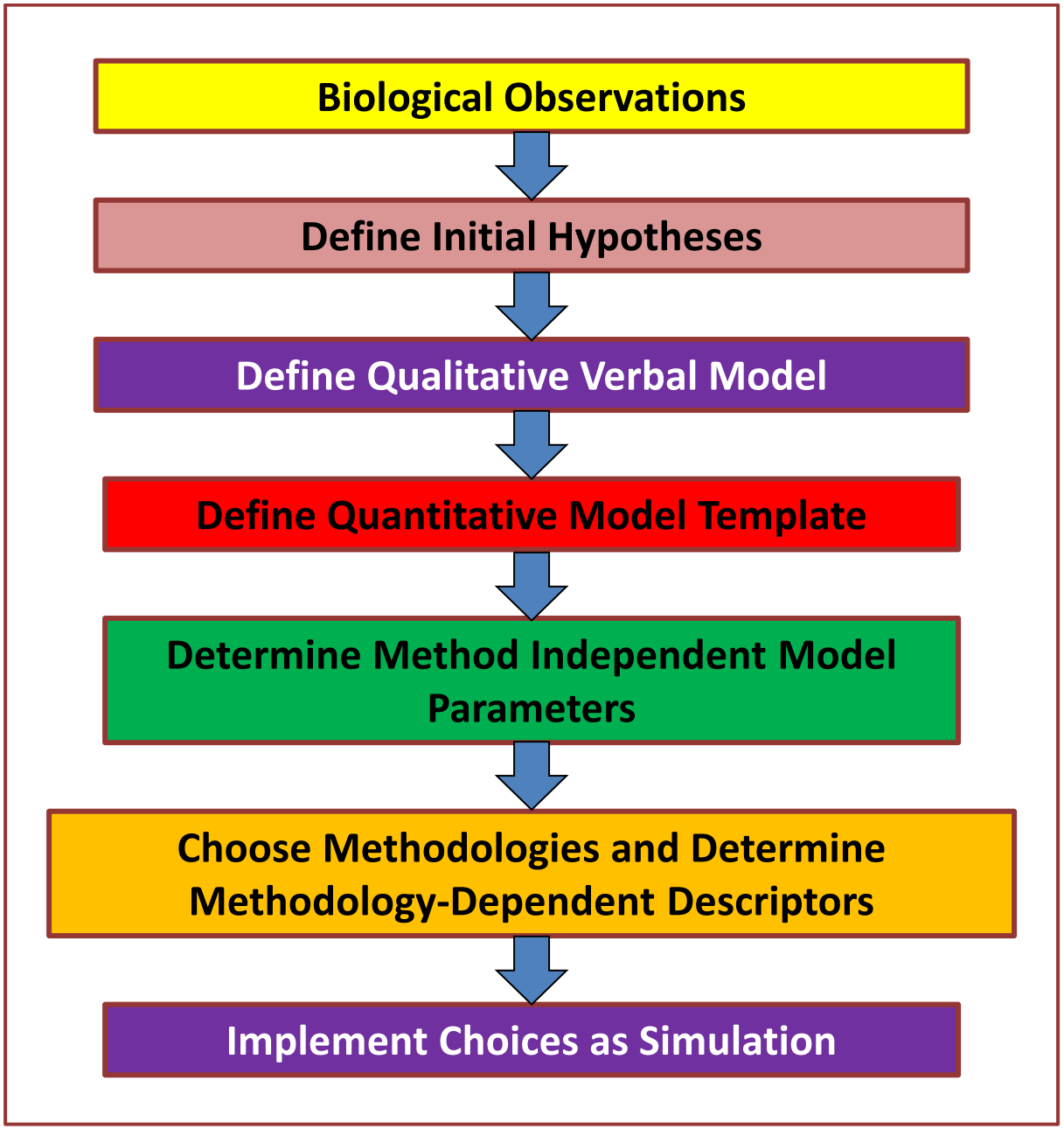
Model Components and Methods



Typical Model / Simulation Workflow

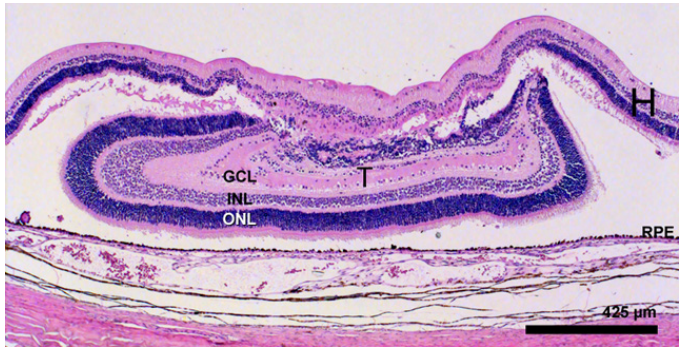


Reproducible Model Development Workflow

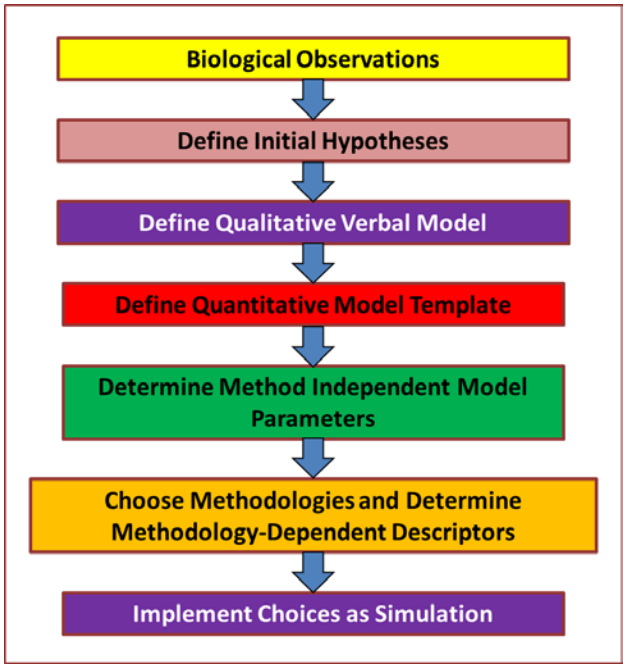


Cell Sorting During Development—Biological Observations

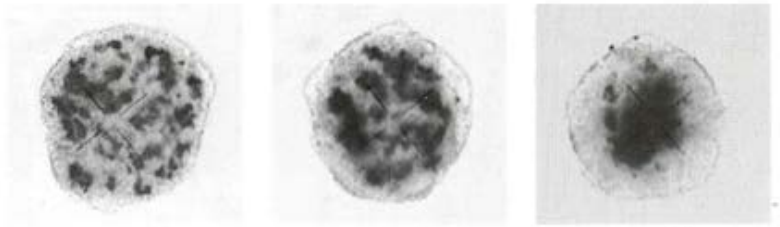
How do different cell types form ordered layered structures in developing embryos? Cells reorganize during early development, then maintain spatial relationships



http://www.djo.harvard.edu/files/3068_362.jpg



Study *in vitro* by making random aggregates of two embryonic cell types and see how they behave.



(Diameter of sorted aggregate is 260 μm.)

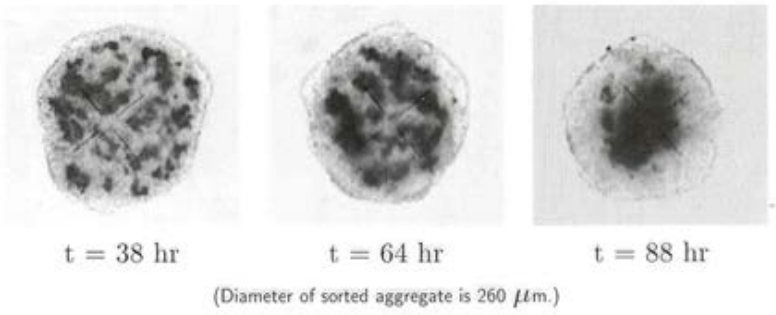
Cell Sorting—References

1. Armstrong PB (1971), “Light and Electron Microscope Studies of Cell Sorting in Combinations of Chick Embryo Neural Retina and Retinal Pigment Epithelium,” *Wilhelm Roux' Archiv* **168**, 125-141.
2. Armstrong PB, Niederman R (1972), “Reversal of Tissue Position after Cell Sorting,” *Developmental Biology* **28**, 518-527.
3. Duguay D, Foty RA, Steinberg MS (2007), “Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants,” *Current Opinion in Genetics & Development* **17**, 281–286.
4. Forgacs G, Foty RA, Shafrir Y, Steinberg MS (1998), “Viscoelastic Properties of Living Embryonic Tissues: a Quantitative Study,” *Biophysical Journal* **74**, 2227–2234. doi:10.1016/S0006-3495(98)77932-9
5. Foty RA, Steinberg MS (2005), “The differential adhesion hypothesis: a direct evaluation,” *Developmental Biology* **278**, 255–263. doi:10.1016/j.ydbio.2004.11.012
6. Steinberg MS (1975), “Adhesion-guided multicellular assembly: a commentary upon the postulates, real and imagined, of the differential adhesion hypothesis, with special attention to computer simulations of cell sorting,” *Journal of Theoretical Biology* **55**, 431–432. doi:10.1016/S0022-5193(75)80091-9
7. Steinberg MS (1996), “Adhesion in Development: An Historical Overview,” *Developmental Biology* **180**, 377–388. doi:10.1006/dbio.1996.0312
8. Steinberg MS (2003), “Differential adhesion in morphogenesis: a modern view,” *Developmental Biology* **253**, 309–323. doi:10.1016/j.gde.2007.05.002

(Available on-line at the Course web site)

Cell Sorting—A Simple Model
Biological Observations

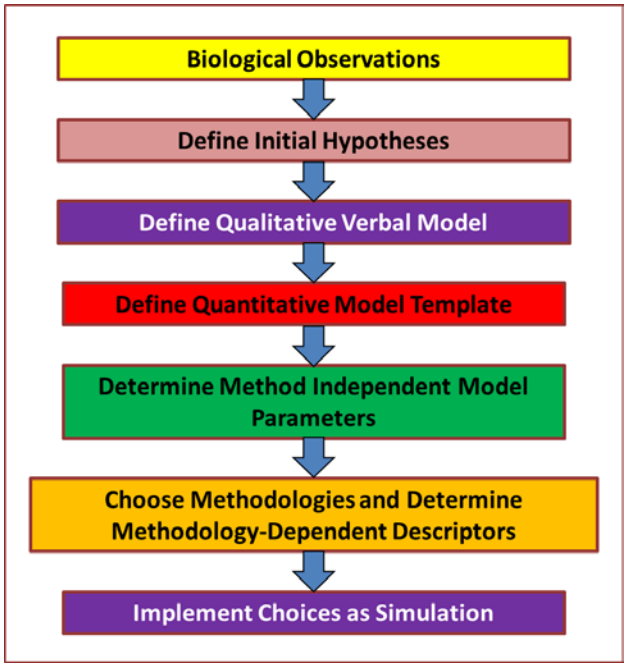
In a randomly mixed aggregate of two embryonic cell types, the aggregate gradually compacts and rounds, cells of the same type form clusters, which gradually merge together until one cell type forms a sphere in the center of the aggregate, while the other type forms a spherical shell around the sphere. We call this phenomenon **cell sorting** and say that the outer cell type **engulfs** the inner cell type.



Mixtures of the same pairs of cell types consistently put the same type to the center.

If cell type A engulfs cell type B and Cell type B engulfs cell type C, then cell type A engulfs cell type C.

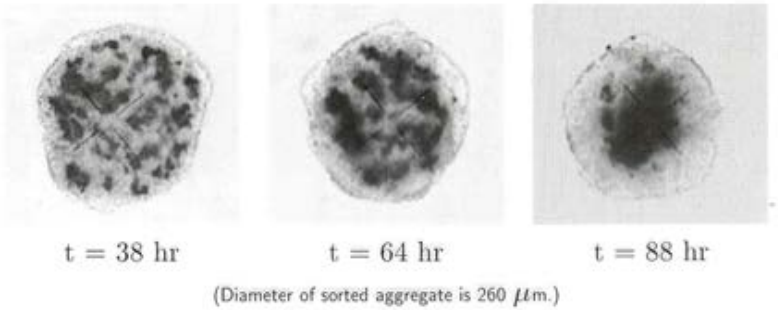
Inhibiting cell movement prevents cell sorting.

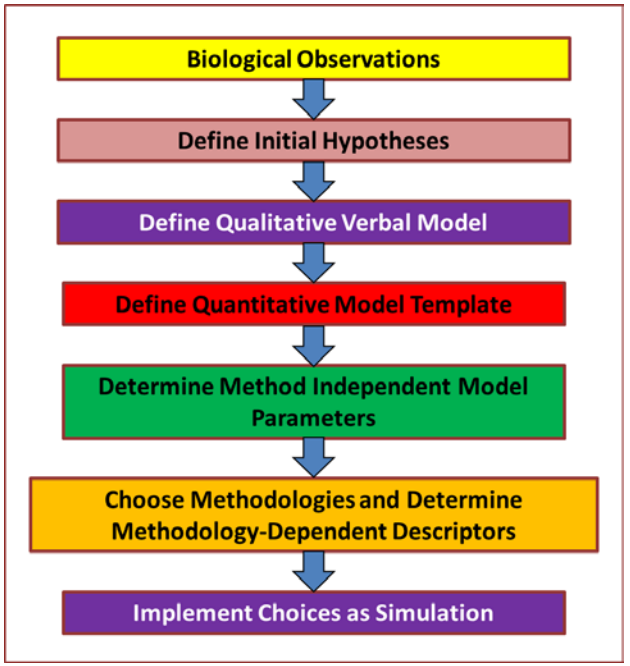


Cell Sorting—Define Initial Hypotheses

Steinberg's Differential Adhesion Hypothesis:

1) Cell types have a consistent hierarchy of adhesion energies when they come into contact with their own and other cell types.
2) The cell's adhesivity depends on the number and type of cell adhesion molecules in its membrane.
3) Cells fluctuate randomly within an aggregate.
4) Cells don't grow or change their properties during sorting
5) Less adhesive cell types engulf more adhesive cell types.

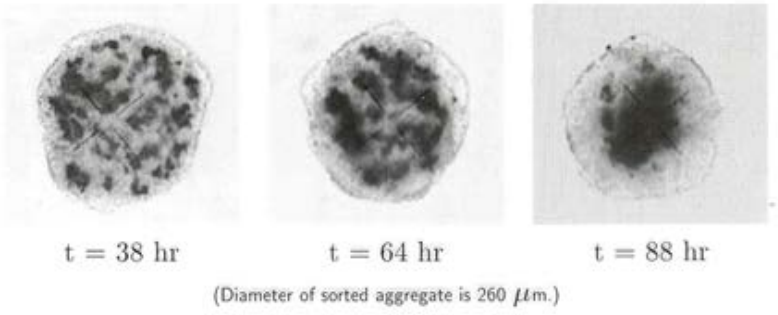




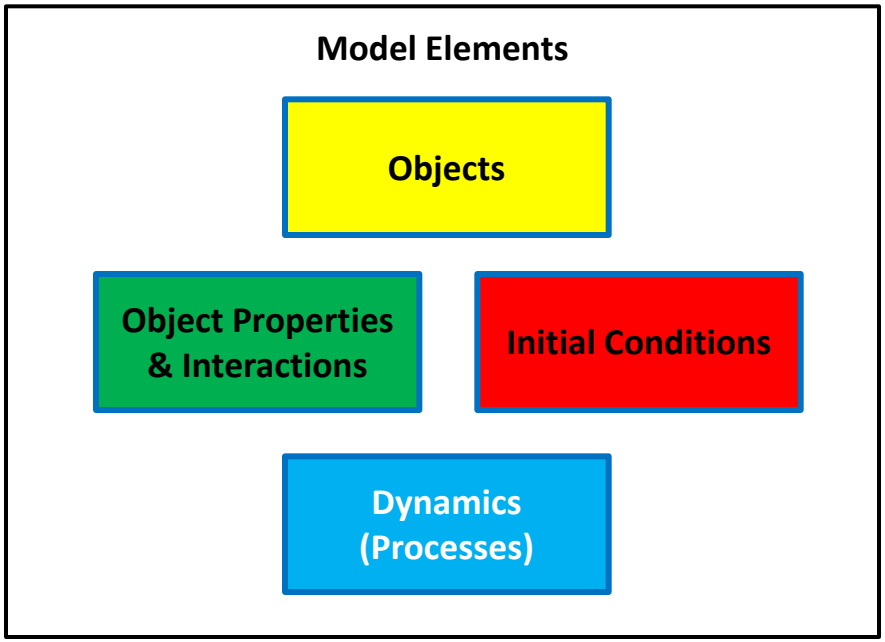
Alternative hypotheses assume: 1) differential contractility in cell membranes, 2) that cells send out a diffusible chemical signal which attracts other cells of the same type by chemotaxis, and 3) cells send out filopodia that directly pull on cells of their own type

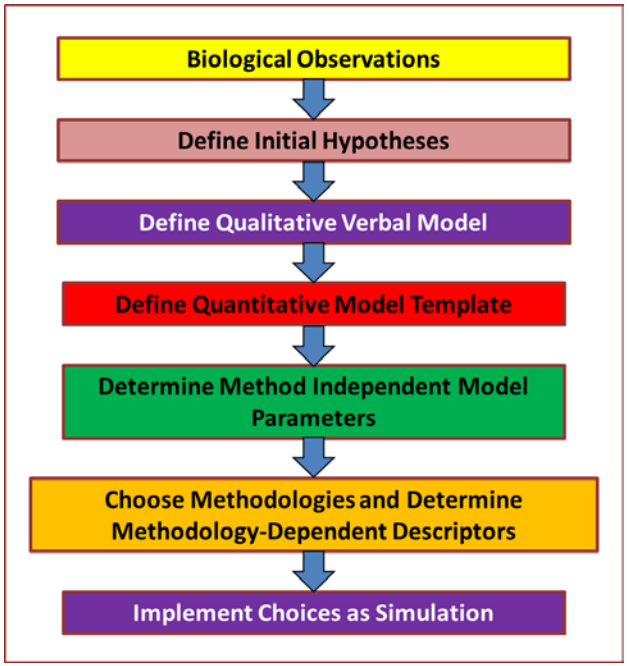
Cell Sorting—Define Model Goals from Hypotheses

Model the evolution of a randomly mixed aggregate of two mesenchymal cell types due to Differential Adhesion and random cell motility.



Questions—how does the outcome depend on the relative adhesion energies between the cell types and between the cells and medium? Can Differential Adhesion and Random Cell Motility explain observed cell sorting?

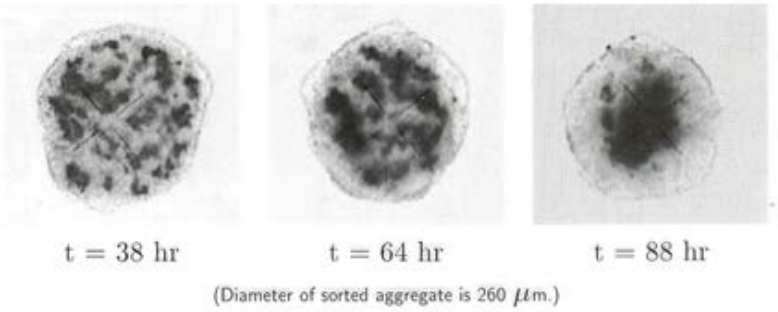


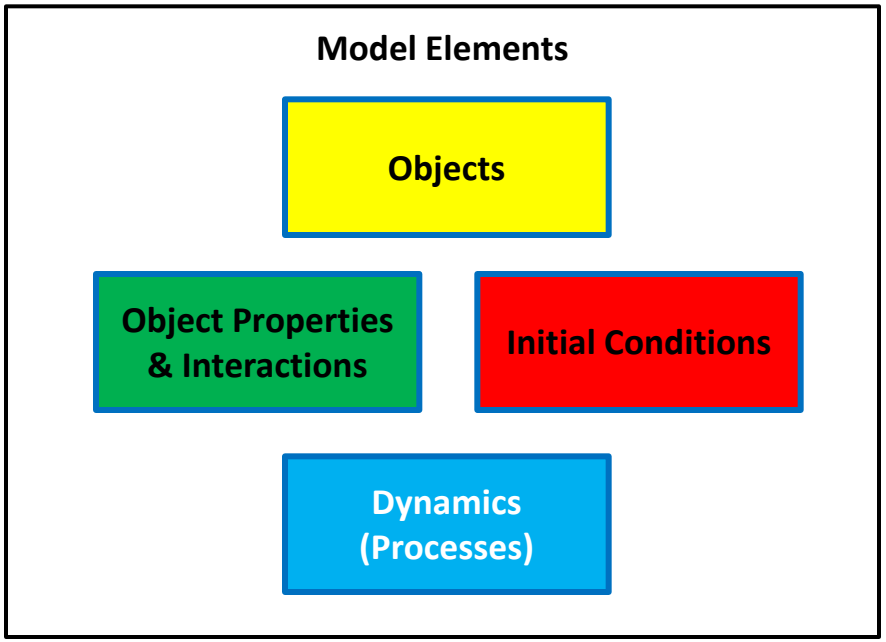


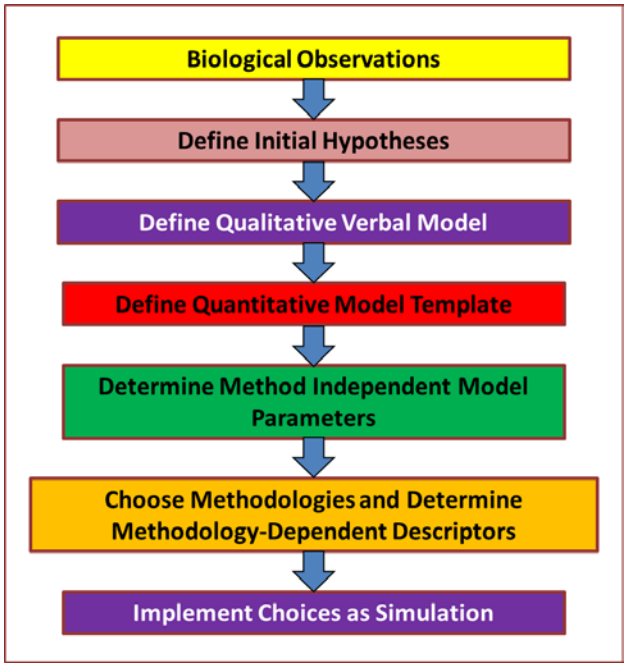
Cell Sorting—Build a Qualitative Verbal Model

What are the **Objects in the model**?

Multiple **Cells** of two **Cell Types** (Call them **Dark** and **Light** as in the image)
Surrounding fluid **Medium**







Cell Sorting—Build a Qualitative Verbal Model

What are the **Object** properties?

Cells do not grow, shrink, divide or die, i.e. they have fixed volumes and surface areas

Cells try to extend and retract their membranes

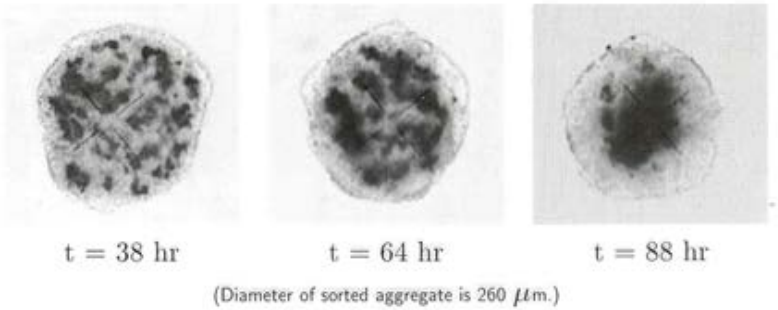
No Medium added or removed

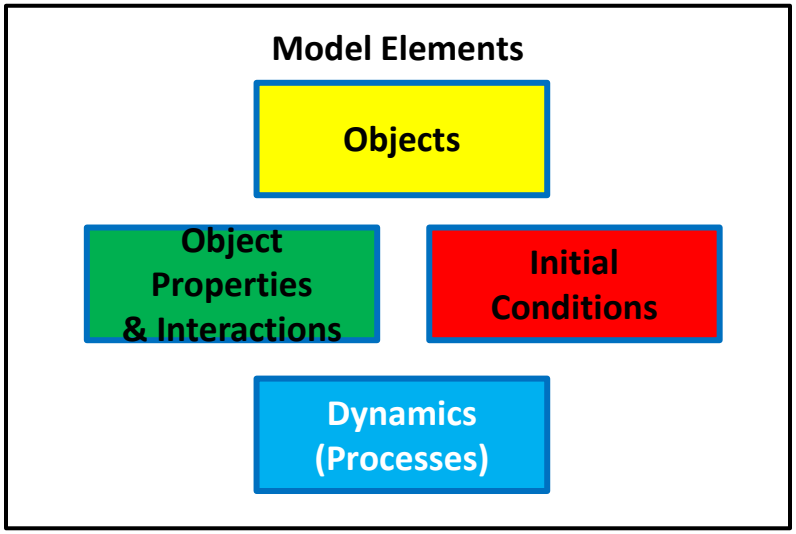
Cells appear isotopic or unpolarized, i.e. no internal structure, uniform surface properties

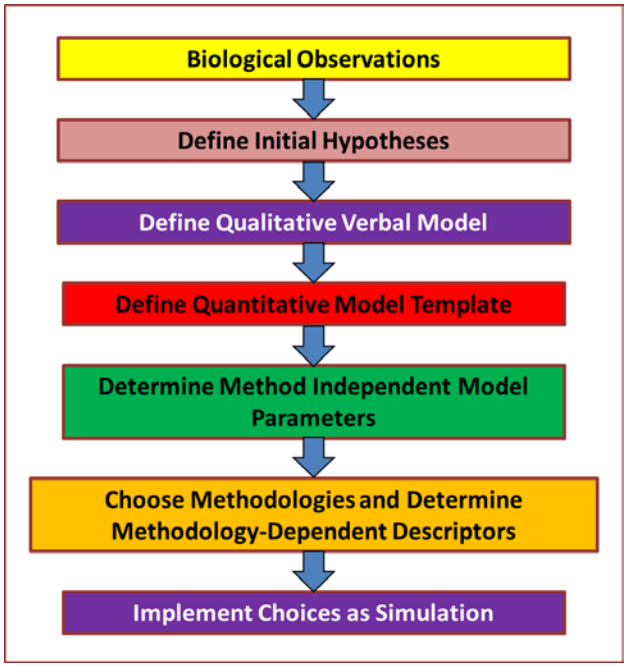
All Dark Cells seem the same

All Light Cells seem the same

Medium Doesn't seem to do much



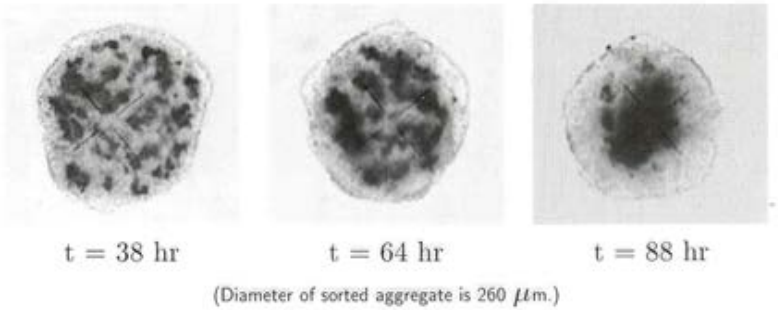


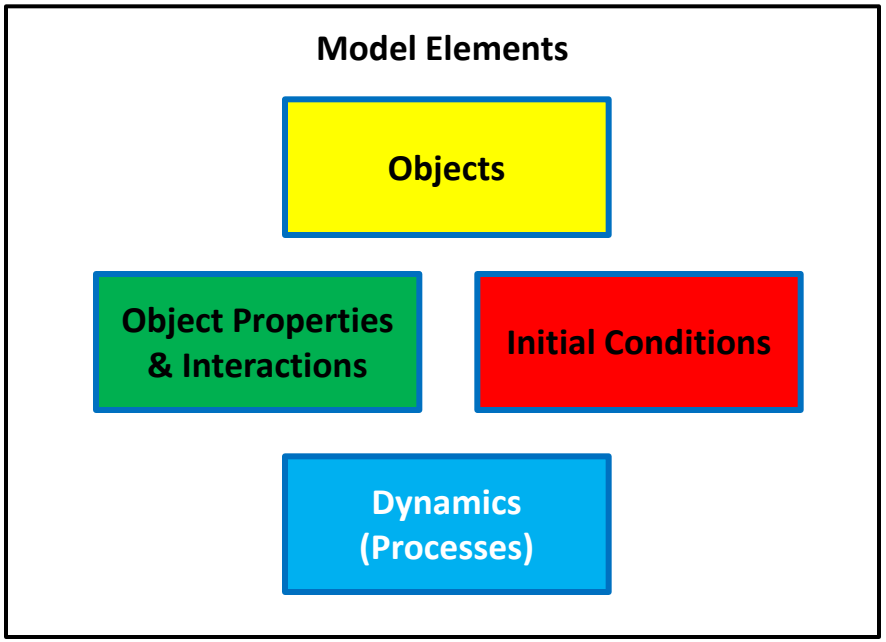


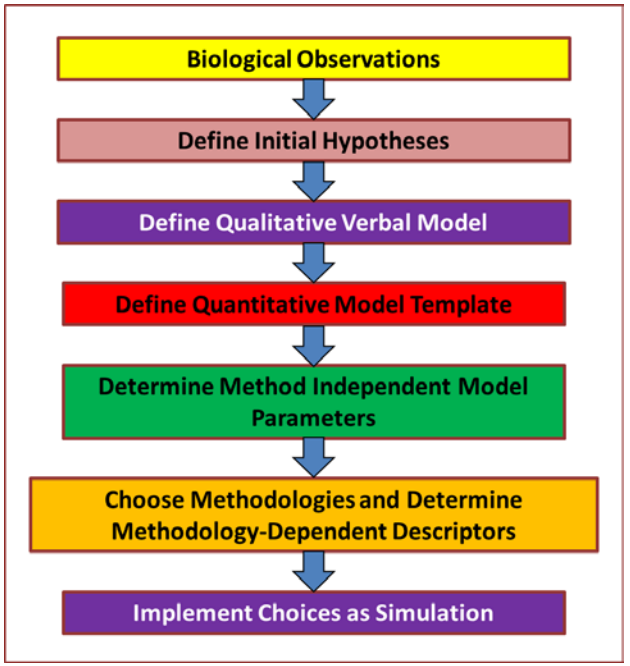
Cell Sorting—Build a Qualitative Verbal Model

What are the object interactions?

- Cells stick to each other, but can rearrange
- Adhesion greater between **Dark** cells and other **Dark** cells, less between **Light** and other **Light** cells
- **Cells** seem to repel **Medium**



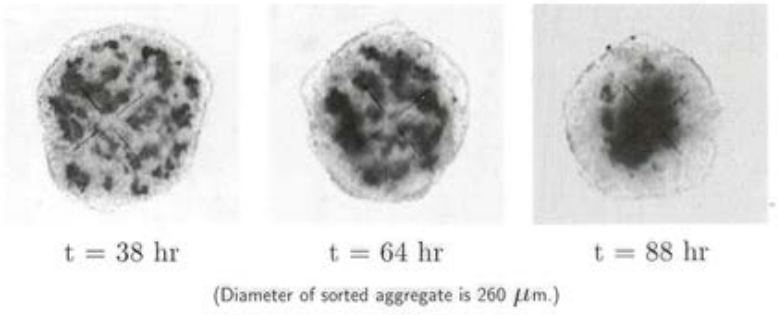


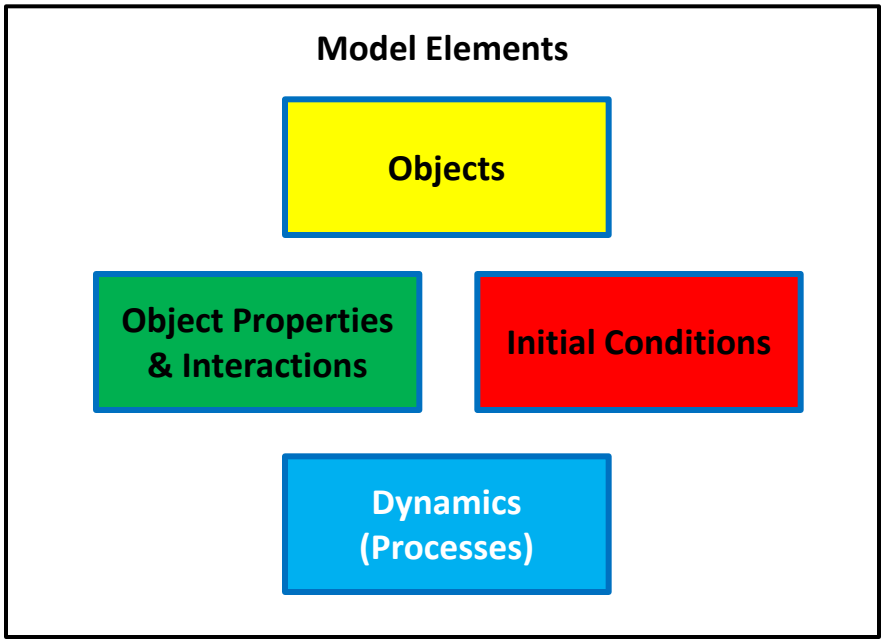


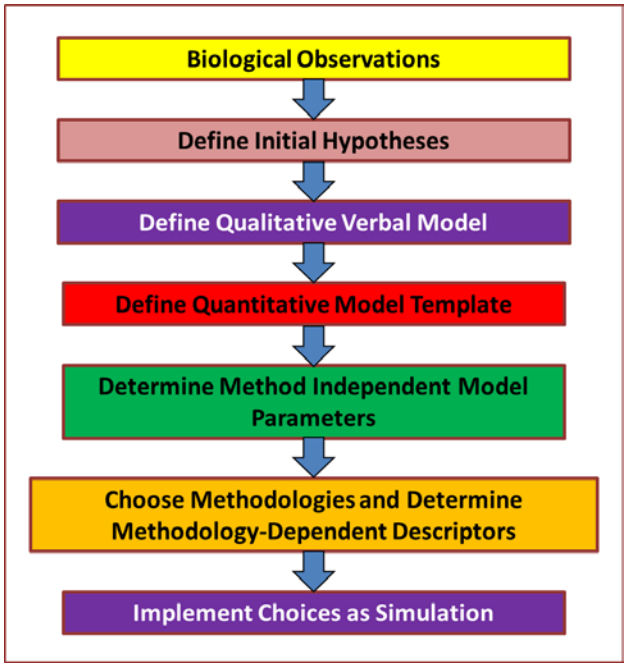
Cell Sorting—Build a Qualitative Verbal Model

What are the object dynamics?

- Cells undergo random amoeboid movements by protruding and retracting their membrane. The net cell movement depends on the forces on the cell
- Medium responds passively to cell movements



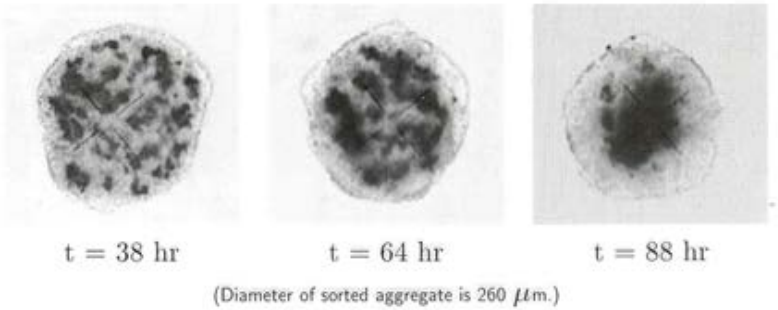


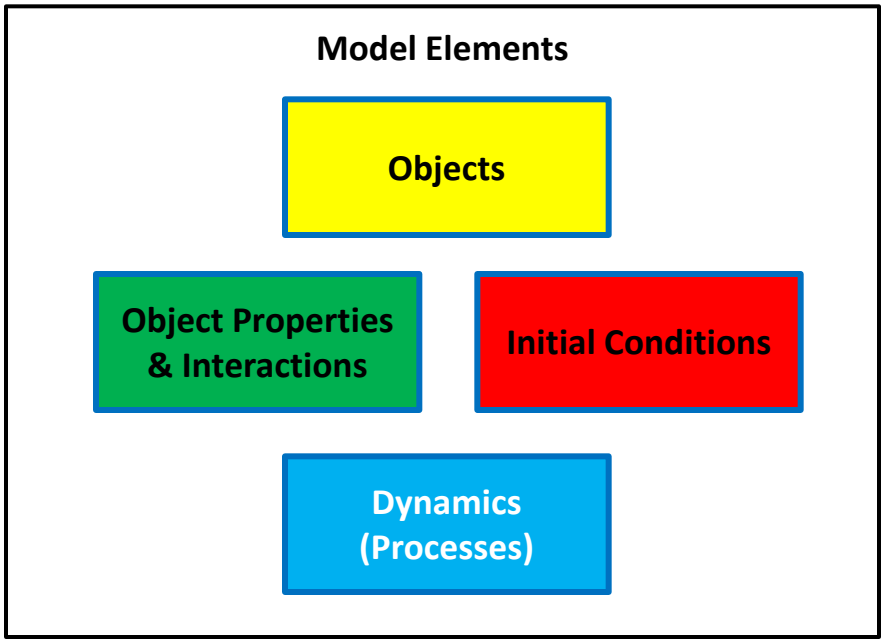


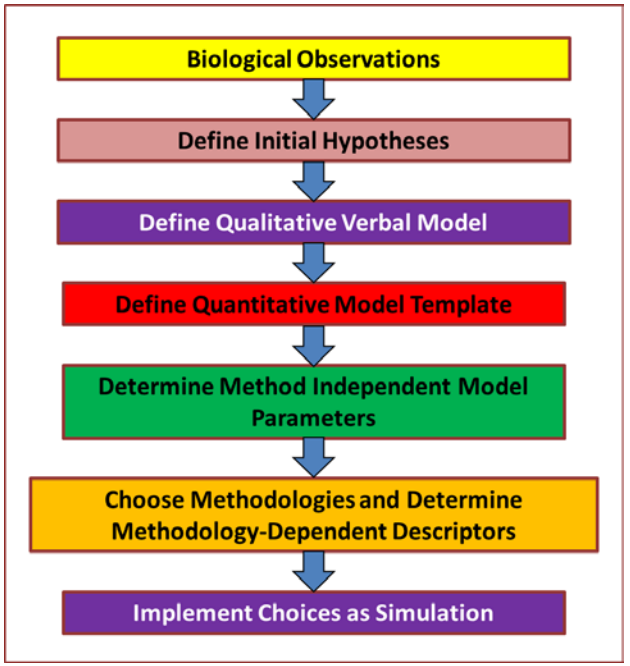
Cell Sorting—Build a Qualitative Verbal Model

What are the initial conditions?

- A randomly mixed sphere of Cells of the two Types surrounded by Medium







Cell Sorting—Building the Qualitative Biological Model

- Objects: Dark Cells, Light Cells, Medium
- Properties, Behaviors:
  - Cells do not grow, shrink, divide or die
  - No Medium added or removed
  - Cells appear isotopic or unpolarized
  - All Dark Cells seem the same; All Light Cells seem the same
- Interactions:
  - Cells stick to each other, but can rearrange
  - Adhesion greater between Dark cells and other Dark cells, less between Light and other Light cells
  - Cells seem to repel Medium
- Dynamics:
  - Cells undergo random amoeboid movements biased by external forces
- Initial Condition:
  - A randomly mixed sphere of Cells surrounded by Medium

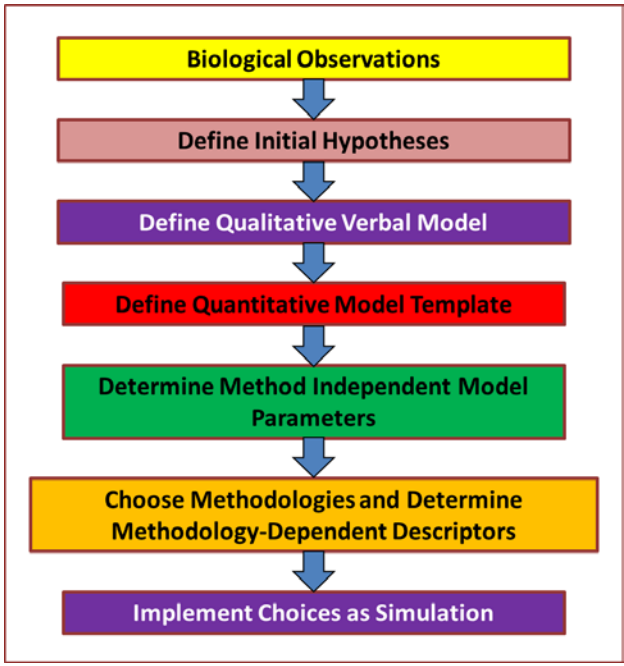
Cell Sorting—From Qualitative Verbal Model to Quantitative Model Template

- Objects: Light Cells, Dark Cells, Medium (Generalized Cell)
- Properties, Behaviors:
  - Cells have Fixed Volumes and Fixed Membrane Areas
  - Medium has Unconstrained Volume and Surface Area
  - Cells are Adhesive
  - Cells have Intrinsic Random Motility
- Interactions:
  - Cells Adhere to each other and to Medium with an Energy/Area which Depends on Cell Type (simulating different types or densities of cadherins on each Cell Type)
- Dynamics:
  - Random Cell Motility modulated by forces from system
- Initial Conditions:
  - Cells in a Blob Surrounded by Medium
  - In Blob, Cells Randomly Mixed

Cell Sorting—Define Quantitative Model Template

Object Type: Cell

Property	Parameters
Center Position	(x, y {, z})
Size	Volume
Surface	Area
Compressibility	Young's Modulus
Membrane Elasticity	Elastic Modulus
Behavior	**Parameters**
Random Motility	Fluctuation Amplitude
Viscous Dissipation	Viscosity

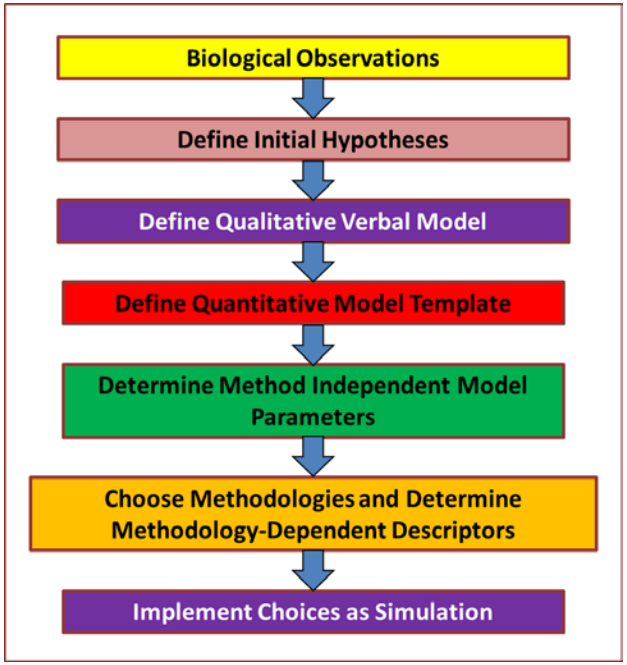


Interactions	Partner Object(s)	Parameters	Conditions
Adhesion	Cells, Medium	Energy/Area, Object Type 1, Object Type 2	Adjacency
Viscous Dissipation	Cells, Medium	Viscous Drag, Object Type 1, Object Type 2	Adjacency

Cell Sorting—Define Quantitative Model Template

Initial Condition of Cells in Aggregate

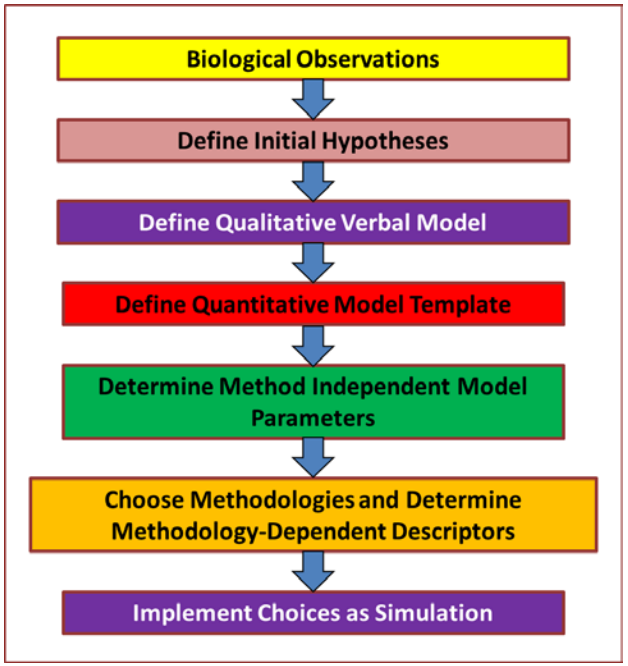
Property	Parameters
Shape	{Round, Square, …}
Size	Volume
Surface	Area
Position of center	(x, y {, z})
Fraction of Cells of Type Dark	Fraction
Fraction of Cells of Type Light	Fraction
Density of Cells in Aggregate	Amount of space between cells
Distribution of Cell Types in Aggregate	{Random, Correlated, ordered,…}



Cell Sorting—Define Quantitative Model Template

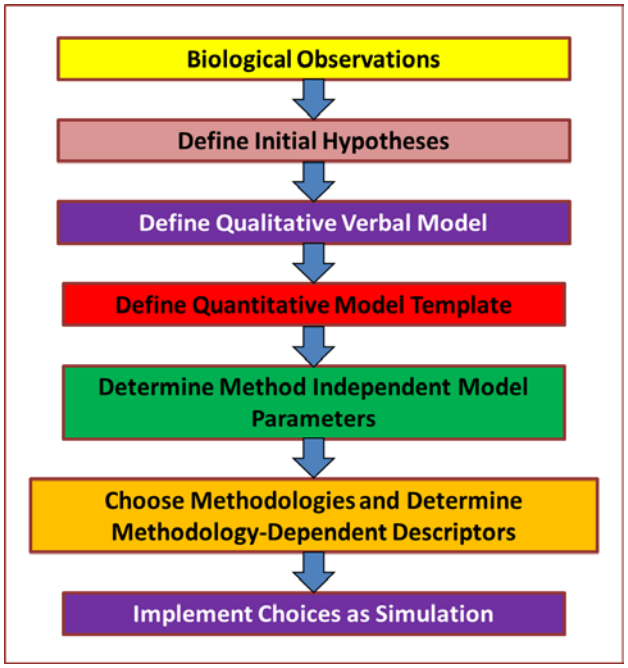
Dynamics of Cells and Medium in Aggregate

Behavior	Parameters
Cells Try to Randomly extend or retract their membranes	Success of movement, change of adhesion energy due to movement, fluctuation amplitude/force/energy of cell



Cell Sorting—Determine Method Independent Model Parameters

Parameter	Parameters
Cell Volume	Volume = 125 microns^3
Cell Surface Area	Area=75 microns^2
Dark-Dark Adhesivity	Adhesivity=100 pN/microns^2
Dark-Dark Adhesivity	Adhesivity=100 pN/microns^2
Light-Light Adhesivity	Adhesivity=30 pN/microns^2
Dark-Light Adhesivity	Adhesivity=50 pN/microns^2
Dark/Light-Medium Adhesivity	Adhesivity=10 pN/microns^2
…	…



Cell Sorting—Choose Methodologies and Methodology-Dependent Parameters

Methodology: CompuCell3D, GGH/CPM cell representation

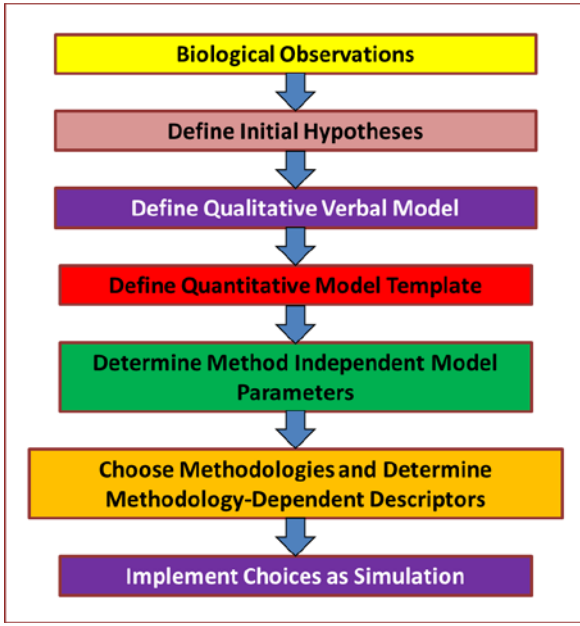
Adhesion energies are called "J" in GGH/CPM and the value of J is minus the adhesion energy

Cells live on a discrete cell lattice, so need to choose lattice element size (*e.g.* 1 micron^3)

Lattice type: Square, 3rd neighbor interactions, periodic boundary conditions (note that in CC3D, parameters don't automatically rescale when you change interaction range)

Montecarlo Dynamics expresses cell fluctuation amplitude in terms of Energy/lattice constant length, called Temperature in GGH/CPM

Blob initializer tool makes spherical blob of cells of specified types with given center position, radius, cell sizes and separation between cells

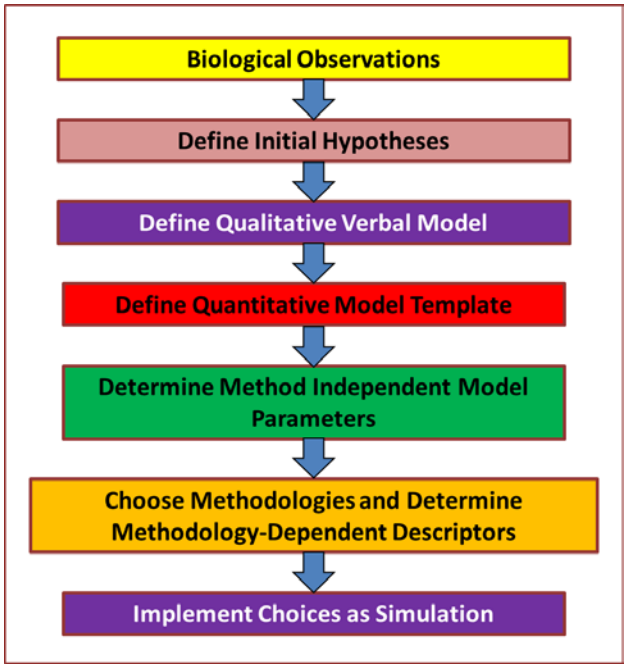


Cell Sorting—Choose Methodologies and Methodology-Dependent Parameters

The GGH/CPM specifies the Young's Modulus and membrane elasticity in terms of the constraint parameters λ_{volume} and $\lambda_{surface}$

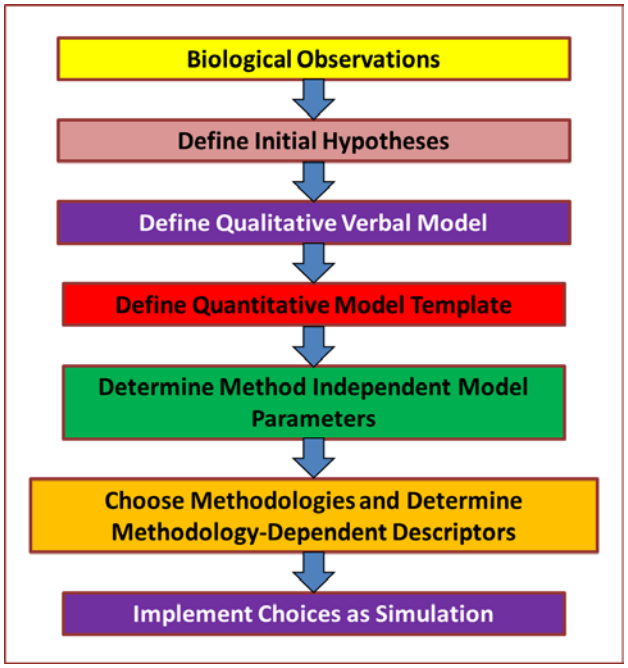
Note that the GGH/CPM does not allow simple user control of viscous terms (all movement is perfectly damped by virtue of the Dynamics)

CompuCell3D calls specifications of Objects, Properties, Behaviors and Interactions, **Plugins** and **Steppables**



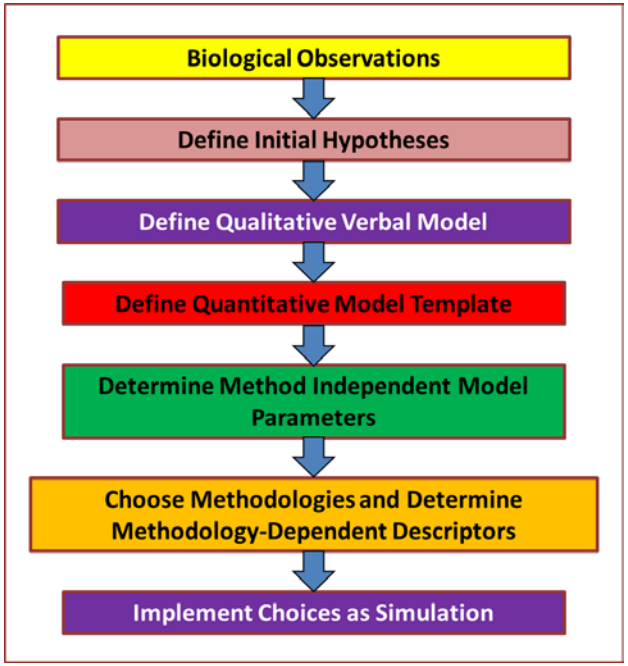
Cell Sorting—Implement as Simulation

```
<CompuCell3D>
 <Potts>
   <Dimensions x="100" y="100" z="100"/>
   <Steps>10000</Steps>
   <Temperature>10</Temperature>
   <NeighborOrder>3</NeighborOrder>
 </Potts>
<Plugin Name="CellType">
    <CellType TypeName="Medium" TypeId="0"/>
    <CellType TypeName="Dark" TypeId="1"/>
    <CellType TypeName="Light" TypeId="2"/>
 </Plugin>
 <Plugin Name="Volume">
   <TargetVolume>125</TargetVolume>
   <LambdaVolume>2.0</LambdaVolume>
 </Plugin>
<Plugin Name="Surface">
   <TargetSurface>75</TargetSurface>
   <LambdaSurface>2.0</LambdaSurface>
 </Plugin>
```



Cell Sorting—Implement as Simulation

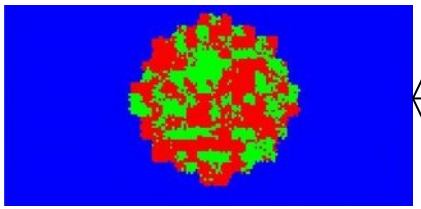
```
<Plugin Name="Contact">
   <Energy Type1="Medium" Type2="Medium">0</Energy>
   <Energy Type1="Light" Type2="Light">16</Energy>
   <Energy Type1="Dark"    Type2="Dark">2</Energy>
   <Energy Type1="Light" Type2="Dark">11</Energy>
   <Energy Type1="Light" Type2="Medium">16</Energy>
   <Energy Type1="Dark"    Type2="Medium">16</Energy>
   <NeighborOrder>3</NeighborOrder>
 </Plugin>

<Steppable Type="BlobInitializer">
   <Region>
      <Center x="50" y="50" z="50"/>
      <Radius>40</Radius>
      <Gap>0</Gap>
      <Width>5</Width>
      <Types>Dark,Light</Types>
   </Region>
</Steppable>

</CompuCell3D>
```

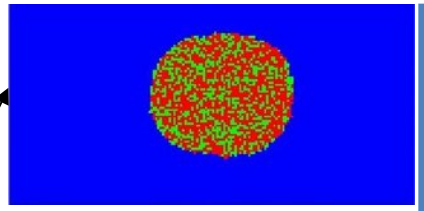


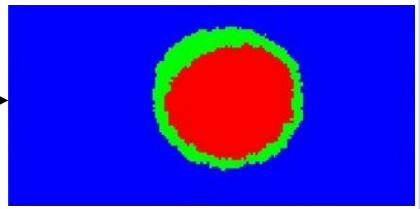
Cell Sorting—Execute Simulation

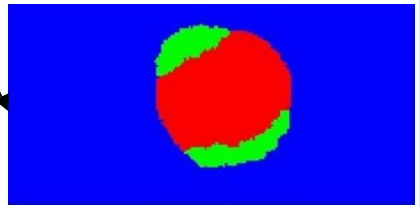
Final Pattern Depends on hierarchy of Js

Initial Condition—
Random Mixture









Cell Mixing with Cell Types with identical behaviors and heterotypic interface favored over homotypic—

$$J_{Dark,Dark} = J_{Light,Light} > J_{Dark,Light}$$

Complete Cell Sorting with Dark Cells more cohesive than Light Cells and Dark-Light contact intermediate in energy—

$$J_{Dark,Dark} < J_{Light,Dark} < J_{Light,Light} < J_{Dark,Medium} = J_{Light,Medium}$$

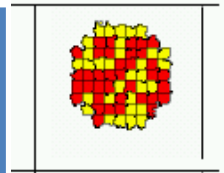
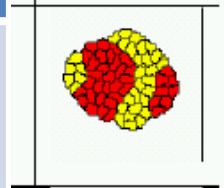
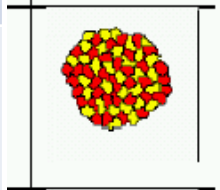
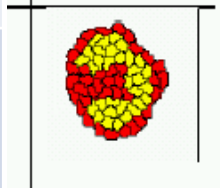
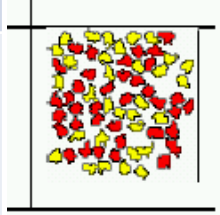
Partial Cell Sorting, with Dark-Light contact very costly—

$$J_{Dark,Dark} < J_{Light,Light} < J_{Dark,Light} < J_{Dark,Medium} = J_{Light,Medium}$$

In this figure, red denotes Dark Cells, green denotes Light Cells. Cell-Cell boundaries not rendered.

Cell Sorting—Execute Simulation

Additional Final Patterns for different J hierarchies

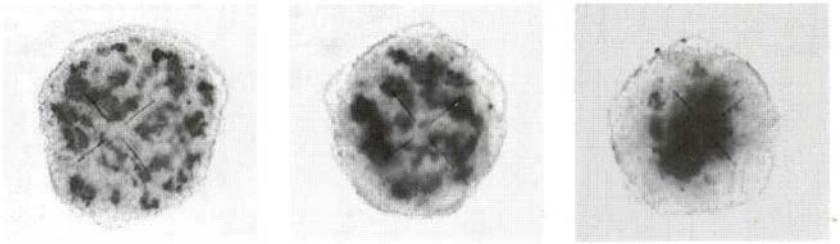
Initial Condition—Random Mixture	
Incomplete Cell Sorting with Cell Types with identical behaviors and homotypic interface favored over heterotypic — $J_{Dark,Dark} = J_{Light,Light} < J_{Dark,Light}$	
Cell Mixing with Cell Types with identical behaviors and heterotypic interface favored over homotypic— $J_{Dark,Dark} = J_{Light,Light} > J_{Dark,Light}$	
Inverse Cell Sorting with Dark Cells more cohesive but preferring Medium contact compared to Light Cells— $J_{Dark,Dark} < J_{Light,Dark} < J_{Light,Light} < J_{Dark,Medium} < J_{Light,Medium}$	
Cell Dispersal, with Cells preferring to contact Medium rather than each other— $\{J_{Dark,Dark}, J_{Dark,Light}, J_{Light,Light}\} > 2 \times \{J_{Dark,Medium}, J_{Light,Medium}\}$	

In this figure, red denotes Dark Cells, yellow denotes Light Cells.

Cell Sorting—Compare Sorting Kinetics in Experiment and Simulation

Growth law in tissues

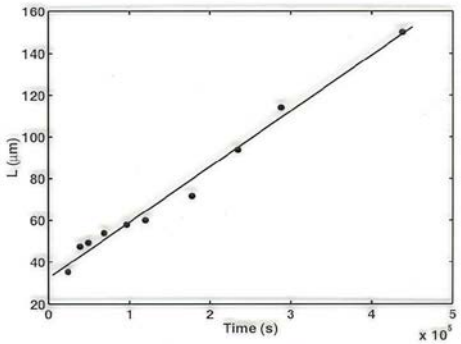
Growth of pigmented clusters in neural retinal tissue



t = 38 hr t = 64 hr t = 88 hr

(Diameter of sorted aggregate is 260 μm.)

For large volume fractions $\rightarrow$ $L \sim (\sigma/\eta)t$



Slope = 2.7 x 10^{-4} μm s^{-1}

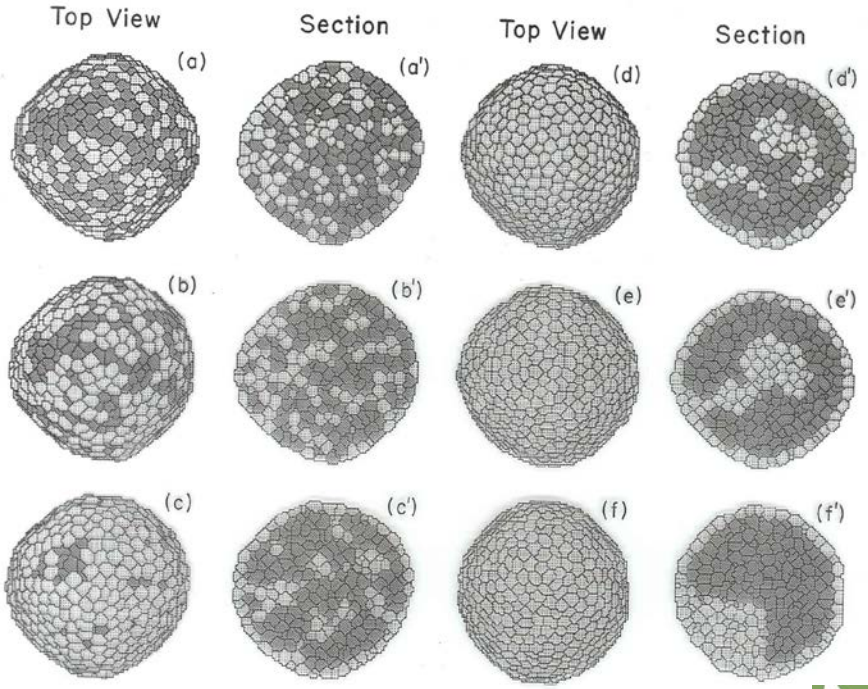


Figure 2.

Sample Questions in Model Definition

- What Structural components (Objects) are important?
- Should Objects be represented individually or as continua?
- What type of Cells are important?
- What is the Chemical and Mechanical Environment around the Cells (ECM)?
- Are the Cells polar or not?
- Do Cells grow and divide?
- Do Cells differentiate or die?
- What Signaling Mechanisms are active within Cells? Between Cells?
- How strongly do Cells of one type adhere to Cells of another type?
- How strongly do Cells of a given type adhere to ECM?
- Is Cell adhesion labile (*e.g*. single molecule pair) or junctional?
- What chemicals do cells secrete and absorb?
- If Chemical Fields diffuse, how rapidly do they diffuse?
- If Chemical Fields do not diffuse, what are their mechanical properties?
- How stable are Chemical Fields (what are their decay rates)?
- What are the mechanical properties of the ECM?
- How do Cells interact mechanically (move in response to and remodel) with ECM?

Things you should do right now

- Separate Models from Simulations
- Use Best Practice Workflows during model development
- Document at all levels of abstraction and maintain lists relating levels
- Modularize Model Components and Simulation Methods
- Annotate at all Scales so Simulations include Models