CPIB SUMMER SCHOOL 2010: INTRODUCTION TO BIOLOGICAL MODELLING

University of Nottingham

5-8 September 2011

Course Outline

Aims:

- To introduce modelling and quantitative approaches to biology
- To explain where equations come from and what they mean, placing the mathematics into a context that is relevant for the life scientist.
- To enable life scientists to gain a better understanding of what a model is, and how to go about building one.

Objectives - By the end of the session, participants will:

- understand key concepts in how to build models of biological systems
- know how to investigate the behaviour of those models
- be able to interpret the results of those models.

Key Themes

The kinds of behaviour that dynamic models can exhibit (e.g. exponential growth or decay, steady states, oscillations, patterns), and their stability.

Single variable differential equation models:

- How to work out their dynamics by sketching one simple graph.
- Applications, including to population growth and gene regulation.

Multi-variable differential equation models:

- Interacting populations, signalling networks, biochemical reactions.
- How to turn reactions into a model with the law of mass action.
- How to work out a lot about their dynamics by sketching two (or more) graphs.

Parameter estimation and sensitivity

Including randomness - stochastic models

Spatial models - signalling, transport, growth and pattern formation

How to create, simulate and analyse models using appropriate software

Course Schedule (1)

Day 1 (Monday 5 September 2011)	
11:00-12:00	Lecture 1.1: *Mathematical modelling and Systems Biology* (Pope C17)
12:00-13:00	Lunch (Mathematical Sciences Atrium)
13:00-14:00	Lecture 1.2: *Introduction to modelling with differential equations* (Pope C17)
14:00-15:00	Practical 1: *Analysis and simulation of single-variable models* (Pope A15)
15:00-15:30	Coffee (Pope A17)
15:30-17:00	Practical 1, continued (Pope A15)

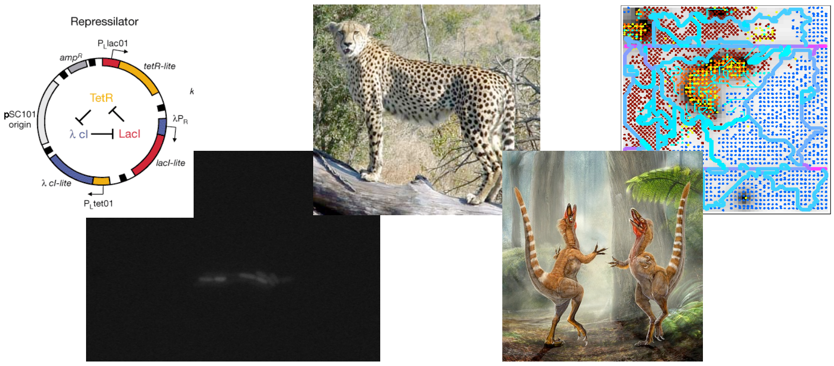
Day 2 (Tuesday 6 September 2011)	
09:00-10:30	Lecture 2.1: *Building multi-variable models* (Pope C17)
10:30-11:00	Coffee (Pope A17)
11:00-12:30	Lecture 2.2: *Analysing multi-variable differential equation models* (Pope C17)
12:30-13:30	Lunch (Mathematical Sciences Atrium)
13:30-15:00	Practical 2: *Building and simulating multi-variable models* (Pope A15)
15:00-15:30	Coffee (Pope A17)
15:30-17:00	Practical 2, continued (Pope A15)

Course Schedule (2)

Day 3 (Wednesday 7 September 2011)	
09:00-10:30	Lecture 3.1: *Parameter estimation and sensitivity analysis* (Pope C17)
10:30-11:00	Coffee (Coates C28)
11:00-12:30	Practical 3.1: *Parameter estimation and sensitivity analysis* (Pope A15)
12:30-13:30	Lunch
13:30-15:00	Lecture 3.2: *Stochastic models* (Pope C17)
15:00-15:30	Coffee (Coates C28)
15:30-17:00	Practical 3.2: *Stochastic models* (Pope A15)

Day 4 (Thursday 8 September 2011)	
09:00-10:30	Lecture 4.1: *Spatial models - signalling, transport and growth* (Pope A17)
10:30-11:00	Coffee (Pope A17)
11:00-12:30	Practical 4: *Spatial modelling* (Pope A15)
12:30-13:30	Lunch
13:30-15:00	Practical 4, continued (Pope A15)
15:00-15:30	Coffee (Coates C28)
15:30-17:00	Lecture 4.2: *Multiscale models and applications* (Pope A17)

Lecture 1.1
Mathematical modelling and Systems Biology



Markus Owen / School of Mathematical Sciences

The University of Nottingham

CMMB
Centre for Mathematical Medicine and Biology

What is Systems Biology?

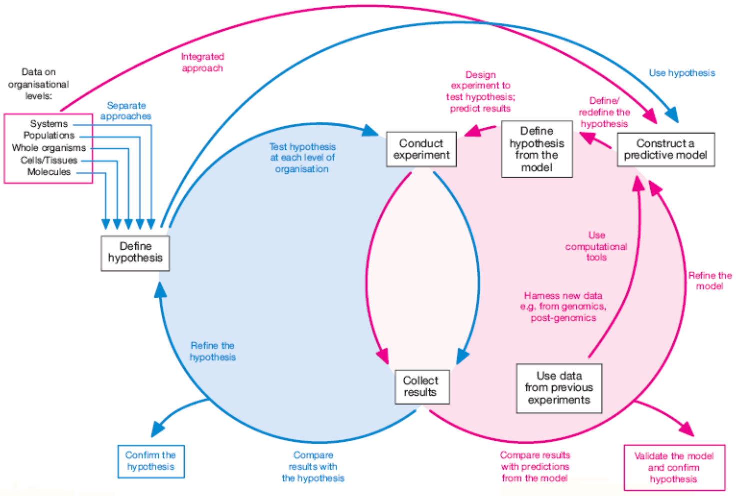
BBSRC bioscience for the future the Biotechnology and Biosciences Research Council says:

"Systems biology is an approach by which biological questions are addressed through integrating experiments with computational modelling and theory, in re-enforcing cycles."

BBSRC funds 6 UK Centres for Integrative Systems Biology: Edinburgh, Imperial, Manchester, Newcastle, Nottingham, Oxford.

- Biological systems: large numbers of components interacting at various scales.
- In the past, life scientists could only study a handful of components at a time.
- This led to an approach assuming a simple chain of cause and effect.
- Most genes, proteins, cells, organisms and other components work within a complex network of interactions, with interlocking positive and negative feedback loops.
- Systems Biology provides a new conceptual framework for understanding biological problems. It combines the mathematical, computational, physical and engineering sciences with biological experiments.

What is Systems Biology?



We'll look at some examples: Gene-regulatory networks, Animal coat patterns, Cancer

Mathematical modelling approaches

- Models help to encode our understanding and assumptions about a system
- Can be used to test hypotheses, make predictions, carry out *in silico* experiments ("What happens if ...?")
- Models are simplifications that can be extended when necessary (ideally in a loop in association with experimental work).

- Compartmental models, e.g. ordinary differential equations

Rate of change of cancer volume = cell division - cell death - killing via therapy

- Spatial models (e.g. partial differential equations, PDEs)

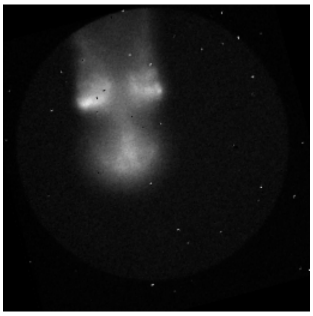
Rate of change of cancer cell density = cell division - cell death - killing via therapy + movement

- Individual-based models, e.g. cellular automaton
- Hybrid multiscale models - combining all of the above.

Back to the beginning...

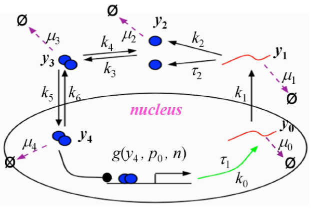
Biological processes understood as emergent properties of complex networks of interacting components.

Question: what are the mechanisms regulating emergence?



Masamizu et al., PNAS. 2006

Hes1 (and other Notch pathway genes) oscillate in the presomitic mesoderm of developing vertebrate embryos.



What kinds of processes?

How to make a switch?

Important for Cell differentiation, quorum sensing, lac operon inducible system, lysis-lysogeny decision by phage Lambda, ... Delta-Notch signalling is a simple example. Feedback + coupling selects a subpopulation of cells for a neuronal fate.

How to make an oscillator?

Cell cycle, circadian rhythms, cardiac action-potential

How to make an organism?

Fate determination + cell movements, proliferation, etc, etc, ...

Population growth and interactions

From bacteria to humans; cancer (mutant cells invading a normal host); epidemiology; ecology; ...

Variables, models and parameters

- The system **state** is a set of measurable properties of the system.
  Examples: mRNA & protein concentration, membrane potential, number of cells, …
- We would like to understand the past and present and predict the future. Given a set of measurements today, what will be the result of making those measurements tomorrow?
- A **model** is a representation of the system that we can use to answer such questions.
  If the state is changing with time (usually denoted *t*), then the model is **dynamical**.
  The time-varying components of the state are **variables** (e.g. denoted $x_1(t), x_2(t), \ldots, x_N(t)$).
- The state of the model at time *t* is just the set of all the variables at time *t*:

$$S(t) = \{x_1(t), x_2(t), \ldots, x_N(t)\}.$$

- The form of model we shall study is:

$$S(t_2) = f\,(S(t_1); p_1, p_2, \ldots, p_M), \qquad t_2 > t_1$$

where *f* is a function encoding our understanding of how the system components affect one another, and $p_1, p_2, \ldots, p_M$ are model **parameters**.

This simply states that the future state of the system is some function of its past state (i.e. that the future is **predictable**).

- **Parameters** are numerical values that encode information about the system that is not included in the dynamic state. E.g. the concentration of an mRNA species is a *variable*; the linear degradation rate of the mRNA is a *parameter*.

Inferring Networks

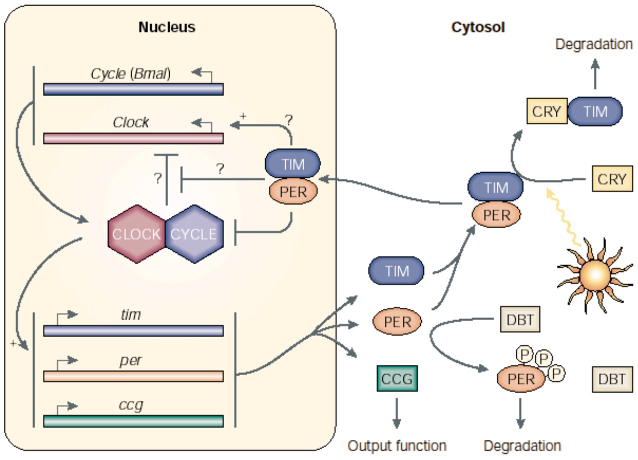
Components ($x_i(t)$) and interactions (encoded in f) can be inferred from a wide range of data sources:

- Genetic screens
- RNAi screens
- mRNA profiling (e.g. microarrays)
- Metabolic profiling
- Protein-protein interaction screens (e.g. yeast-two-hybrid, TAP mass spec.)
- ChIP-on-chip analysis of transcription factor binding
- Biochemistry
- Population data (e.g. on predator-prey or epidemiological interactions)

Each has strengths and limitations

Integration of multiple data sources is important for reliable inference.

The Circadian Oscillator Network



Cermakian & Sassone-Corsi, Nature Rev. Cell Molec. Biol. 1, 59–67 (2000).

Continuous Process Models

- In reality, the state of each network component should be represented by a discrete quantity — an integer (e.g. the number of molecules of a particular mRNA in a cell, the number of individuals in a population).
- Also, changes in state over time are discrete events (production or degradation of a network component, births/deaths in a population).
- In practice, if the amount of each component is sufficiently large, then its state can be approximated by a continuous variable that changes smoothly and continuously in time (e.g. concentration, population density).
- In doing this, we are essentially representing a continuous process rather than a set of events.

Ordinary differential equations (ODEs)

- If we represent the network state $S(t)$ as continuous, then it has a well-defined **rate of change**: $\frac{dS}{dt}(t)$
- An **ordinary differential equation (ODE)** model gives the rates of change of the variables $x_i(t)$ as functions of the state at that time:

$$\boxed{\frac{dx_i}{dt}(t) = f_i\left(\left\{x_{\langle i\rangle}(t)\right\}; p_1, p_2, \ldots, p_M\right), \quad i = 1,2,\ldots,n}$$

  where $<i>$ is the set of variables that affect $x_i(t)$.
- The functions f_i encode the form of the interactions between components. These functions are hard to determine. There are no high-throughput methodologies for getting them.
- In practice, models are often based on a small set of standard representative forms for the f_i (see later for examples).

  For chemical reactions, an important concept is the law of mass action...

The law of Mass Action

- The **law of mass action** states that the rate of a chemical reaction is proportional to the product of the concentrations of the reactants.
- Used to develop Stochastic and Differential equation models for biochemical reactions.
- **Assumptions: i)** a well stirred solution and **ii)** low molecular concentrations, where the probability of diffusing molecules to get close enough, for a reaction to occur, is proportional to the concentrations.
- A rate parameter is used to define the 'probability' of a reaction to occur if two molecules approach each other.
- The mass action formalism has been validated in many experimental settings.
- Given: $n_1 S_1 + n_2 S_2 + \cdots \xrightarrow{k_f} m_1 P_1 + m_2 P_2 + \cdots$
- The reaction rate is $k_f\left(S_1^{n_1} \cdot S_2^{n_2} \cdots\right)$
- The rate of change of a species depends on the rate of reaction and **the net change in the number of molecules** of that species.
- In reality, all reactions should be broken down into bimolecular steps.

Michaelis-Menten enzyme kinetics

- S, substrate; E, enzyme; P, product: $S + E \underset{k_2\,\boxed{2}}{\overset{k_1\,\boxed{1}}{\rightleftharpoons}} SE \xrightarrow[\boxed{3}]{k_3} P + E$

$\boxed{1}$ Reactants S and E, rate $k_1[S][E]$.
Consumes one molecule of S and E,
and produces one molecule of SE.

$\boxed{2}$ Single reactant SE, rate $k_2[SE]$.
Consumes one molecule of SE,
and produces one molecule of S and E.

$\boxed{3}$ Single reactant SE, rate $k_3[SE]$.
Consumes one molecule of SE,
and produces one molecule of P and E.

$$\frac{d[S]}{dt} = \overset{\boxed{1}}{-k_1[S][E]} + \overset{\boxed{2}}{k_2[SE]}$$

$$\frac{d[E]}{dt} = \overset{\boxed{1}}{-k_1[S][E]} + \overset{\boxed{2}}{k_2[SE]} + \overset{\boxed{3}}{k_3[SE]}$$

$$\frac{d[SE]}{dt} = \overset{\boxed{1}}{k_1[S][E]} - \overset{\boxed{2}}{k_2[SE]} - \overset{\boxed{3}}{k_3[SE]}$$

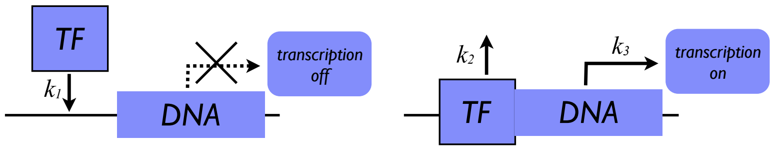
$$\frac{d[P]}{dt} = \overset{\boxed{3}}{k_3[SE]}$$

- Constant total enzyme
- Substrate assumed in excess
- *[SE]* at **quasi-steady state**

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_m + [S]} \qquad V_{max} = k_3 E_0 \qquad K_m = (k_2 + k_3)/k_1$$

Transcriptional/translational activation

- TF binds to DNA, this complex activates production of protein P.

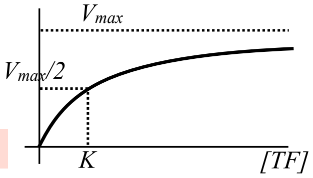


$$TF + DNA \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TFDNA \xrightarrow{k_3} P + TFDNA$$

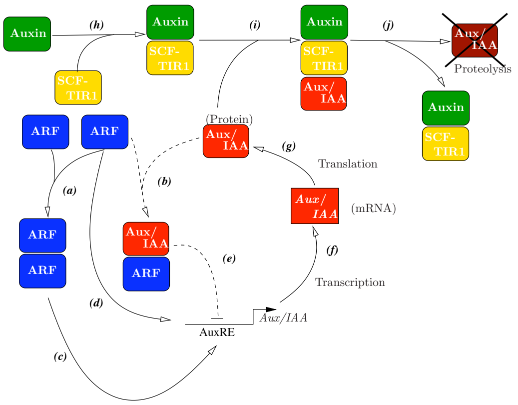
- Assuming TF binding is fast enables use of Michaelis-Menten approach.
- DNA acts as *enzyme*, $[DNA] + [TF\text{-}DNA] = 1$

$$\frac{d[P]}{dt} = V_{max}\frac{[TF]}{K + [TF]}$$

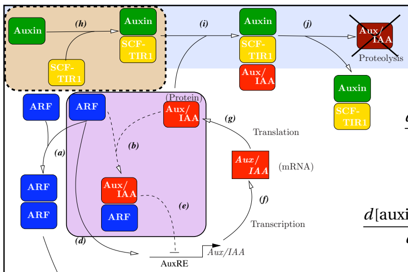
This and similar forms are used in "gene network" models



ODE Example - Auxin signalling



- Auxin is a plant hormone, which stimulates degradation of Aux/IAAs.
- Aux/IAAs repress their own transcription.
- Hence Auxin stimulates Aux/IAA transcription.



Mass action:

$$\frac{d[\text{auxin}]}{dt} = \omega + \boxed{k_\text{d}[\text{auxin–TIR1}] - k_\text{a}[\text{auxin}][\text{TIR1}]} - \mu_\text{auxin}[\text{auxin}],$$

$$\frac{d[\text{TIR1}]}{dt} = \boxed{-k_\text{a}[\text{auxin}][\text{TIR1}] + k_\text{d}[\text{auxin–TIR1}],}$$

$$\frac{d[\text{auxin–TIR1}]}{dt} = \boxed{k_\text{a}[\text{auxin}][\text{TIR1}] - k_\text{d}[\text{auxin–TIR1}]}$$
$$+ (l_\text{d} + l_\text{m})[\text{auxin–TIR1–IAA}] - l_\text{a}[\text{auxin–TIR1}][\text{IAA}_\text{p}]$$

$$\frac{d[\text{IAA}_\text{p}]}{dt} = \delta[\text{IAA}_\text{m}] - l_\text{a}[\text{IAA}_\text{p}][\text{auxin–TIR1}] + l_\text{d}[\text{auxin–TIR1–IAA}]\boxed{-p_\text{a}[\text{IAA}_\text{p}][\text{ARF}] + p_d[\text{ARF-IAA}],}$$

$$\frac{d[\text{auxin–TIR1–IAA}]}{dt} = l_\text{a}[\text{IAA}_\text{p}][\text{auxin–TIR1}] - (l_d + l_\text{m})[\text{auxin–TIR1–IAA}],$$

$$\frac{d[\text{IAA}^*]}{dt} = l_\text{m}[\text{auxin–TIR1–IAA}] - \mu_{\text{IAA}^*}[\text{IAA}^*],$$

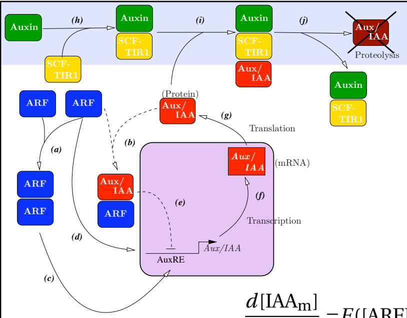
$$\frac{d[\text{ARF}]}{dt} = -2q_\text{a}[\text{ARF}]^2 + 2q_\text{d}[\text{ARF}_2]\boxed{-p_\text{a}[\text{ARF}][\text{IAA}_\text{p}] + p_\text{d}[\text{ARF-IAA}],}$$

$$\frac{d[\text{ARF-IAA}]}{dt} = \boxed{p_\text{a}[\text{ARF}][\text{IAA}_\text{p}] - p_\text{d}[\text{ARF-IAA}],}$$

$$\frac{d[\text{ARF}_2]}{dt} = q_\text{a}[\text{ARF}]^2 - q_\text{d}[\text{ARF}_2],$$

$$\frac{d[\text{IAA}_\text{m}]}{dt} = F([\text{ARF}], [\text{IAA}_\text{p}], [\text{ARF-IAA}], [\text{ARF}_2]) - \mu_{\text{IAA}_\text{m}}[\text{IAA}_\text{m}],$$

Transcriptional regulation:



$$\frac{d[\mathrm{IAA_m}]}{dt} = F([\mathrm{ARF}], [\mathrm{IAA_p}], [\mathrm{ARF\text{-}IAA}], [\mathrm{ARF_2}]) - \mu_{\mathrm{IAA_m}}[\mathrm{IAA_m}],$$

$$F([\mathrm{ARF}], [\mathrm{IAA_p}], [\mathrm{ARF\text{-}IAA}], [\mathrm{ARF_2}]) = \boxed{\frac{\lambda_1 \dfrac{[\mathrm{ARF}]}{\theta_{\mathrm{ARF}}} + \lambda_2 \left(\dfrac{[\mathrm{ARF_2}]}{\theta_{\mathrm{ARF_2}}} + \dfrac{[\mathrm{ARF}]^2}{\psi_{\mathrm{ARF}}} \right)}{1 + \dfrac{[\mathrm{ARF}]}{\theta_{\mathrm{ARF}}} + \dfrac{[\mathrm{ARF_2}]}{\theta_{\mathrm{ARF_2}}} + \dfrac{[\mathrm{ARF\text{-}IAA}]}{\theta_{\mathrm{ARF\text{-}IAA}}} + \dfrac{[\mathrm{ARF}][\mathrm{IAA_p}]}{\psi_{\mathrm{ARF\text{-}IAA}}} + \dfrac{[\mathrm{ARF}]^2}{\psi_{\mathrm{ARF}}}}}$$

Equivalent to Shea-Ackers formulation

ODE models: Basic assumptions

Assumption	Relaxation
The numbers of each molecular species are large enough to represent as continuous variables	Discrete models
Production and degradation processes are continuous	Discrete models
Outputs of processes begin to change as soon as the inputs change	Delay differential equations
Processes are deterministic (non-random)	Stochastic models
Spatial distribution in a cellular compartment is not important	Partial differential equations

Summary

The properties of a system can be represented by a set of variables that collectively constitute the state of a model

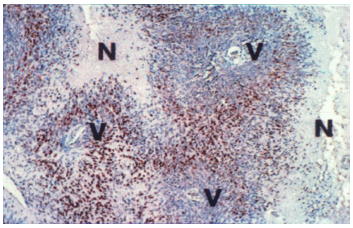
In dynamic models, the state is a dynamical variable (i.e. changes in time)

State evolution models encode mathematically the way that the state changes over time

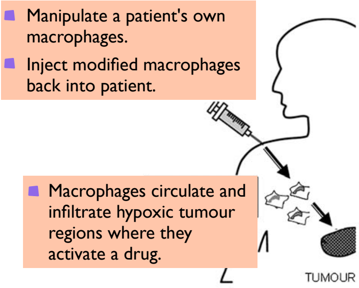
ODEs are based on the assumption that the state changes continuously, at a rate that depends only on the current state

Modelling cancer growth and therapy

- Lots of medical research into cancer, but mathematics is also playing its part.
- Hypoxic (nutrient starved) cancer cells are resistant to many therapies.
- Macrophages 'home-in' on hypoxic regions.
- Can we use macrophages to target hypoxic cancer cells?

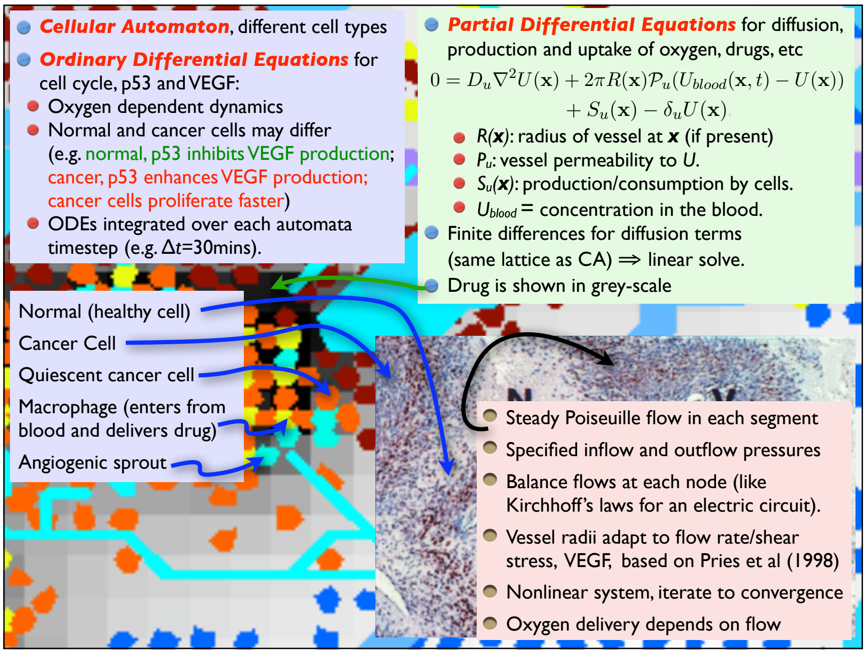


- V: vessel (plenty of food nearby)
- Brown: hungry cancer cells
- N: necrosis (cell death)



- Manipulate a patient's own macrophages.
- Inject modified macrophages back into patient.

- Macrophages circulate and infiltrate hypoxic tumour regions where they activate a drug.



- ***Cellular Automaton***, different cell types
- ***Ordinary Differential Equations*** for cell cycle, p53 and VEGF:
  - Oxygen dependent dynamics
  - Normal and cancer cells may differ (e.g. normal, p53 inhibits VEGF production; cancer, p53 enhances VEGF production; cancer cells proliferate faster)
  - ODEs integrated over each automata timestep (e.g. Δt=30mins).

- ***Partial Differential Equations*** for diffusion, production and uptake of oxygen, drugs, etc

$$0 = D_u\nabla^2U(\mathbf{x}) + 2\pi R(\mathbf{x})\mathcal{P}_u(U_{blood}(\mathbf{x},t) - U(\mathbf{x})) + S_u(\mathbf{x}) - \delta_u U(\mathbf{x})$$

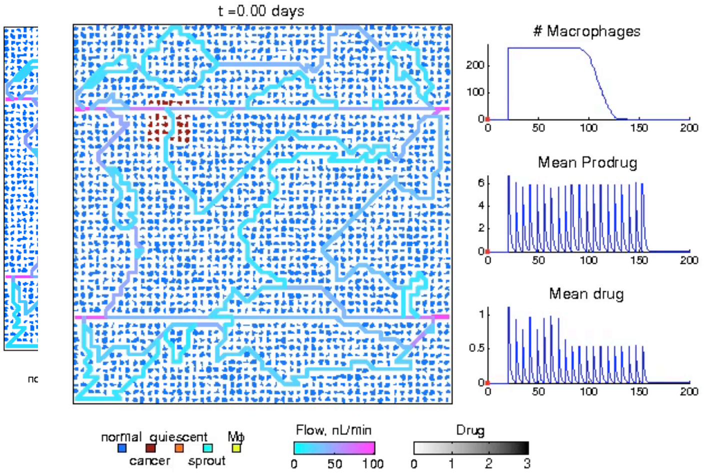
  - $R(\mathbf{x})$: radius of vessel at $\mathbf{x}$ (if present)
  - P_u: vessel permeability to U.
  - $S_u(\mathbf{x})$: production/consumption by cells.
  - U_{blood} = concentration in the blood.
- Finite differences for diffusion terms (same lattice as CA) ⟹ linear solve.
- Drug is shown in grey-scale

Normal (healthy cell)
Cancer Cell
Quiescent cancer cell
Macrophage (enters from blood and delivers drug)
Angiogenic sprout

- Steady Poiseuille flow in each segment
- Specified inflow and outflow pressures
- Balance flows at each node (like Kirchhoff's laws for an electric circuit).
- Vessel radii adapt to flow rate/shear stress, VEGF, based on Pries et al (1998)
- Nonlinear system, iterate to convergence
- Oxygen delivery depends on flow

Combined Macrophage & Conventional therapy

- What happens when both therapies have small individual efficacy?



.I

every 7
day 21.

rom
xic
d by
e liver.

Opportunities in Systems Biology

- Heart (electrical activity, muscle mechanics, blood flow)
- Lungs (air flow, asthma, ...)
- Brain (single neurons, whole brain, Parkinson's disease, ...)
- Developmental biology (how organisms grow)
- Cancer
- Immunology (how we fight infections, how it can go wrong - e.g. Rheumatoid Arthritis, HIV)
- Bacterial infections (managing infections in hospitals)
- Ecology (control of invasive weeds, management of fisheries, ...)
- Plants (how to improve food crops)
- ...

- **Maths:** Model development, model analysis, simulation, statistical analysis of data, ...
- **Biology:** High throughput techniques, genetics, proteomics, epigenetic regulation, ...
- **Computer Science:** Image analysis, algorithm and software development, data mining, optimisation, model sharing and markup languages, ...
- **Engineering, Physics, Chemistry:** Bioengineering, tissue engineering, nanotechnology, MRI, new microscopy and measurement techniques, ...

Success stories in Mathematical Biology

A.L. Hodgkin, A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117:500–544 (1952).

R.M. Anderson, G.F. Medley, R.M. May, A.M. Johnson: A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS.
Math. Med. Biol. 3:229–263 (1986).

J.J. Tyson, B. Novak. Regulation of the eukaryotic cell cycle: Molecular antagonism, hysteresis, and irreversible transitions. *J. Theor. Biol.* 210:249–263 (2001).

G. Dupont, A. Goldbeter. One-pool model for Ca2+ oscillations involving Ca2+ and inositol 1,4,5-trisphosphate as co-agonists for Ca2+ release. *Cell Calcium* 14:311–322 (1993).

P. Hahnfeldt, D. Panigrahy, J. Folkman, L. Hlatky. Tumor development under angiogenic signaling: A dynamical theory of tumor growth, treatment response, and postvascular dormancy.
Cancer Research 59:4770–4775 (1999).

A.M. Turing: The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond. B*, 237:37–72 (1952).

M. Mackey, L. Glass. Oscillation and chaos in physiological control systems. *Science* 197:287–289 (1977).

N. Barkai, S. Leibler. Robustness in simple biochemical networks. *Nature* 387:913–917 (1997).

A.D. Lander, Q. Nie, F.W.M. Wan. Do morphogen gradients arise by diffusion? *Dev. Cell.* 2:785–796 (2002).

... AND MANY MORE ...