



Celebrating **100000** Years of Model Checking

Parameter Identification using Model Checking for Systems Biology

Bing Liu

Ed's postdoc 2012-2013

Joint work with

P.S. Thiagarajan's group and Ed Clarke's group



**University of Pittsburgh
School of Medicine**

Parameter Identification using Model Checking for Systems Biology

Bing Liu

Ed's postdoc 2012-2013

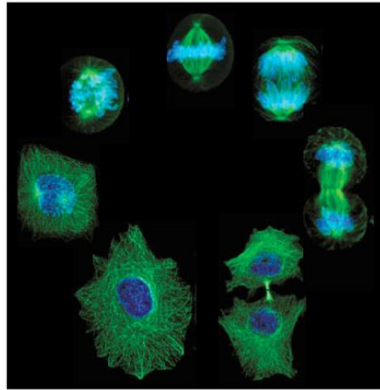
Joint work with

P.S. Thiagarajan's group and Ed Clarke's group

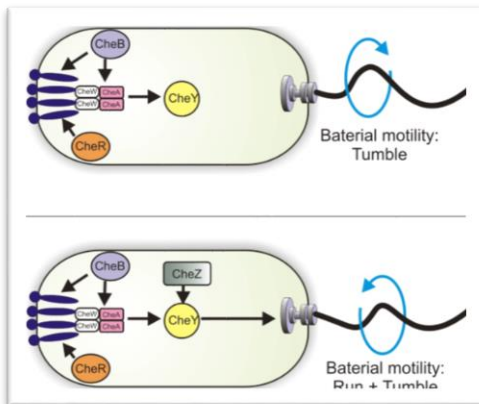


**University of Pittsburgh
School of Medicine**

Biological Machines



Cell cycle



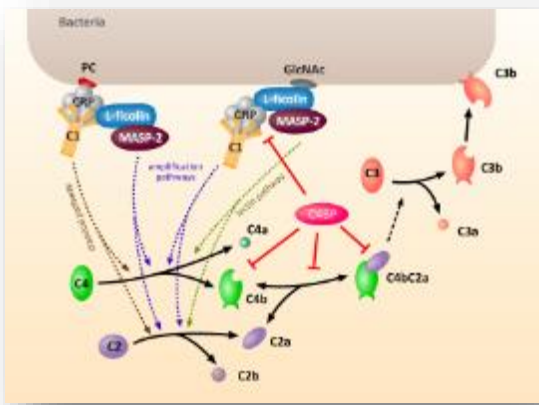
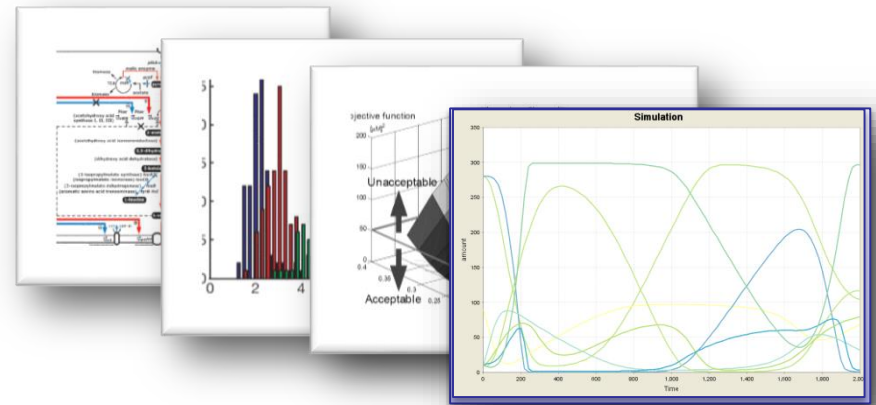
E. coli chemotaxis



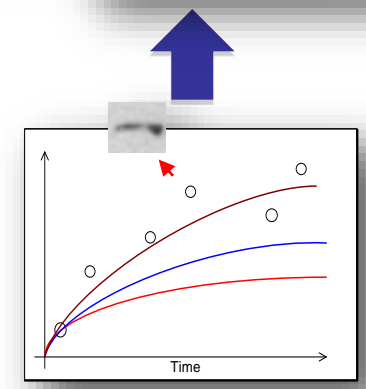
Microtubule assembly, vesicle transport driven by motor proteins, protein synthesis by ribosome, power station mitochondria

Computational Modeling

- Build a model
 - Formalisms:
 - ODE, Boolean, rule-based, hybrid automata ...
 - Calibration & Validation
- Perform analysis



$$\begin{aligned}
 \frac{dc_1}{dt} &= -k_1 \cdot c_1 \cdot c_2 + k_2 \cdot c_3 \\
 \frac{dc_2}{dt} &= -k_1 \cdot c_1 \cdot c_2 + k_2 \cdot c_3 + k_{17} \cdot c_{18} + k_{11} \cdot c_{11} \\
 \frac{dc_3}{dt} &= k_1 \cdot c_1 \cdot c_2 - k_2 \cdot c_3 - k_3 \cdot c_3 \cdot c_4 + k_4 \cdot c_5 \\
 \frac{dc_4}{dt} &= -k_3 \cdot c_3 \cdot c_4 + k_4 \cdot c_5 + k_{11} \cdot c_{11} + k_{20} \cdot c_{21} \\
 \frac{dc_5}{dt} &= k_3 \cdot c_3 \cdot c_4 - k_4 \cdot c_5 - k_5 \cdot c_5 \cdot c_6 + k_6 \cdot c_7 \\
 \frac{dc_6}{dt} &= -k_5 \cdot c_5 \cdot c_6 + k_6 \cdot c_7 + k_{11} \cdot c_{11} + k_{20} \cdot c_{21} \\
 \frac{dc_7}{dt} &= k_5 \cdot c_5 \cdot c_6 - k_6 \cdot c_7 - k_7 \cdot c_7 \cdot c_8 + k_8 \cdot c_9 \\
 \frac{dc_8}{dt} &= -k_7 \cdot c_7 \cdot c_8 + k_8 \cdot c_9 + k_{11} \cdot c_{11} + k_{20} \cdot c_{21} \\
 \frac{dc_9}{dt} &= k_7 \cdot c_7 \cdot c_8 - k_8 \cdot c_9 - k_9 \cdot c_9 \cdot c_{10} + k_{10} \cdot c_{11} \\
 &\quad - k_{15} \cdot c_9 \cdot c_{17} + k_{16} \cdot c_{18}
 \end{aligned}$$



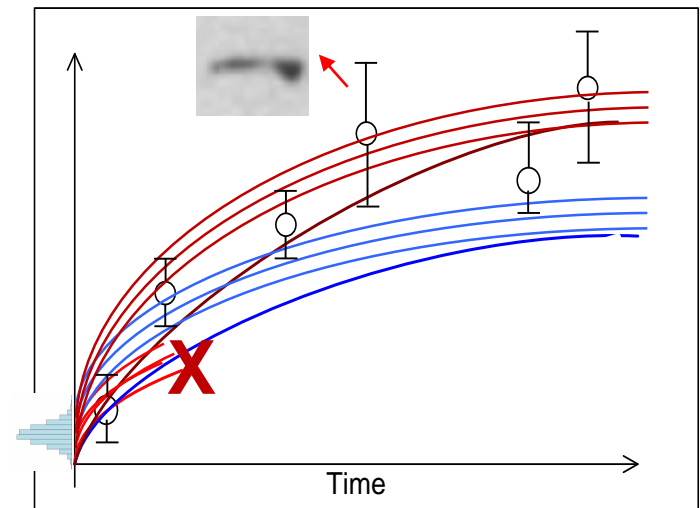
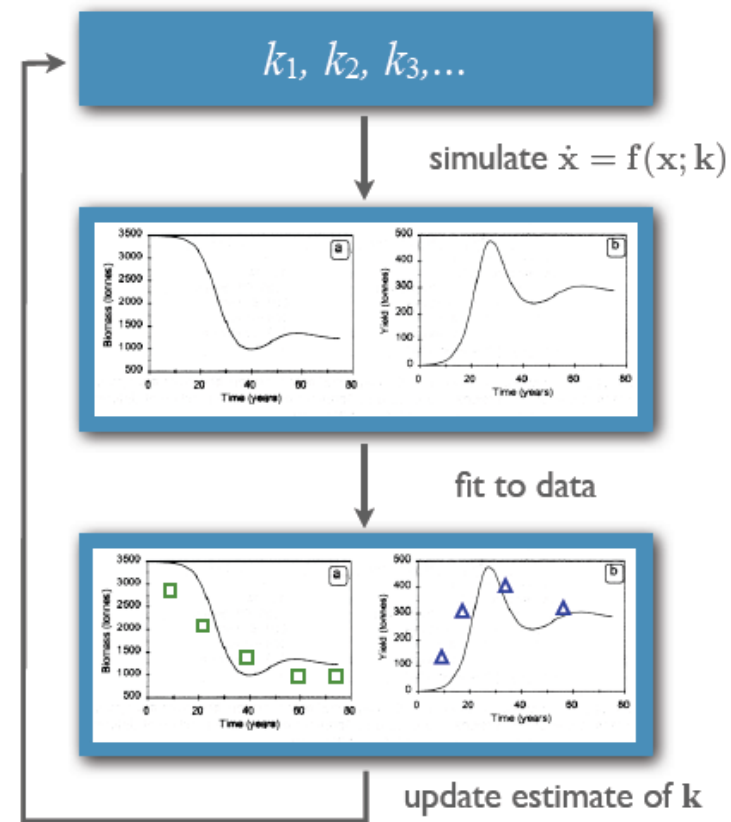
Problems Faced

- Which model structure is more plausible?
- How to estimate unknown model parameters?
- Which reactions are critical to the dynamics?
- Which conditions will lead to desired states?
- How to control the system to avoid bad state?

- Parameter identification is hard
 - Curse of dimensionality, noisy data,
- **Model checking can help!**

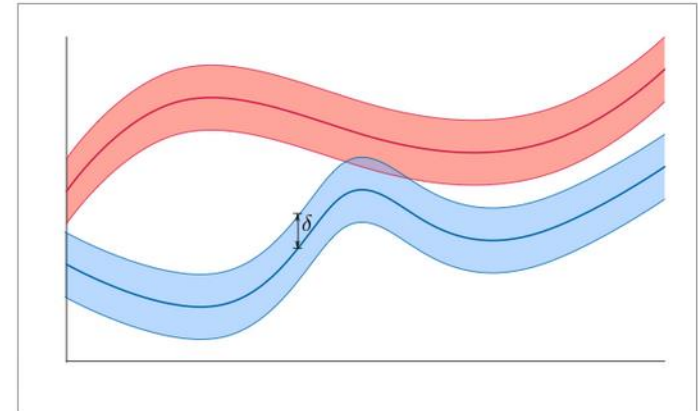
Our techniques

- Statistical model checking (SMC) based parameter estimation and global sensitivity analysis
 - Encode training data as a **BLTL** formula
 - Evaluate candidate parameters using SMC
 - Perform global optimization (e.g. SRES)
- Advantage:
 - Utilize both **quantitative** and **qualitative** knowledge
 - Deal with uncertainty of the initial states and the noisy cell-population data
 - Good scalability due to the power of statistical testing



Our techniques

- Parameter synthesis for nonlinear hybrid systems using delta-decisions
 - Encode a parameter synthesis problem as a first-order formula over the reals
 - Perform bounded model checking
 - Employ an interval constraint propagation (ICP) based algorithm to identify the resulting parameters

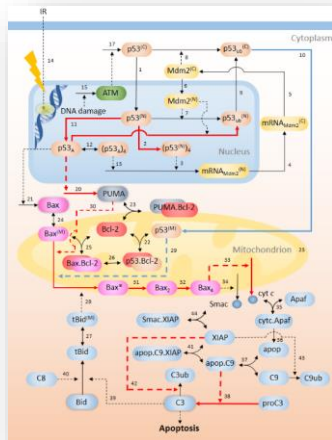


δ -Weakening of ϕ

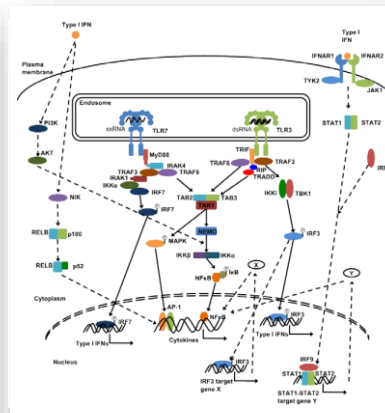
$$\phi^\delta := \exists \mathbf{x} \bigwedge_{i=1}^m \bigvee_{j=1}^k |f_{ij}(\mathbf{x})| \leq \delta$$

Applications

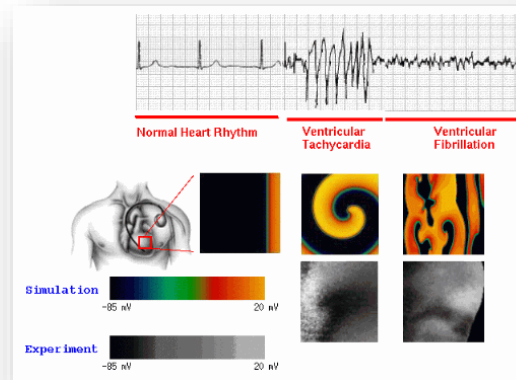
- Validate the dynamical behaviors of a p53 pathway model
- Estimate parameters for a Toll-like receptor pathway model
- Identify cardiac diseases-related parameter ranges
- Design personalized therapy scheme for prostate cancer



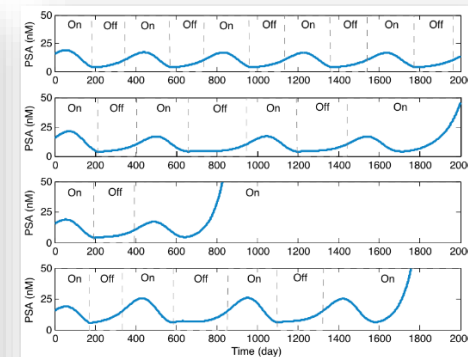
p53



TLR

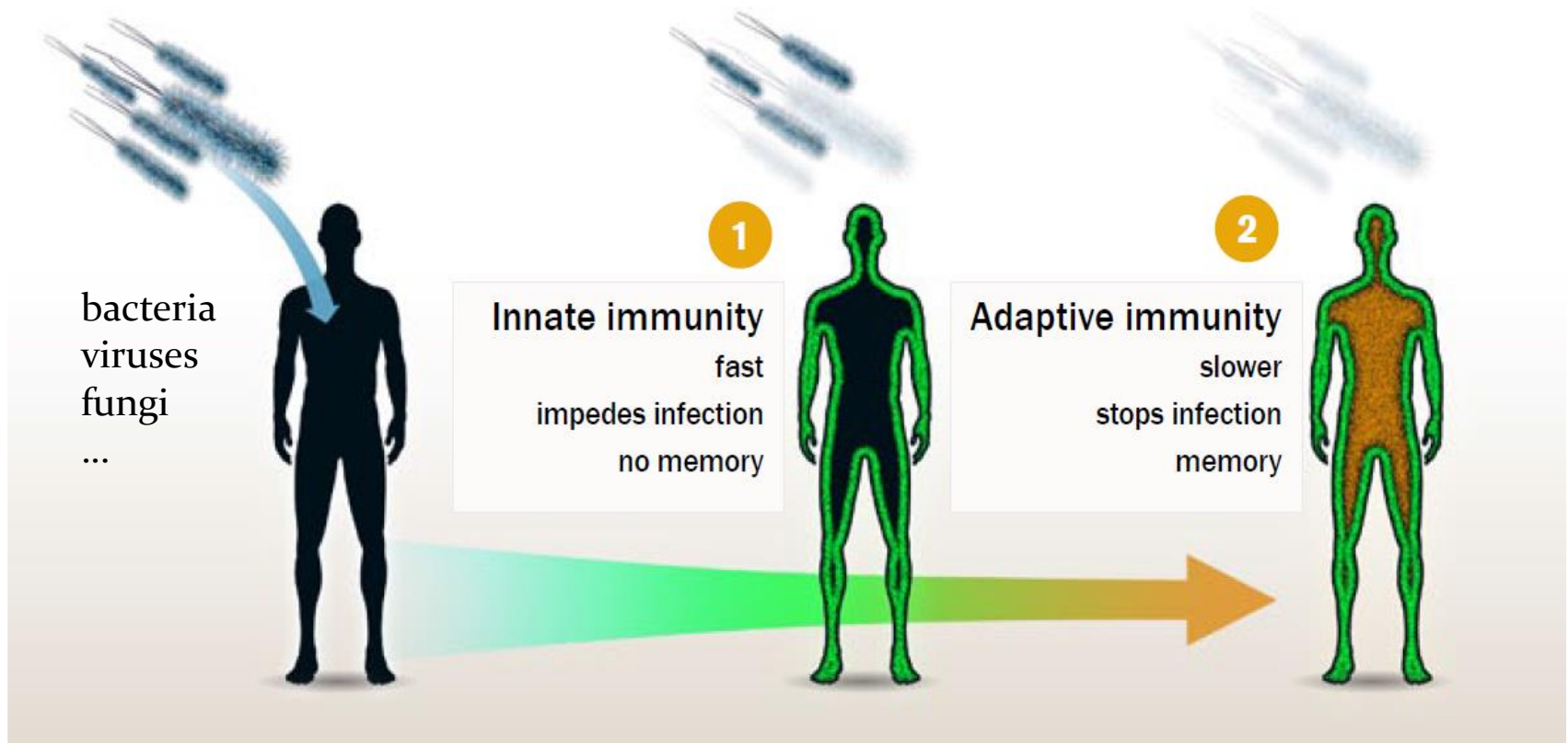


Cardiac cell



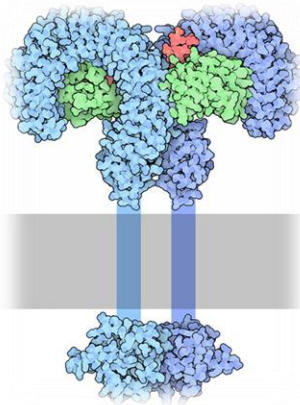
Prostate cancer

Immune System – two lines of defense

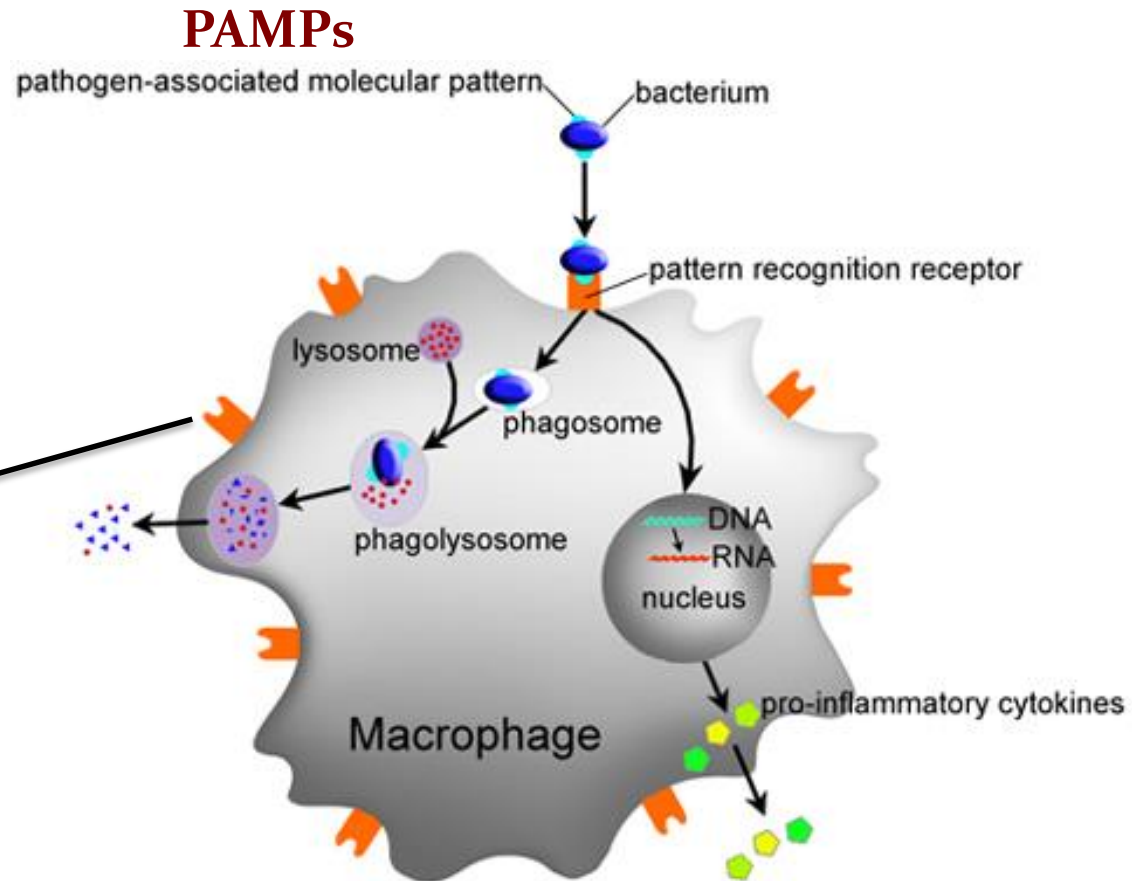


Innate Immune Response

- Recognize pathogens
- Recruit immune cells
- Activate adaptive immune system



Toll-like Receptors (TLR)



Toll-like receptors

- Why *'Toll-like'* ?



Toll ?

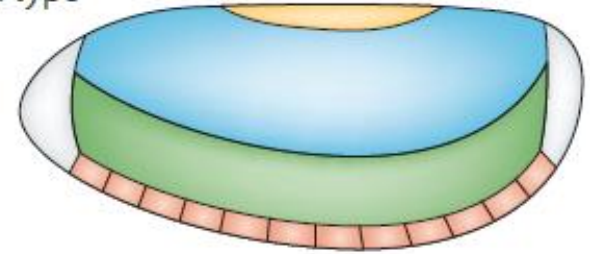
Toll-like Receptors

- *Toll* gene

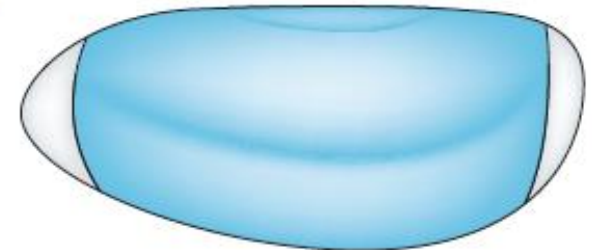


Christiane Nüsslein-Volhard, Nobel Prize 1995

Wild type



Toll⁻



Mesoderm
Amnioserosa

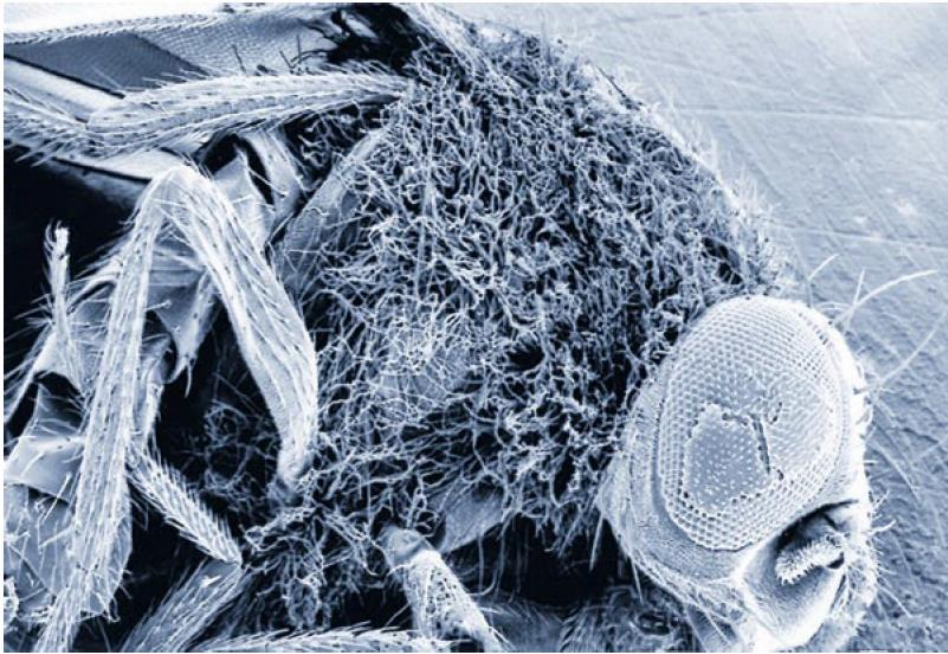
Dorsal
ectoderm

Neurogenic
ectoderm

Sommer, *Nat Rev Genetics*, 2009

Toll-like Receptors

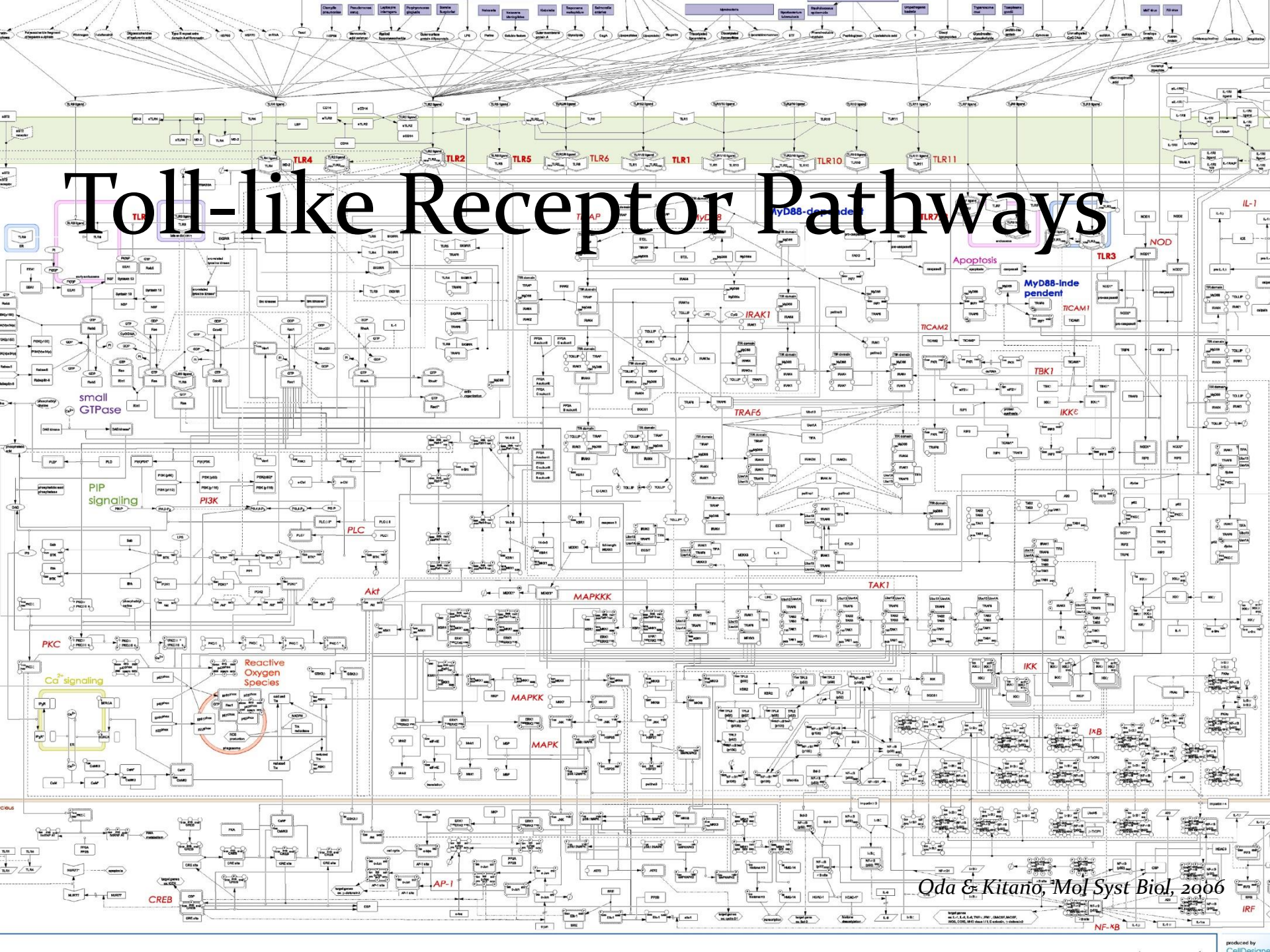
- *Toll* gene



Jules A. Hoffmann, Nobel Prize 2011

Fruit flies with mutant *Toll* cannot fight off fungal infection

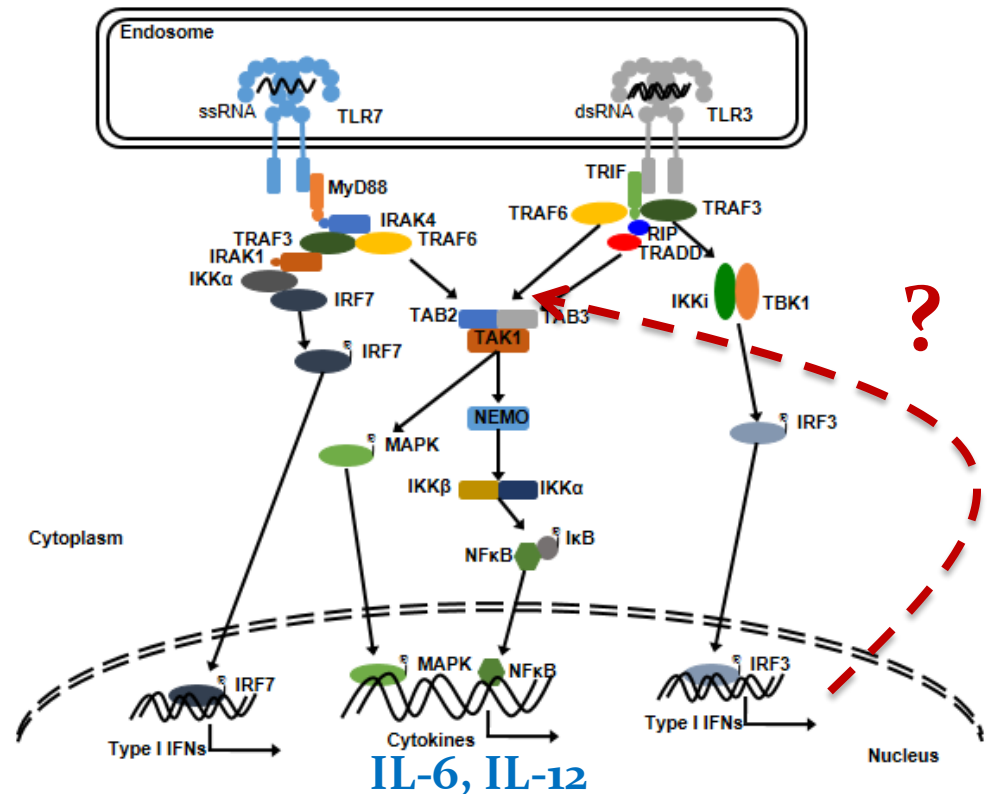
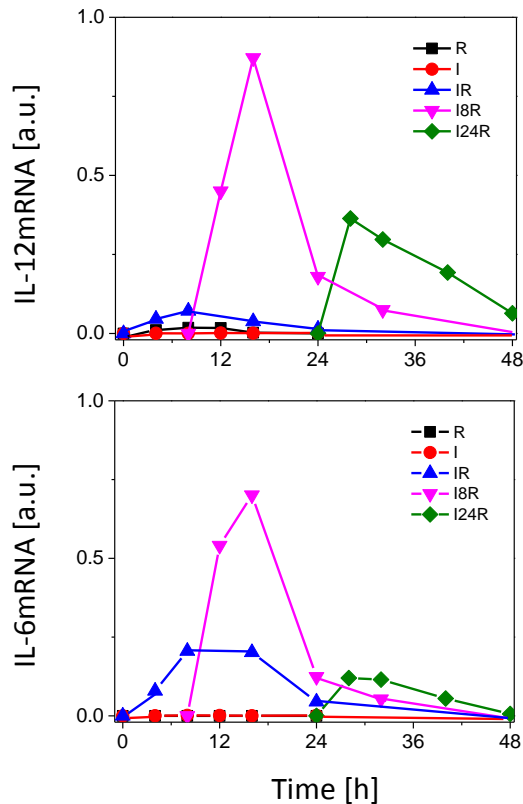
Toll-like Receptor Pathways



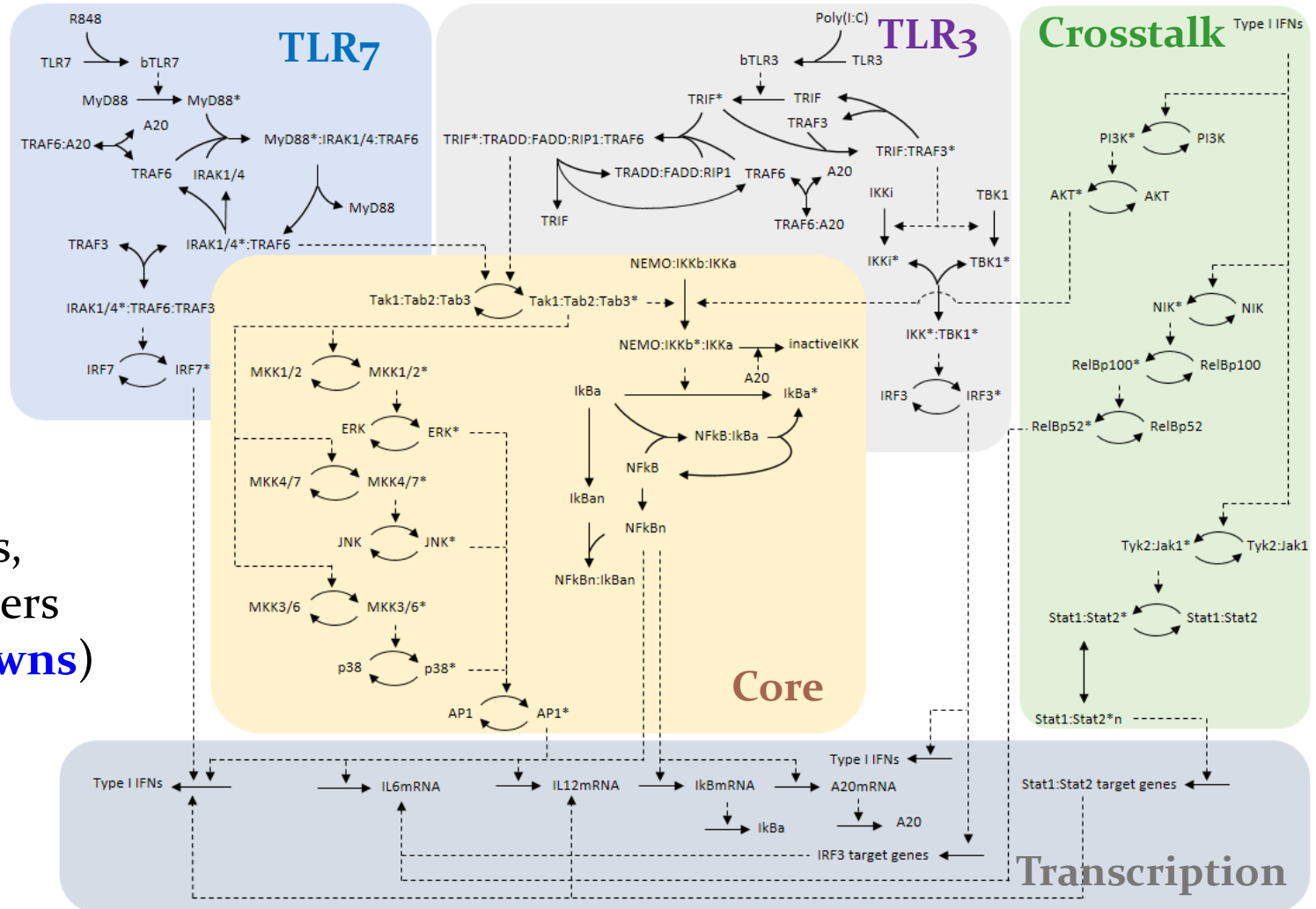
Oda & Kitano, Mol Syst Biol, 2006

Modeling TLR3-TLR7 pathways

- Key experimental observation: **a time-dependent synergy**
- How macrophage create such '*innate immune memory*'?

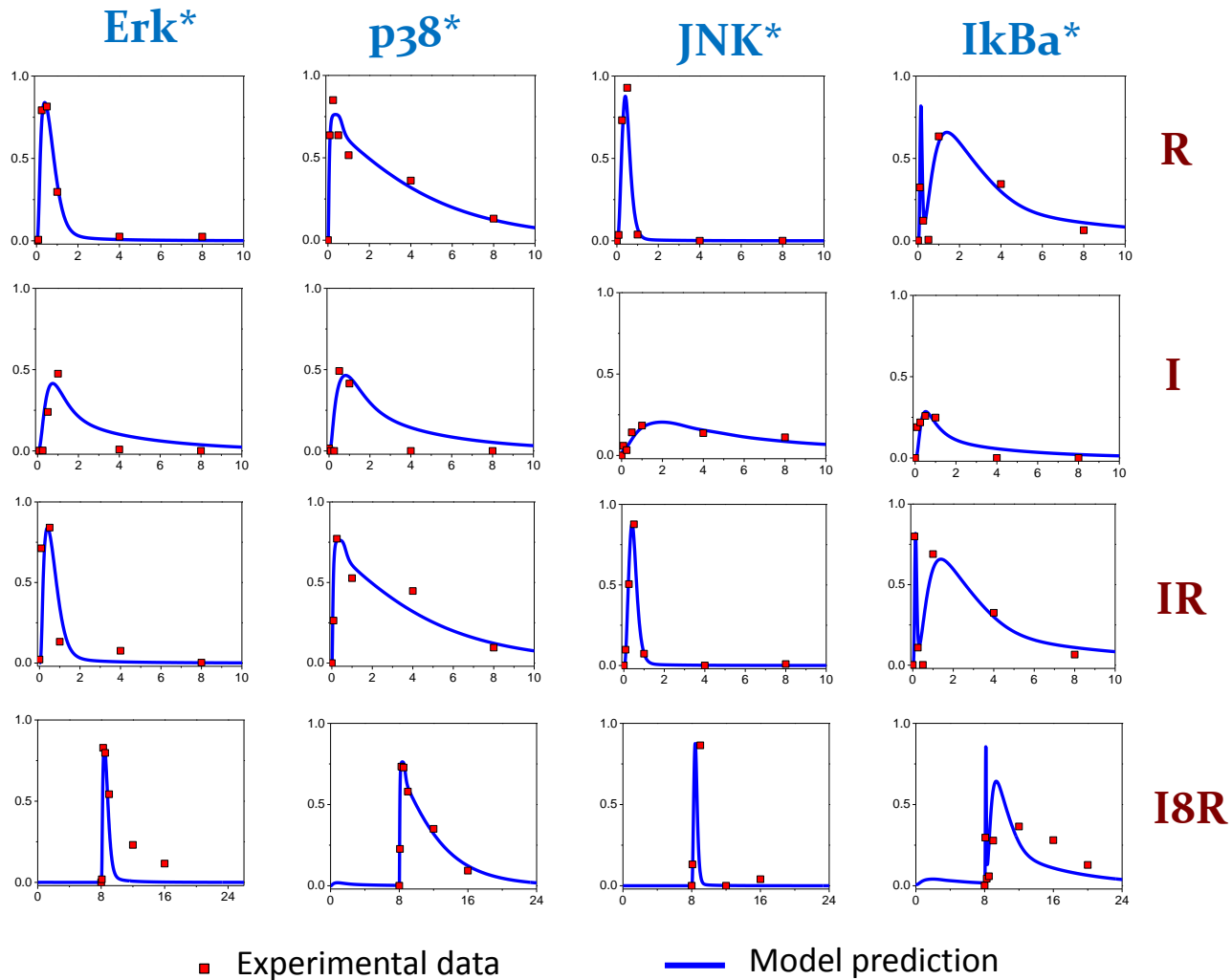


Our ODE based Model

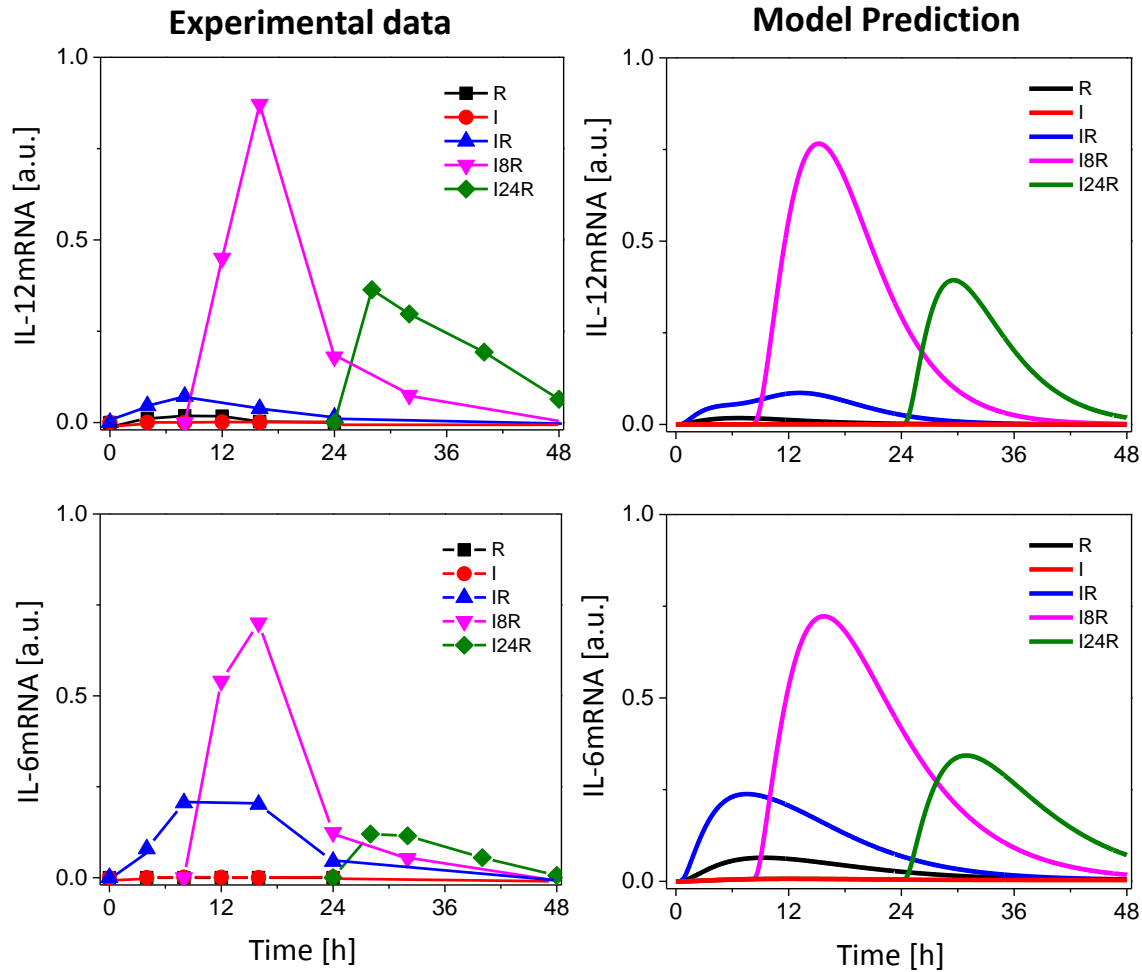


91 species
 114 reactions,
 139 parameters
 (112 unknowns)

Estimate parameters using SMC

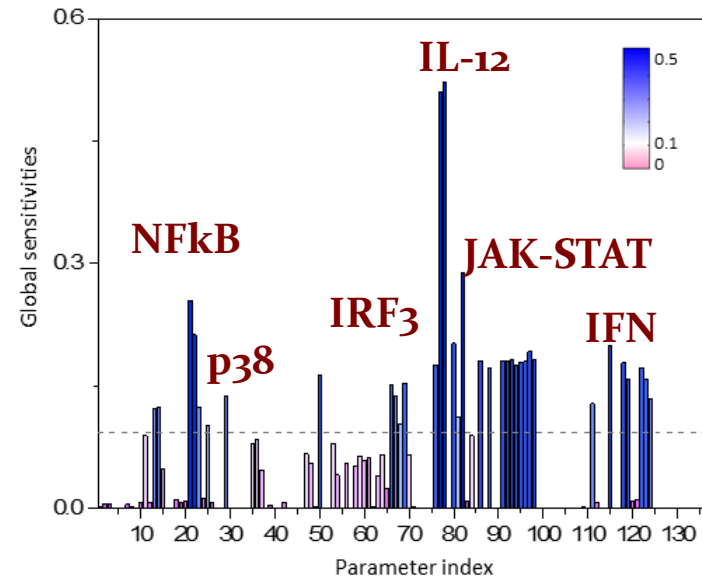


Model Prediction vs. Experimental Data



Identify critical reactions using SMC

- Our combined computational and experimental study reveals a STAT1-mediated synergistic mechanism

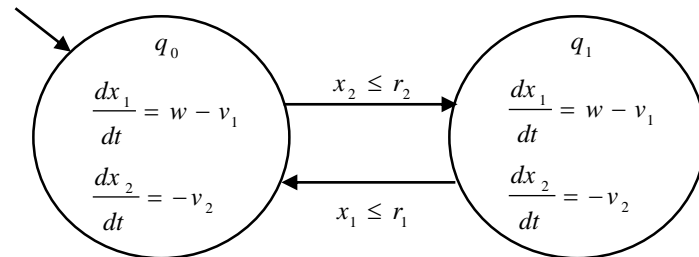
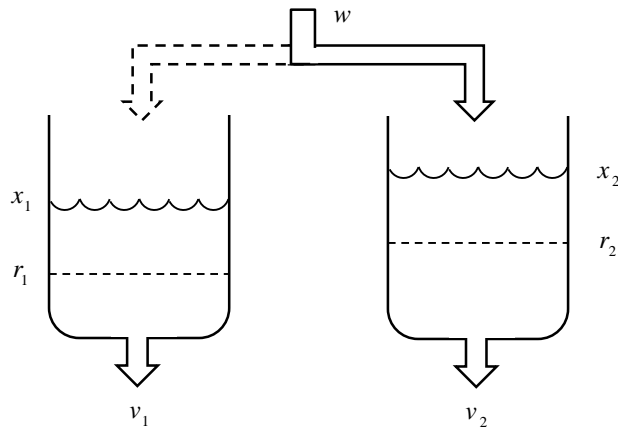


Sensitive kinetic parameters	Function
$k_{77}, k_{78}, k_{115}, k_{80}, k_{76}, k_{111}$	IL12mRNA transcription/degradation
$k_{82}, k_{21}, k_{22}, k_{50}, k_{81}$	NEMO:IKKb:IKKa activation/degradation
$k_{91}, k_{92}, k_{93}, k_{94}, k_{95}, k_{96}, k_{97}, k_{98}$	JAK-STAT signaling cascade
$k_{86}, k_{88}, k_{119}, k_{122}, k_{123}$	Type I INF transcription
$k_{29}, k_{13}, k_{69}, k_{11}, k_{23}, k_{25}, k_{84}$	p38 signaling cascade
k_{66}, k_{67}, k_{69}	IRF3 activation
k_{68}	TLR3 degradation

Hybrid Systems

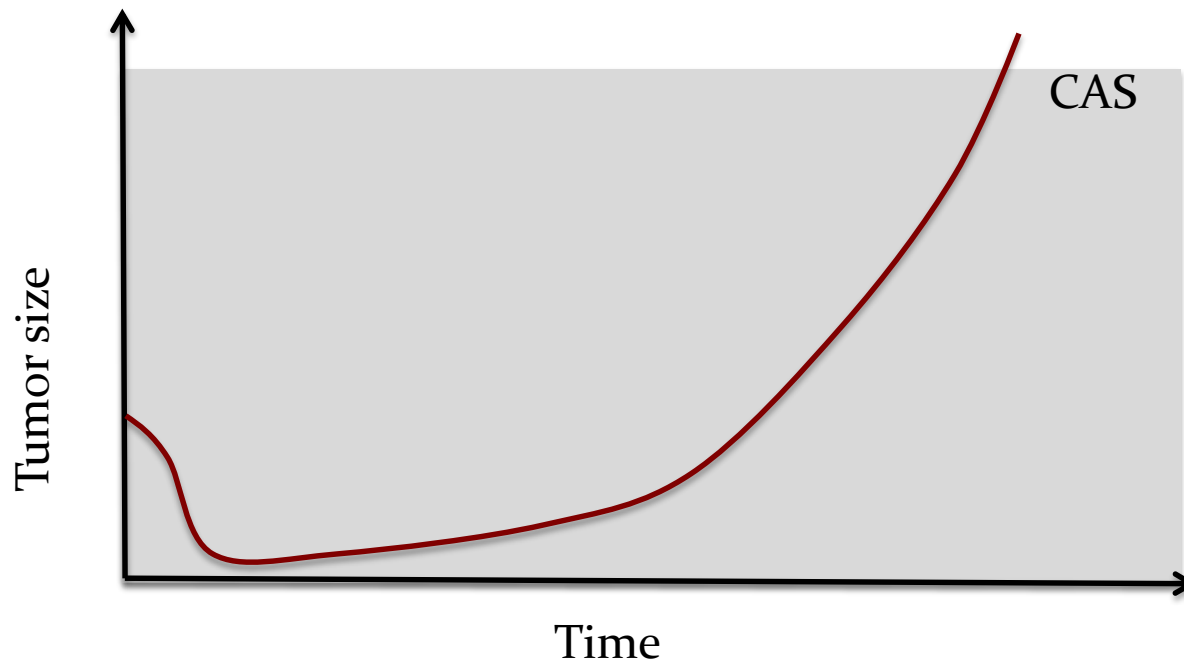
- Evolve in a **continuous** way in each mode
- Ruled by **discrete** transitions
- Hybrid automata

$$H = \langle X, Q, \text{Init}, \text{Flow}, \text{Jump} \rangle$$



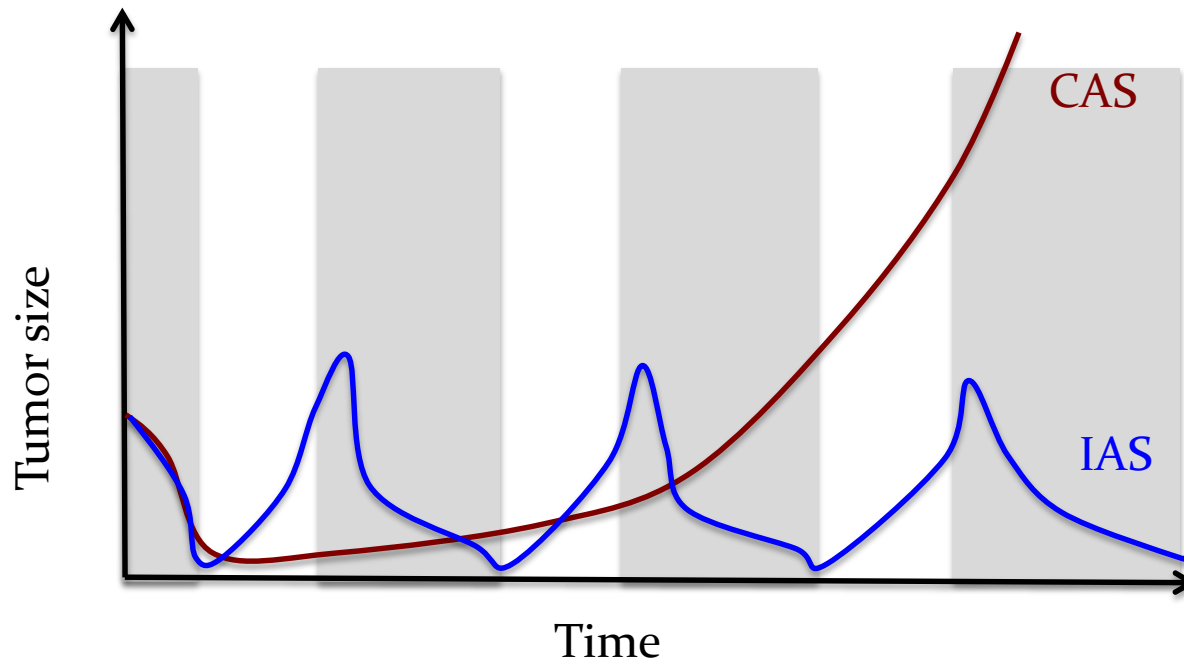
Hormone therapy for prostate cancer

- Continuous Androgen Suppression
 - Side effects: anemia, osteoporosis, impotence, etc.
 - Relapse after a median duration of 18-24 months, due to the proliferation of androgen independent (AI) cancer cells.



Intermittent Androgen Suppression

- Reduce side effects
- May delay the time to relapse
 - Avoid emergence of AI cells
- The efficacy varies for individual patients



Model

- Population of AD cells, AI cells, serum androgen concentration, PSA level (Ideta et al. 2008)

jump $_{2 \rightarrow 1}$:

$$x + y \geq r_1 \wedge \frac{dx}{dt} + \frac{dy}{dt} > 0$$

flow $_1$:

Mode 1 (on-treatment)

$$\frac{dx}{dt} = \left(\alpha_x \left(k_1 + \frac{(1-k_1)z}{z+k_2} \right) - \beta_x \left(k_3 + \frac{(1-k_3)z}{z+k_4} \right) - m_1 \left(1 - \frac{z}{z_0} \right) \right) x$$

$$\frac{dy}{dt} = m_1 \left(1 - \frac{z}{z_0} \right) x + \left(\alpha_y \left(1 - d \frac{z}{z_0} \right) - \beta_y \right) y$$

$$\frac{dz}{dt} = \frac{-z}{\tau}$$

$$\frac{dv}{dt} = \left(\alpha_x \left(k_1 + \frac{(1-k_1)z}{z+k_2} \right) - \beta_x \left(k_3 + \frac{(1-k_3)z}{z+k_4} \right) - m_1 \left(1 - \frac{z}{z_0} \right) \right) x$$

$$+ m_1 \left(1 - \frac{z}{z_0} \right) x + \left(\alpha_y \left(1 - d \frac{z}{z_0} \right) - \beta_y \right) y$$

flow $_2$:

Mode 2 (off-treatment)

$$\frac{dx}{dt} = \left(\alpha_x \left(k_1 + \frac{(1-k_1)z}{z+k_2} \right) - \beta_x \left(k_3 + \frac{(1-k_3)z}{z+k_4} \right) - m_1 \left(1 - \frac{z}{z_0} \right) \right) x$$

$$\frac{dy}{dt} = m_1 \left(1 - \frac{z}{z_0} \right) x + \left(\alpha_y \left(1 - d \frac{z}{z_0} \right) - \beta_y \right) y$$

$$\frac{dz}{dt} = \frac{z_0 - z}{\tau}$$

$$\frac{dv}{dt} = \left(\alpha_x \left(k_1 + \frac{(1-k_1)z}{z+k_2} \right) - \beta_x \left(k_3 + \frac{(1-k_3)z}{z+k_4} \right) - m_1 \left(1 - \frac{z}{z_0} \right) \right) x$$

$$+ m_1 \left(1 - \frac{z}{z_0} \right) x + \left(\alpha_y \left(1 - d \frac{z}{z_0} \right) - \beta_y \right) y$$

jump $_{1 \rightarrow 2}$:

$$x + y \leq r_0 \wedge \frac{dx}{dt} + \frac{dy}{dt} < 0$$

Personalized therapy design using dReach

- Personalized parameters (e.g. the mutation rate from AD to AI cells)
- Estimate personalized parameters by fitting clinical PSA data
- Given r_o in $[0,8)$ and r_i in $[8,15]$, verify if invariant $v \leq 30$ holds (i.e. no relapse)
 - **False**: androgen suppression does not work
 - **True**: feasible values for r_o and r_i will be returned.

Patient	α_y	β_y	m_i	$z(o)$	Suggested scheme
10	0.019	0.009	5.9e-5	9.44	$r_o = 4.1$ and $r_i = 9.4$
45	0.012	0.041	1.0e-5	12.61	$r_o = 3.8$ and $r_i = 12.2$
97	0.031	0.015	2.3e-5	10.61	relapse

Conclusion

- Model checking can be used to tackle biological problems
- Model checking groups become active in systems biology domain
- Model checking will contribute to more fields, e.g. synthetic biology

