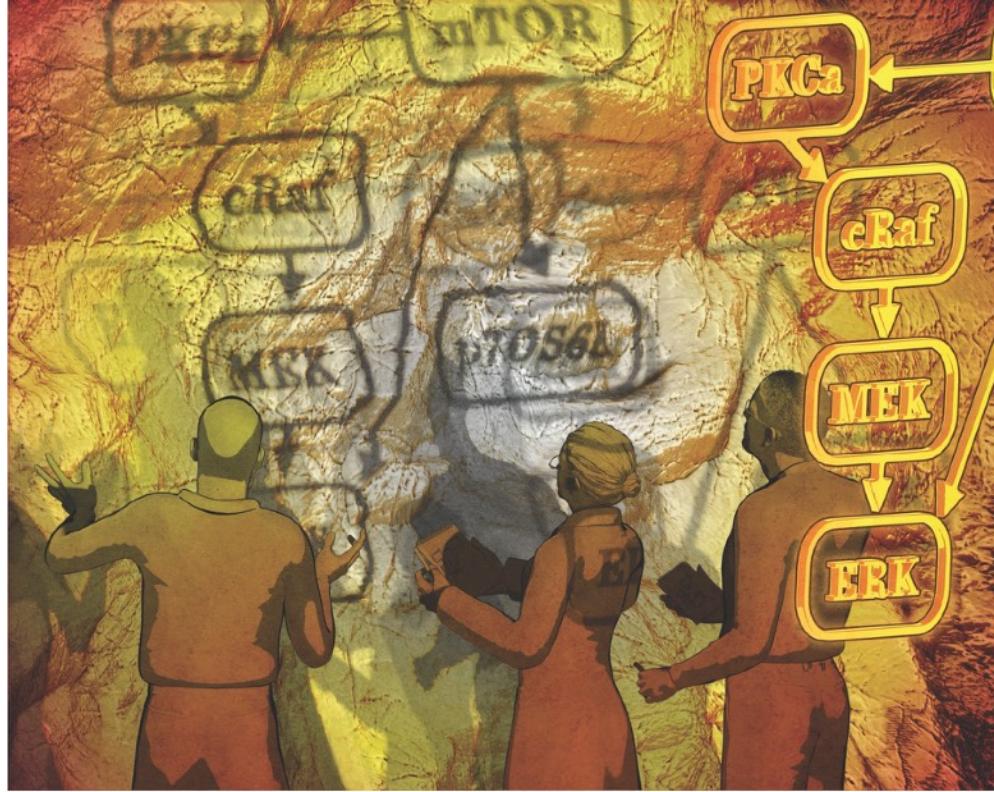


# Logic models of signalling networks & their training to experimental data with CellNOpt



Artwork by S. Philips on idea of J. Saez-Rodriguez; appeared in cover of *Nat Meth*, 13:4, 2016

**Julio Saez-Rodriguez**

Institute for Computational Biomedicine,  
University Hospital Heidelberg

Joint Research Centre for Computational  
Biomedicine - RWTH Aachen, Germany



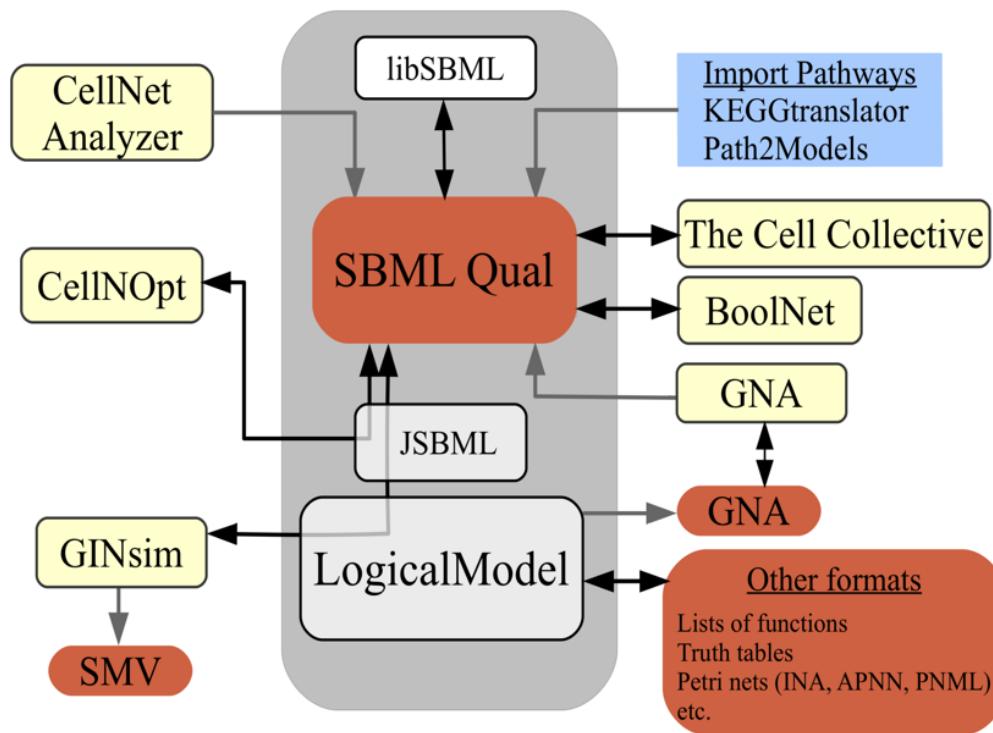
European Bioinformatics Institute  
European Molecular Biology Laboratory  
Hinxton, UK

[www.saezlab.org](http://www.saezlab.org)  
 [@sysbiomed](https://twitter.com/sysbiomed)



# CellNOpt in CoLoMoTo ecosystem: fit knowledge to data to build logic model

Complementary to other tools - models can be output to other tools



Consortium for Logical Models and Tools  
(CoLoMoTo; Naldi et al, *Bioinformatics*, 2015; [www.colomoto.org](http://www.colomoto.org))  
Exchange via SBML-qual (Chaouiya et al, *BMC Sys Bio*, 2013)



# CellNOpt: Building logic models by training signalling networks to perturbation data

freely available at [www.cellnopt.org](http://www.cellnopt.org)

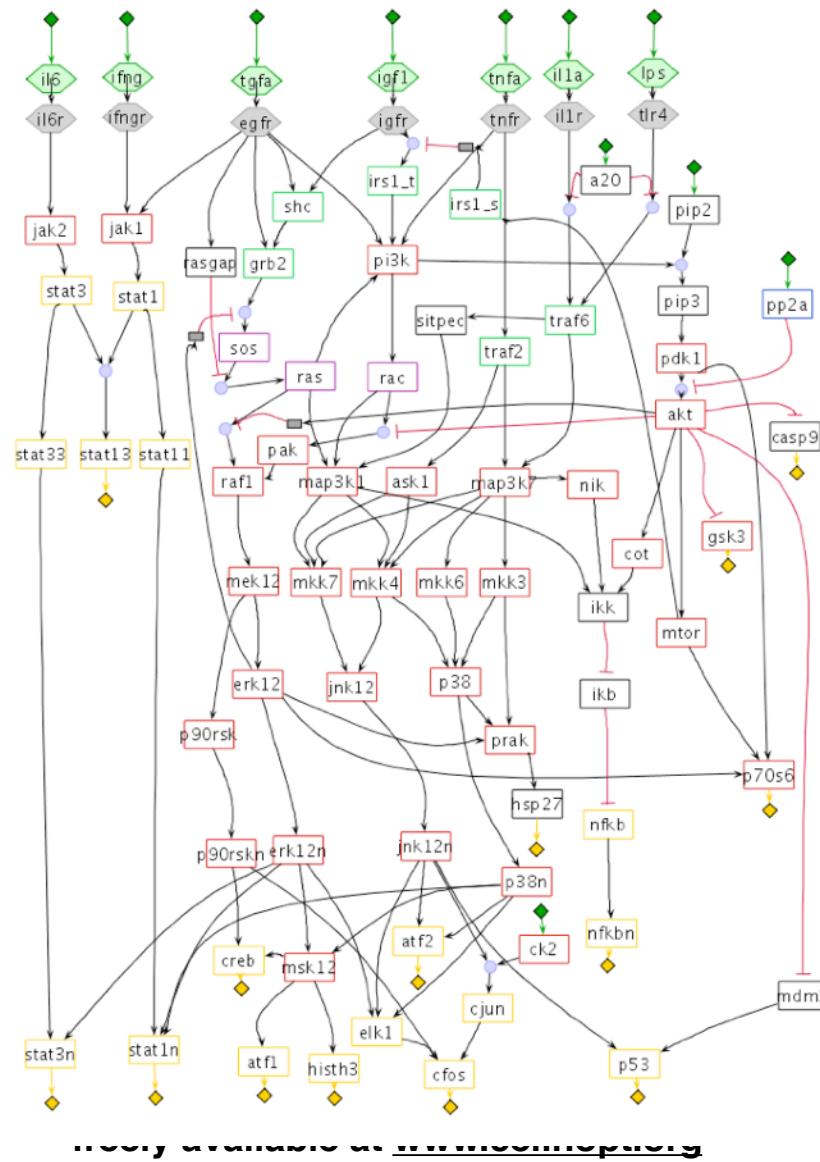
Saez-Rodriguez, et al. *Mol. Syst. Biol.*, 2009  
Eduati et al. *J Bioinformatics*, 2012

Terfve et al *BMC Sys Bio*, 2012  
MacNamara et al. *Phys Biol*, 2012

Morris et al PLoS CB 2011  
Traynard et al., *CPT:PSP*, 2017



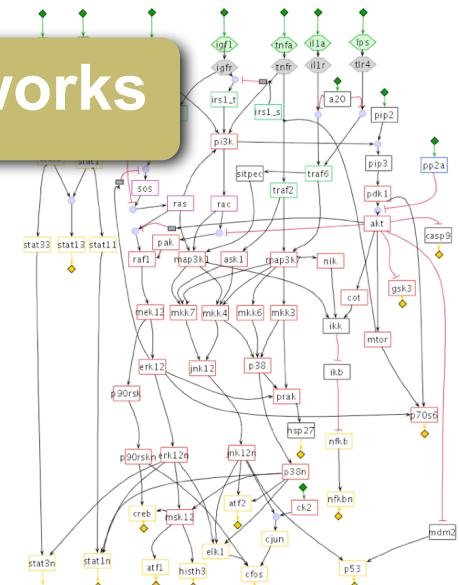
# CellNOpt: Building logic models by training signalling networks to perturbation data





# CellNOpt: Building logic models by training signalling networks to perturbation data

## Networks



freely available at [www.cellnopt.org](http://www.cellnopt.org)

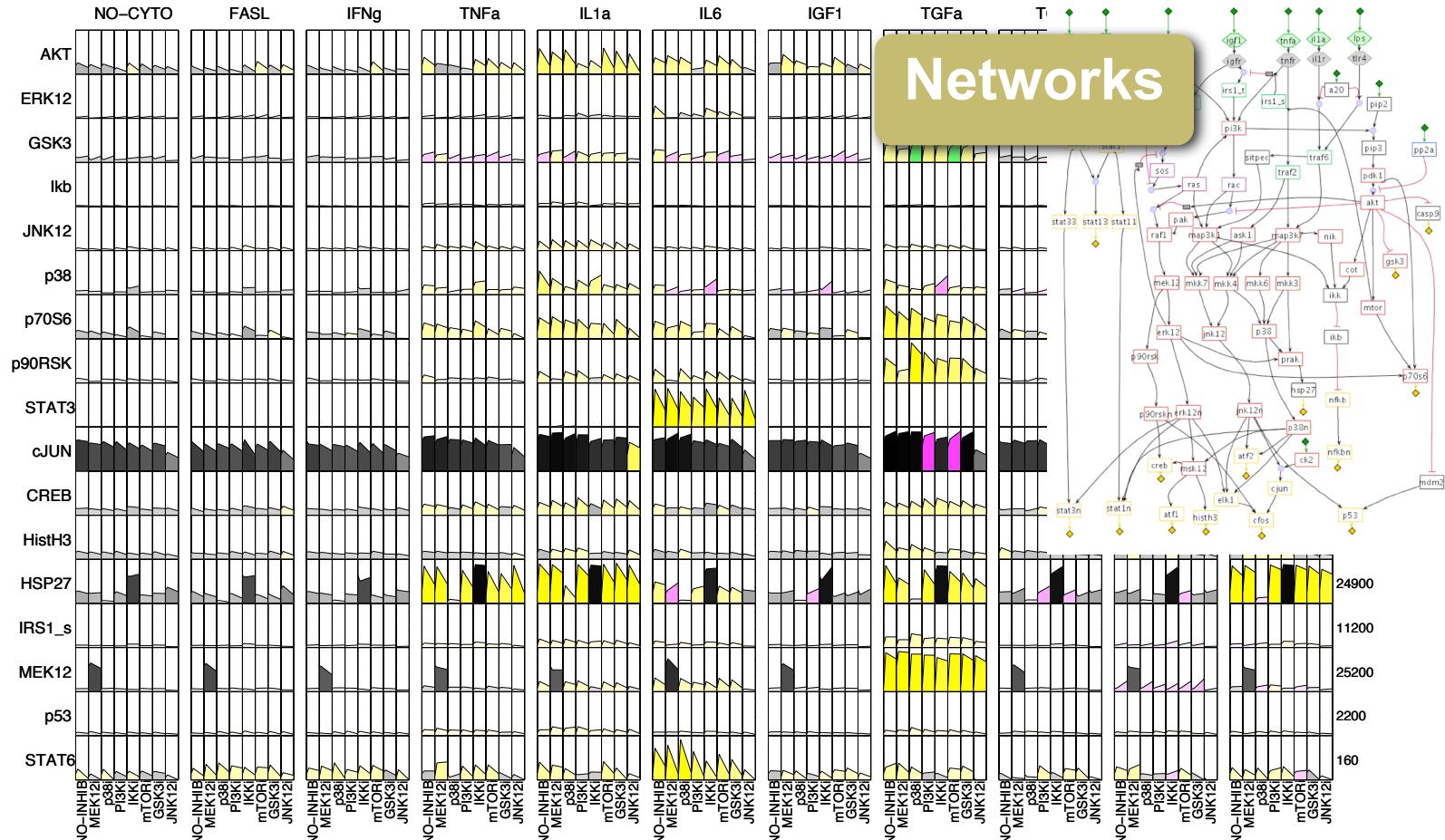
Saez-Rodriguez, et al. *Mol. Syst. Biol.*, 2009  
Eduati et al. *J Bioinformatics*, 2012

Terfve et al *BMC Sys Bio*, 2012  
MacNamara et al. *Phys Biol*, 2012

Morris et al *PLoS CB* 2011  
Traynard et al., *CPT:PSP*, 2017



# CellNOpt: Building logic models by training signalling networks to perturbation data



freely available at [www.cellnopt.org](http://www.cellnopt.org)

Saez-Rodriguez, et al. *Mol. Syst. Biol.*, 2009  
Eduati et al. *J Bioinformatics*, 2012

Terfve et al *BMC Sys Bio*, 2012  
MacNamara et al. *Phys Biol*, 2012

Morris et al *PLoS CB* 2011  
Traynard et al., *CPT:PSP*, 2017

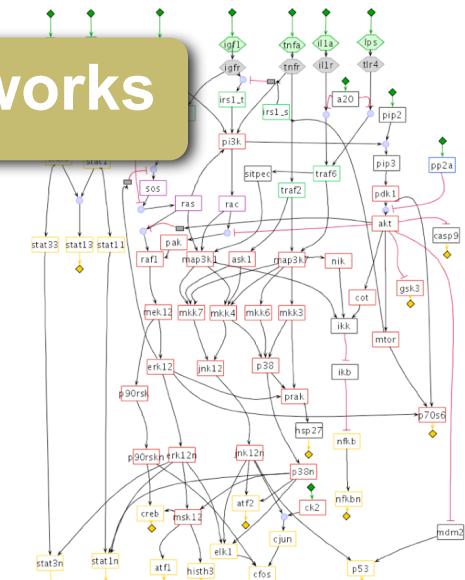


# CellNOpt: Building logic models by training signalling networks to perturbation data



Data

Networks



freely available at [www.cellnopt.org](http://www.cellnopt.org)

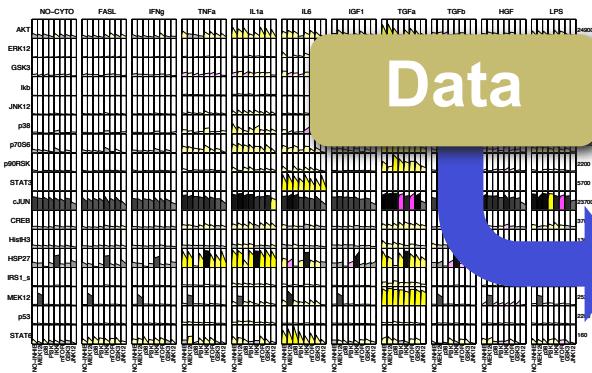
Saez-Rodriguez, et al. *Mol. Syst. Biol.*, 2009  
Eduati et al. *J Bioinformatics*, 2012

Terfve et al *BMC Sys Bio*, 2012  
MacNamara et al. *Phys Biol*, 2012

Morris et al *PLoS CB* 2011  
Traynard et al., *CPT:PSP*, 2017



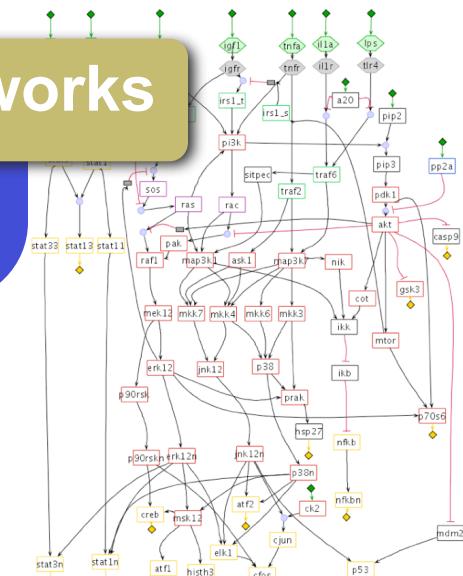
# CellNOpt: Building logic models by training signalling networks to perturbation data



Data

CellNOpt

Networks



freely available at [www.cellnopt.org](http://www.cellnopt.org)

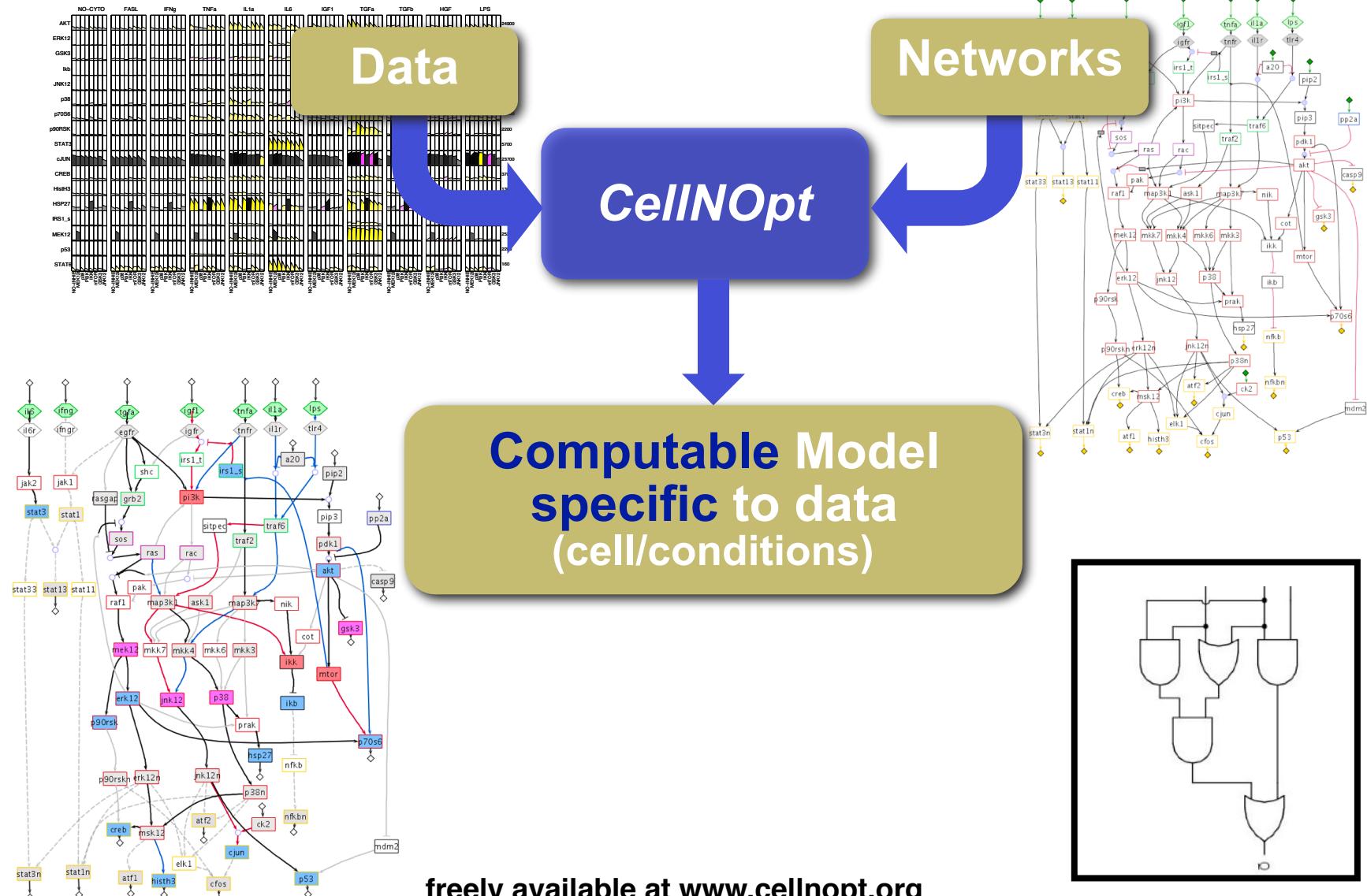
Saez-Rodriguez, et al. *Mol. Syst. Biol.*, 2009  
Eduati et al. *J Bioinformatics*, 2012

Terfve et al *BMC Sys Bio*, 2012  
MacNamara et al. *Phys Biol*, 2012

Morris et al *PLoS CB* 2011  
Traynard et al., *CPT:PSP*, 2017



# CellNOpt: Building logic models by training signalling networks to perturbation data



freely available at [www.cellnopt.org](http://www.cellnopt.org)

Saez-Rodriguez, et al. *Mol. Syst. Biol.*, 2009  
Eduati et al. *J Bioinformatics*, 2012

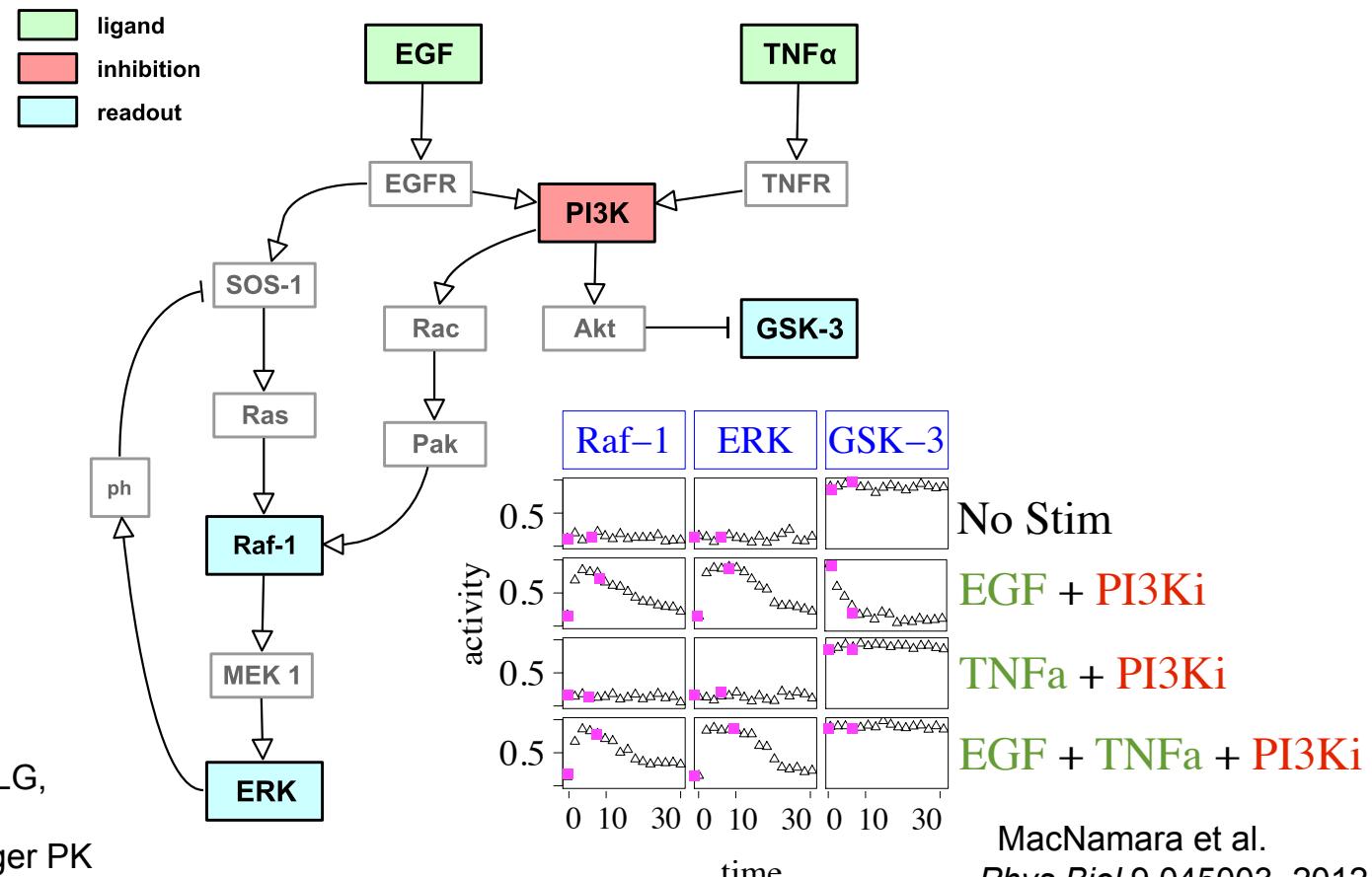
Terfve et al *BMC Sys Bio*, 2012  
MacNamara et al. *Phys Biol*, 2012

Morris et al *PLoS CB* 2011  
Traynard et al., *CPT:PSP*, 2017



# Logic modelling to link protein signalling networks with functional analysis of signal transduction

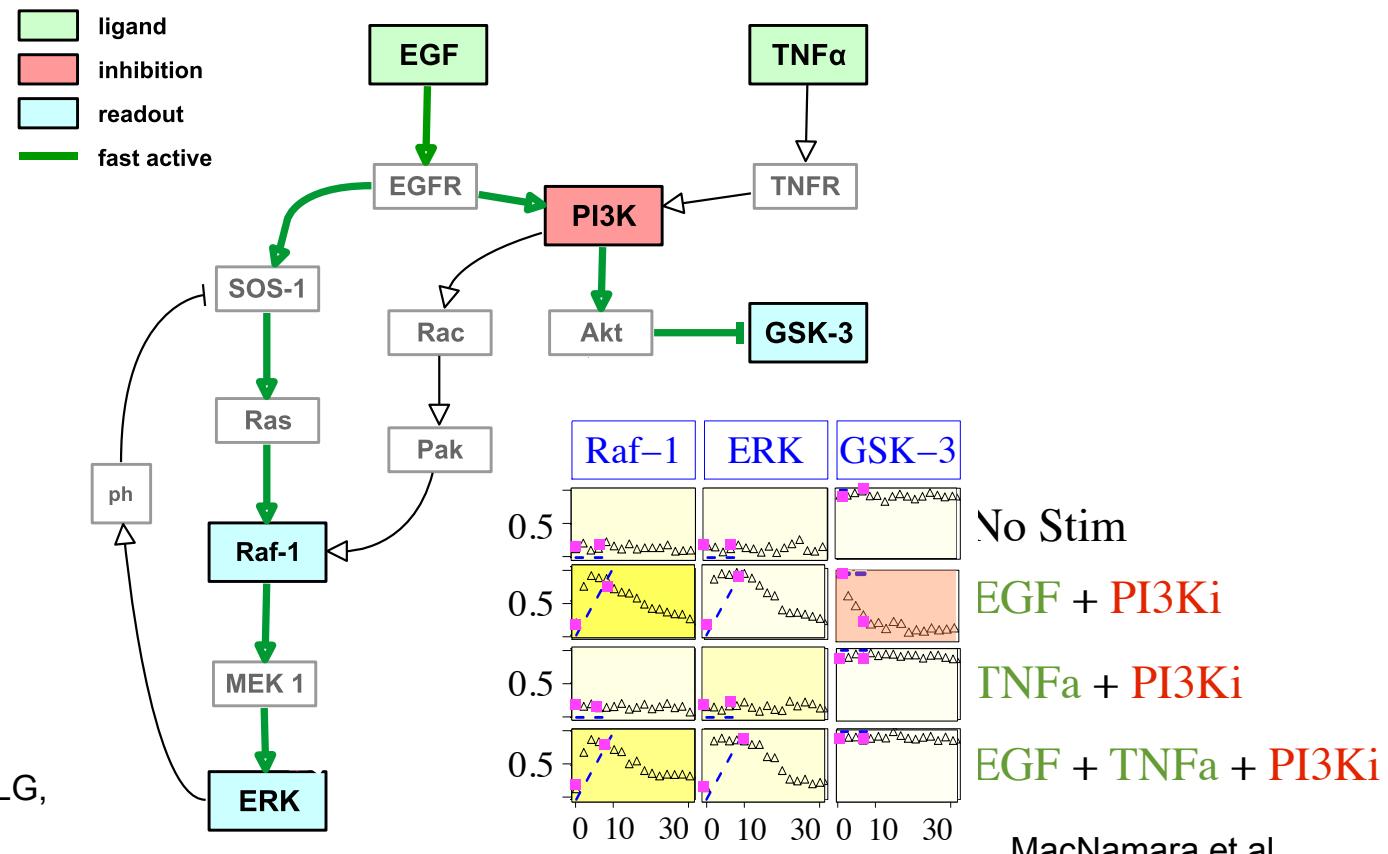
- 1 - Build general signalling network
- 2 - Perform stimulation experiments followed by phosphoproteomic measurements
- 3 - Find the combination of edges+logic gates (AND/OR) that best describes the experimental data (optimization)





# Logic modelling to link protein signalling networks with functional analysis of signal transduction

- 1 - Build general signalling network
- 2 - Perform stimulation experiments followed by phosphoproteomic measurements
- 3 - Find the combination of edges+logic gates (AND/OR) that best describes the experimental data (optimization)



Saez-Rodriguez J, Alexopoulos LG,  
Epperlein J, Samaga R,  
Lauffenburger DA, Klamt S, Sorger PK  
*Mol Sys Bio* 5:331,2009

MacNamara et al.  
*Phys Biol* 9 045003, 2012



# Logic modelling to link protein signalling networks with functional analysis of signal transduction

- 1 - Build general signalling network
- 2 - Perform stimulation experiments followed by phosphoproteomic measurements
- 3 - Find the combination of edges+logic gates (AND/OR) that best describes the experimental data (optimization)

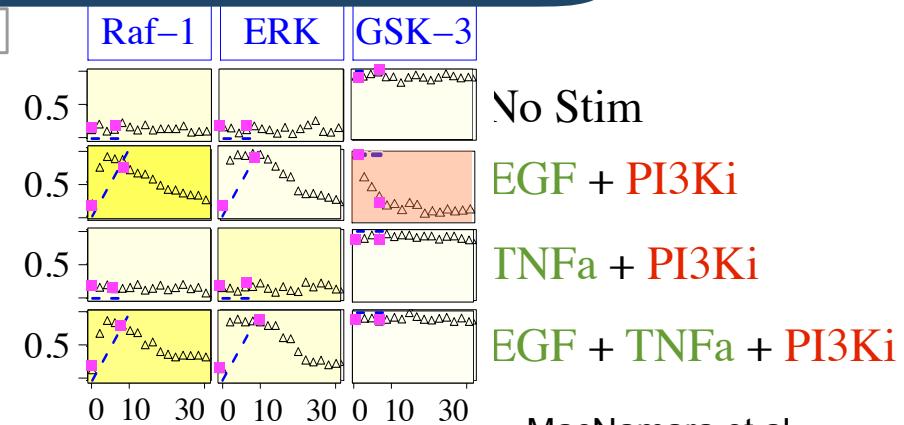
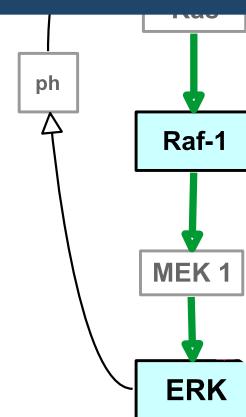
— ligand  
— EGF  
— TNF $\alpha$

CellNOpt: [www.cellnopt.org](http://www.cellnopt.org)

free open-source R/Python/Matlab/Cytoscape

Terfve C et al. *BMC Syst Biol*, 6:133, 2012

Morris MK, et al., *Methods Mol. Biol.*, 930:179-214, 2013



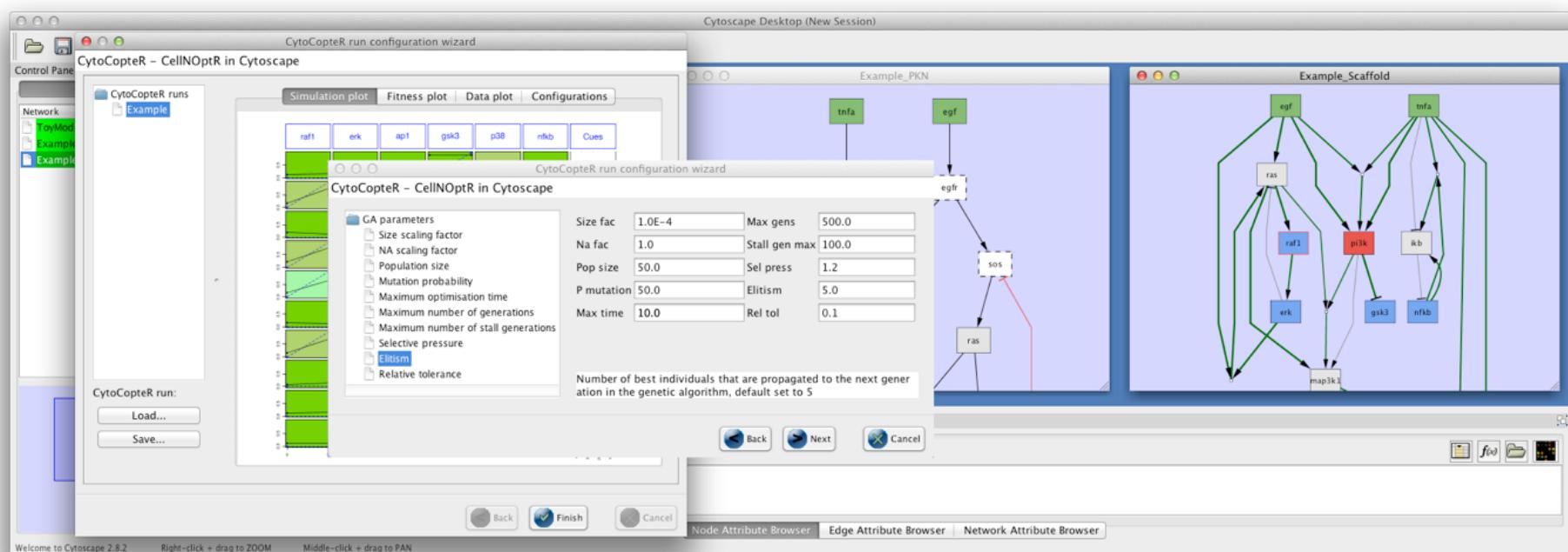
Saez-Rodriguez J, Alexopoulos LG,  
Epperlein J, Samaga R,  
Lauffenburger DA, Klamt S, Sorger PK  
*Mol Sys Bio* 5:331,2009

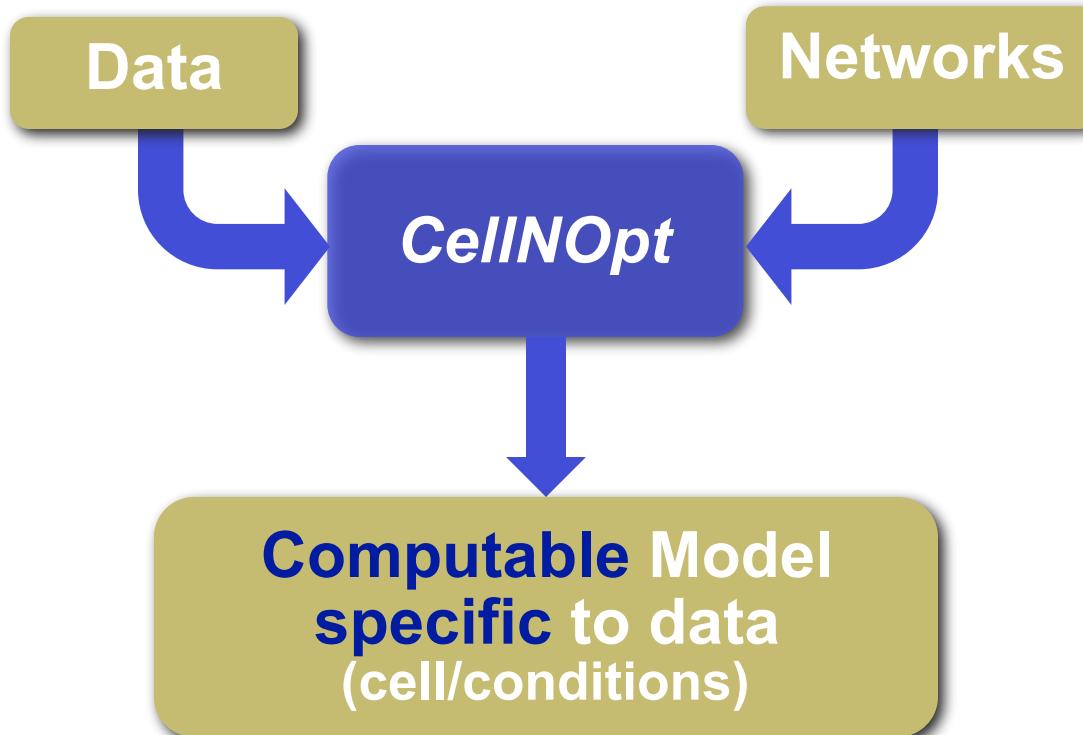
MacNamara et al.  
*Phys Biol* 9 045003, 2012



# CytoCopter - CellNOpt in Cytoscape

A Cytoscape plugin to run CellNOptR  
(<http://apps.cytoscape.org/apps/cytocopter>)





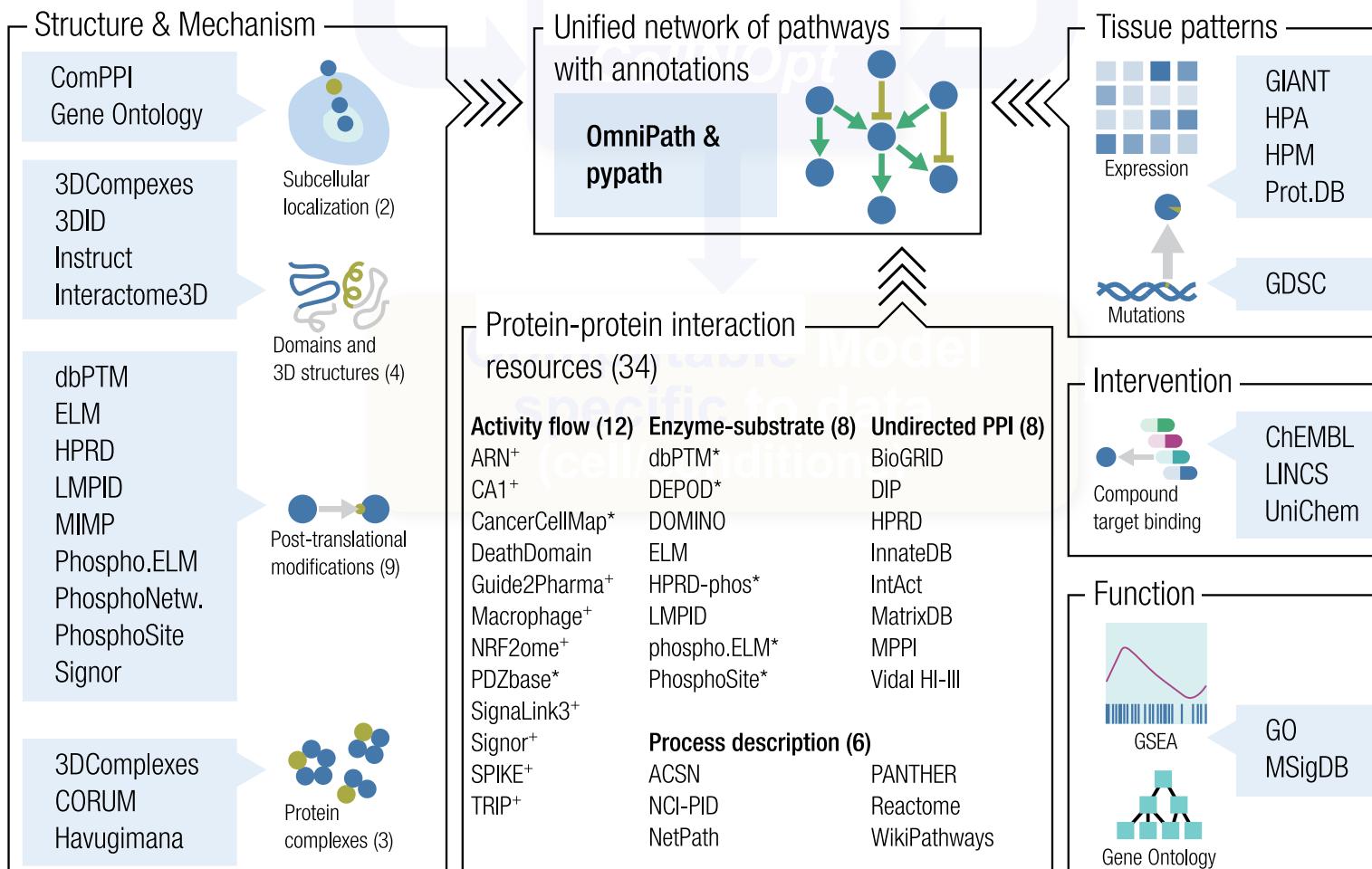


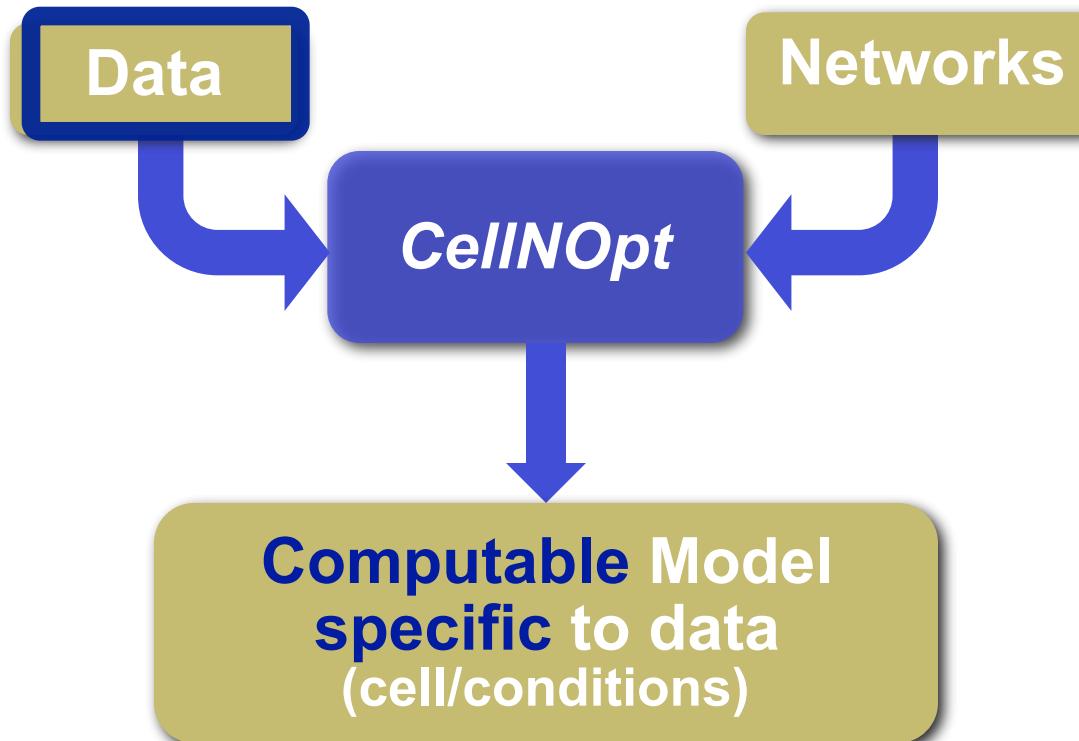
# Omnipath: Integration of existing pathway resources to improve modelling

P

[www.omnipathdb.org](http://www.omnipathdb.org)

## Networks





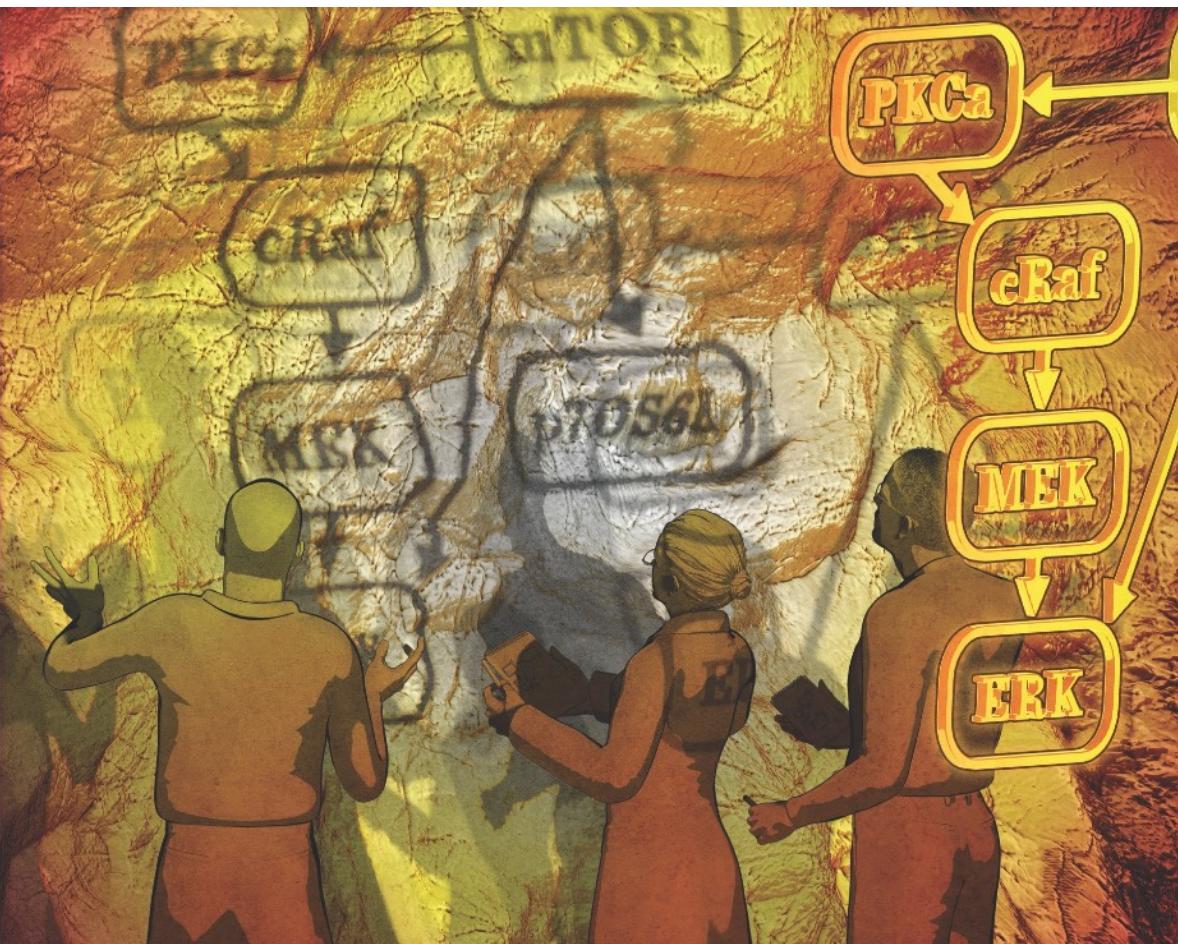


# Plato's allegory of the cave





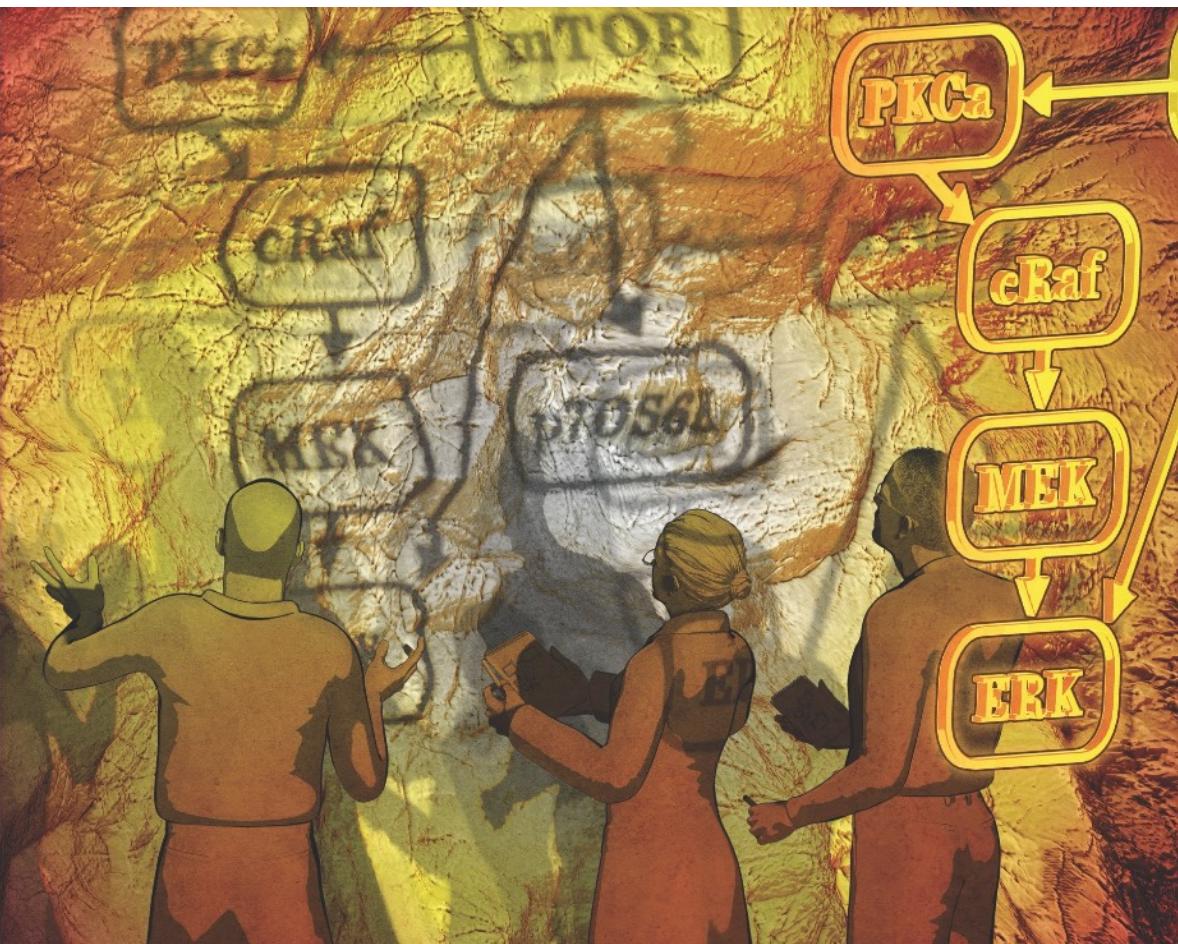
# Challenges modelling signalling networks



Artwork by S. Philips on idea of J. Saez-Rodriguez; appeared in cover of *Nat Meth*, 13:4, 2016



# Challenges modelling signalling networks

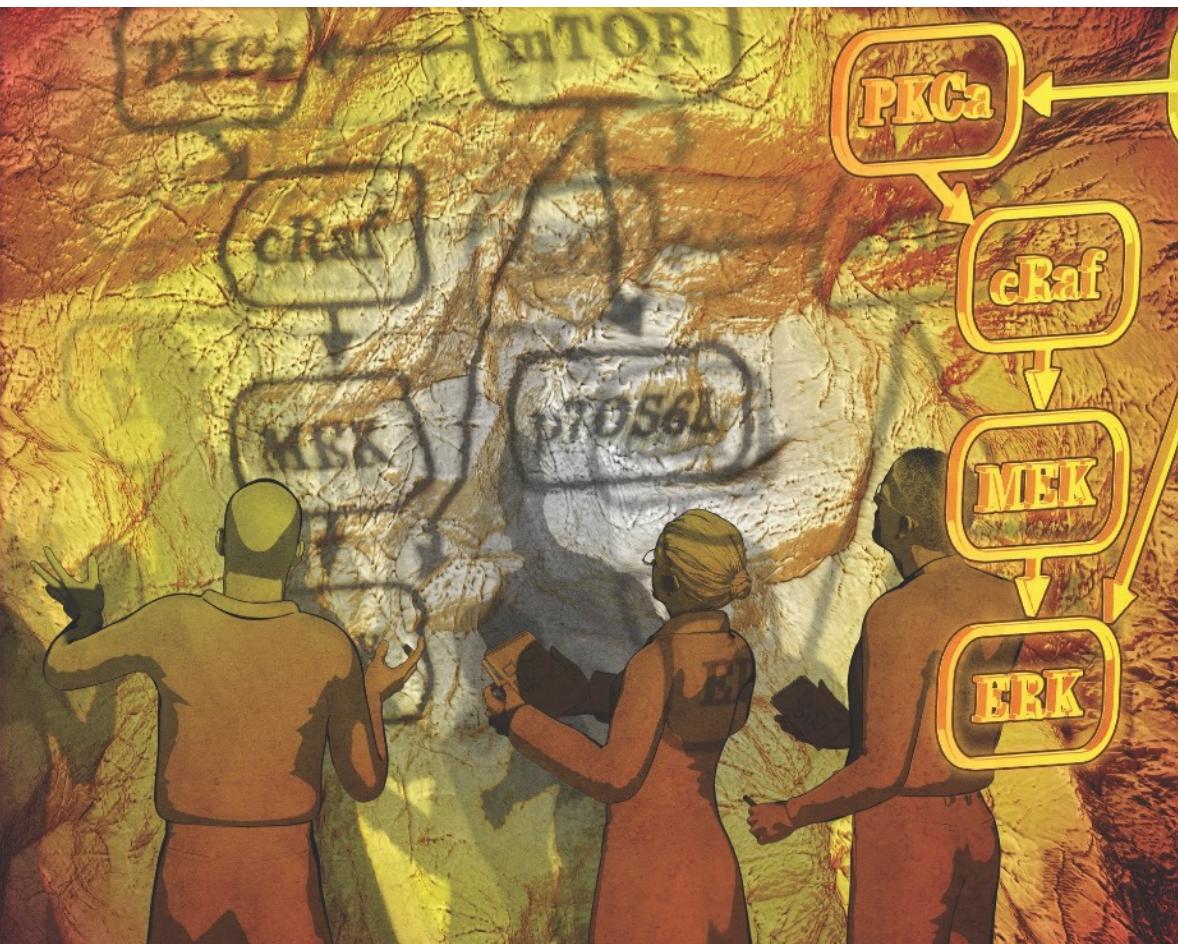


- **Cues are lights**

Artwork by S. Philips on idea of J. Saez-Rodriguez; appeared in cover of *Nat Meth*, 13:4, 2016



# Challenges modelling signalling networks

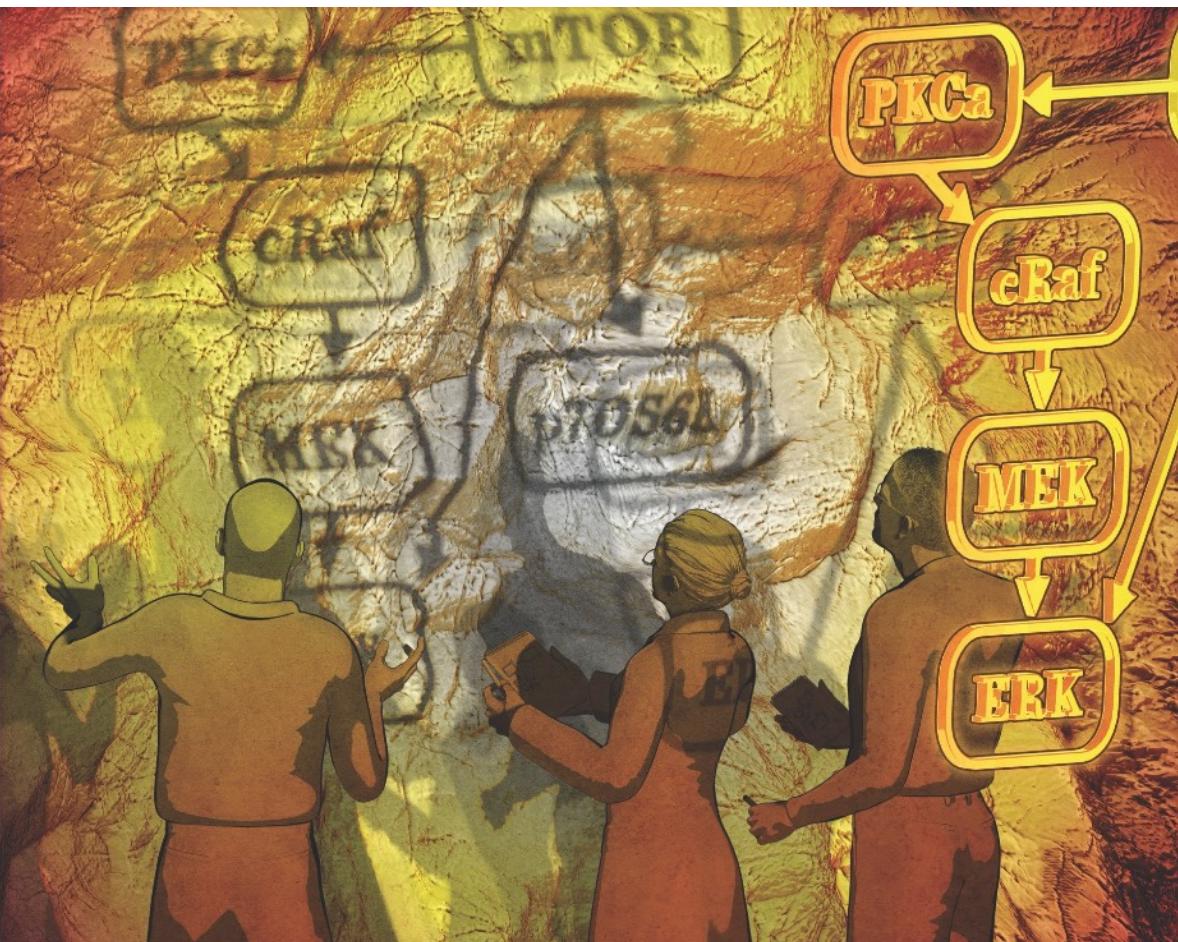


Artwork by S. Philips on idea of J. Saez-Rodriguez; appeared in cover of *Nat Meth*, 13:4, 2016

- **Cues are lights**
- **Measurements are shadows:**
  - Phosphorylation = Activation?  
Which site? How does it affect the regulation of the protein?



# Challenges modelling signalling networks

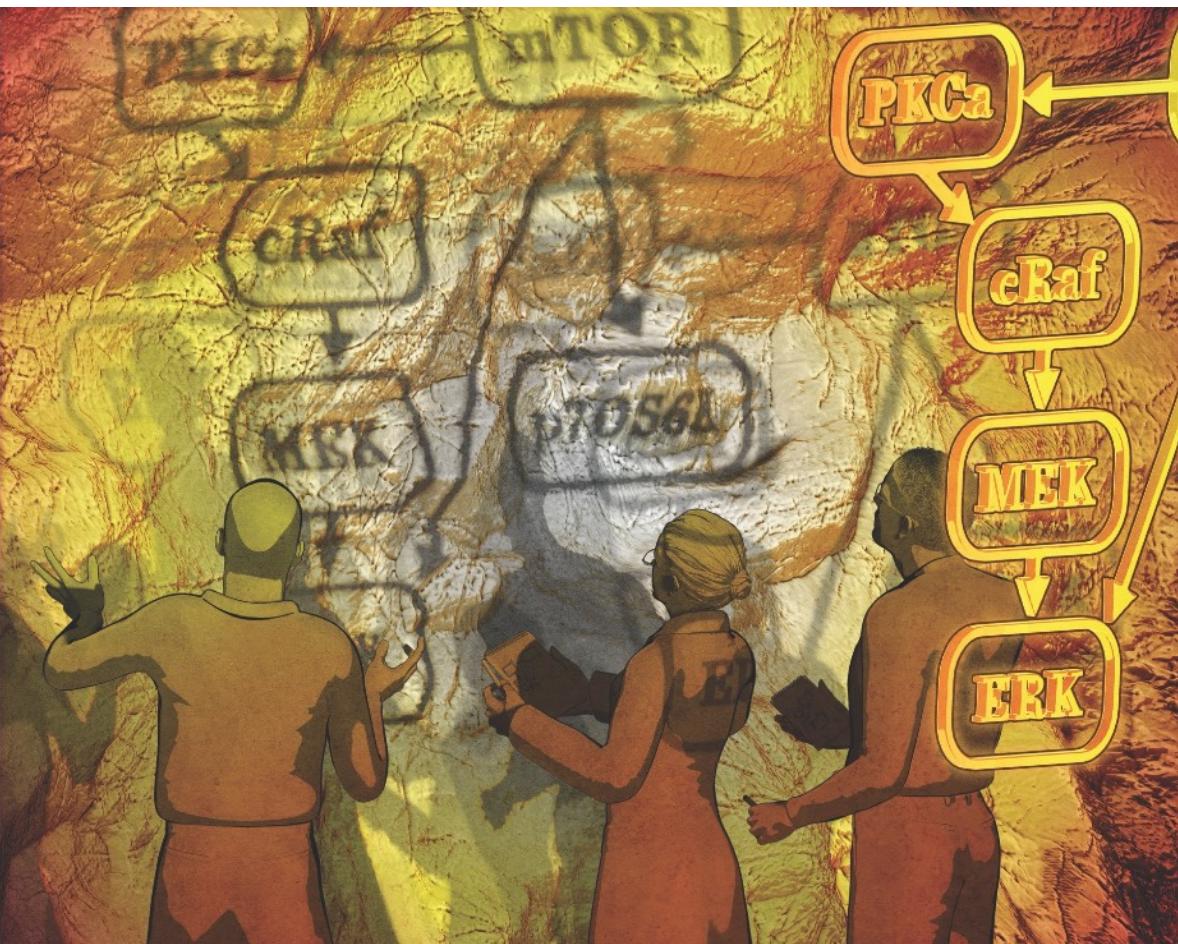


Artwork by S. Philips on idea of J. Saez-Rodriguez; appeared in cover of *Nat Meth*, 13:4, 2016

- **Cues are lights**
- **Measurements are shadows:**
  - Phosphorylation = Activation?  
Which site? How does it affect the regulation of the protein?
  - Fluorescence=phosphorylation?  
Signal saturated?  
Below detection level?



# Challenges modelling signalling networks



Artwork by S. Philips on idea of J. Saez-Rodriguez; appeared in cover of *Nat Meth*, 13:4, 2016

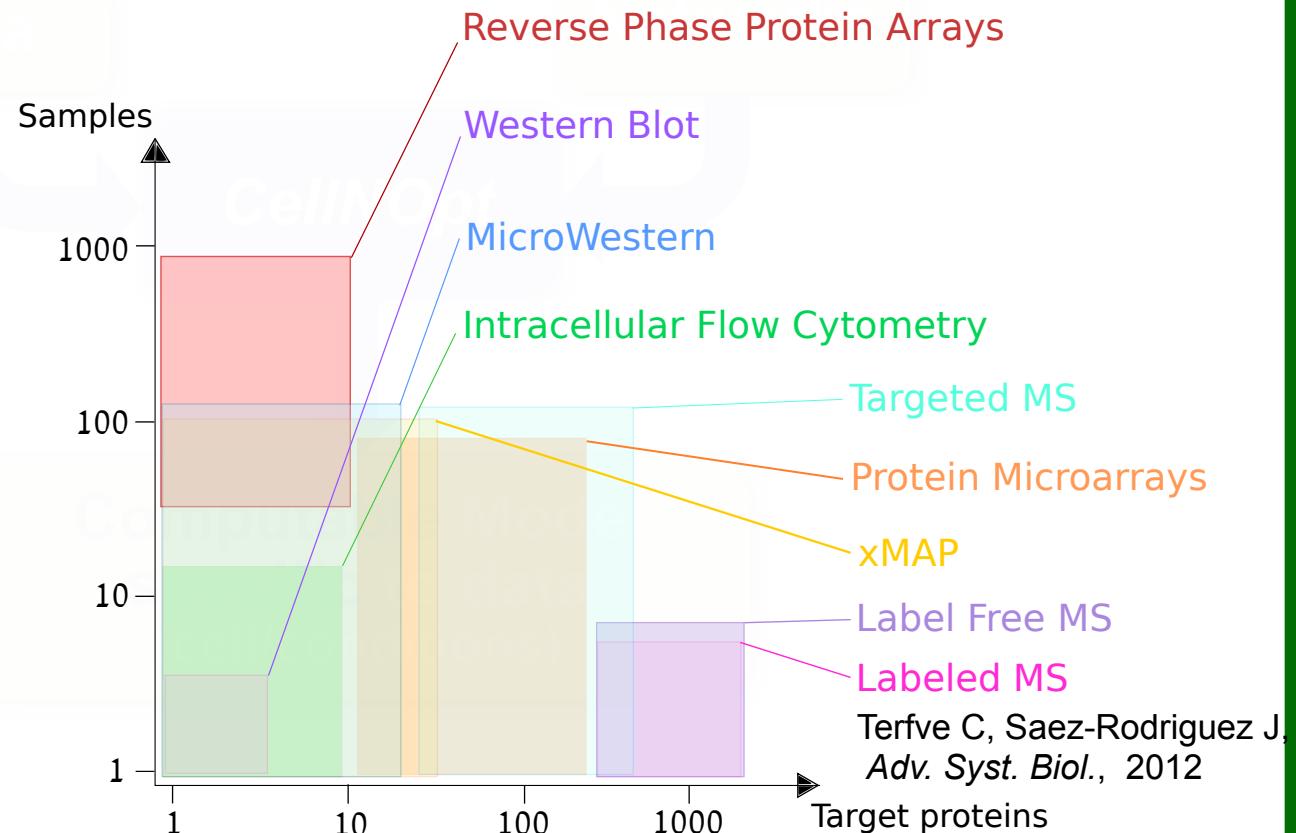
- **Cues are lights**
- **Measurements are shadows:**
  - Phosphorylation = Activation?  
Which site? How does it affect the regulation of the protein?
  - Fluorescence=phosphorylation?  
Signal saturated?  
Below detection level?

- Different 'lights' (stimulation/perturbation) provide complementary information



# Leveraging different proteomic platforms

Different data types



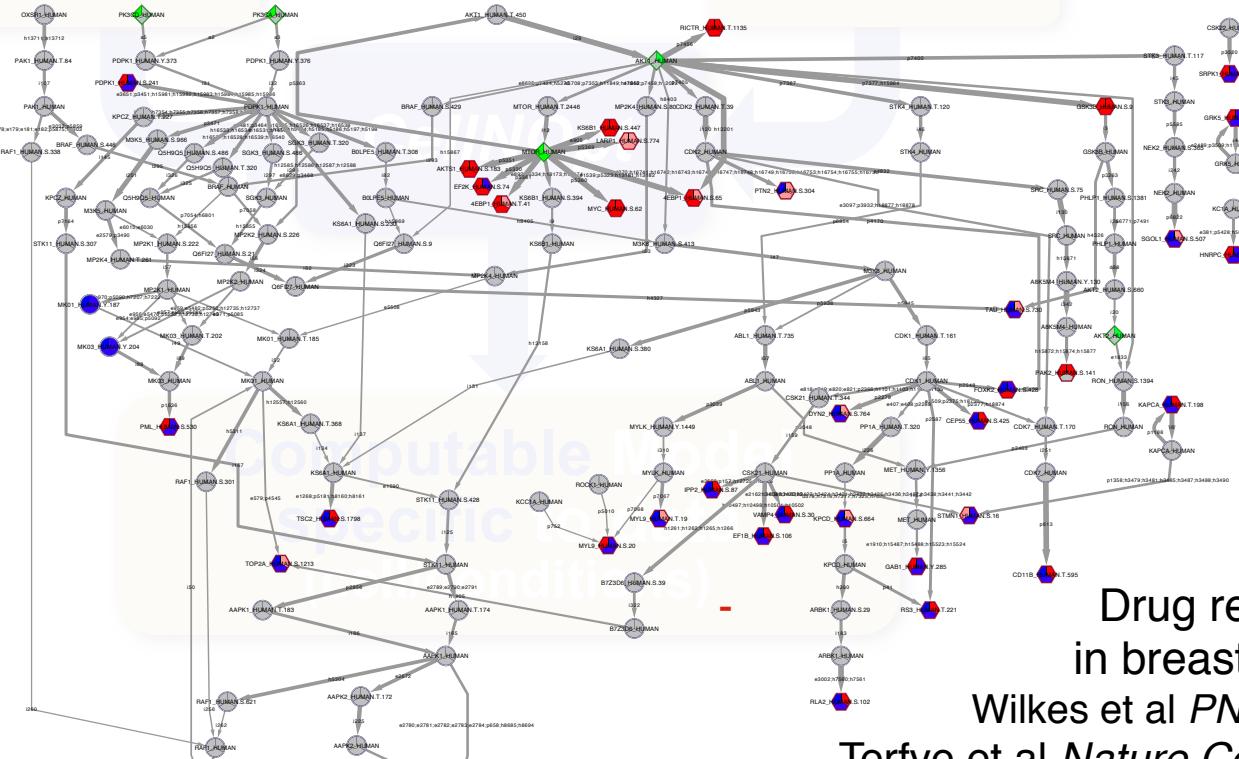


# Leveraging different proteomic platforms

Different data types

- Mass spectrometry phospho-proteomics for high coverage of signalling networks

From ~ 10s (antibody-based) to ~ 1000s proteins  
w. R. Aebersold (ETH Zurich), P. Cutillas (Barts London)



Drug response  
in breast cancer  
Wilkes et al *PNAS* 2015  
Terfve et al *Nature Com* 2015

Tool for modeling MS P-proteomics:  
[www.cellnopt.org/PHONEMeS](http://www.cellnopt.org/PHONEMeS)

Crosstalk in yeast  
Vaga et al, *Mol Syst Bio* 2014



# Leveraging different proteomic platforms

Different data types

- **Mass spectrometry** phospho-proteomics for high coverage of signalling networks  
From ~ 10s (antibody-based) to ~ 1000s proteins  
w. R. Aebersold (ETH Zurich), P. Cutillas (Barts London)
- **Single cell signaling:**
  - Imaging (w. C. Schultz, EMBL),
  - CytoF (w. B. Bodenmiller, U. Zurich)
- **Combination of proteomic and metabolomics**  
(Blattmann et al *Cell Systems* 2017)
- **Transcriptomics** (complementary tool: **CARNIVAL**)  
[saezlab.github.io/CARNIVAL/](http://saezlab.github.io/CARNIVAL/)



# How to choose model: balance of fit of data and size of model

A good model should describe (and predict) data well and be as simple as possible

Metric

$$\theta = \theta_f + \alpha \cdot \theta_S$$

**Fit to data**

$$\theta_f = \sum_{l=1}^S \sum_{K=1}^M (Bi_{kl}^M - Bi_{kl}^E)^2$$

$\in \{0,1\}$     $\in [0,1]$



# How to choose model: balance of fit of data and size of model

A good model should describe (and predict) data well and be as simple as possible

Metric

$$\theta = \theta_f + \alpha \cdot \theta_S$$

**Fit to data**

$$\theta_f = \sum_{l=1}^S \sum_{K=1}^M (Bi_{kl}^M - Bi_{kl}^E)^2$$

$\in \{0,1\}$     $\in [0,1]$

**Size of model**

$$\theta_S = \sum_{k=1}^n v_k P_k$$



# How to choose model: balance of fit of data and size of model

A good model should describe (and predict) data well and be as simple as possible

Metric

$$\theta = \theta_f + \alpha \cdot \theta_S$$

**Fit to data**

$$\theta_f = \sum_{l=1}^S \sum_{K=1}^M (Bi_{kl}^M - Bi_{kl}^E)^2$$

$\in \{0,1\}$     $\in [0,1]$

**Relative  
importance  
Fit vs. Size**

**Size of model**

$$\theta_S = \sum_{k=1}^n v_k P_k$$



# How to choose model: balance of fit of data and size of model

A good model should describe (and predict) data well and be as simple as possible

Metric

$$\theta = \theta_f + \alpha \cdot \theta_S$$

**Fit to data**

$$\theta_f = \sum_{l=1}^S \sum_{K=1}^M (Bi_{kl}^M - Bi_{kl}^E)^2$$

$\in \{0,1\}$     $\in [0,1]$

**Relative  
importance  
Fit vs. Size**

**Size of model**

$$\theta_S = \sum_{k=1}^n v_k P_k$$

In practice: small value ( $\sim 0.0001$ ) to prioritize fitness of data



# How to choose model: balance of fit of data and size of model

A good model should describe (and predict) data well and be as simple as possible

Metric

$$\theta = \theta_f + \alpha \cdot \theta_S$$

**Fit to data**

$$\theta_f = \sum_{l=1}^S \sum_{K=1}^M (Bi_{kl}^M - Bi_{kl}^E)^2$$

Data is normalized  
between 0 and 1

**Relative  
importance  
Fit vs. Size**

**Size of model**

$$\theta_S = \sum_{k=1}^n v_k P_k$$

In practice: small value ( $\sim 0.0001$ ) to prioritize fitness of data



# How to choose model: balance of fit of data and size of model

A good model should describe (and predict) data well and be as simple as possible

Metric

$$\theta = \theta_f + \alpha \cdot \theta_S$$

**Fit to data**

$$\theta_f = \sum_{l=1}^S \sum_{K=1}^M (Bi_{kl}^M - Bi_{kl}^E)^2$$

Data is normalized  
between 0 and 1

**Relative  
importance  
Fit vs. Size**

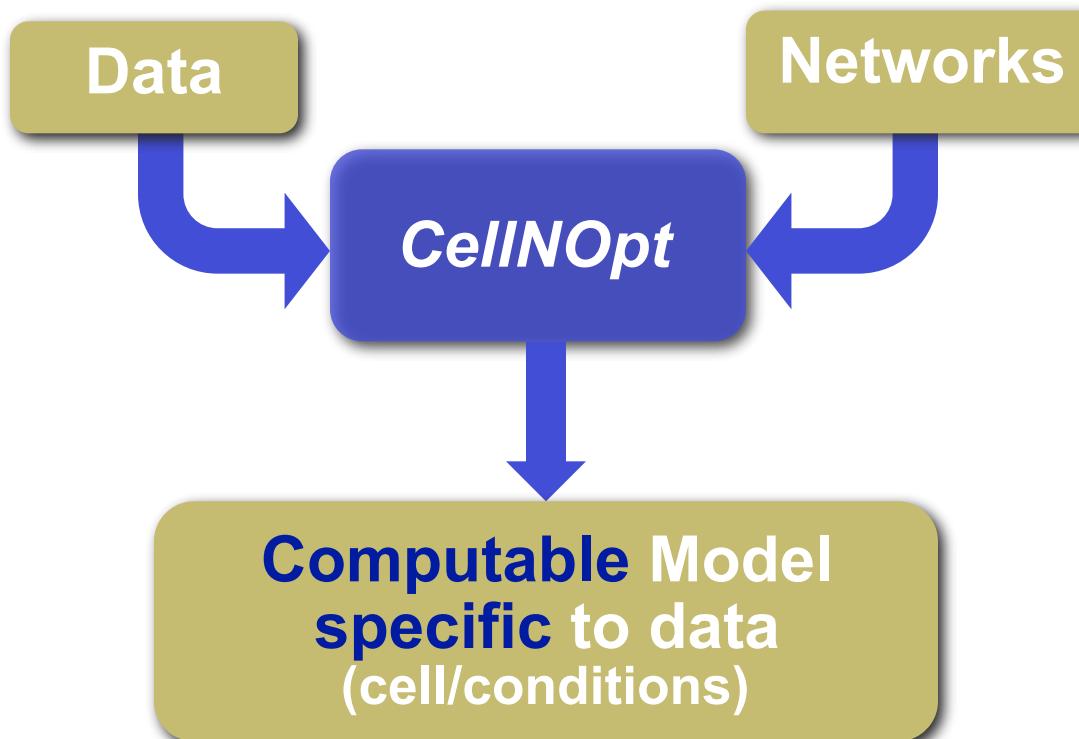
**Size of model**

$$\theta_S = \sum_{k=1}^n v_k P_k$$

Best model ~ minimum metric  
(optimization problem) - can be solved algorithmically

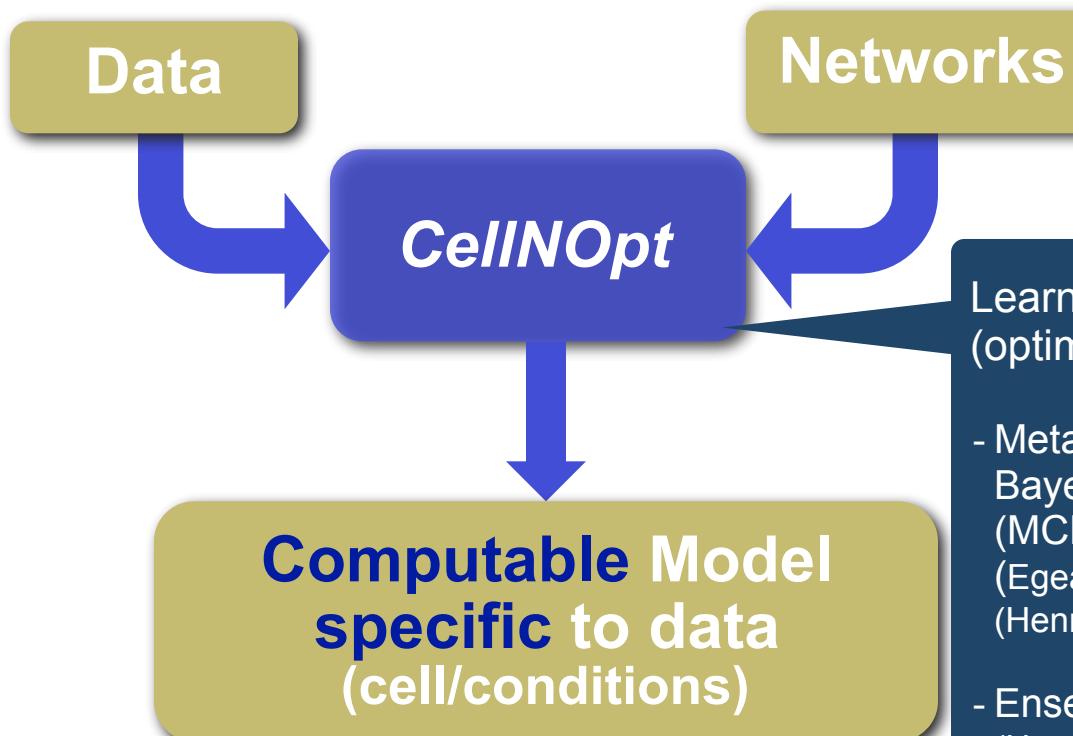


# Fitting: Solving optimisation problem





# Fitting: Solving optimisation problem

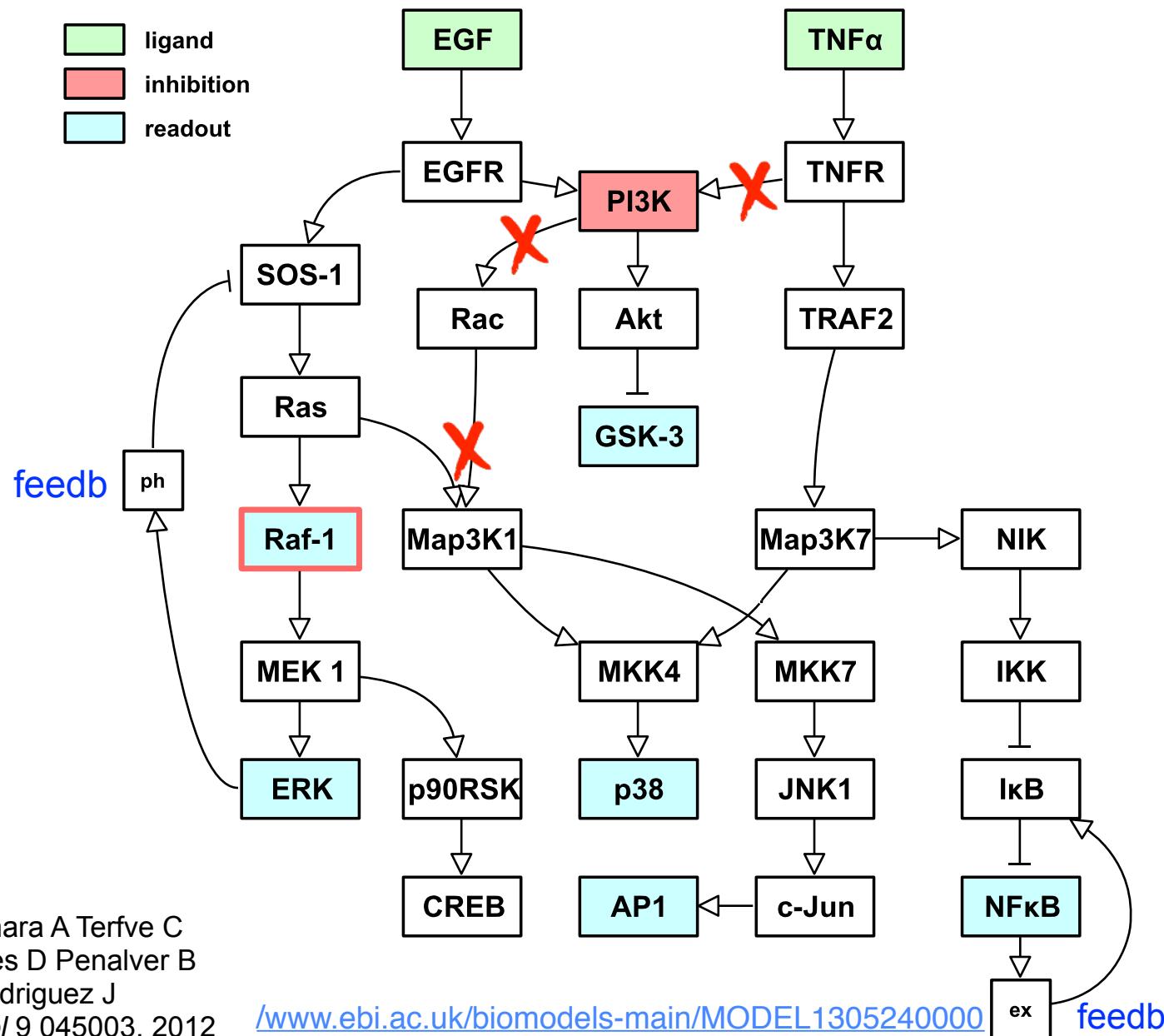


Learning algorithms  
(optimization):

- Metaheuristics & Bayesian Inference (MCMC)  
(Egea et al. *BMC Bioinf* 2014;  
(Henriques et al. *Bioinf* 2015)
- Ensembles of models  
(Henriques et al. *PLoS CB*, 2017)
- Use of Answer Set Programming (Guziowski et al. *Bioinf* 2013, Videla et al. *Bioinf* 2017) and Integer Linear Programming (Mitsos et al *PLoS CB* 2009)

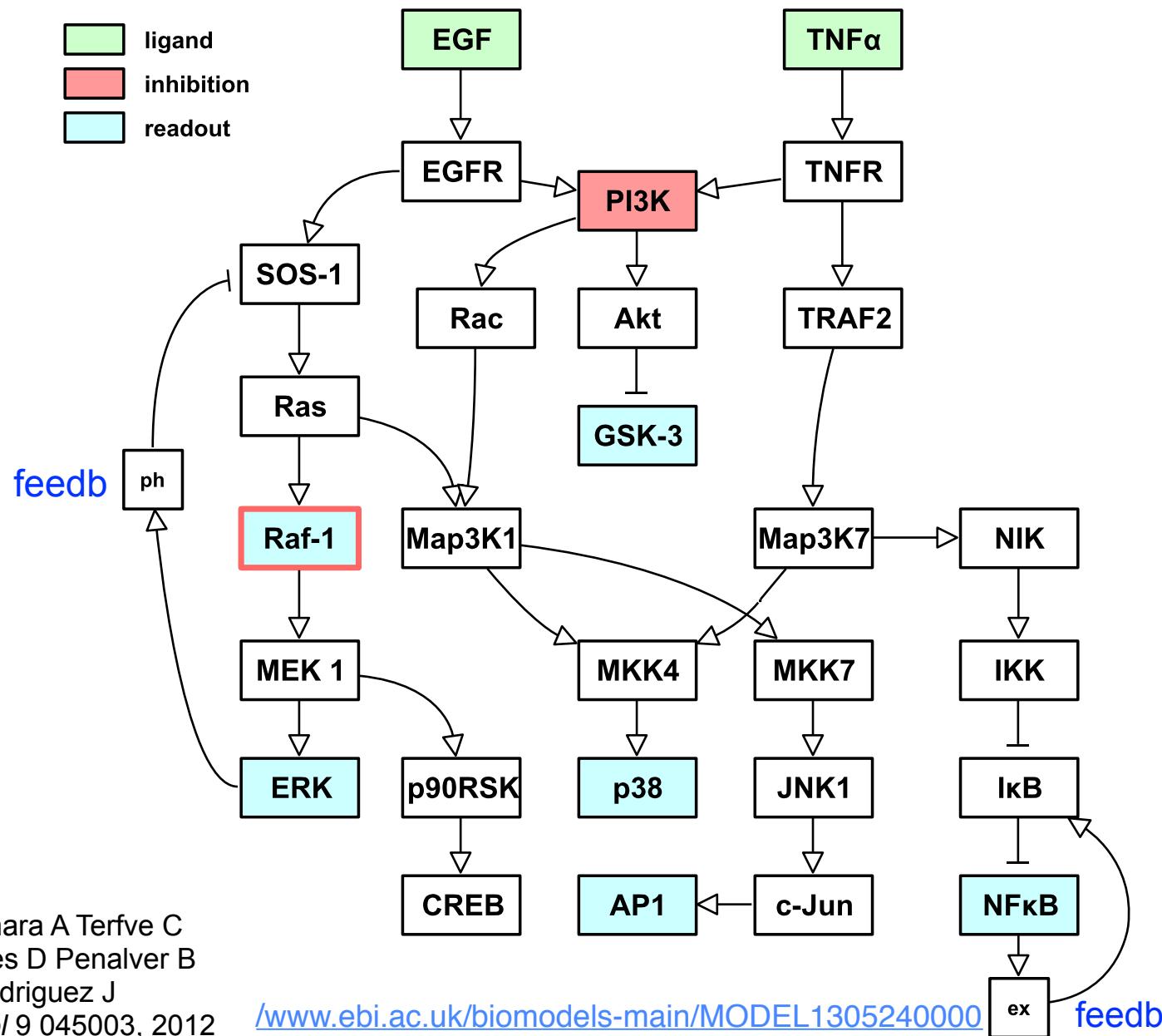


# A Toy model



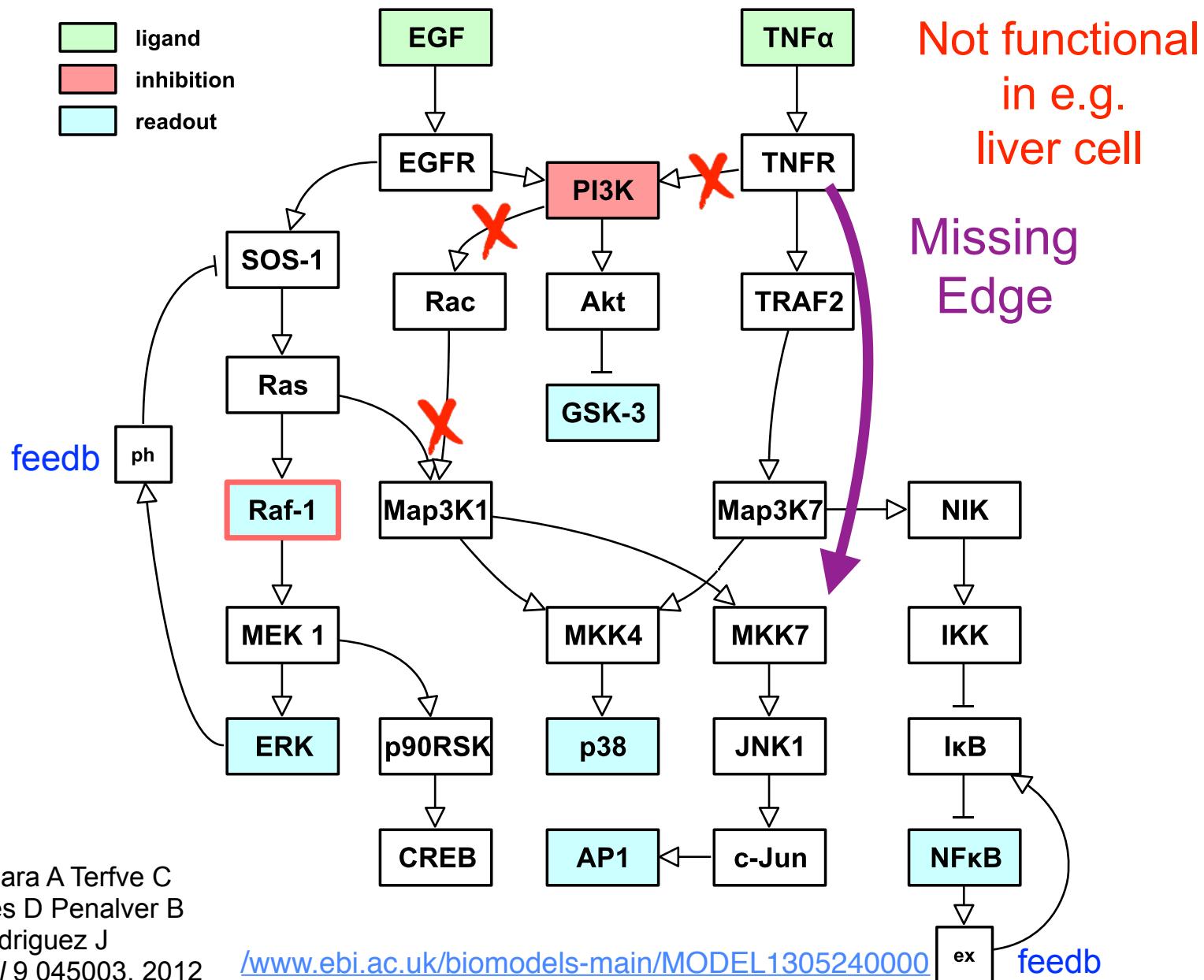


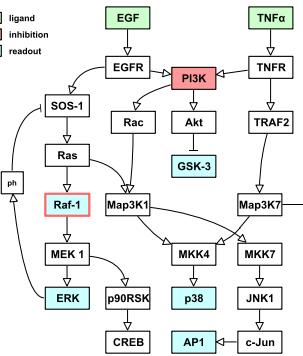
# A Toy model



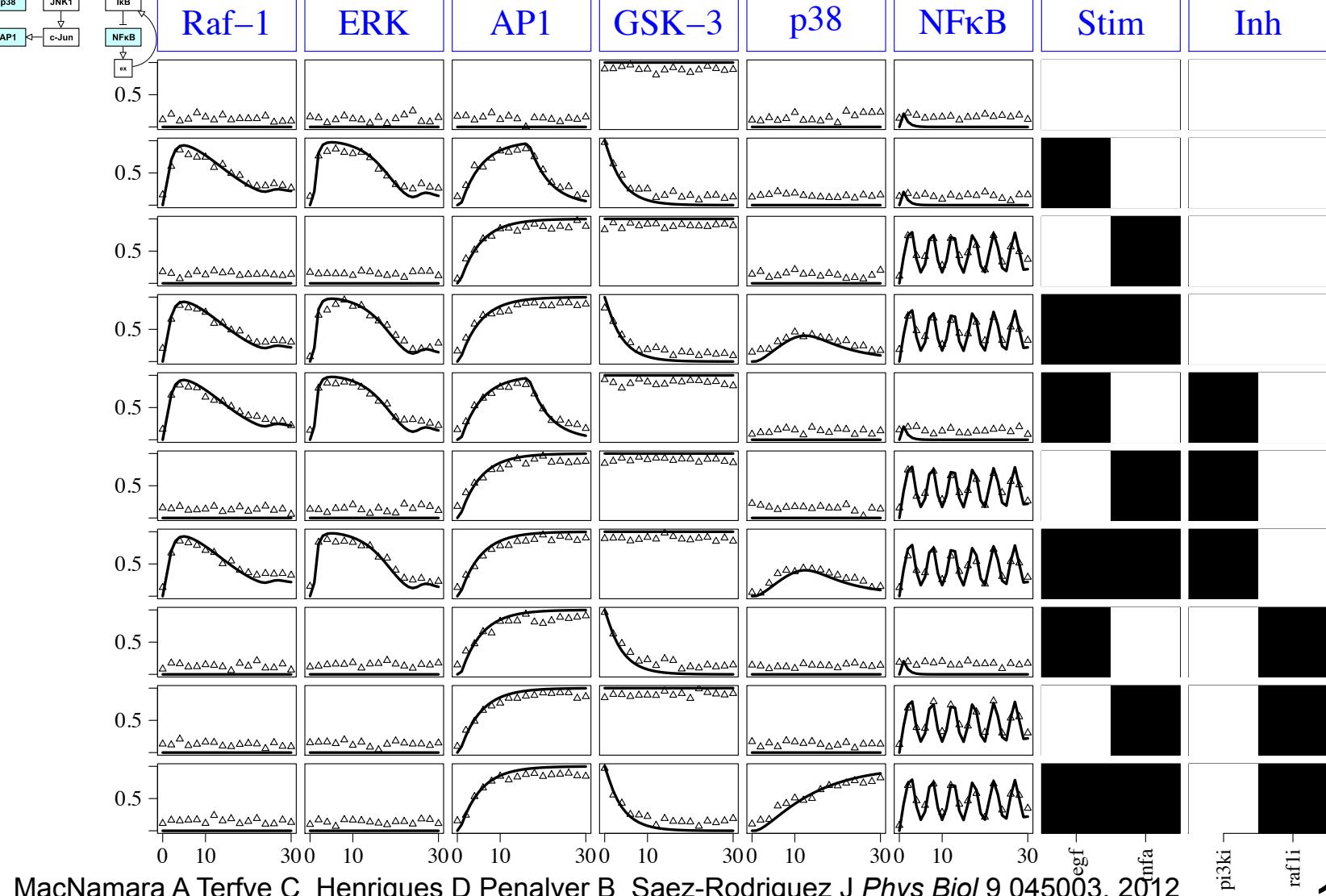


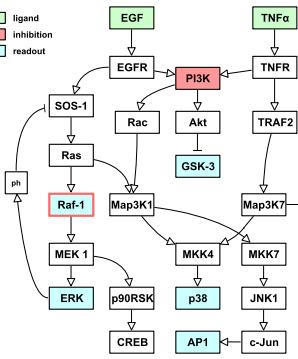
# A Toy model





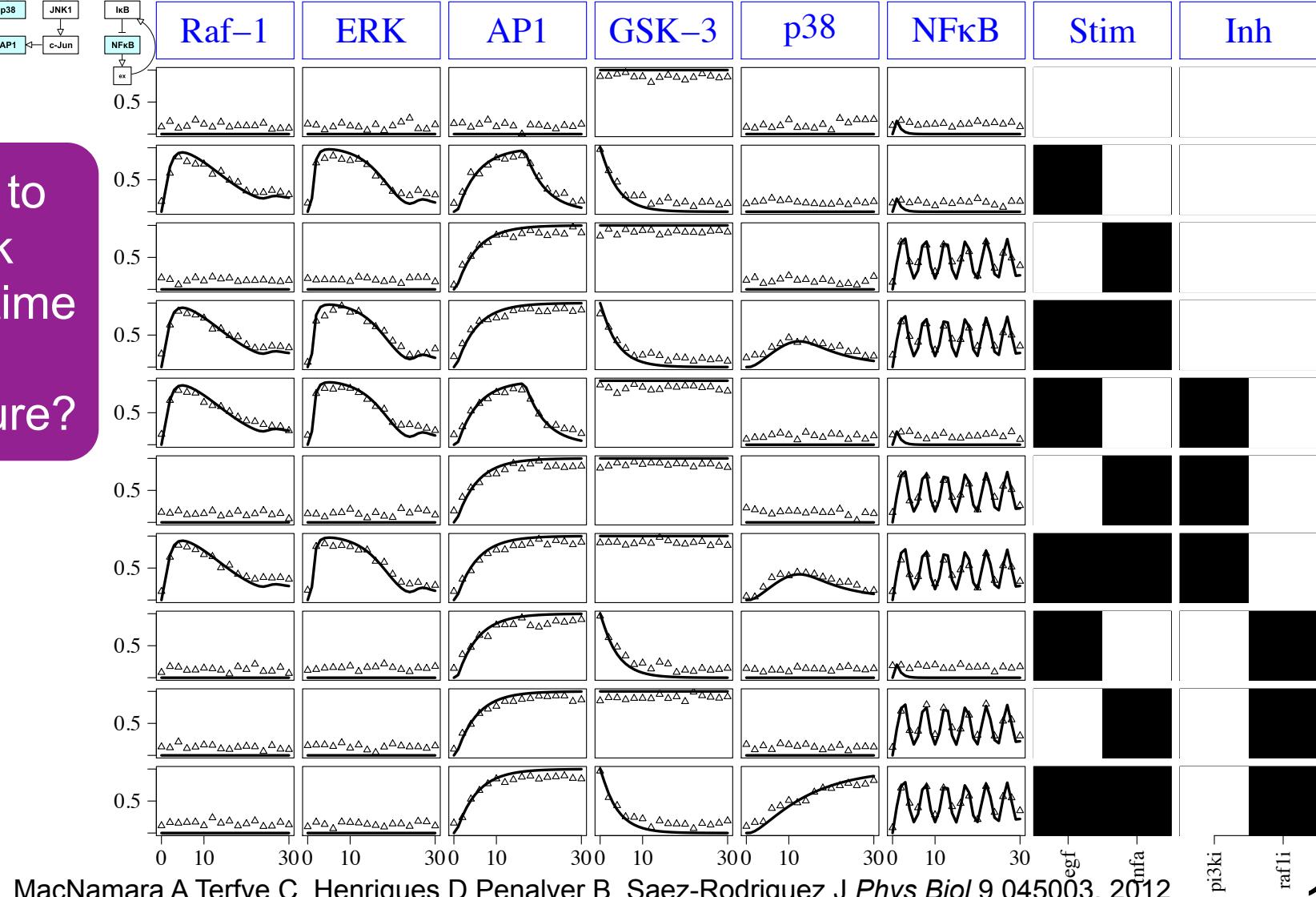
# The ‘real’ data

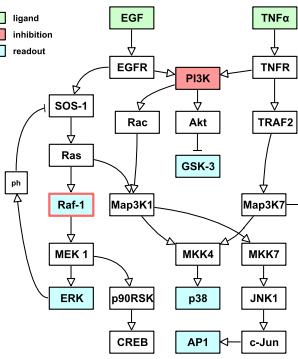




# The ‘real’ data

How to  
pick  
right time  
to  
measure?

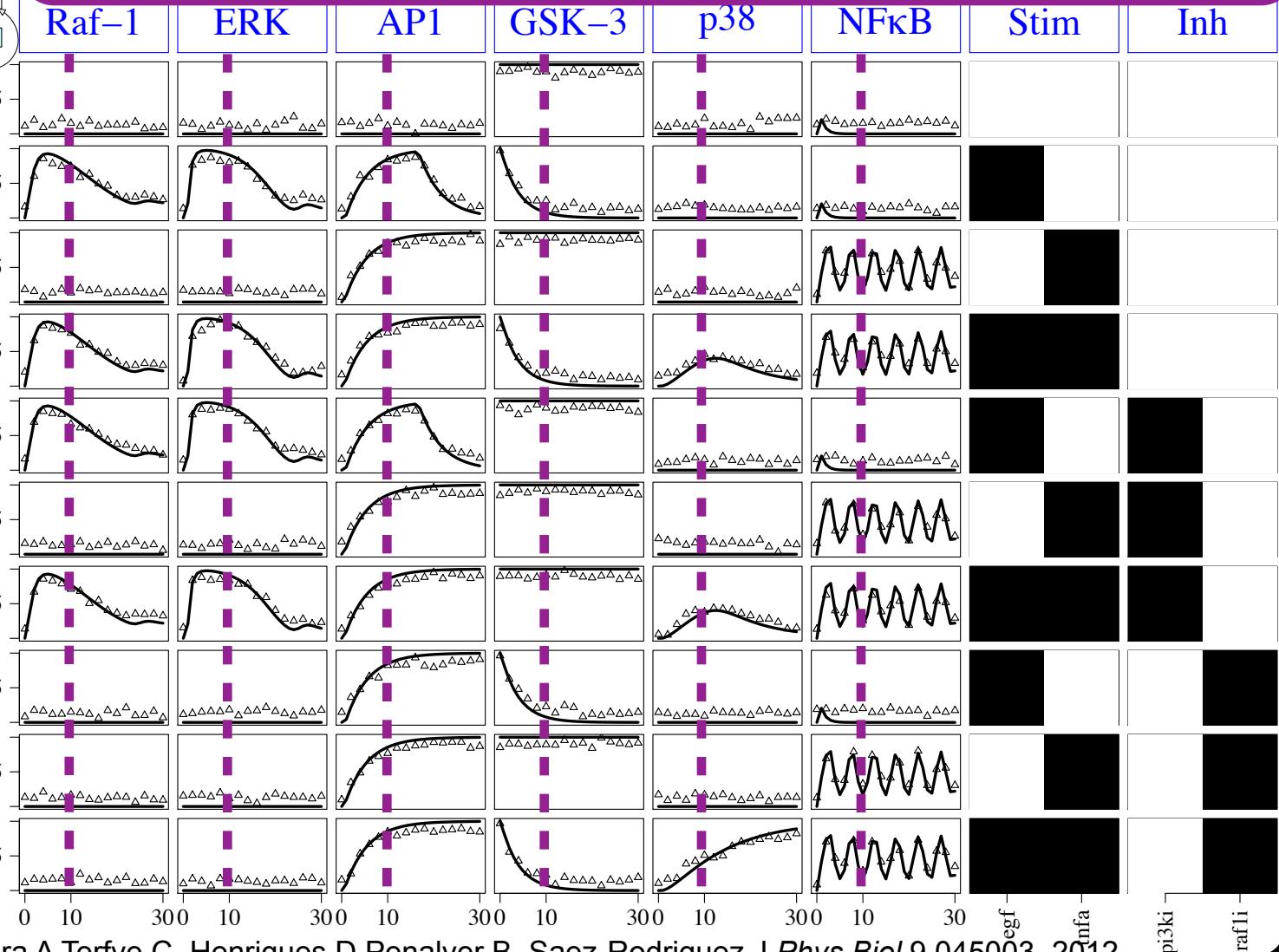




# The ‘real’ data

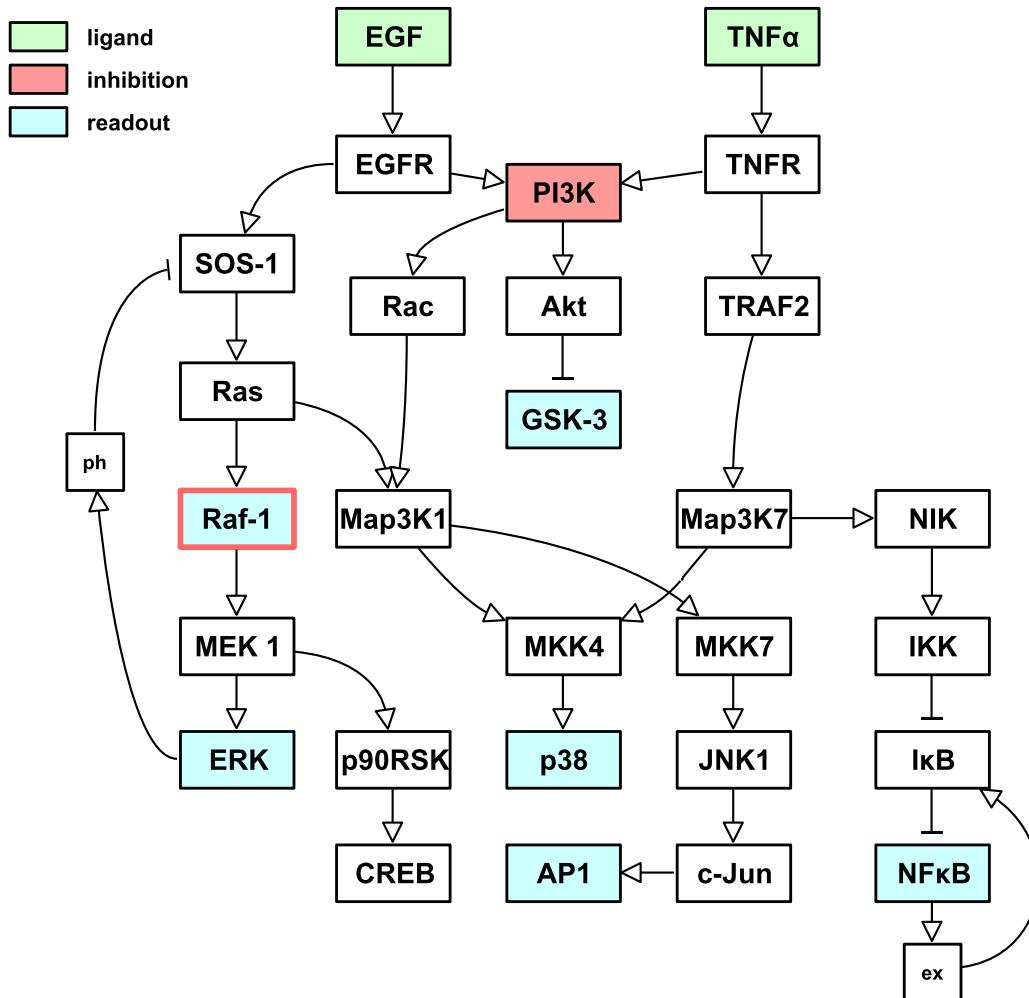
If you can only pick one (\$\$), choose one representative of a ‘time scale’

How to  
pick  
right time  
to  
measure?



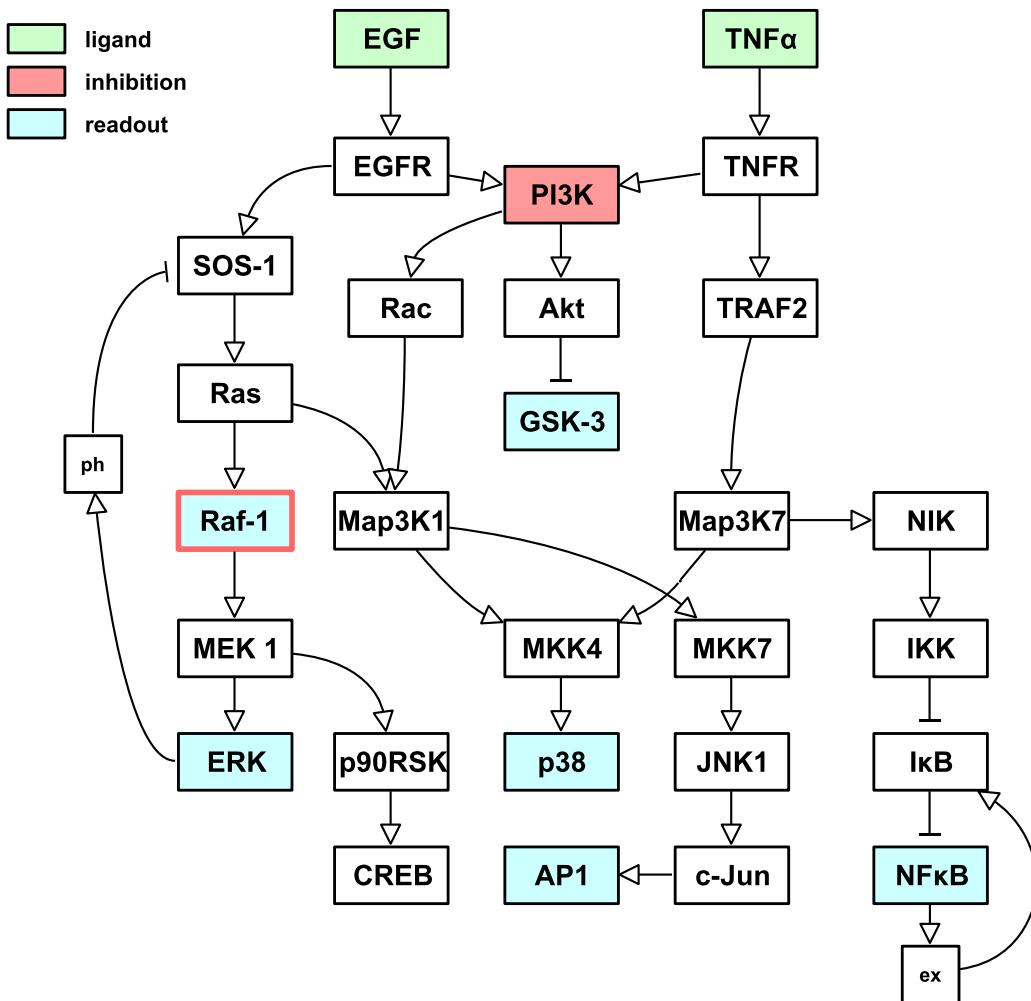


# Model preprocessing:

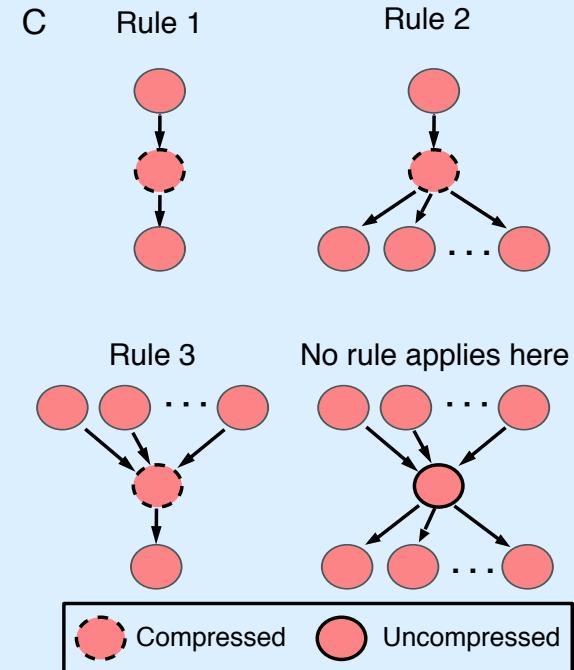




# Model preprocessing:



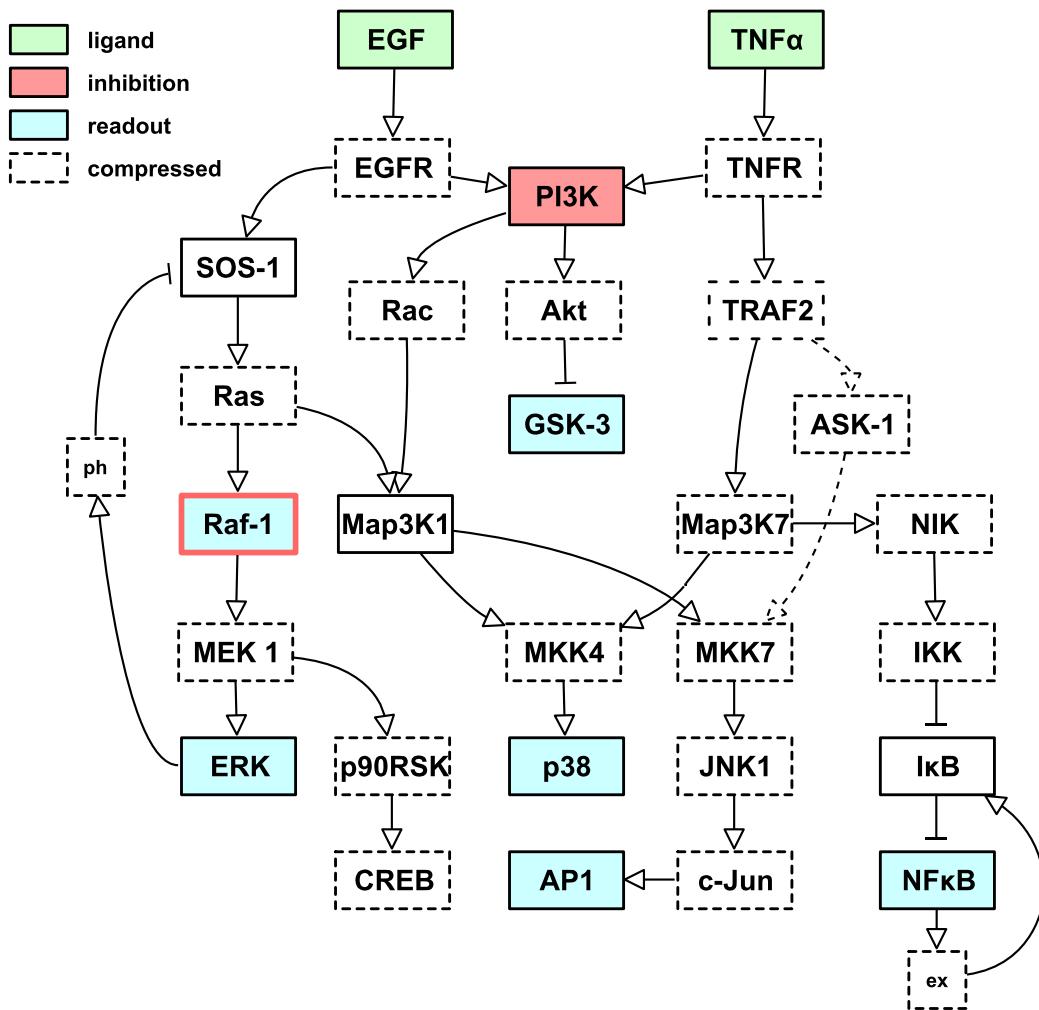
Model compression & removal of non-controlable & non-observable branches



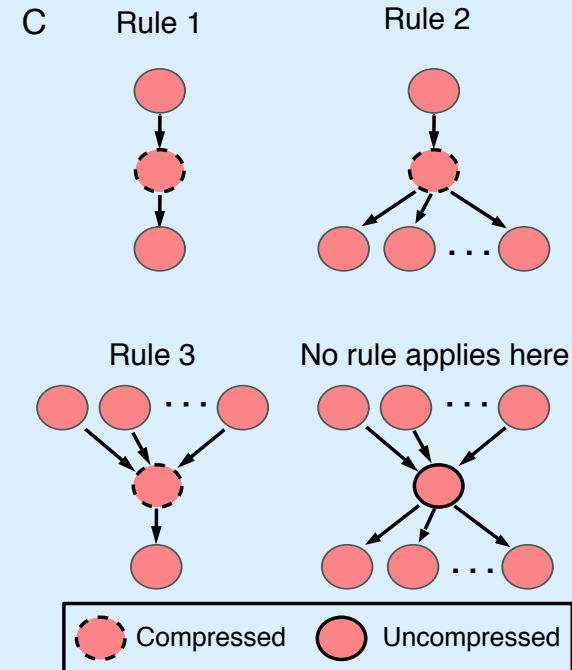
Saez-Rodriguez J, et al., Mol. Syst. Biol, 2009



# Model preprocessing:



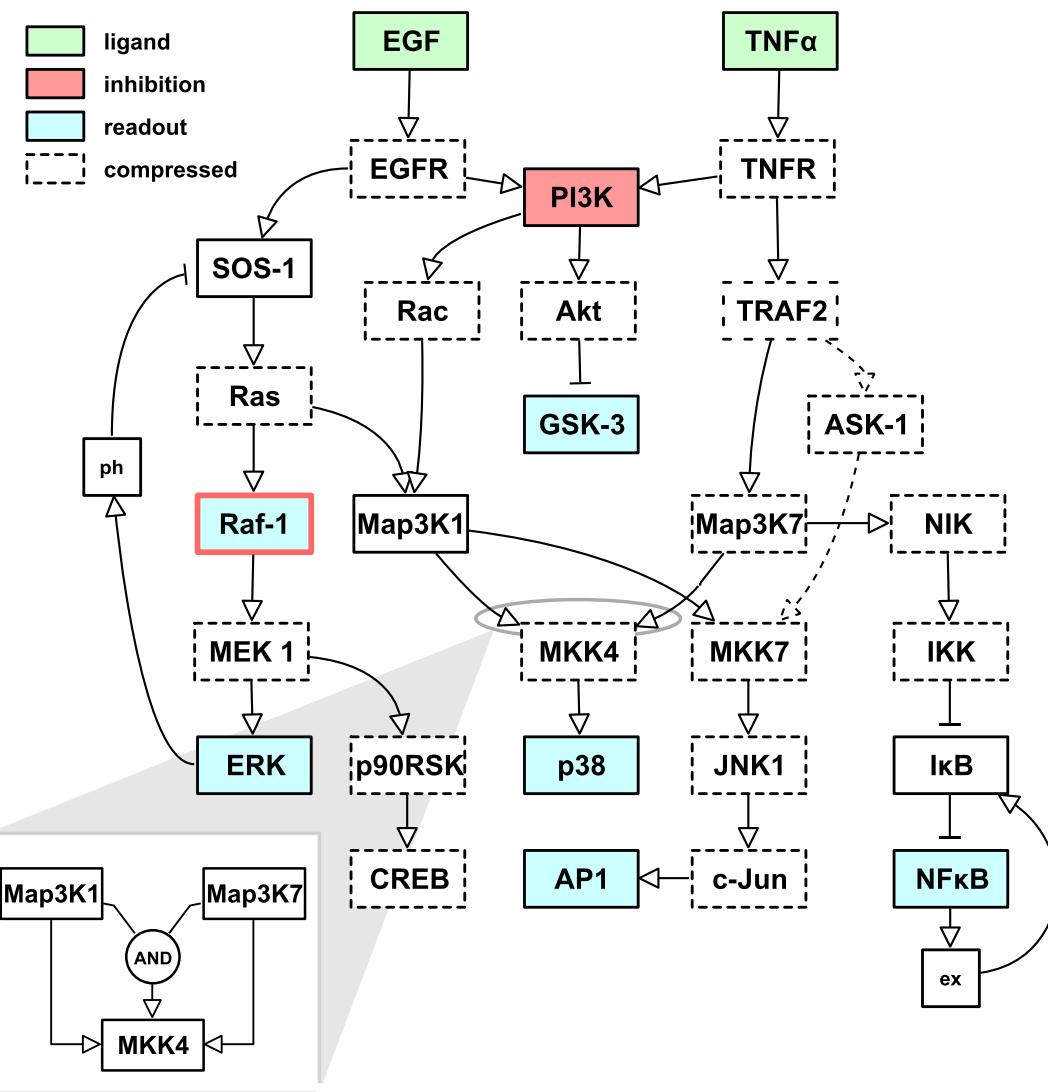
Model compression&  
removal of non-controlable  
& non-observable branches



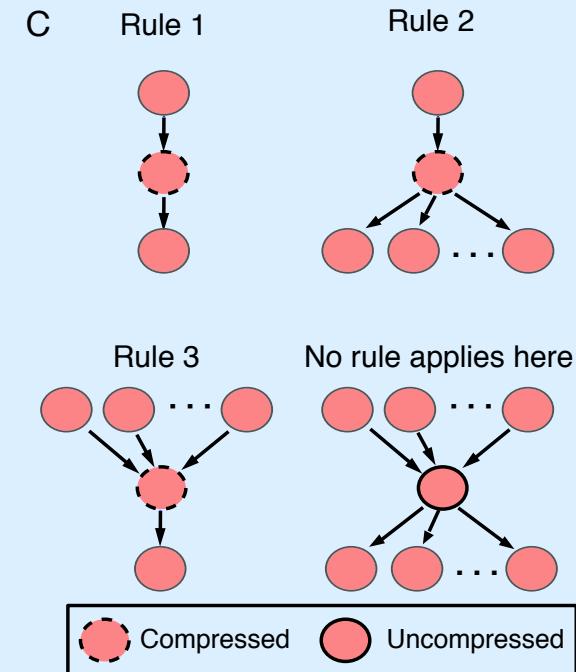
Saez-Rodriguez J, et al., Mol. Syst. Biol, 2009



# Model preprocessing:



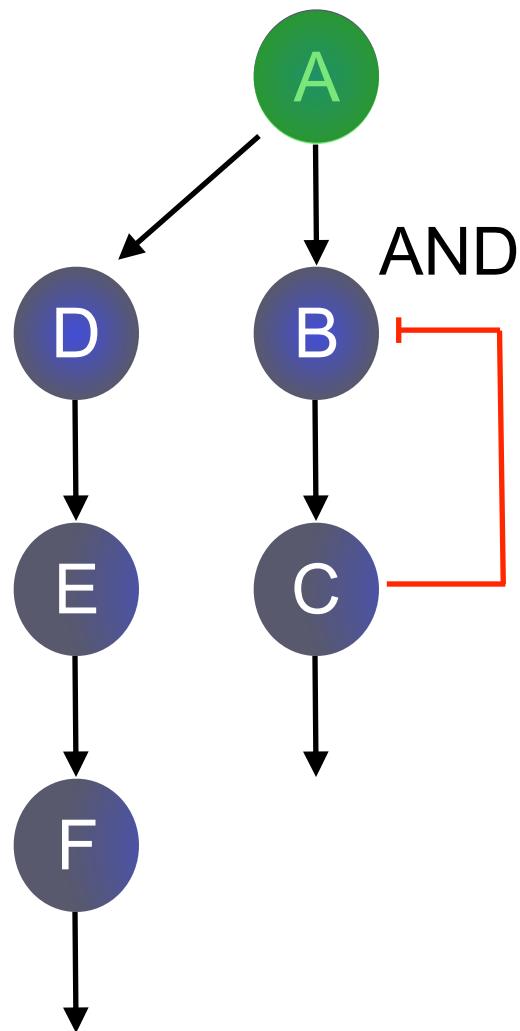
Model compression & removal of non-controlable & non-observable branches



Saez-Rodriguez J, et al., Mol. Syst. Biol, 2009

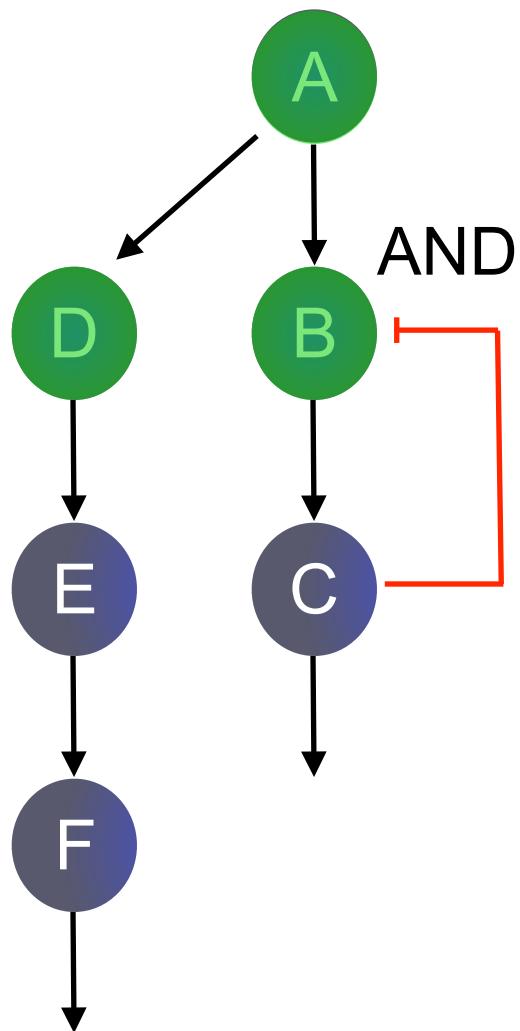


# Boolean simulation performed using pseudo-steady state



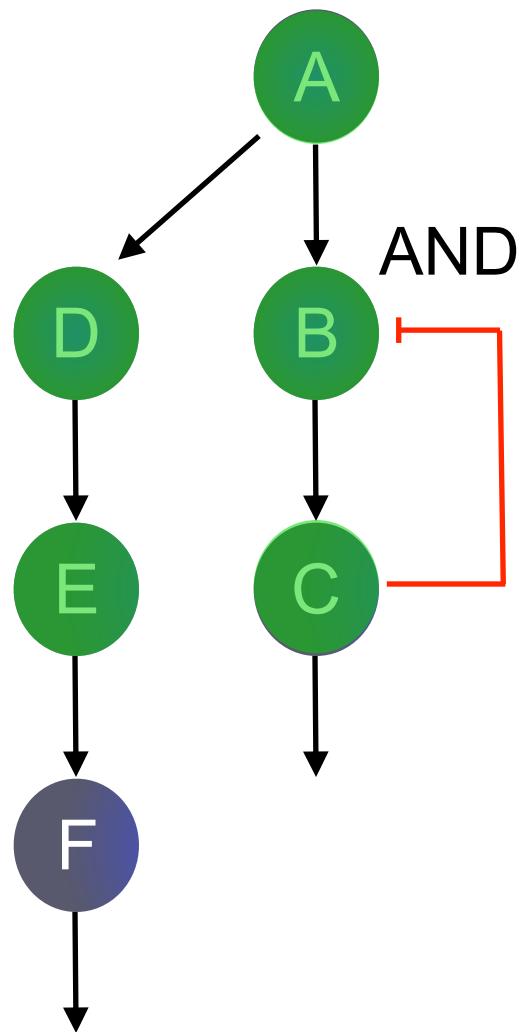


# Boolean simulation performed using pseudo-steady state



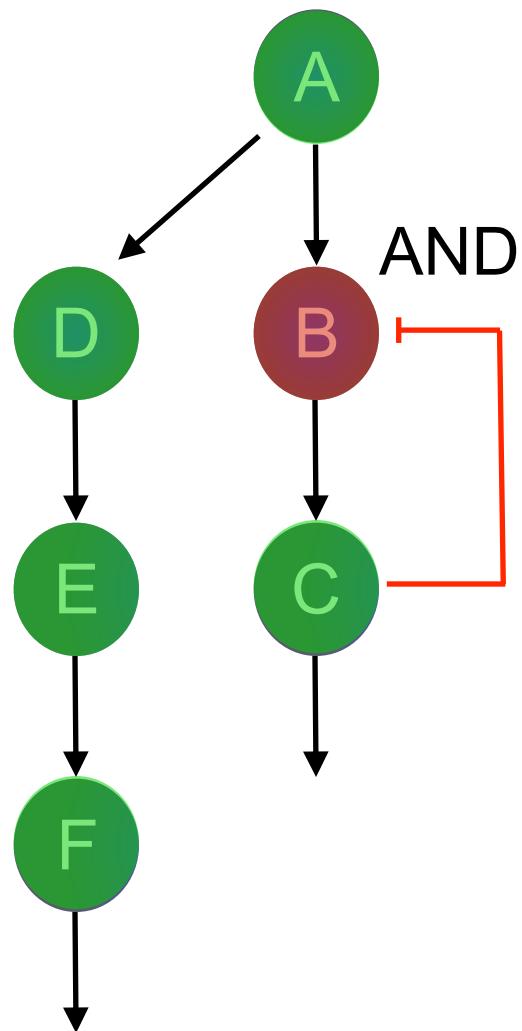


# Boolean simulation performed using pseudo-steady state



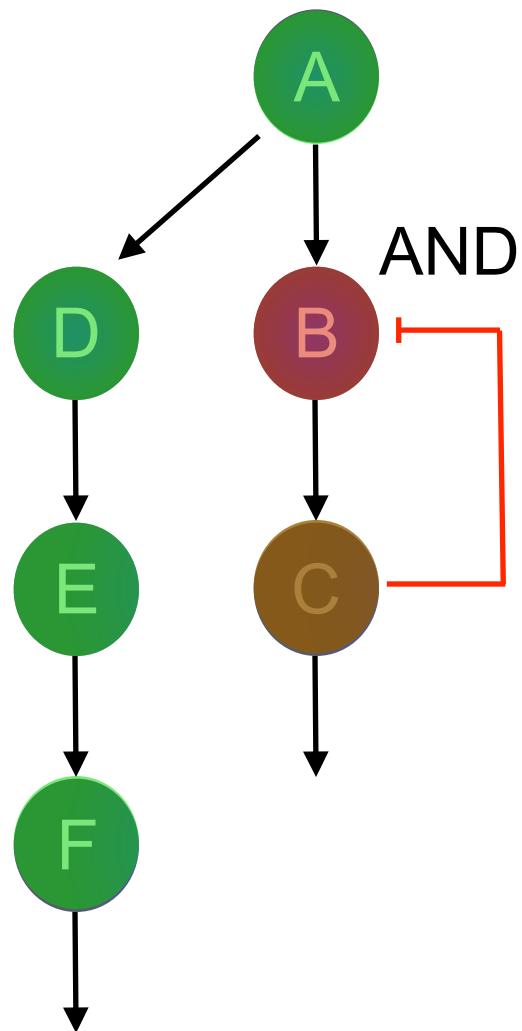


# Boolean simulation performed using pseudo-steady state



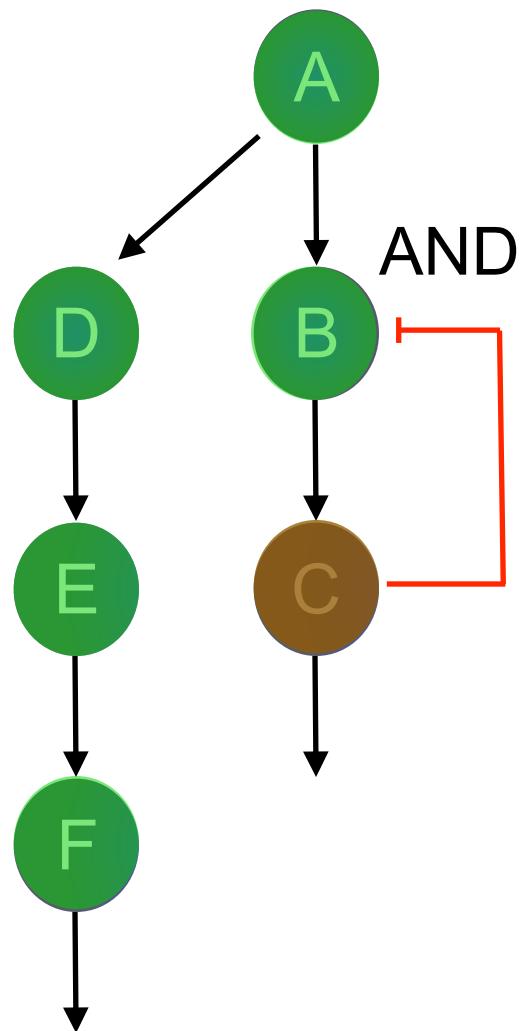


# Boolean simulation performed using pseudo-steady state



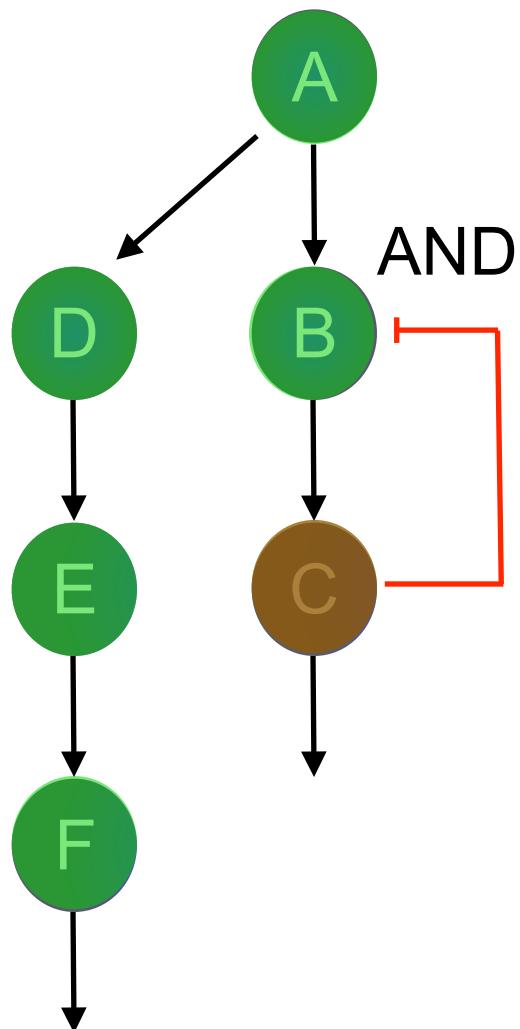


# Boolean simulation performed using pseudo-steady state





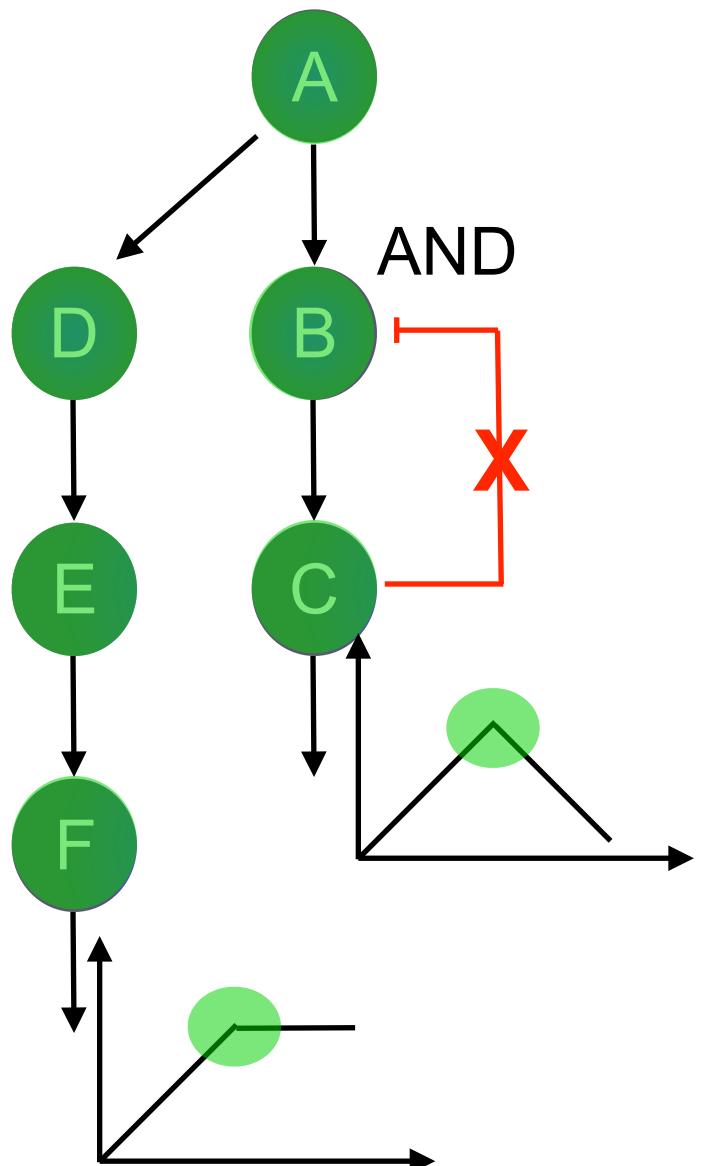
# Boolean simulation performed using pseudo-steady state



Algorithm penalizes lack of steady state,  
only effective for one 'early' time



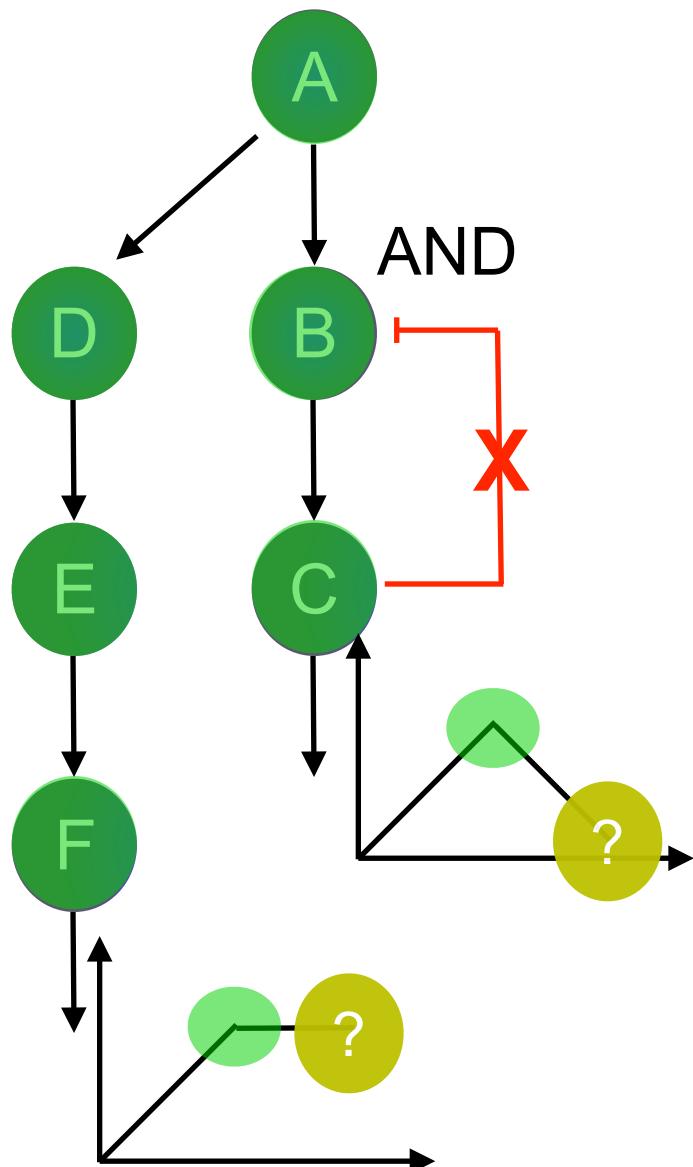
# Boolean simulation performed using pseudo-steady state



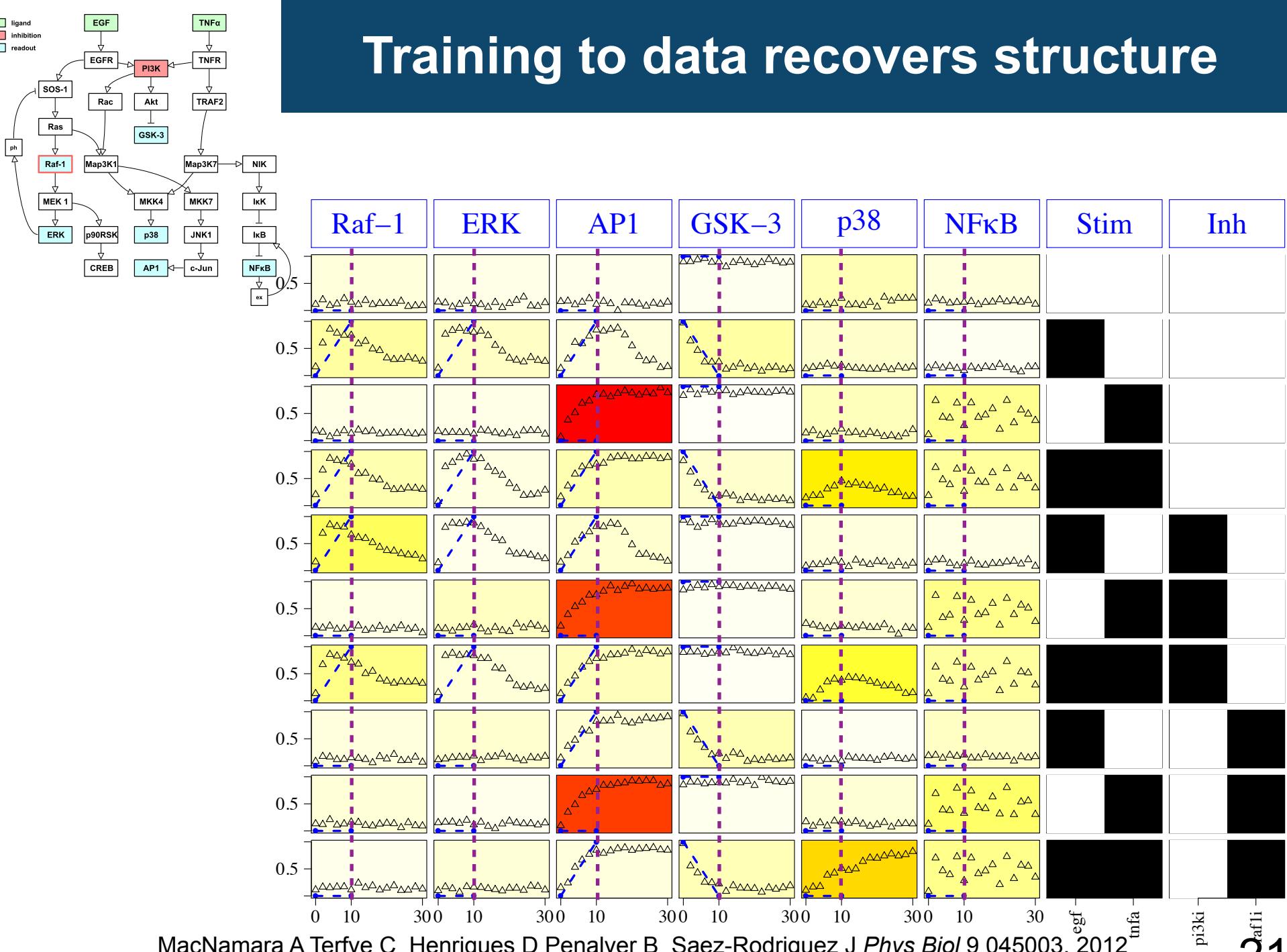
Algorithm penalizes lack of steady state,  
only effective for one 'early' time



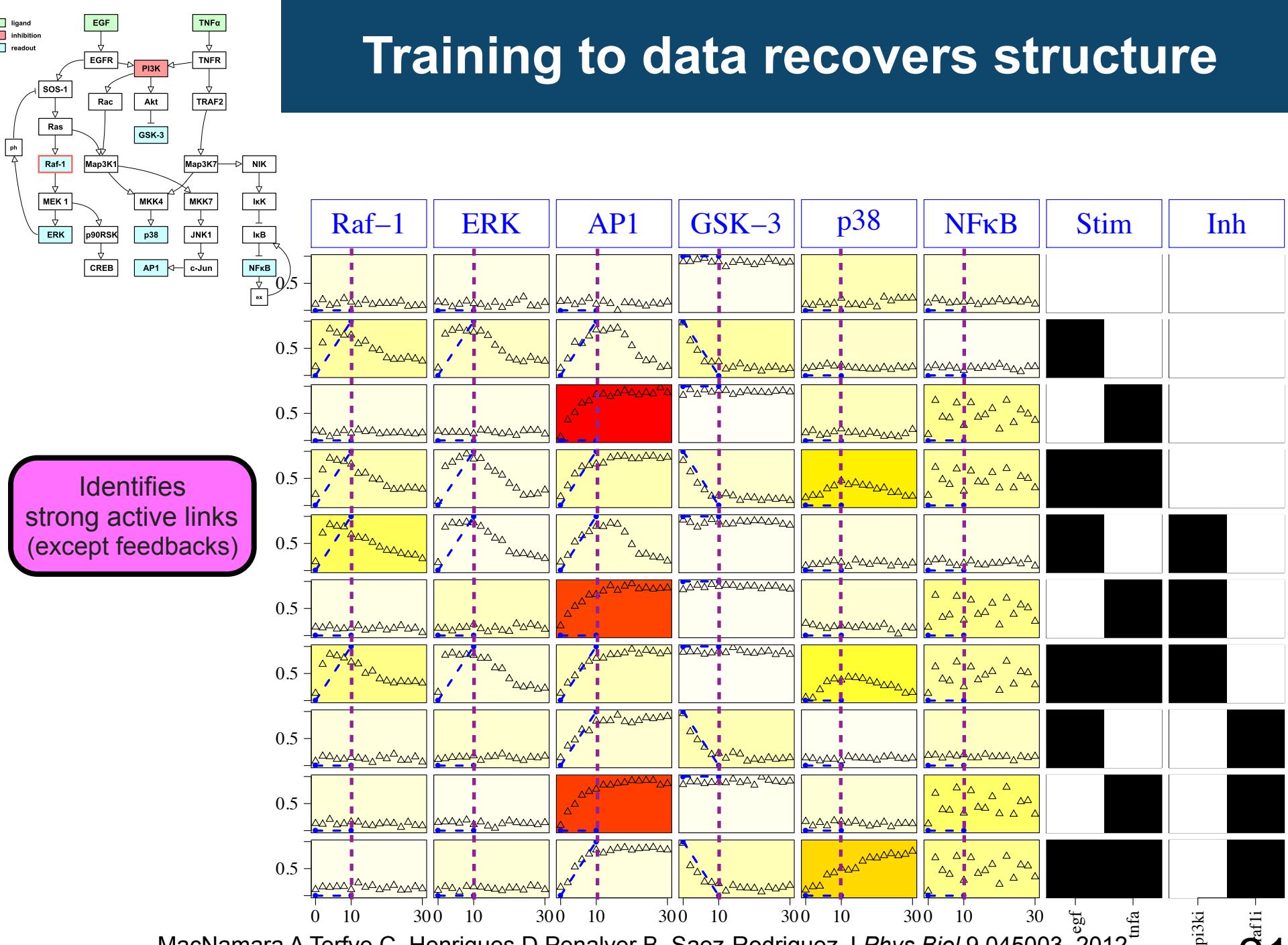
# Boolean simulation performed using pseudo-steady state



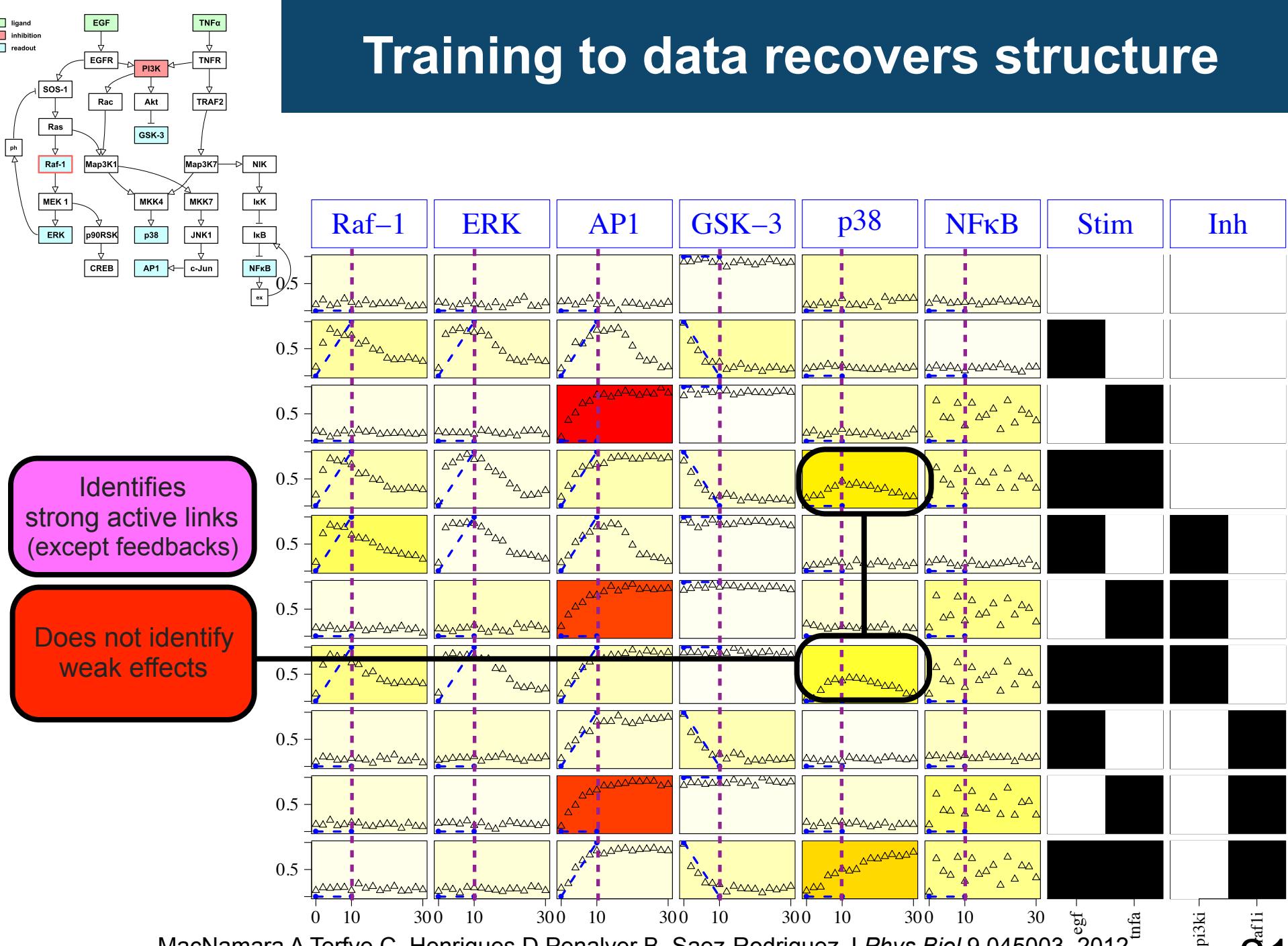
Algorithm penalizes lack of steady state,  
only effective for one 'early' time



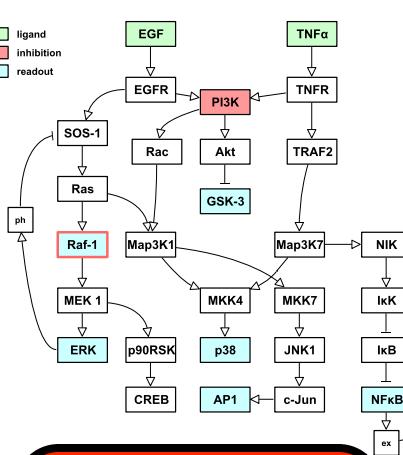
# Training to data recovers structure



# Training to data recovers structure



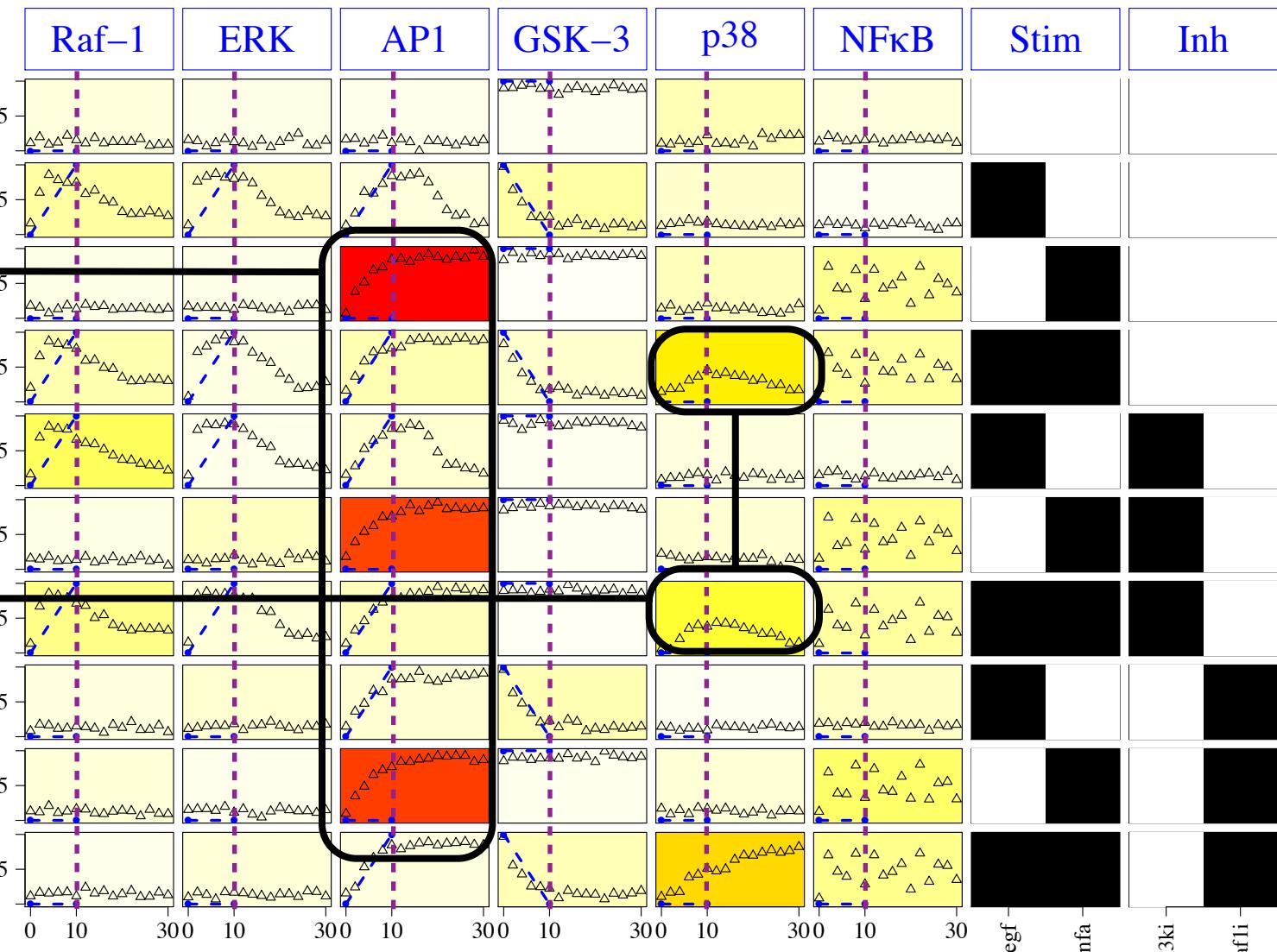
# Training to data recovers structure



Can not explain data due to missing links

Identifies strong active links (except feedbacks)

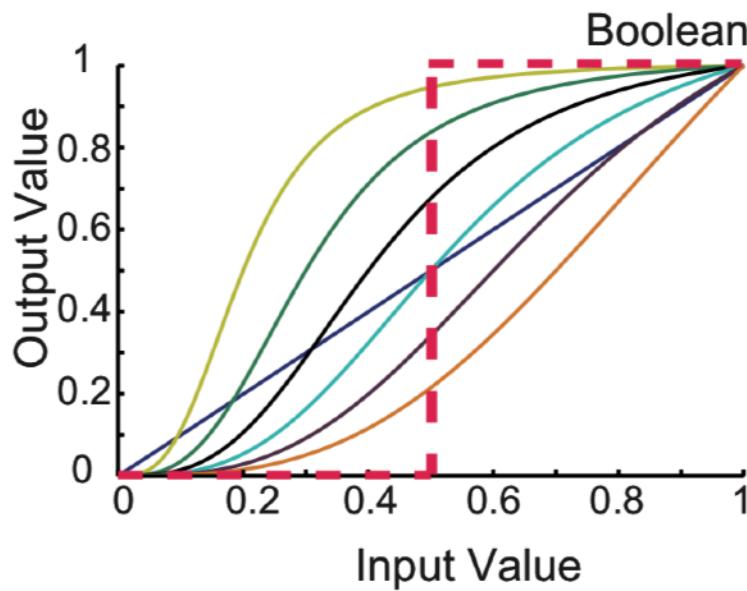
Does not identify weak effects



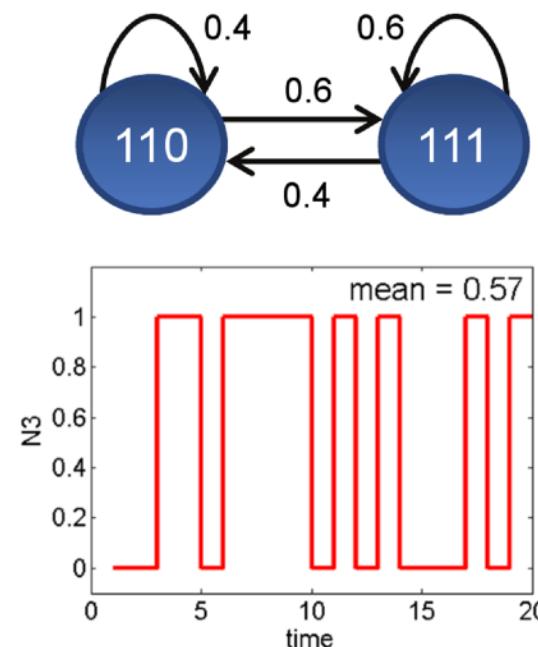


# Constrained Fuzzy logic and Probabilistic Logic can handle quantitative differences

- Boolean modeling can **not** describe **quantitative** aspect (e.g. intermediate activation)
- Fuzzy logic (Aldridge et al. Plos Comp. Bio. 2009; Morris et al. Plos Comp. Bio. 2011) and Probabilistic Logic (Trairatphisan et al. Plos One 2014) can model quantitative signalling data



- **Constrained Fuzzy logic**

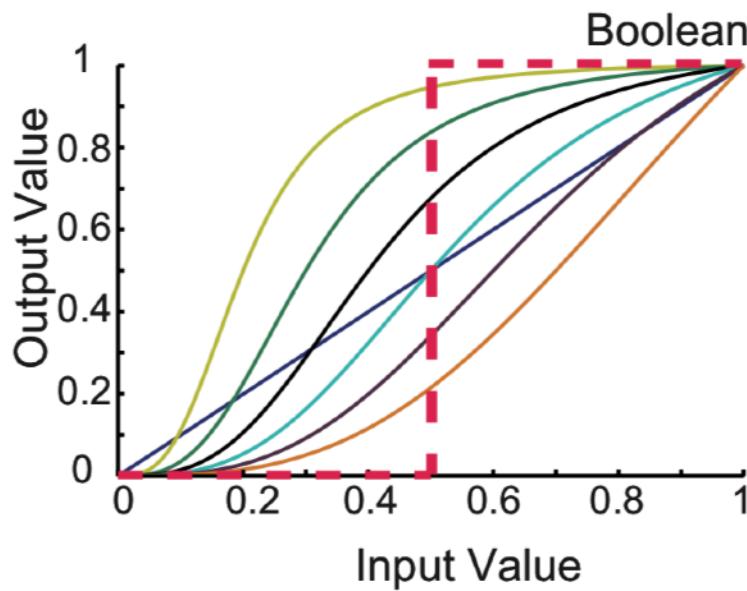


- **Probabilistic Logic**

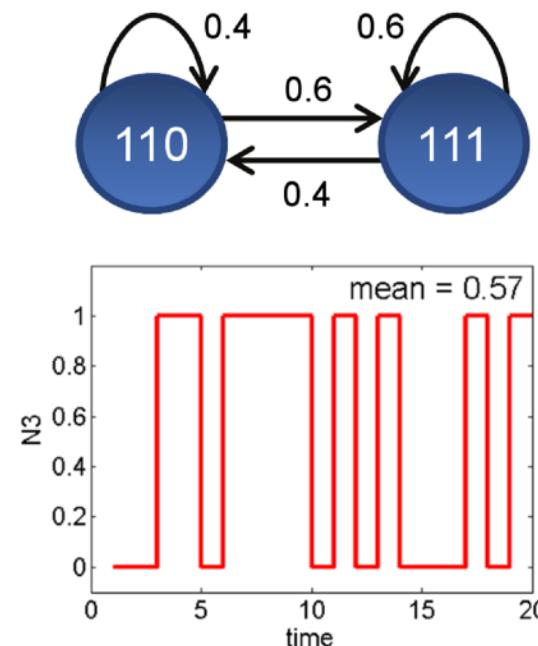


# Constrained Fuzzy logic and Probabilistic Logic can handle quantitative differences

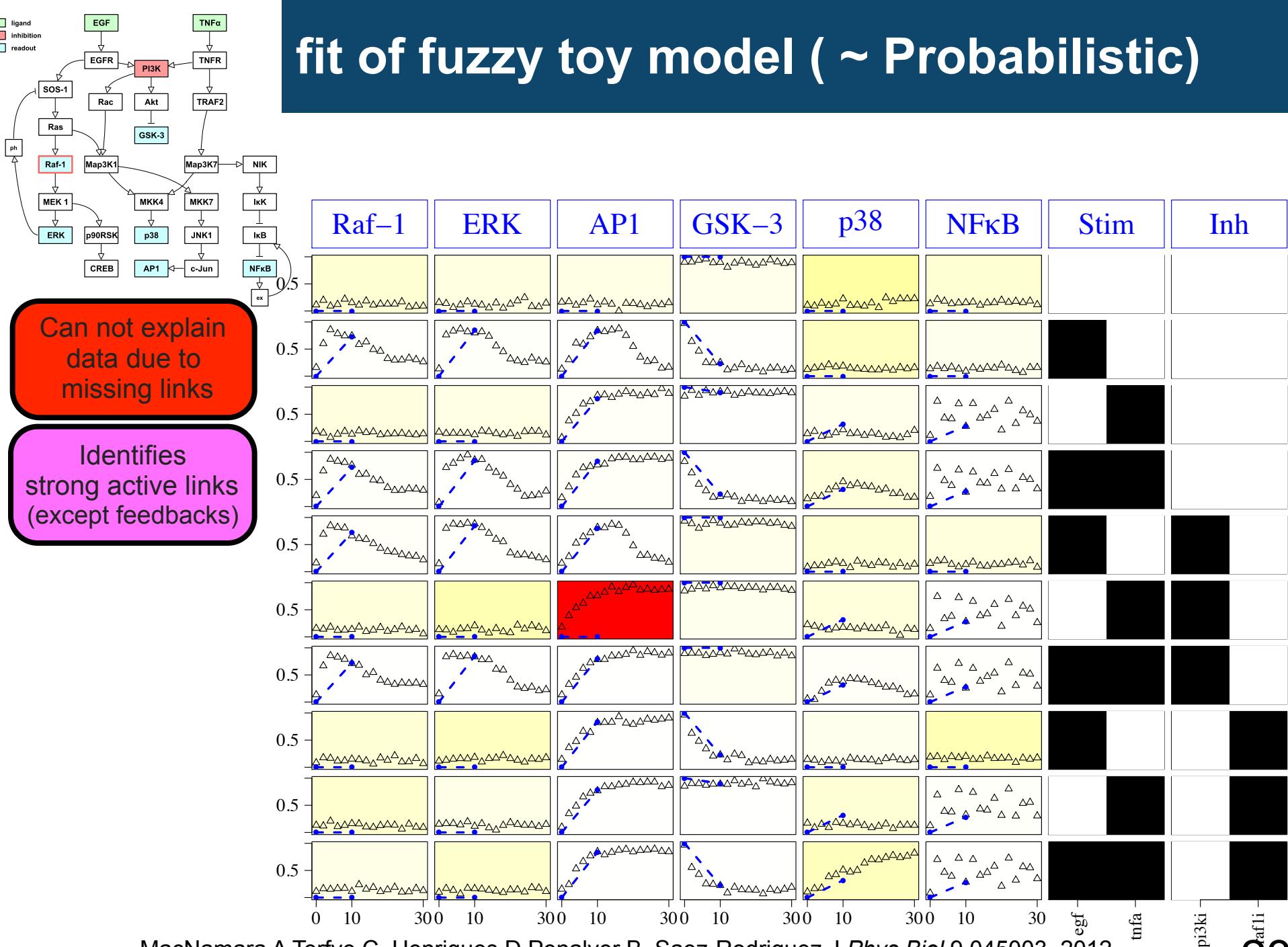
- Boolean modeling can **not** describe **quantitative** aspect (e.g. intermediate activation)
- Fuzzy logic (Aldridge et al. Plos Comp. Bio. 2009; Morris et al. Plos Comp. Bio. 2011) and Probabilistic Logic (Trairatphisan et al. Plos One 2014) can model quantitative signalling data

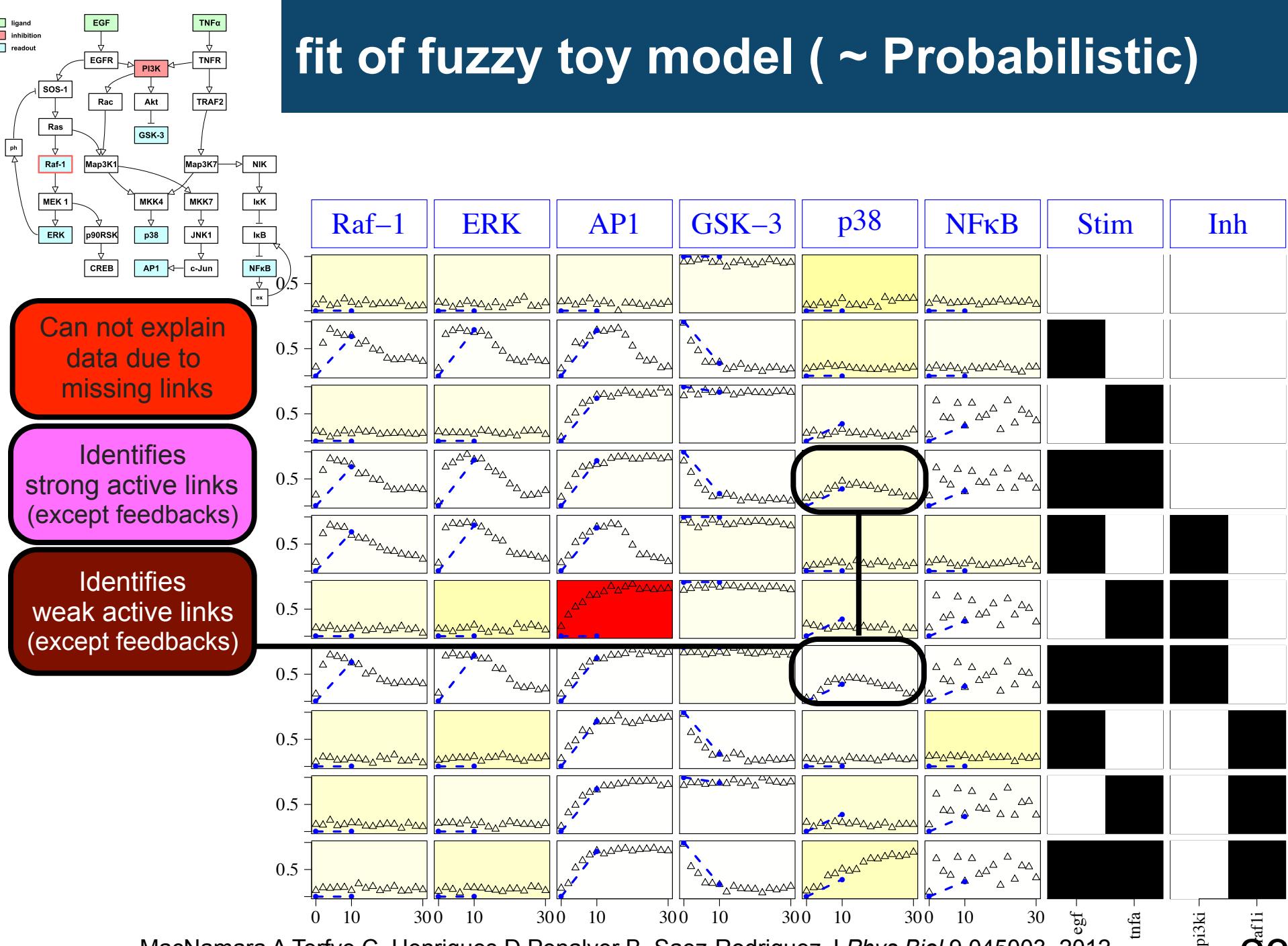


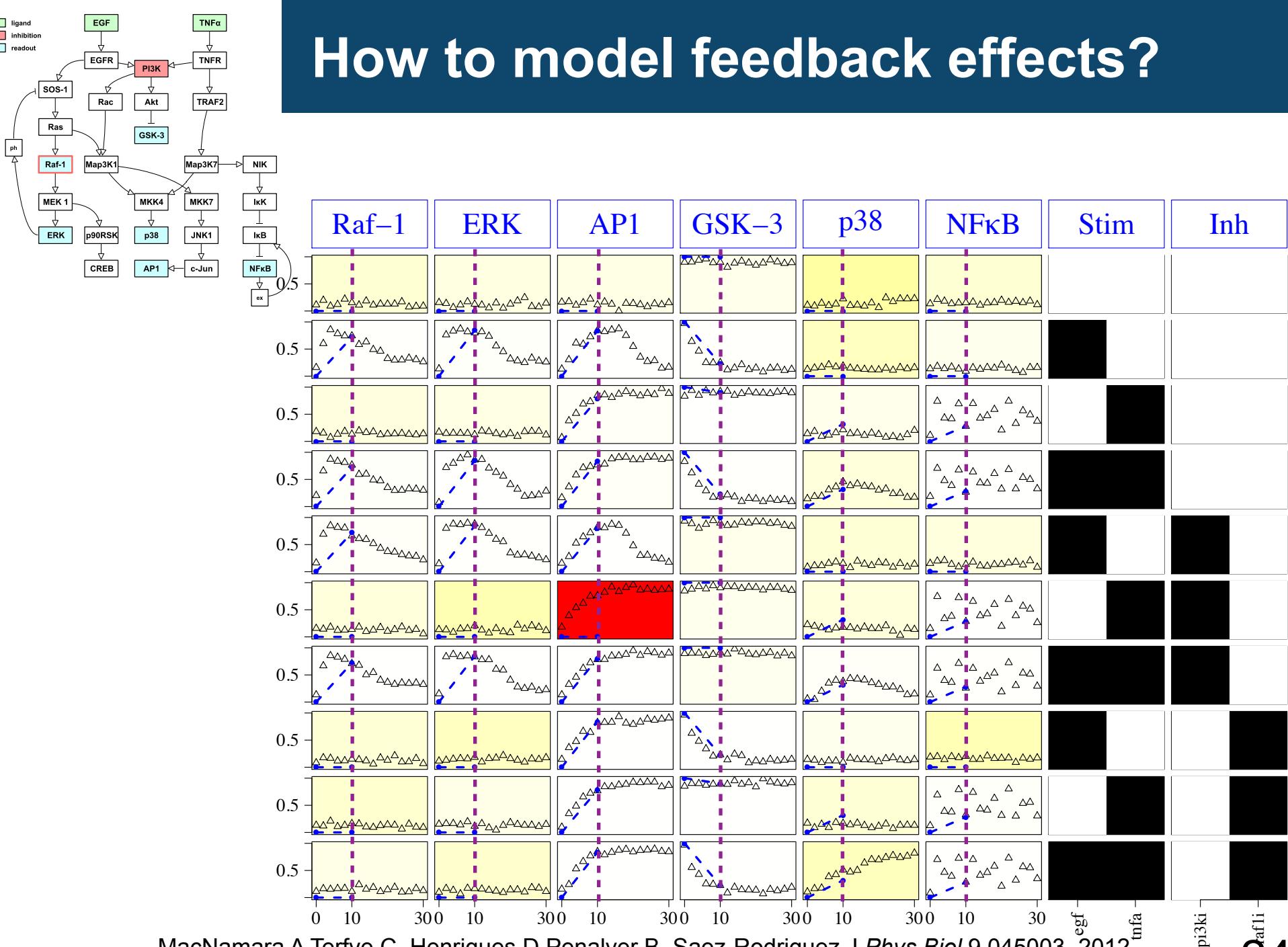
- **Constrained Fuzzy logic**

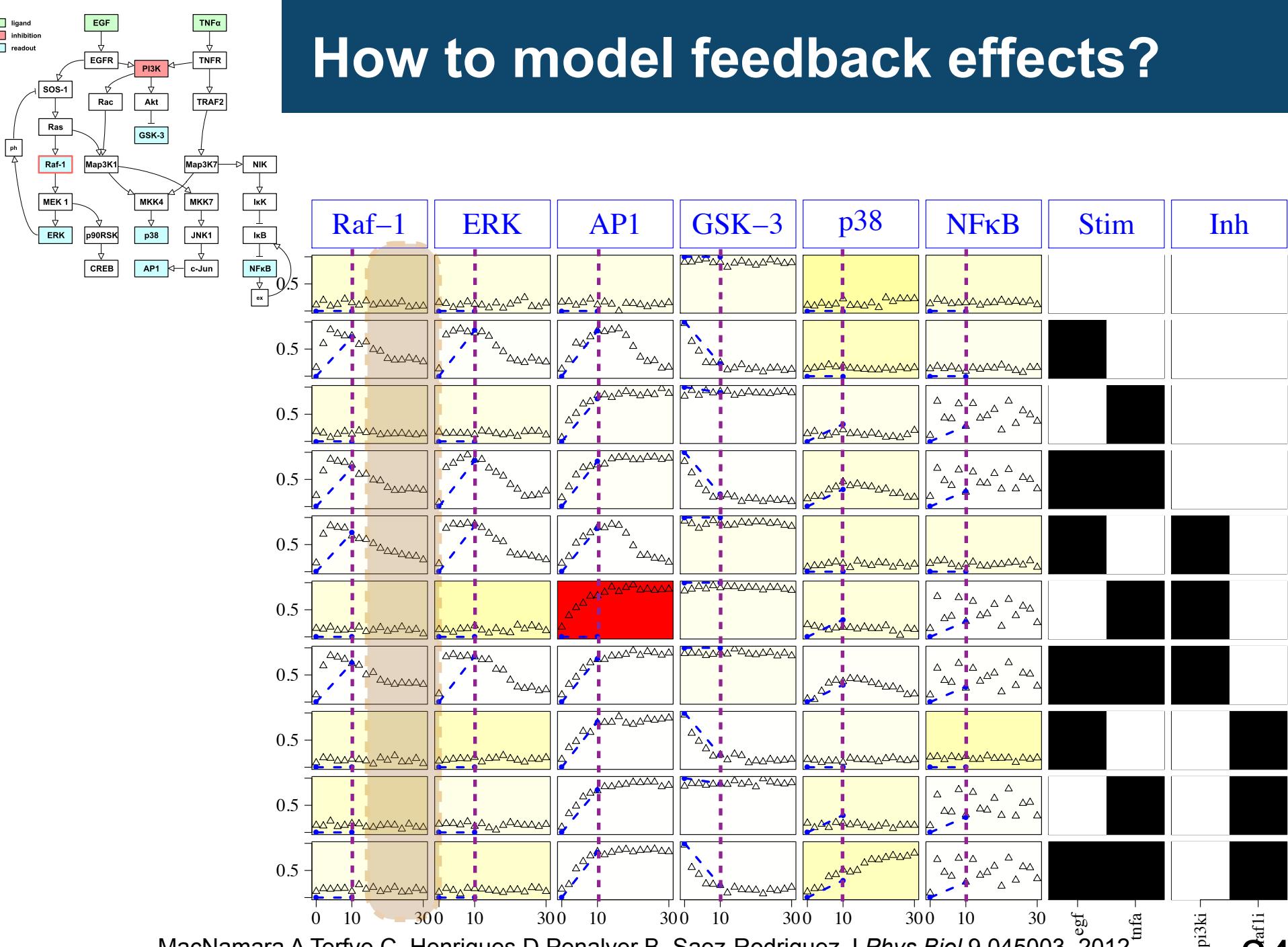


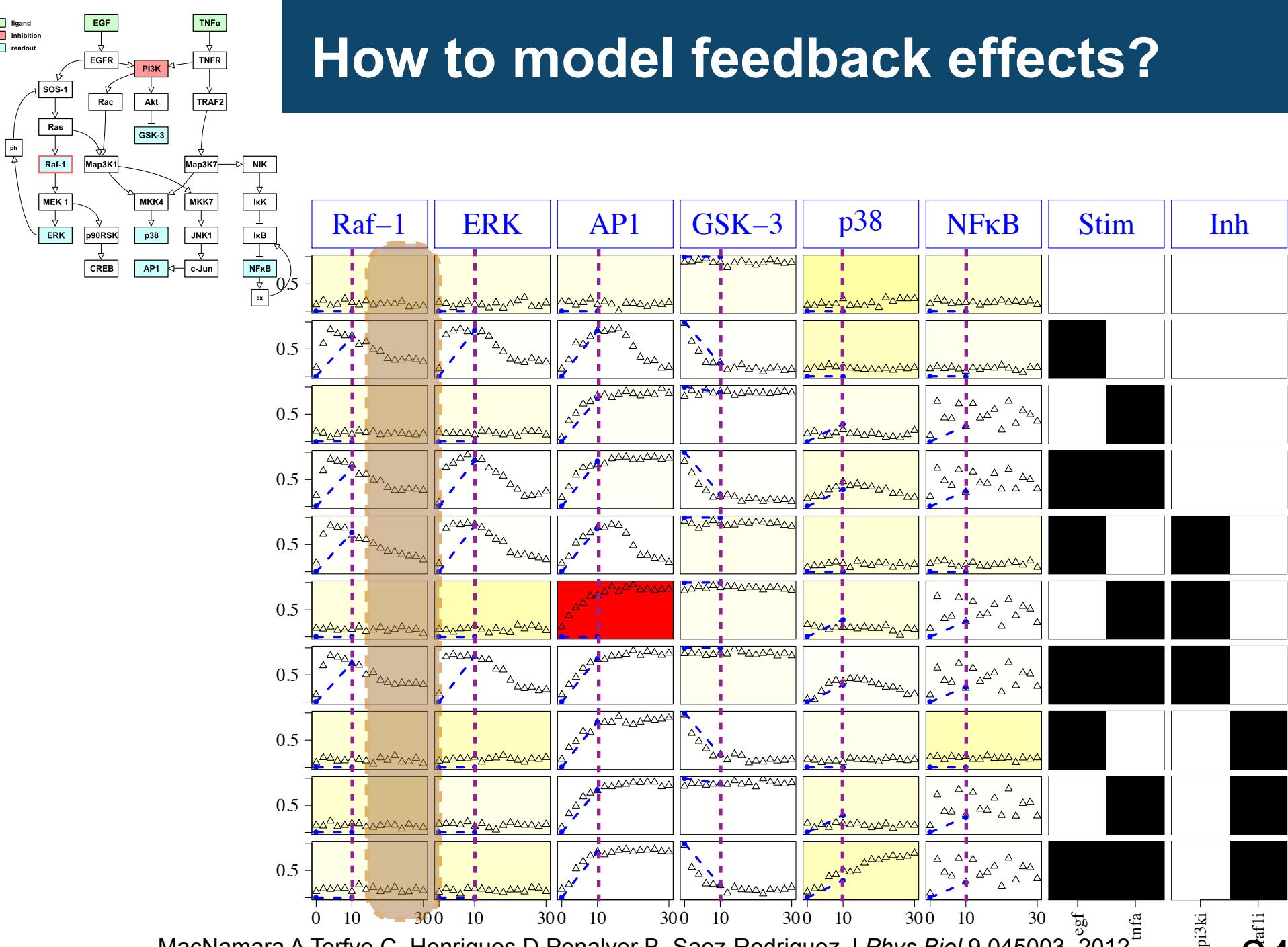
- **Probabilistic Logic**





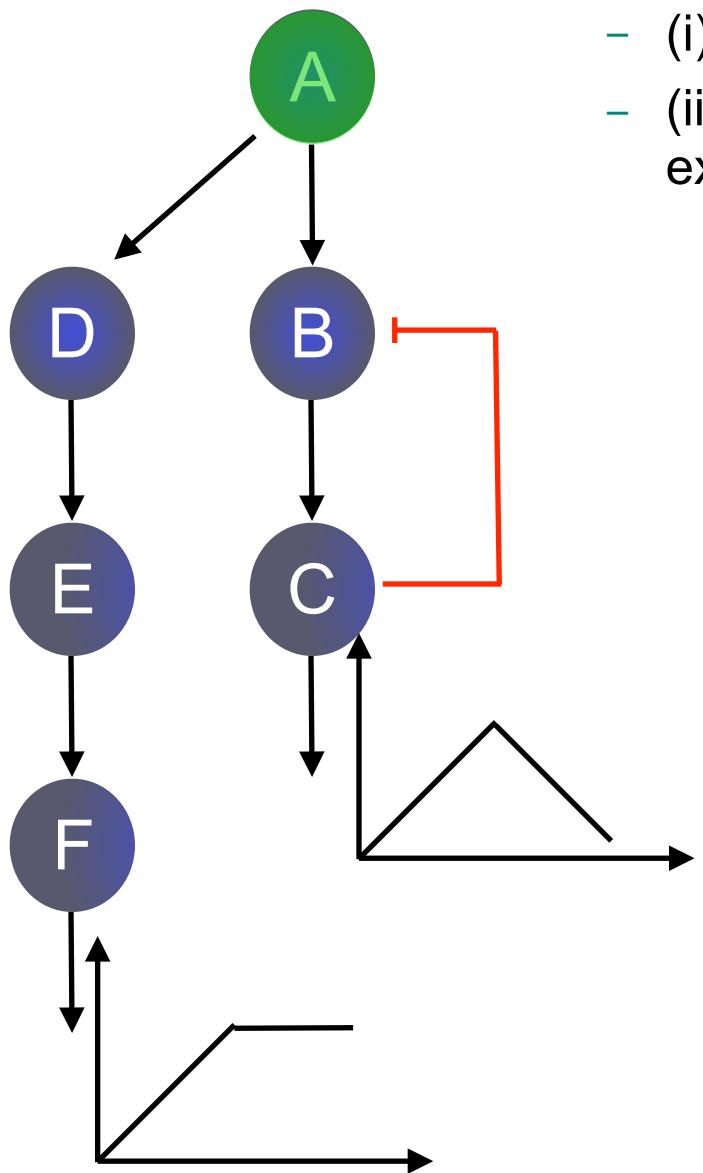








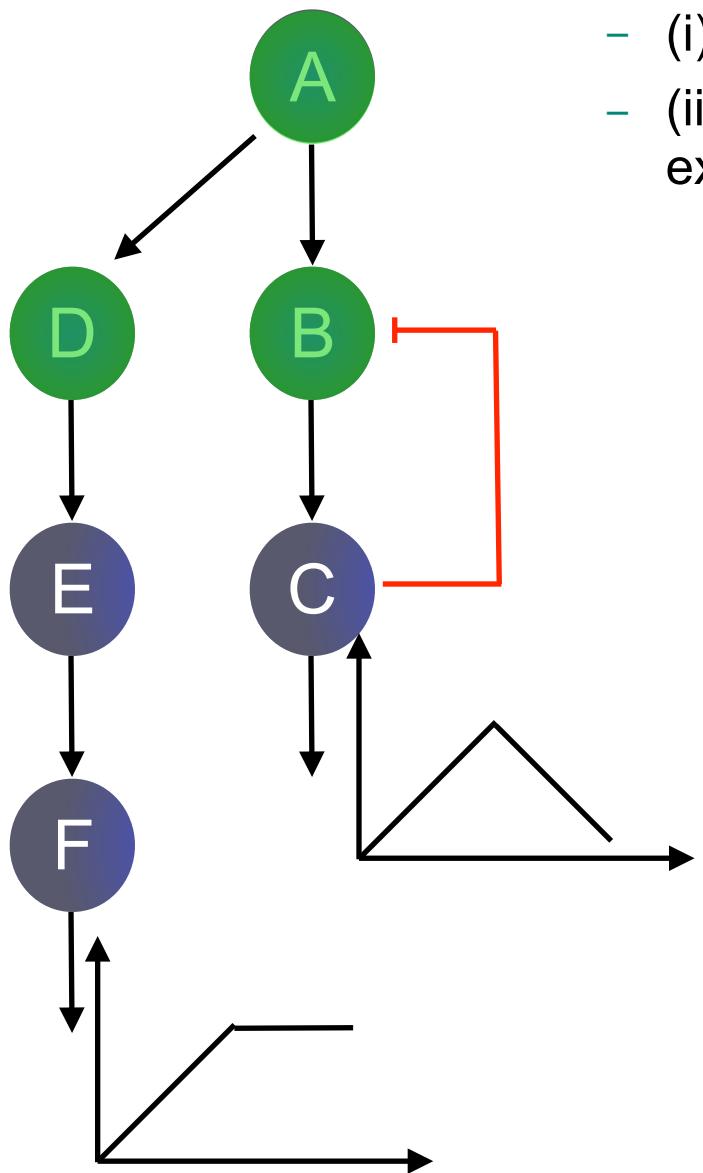
# Approximation of transient behaviour using multiple time-scales



- (i) Train  $\tau = 1 \rightarrow$  get early events
- (ii) Train  $\tau = 2 \rightarrow$  find gates not active at  $\tau = 1$  that explain evolution from  $\tau = 1$  to  $\tau = 2$



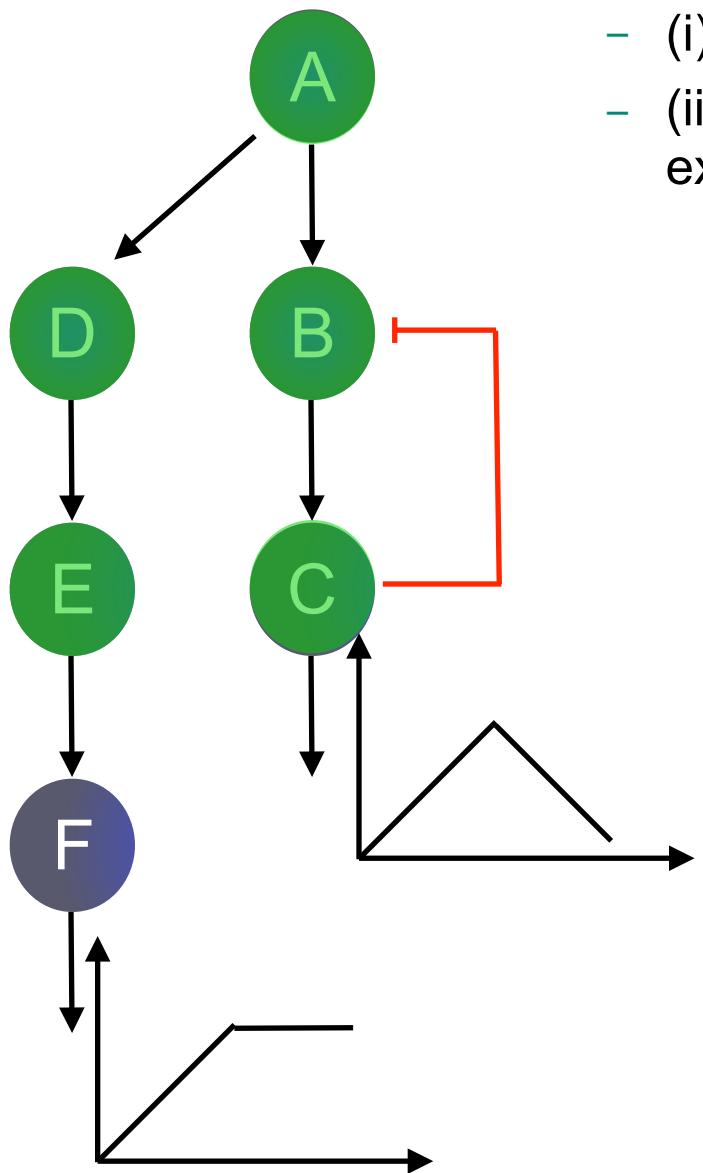
# Approximation of transient behaviour using multiple time-scales



- (i) Train  $\tau = 1 \rightarrow$  get early events
- (ii) Train  $\tau = 2 \rightarrow$  find gates not active at  $\tau = 1$  that explain evolution from  $\tau = 1$  to  $\tau = 2$



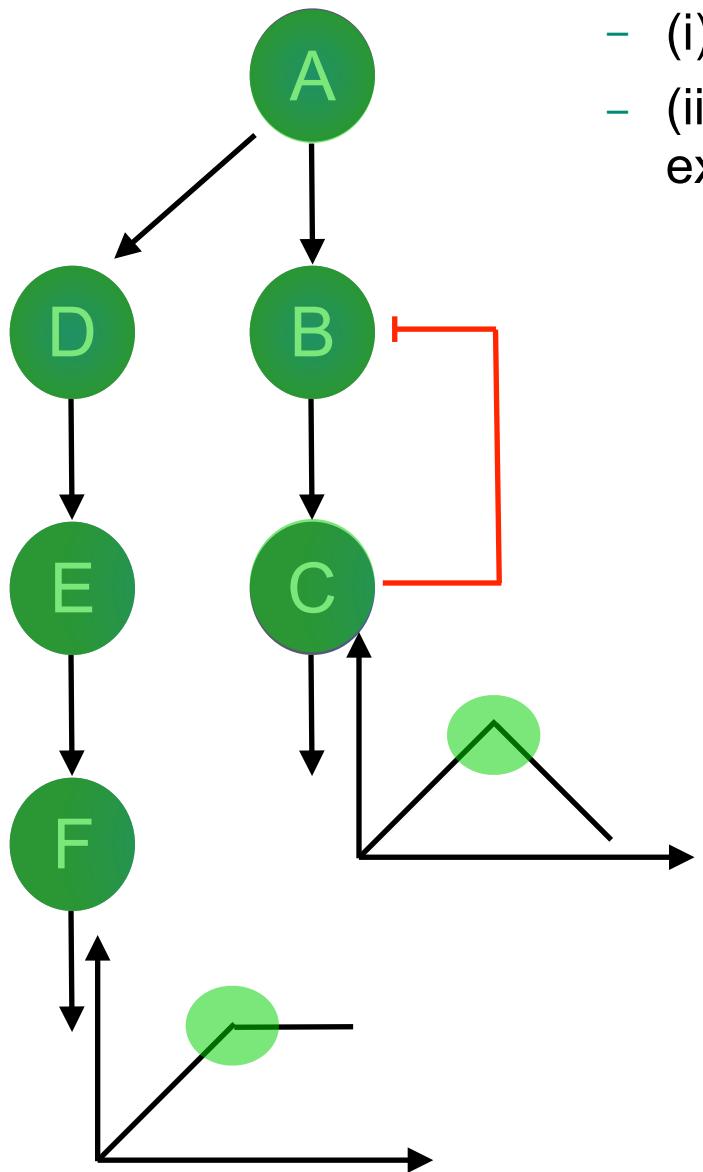
# Approximation of transient behaviour using multiple time-scales



- (i) Train  $\tau = 1 \rightarrow$  get early events
- (ii) Train  $\tau = 2 \rightarrow$  find gates not active at  $\tau = 1$  that explain evolution from  $\tau = 1$  to  $\tau = 2$



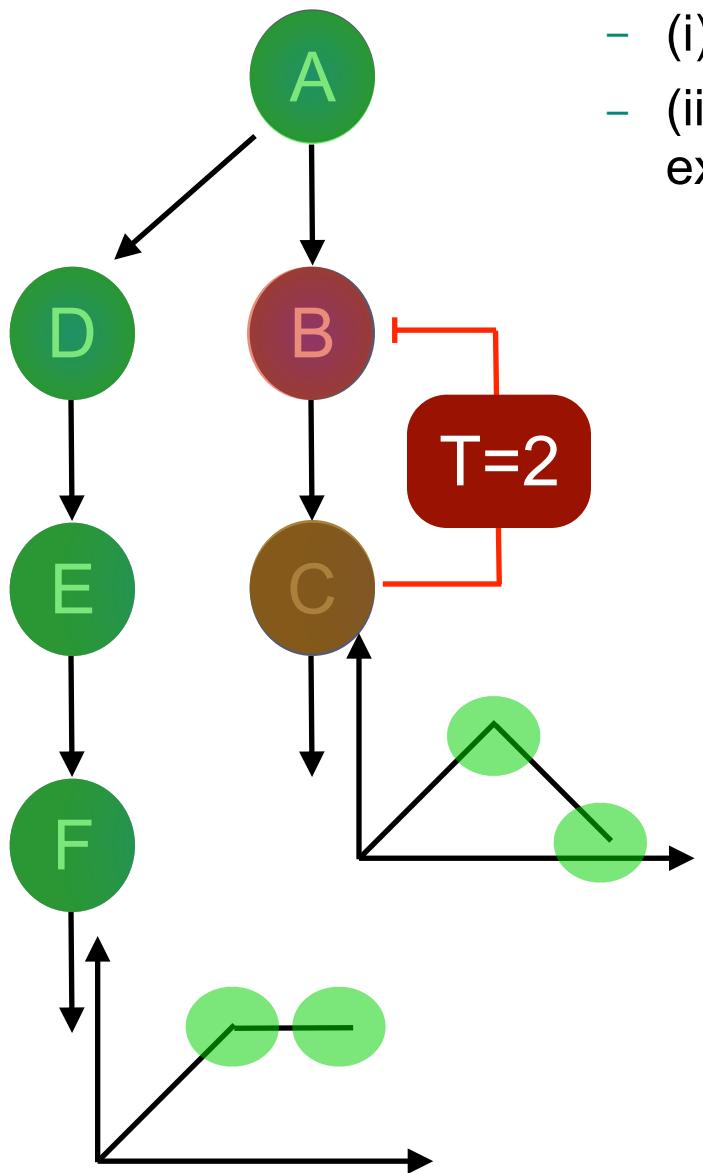
# Approximation of transient behaviour using multiple time-scales



- (i) Train  $\tau = 1 \rightarrow$  get early events
- (ii) Train  $\tau = 2 \rightarrow$  find gates not active at  $\tau = 1$  that explain evolution from  $\tau = 1$  to  $\tau = 2$



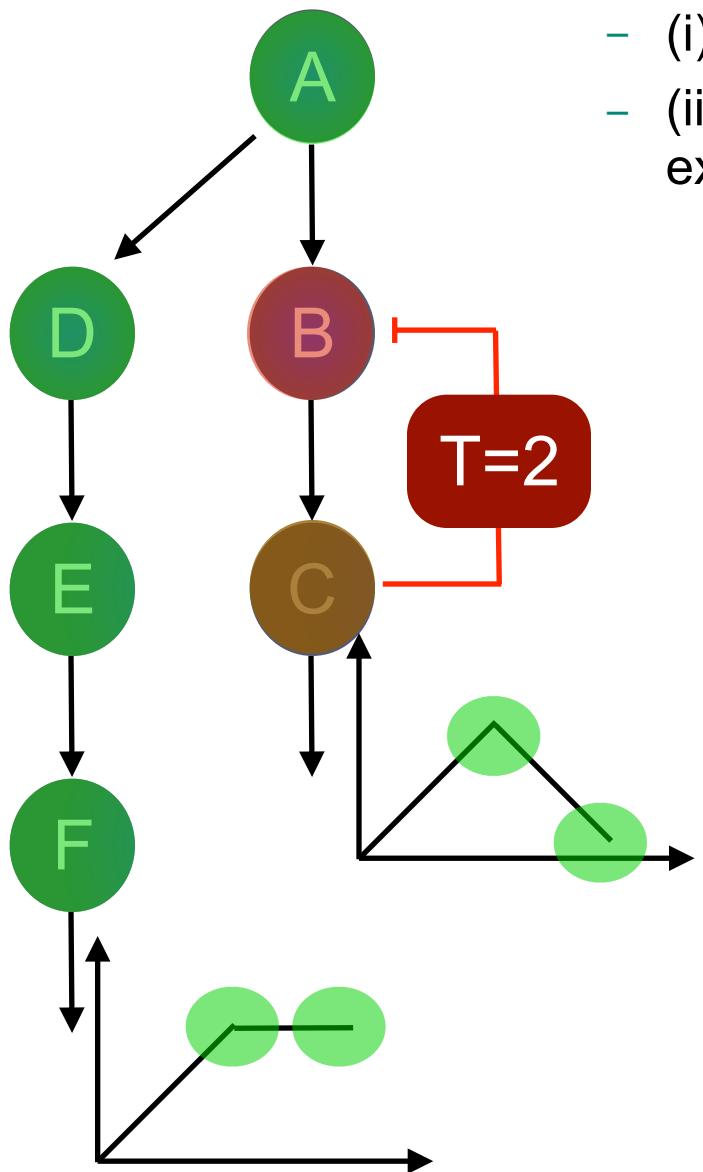
# Approximation of transient behaviour using multiple time-scales



- (i) Train  $\tau = 1 \rightarrow$  get early events
- (ii) Train  $\tau = 2 \rightarrow$  find gates not active at  $\tau = 1$  that explain evolution from  $\tau = 1$  to  $\tau = 2$

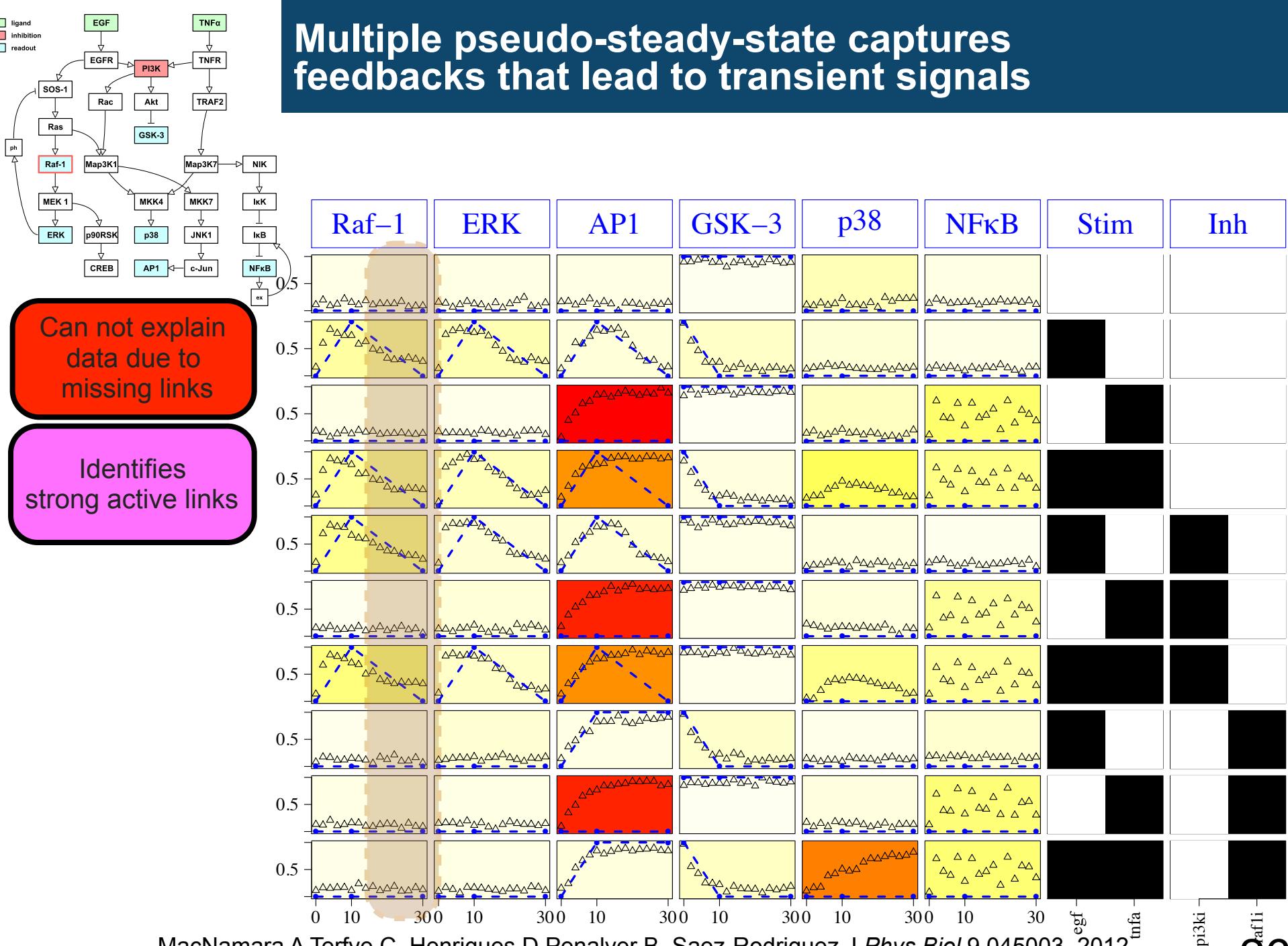


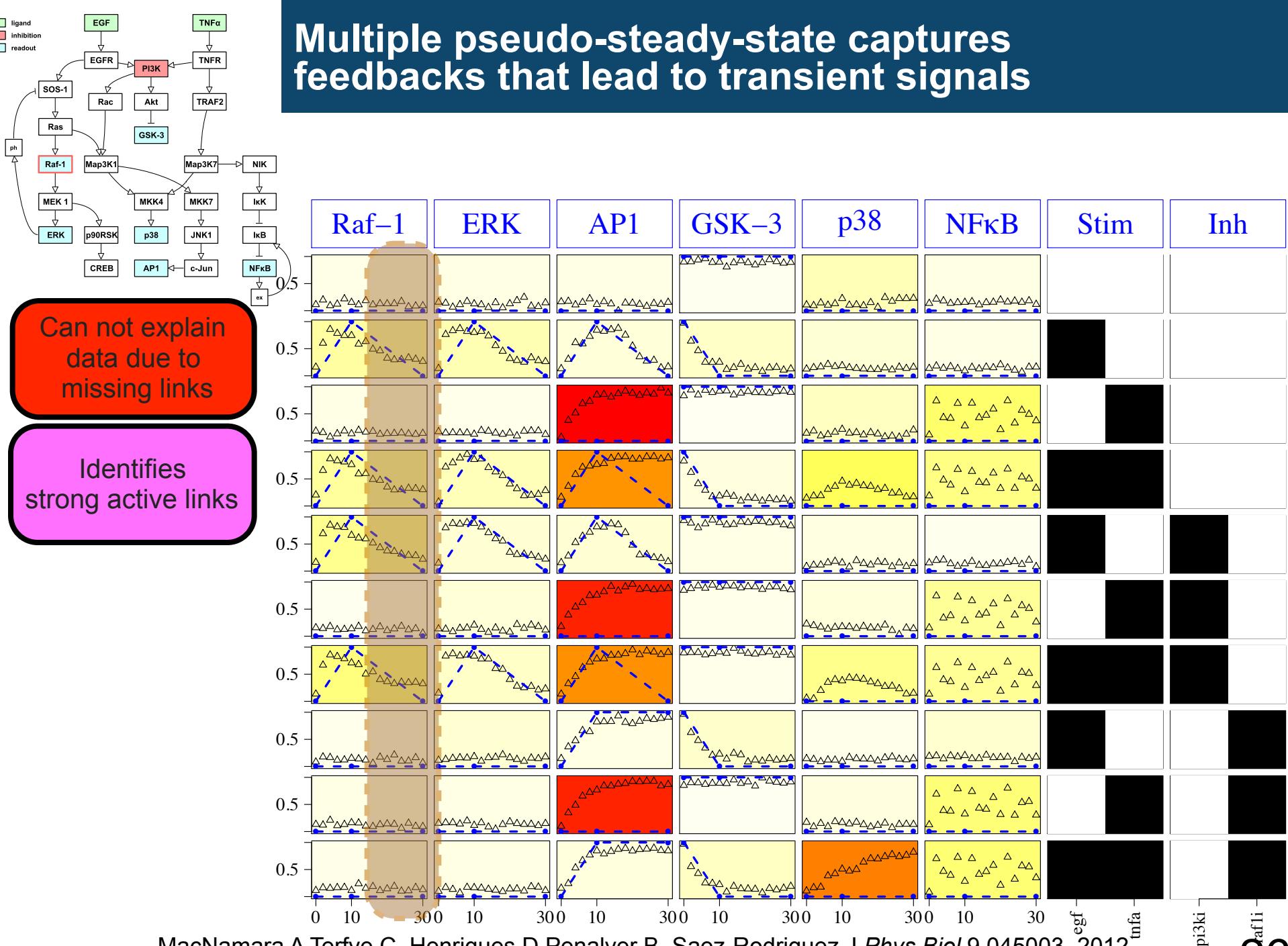
# Approximation of transient behaviour using multiple time-scales

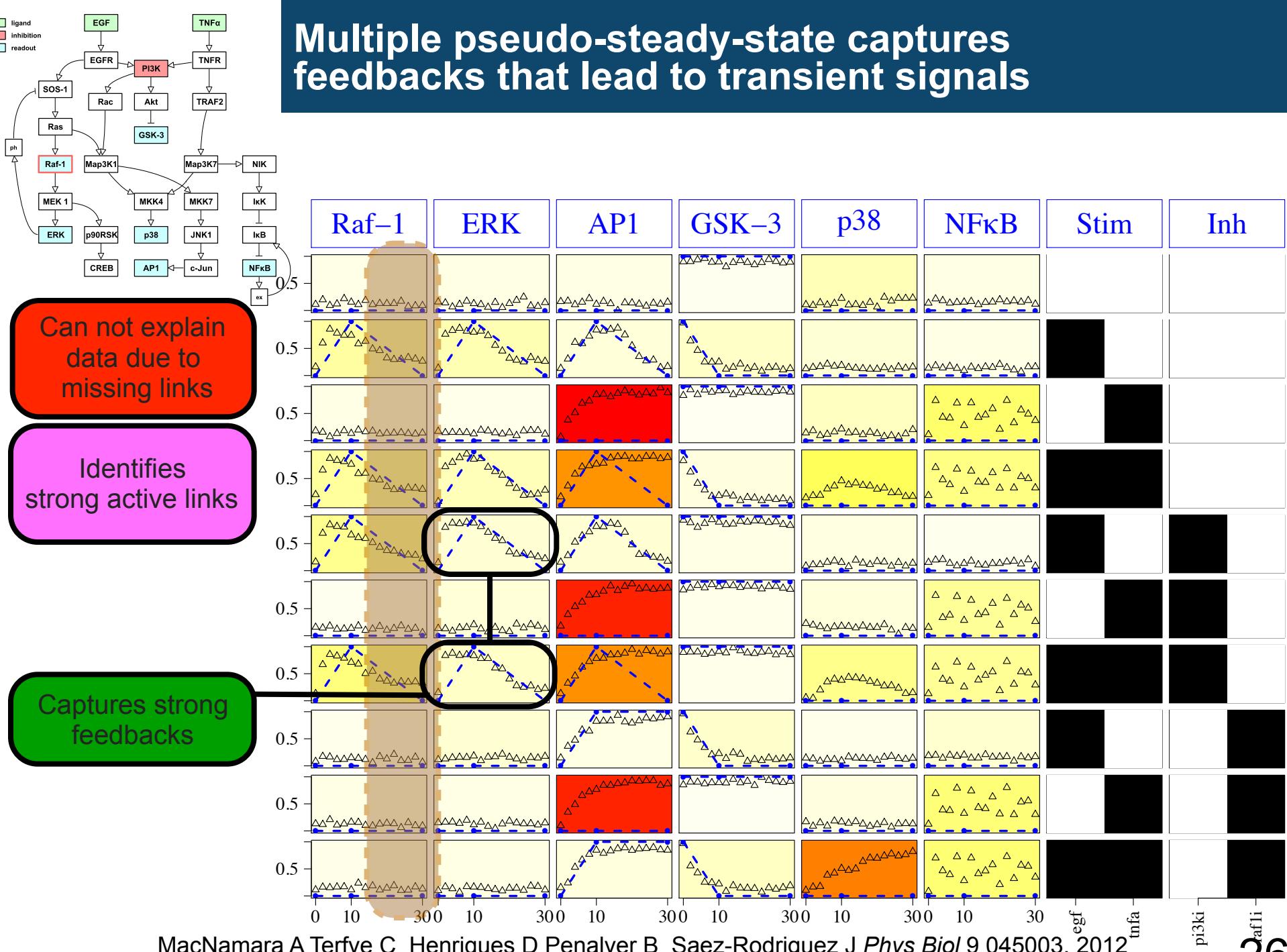


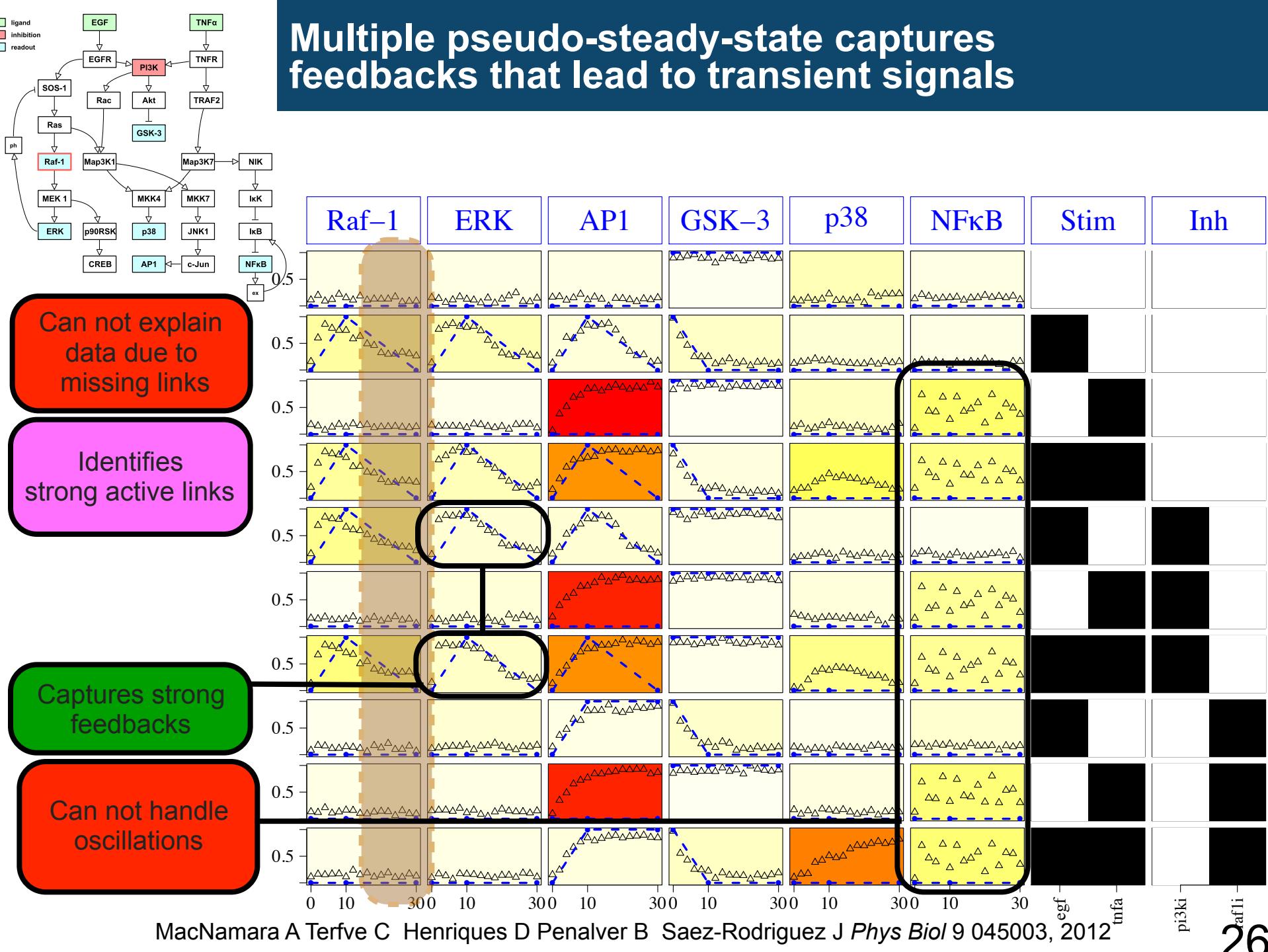
- (i) Train  $\tau = 1 \rightarrow$  get early events
- (ii) Train  $\tau = 2 \rightarrow$  find gates not active at  $\tau = 1$  that explain evolution from  $\tau = 1$  to  $\tau = 2$

Rough approximation of dynamics,  
still computationally efficient



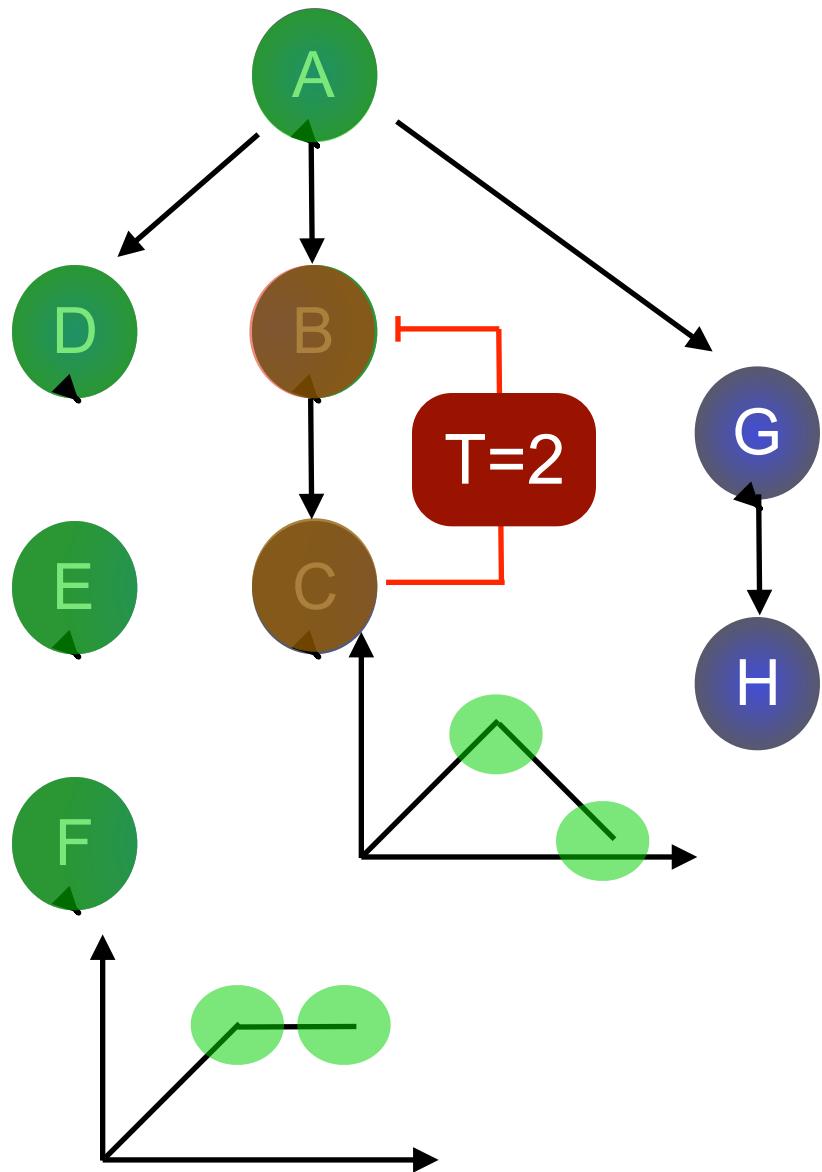




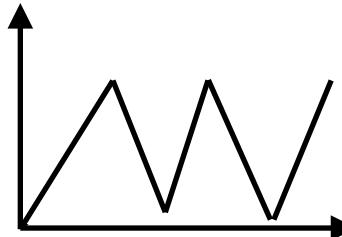




# Approximation of dynamics using synchronous simulation & multiple time-scales

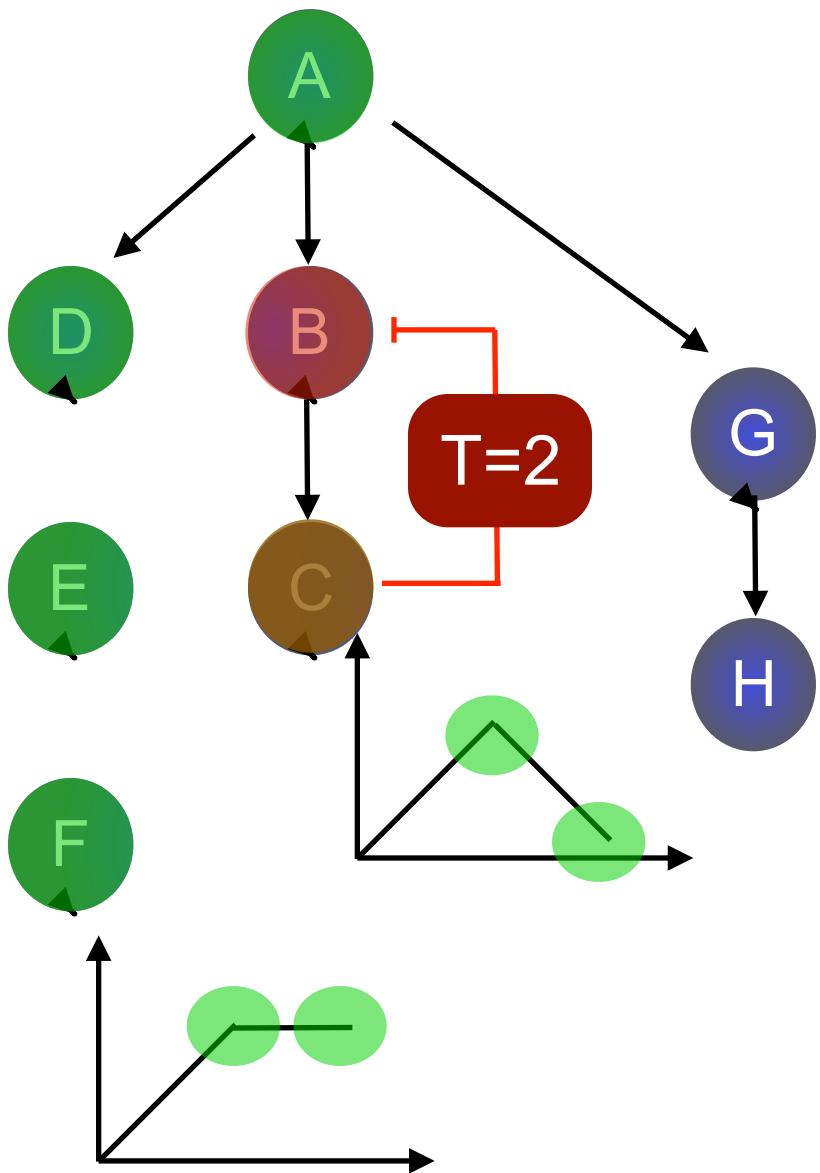


More recently: update to link to MaBoss

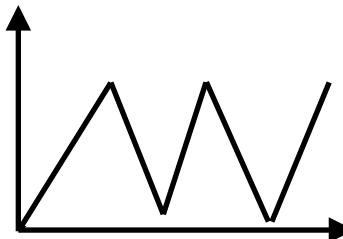




# Approximation of dynamics using synchronous simulation & multiple time-scales

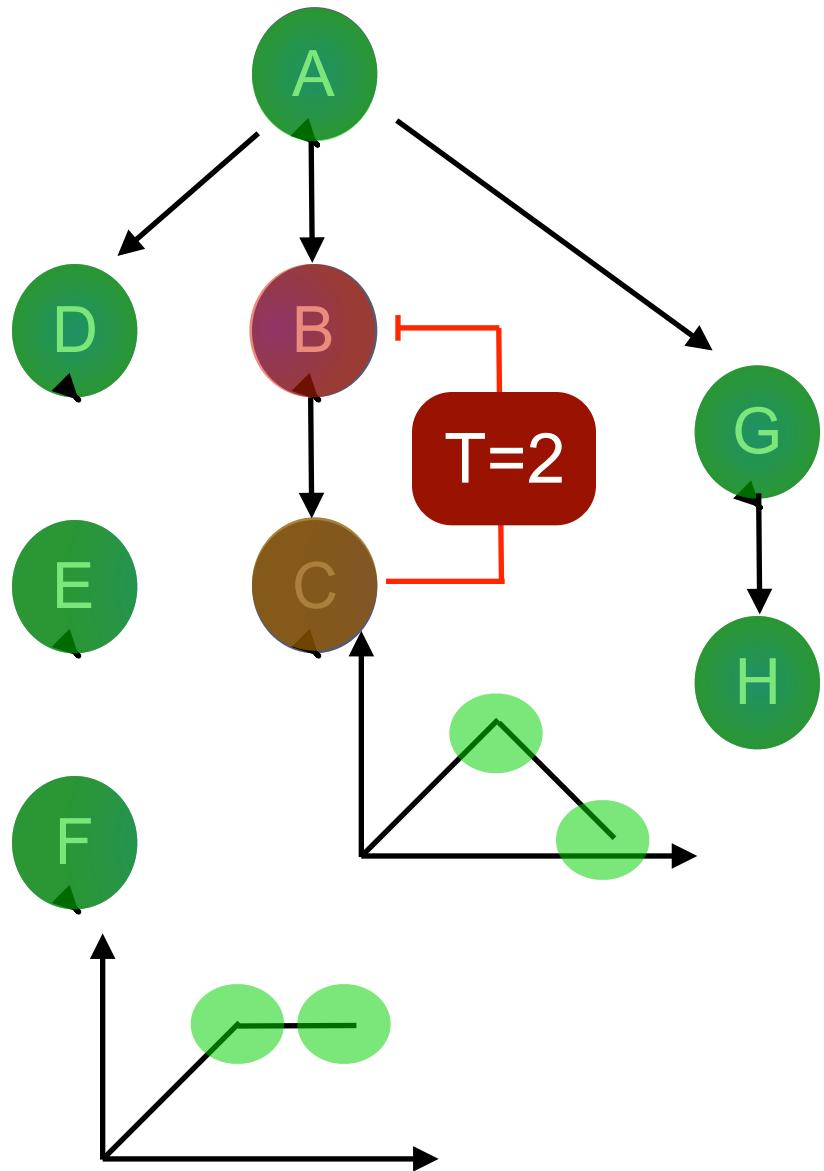


More recently: update to link to MaBoss

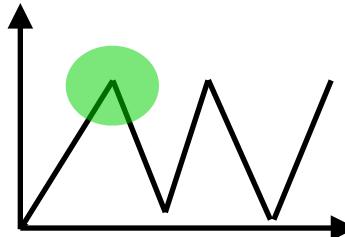




# Approximation of dynamics using synchronous simulation & multiple time-scales

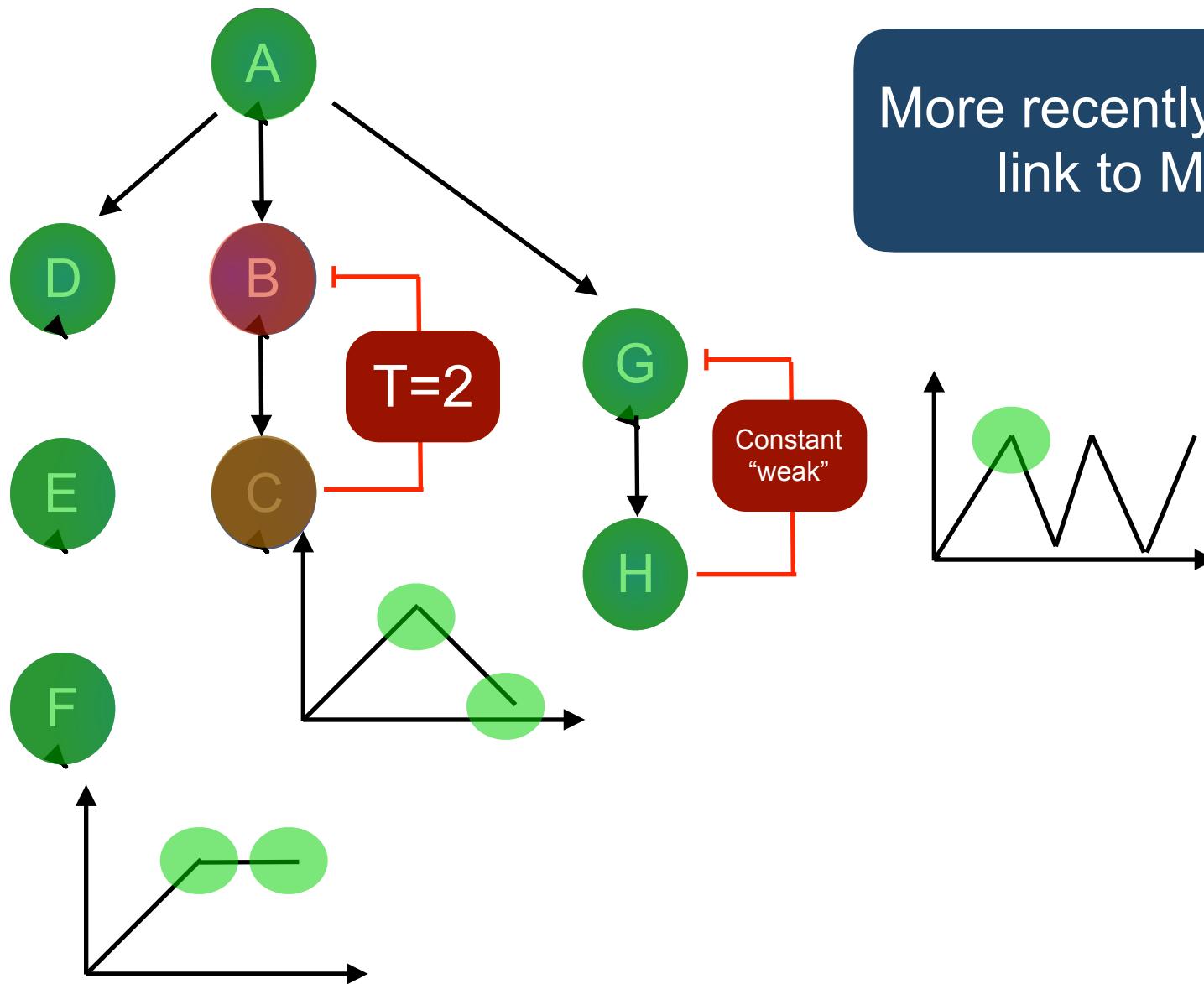


More recently: update to link to MaBoss





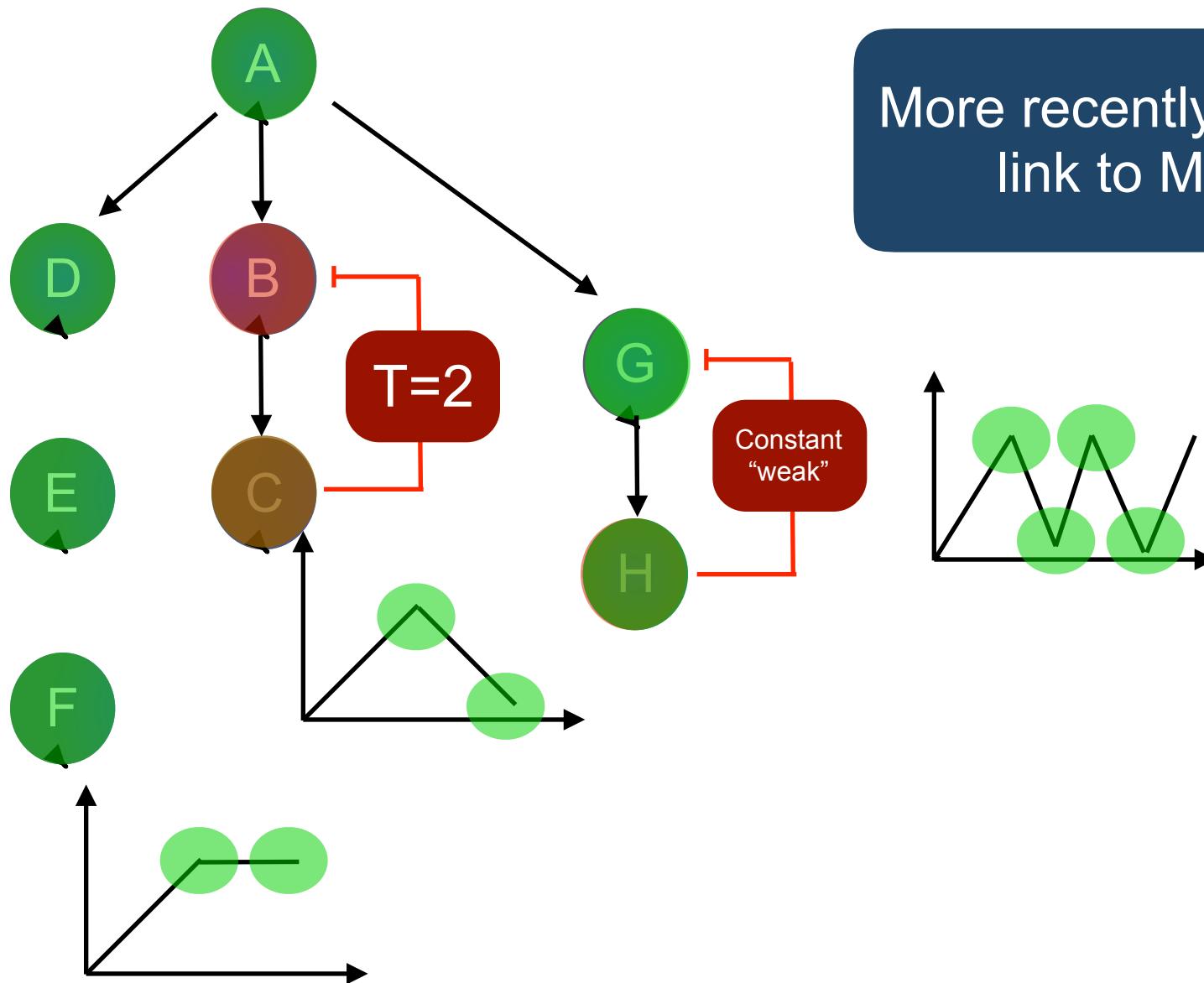
# Approximation of dynamics using synchronous simulation & multiple time-scales



More recently: update to link to MaBoss



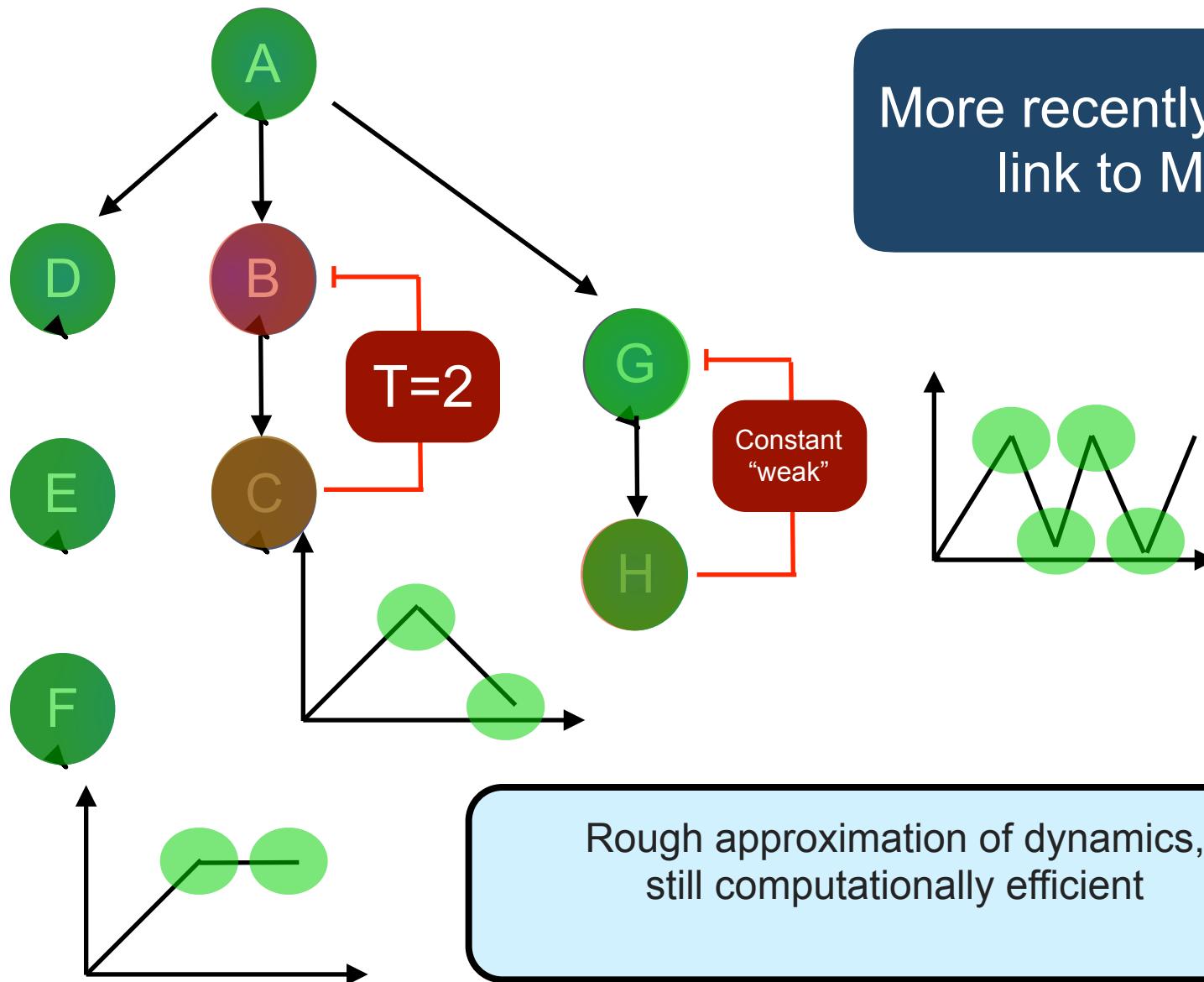
# Approximation of dynamics using synchronous simulation & multiple time-scales



More recently: update to link to MaBoss

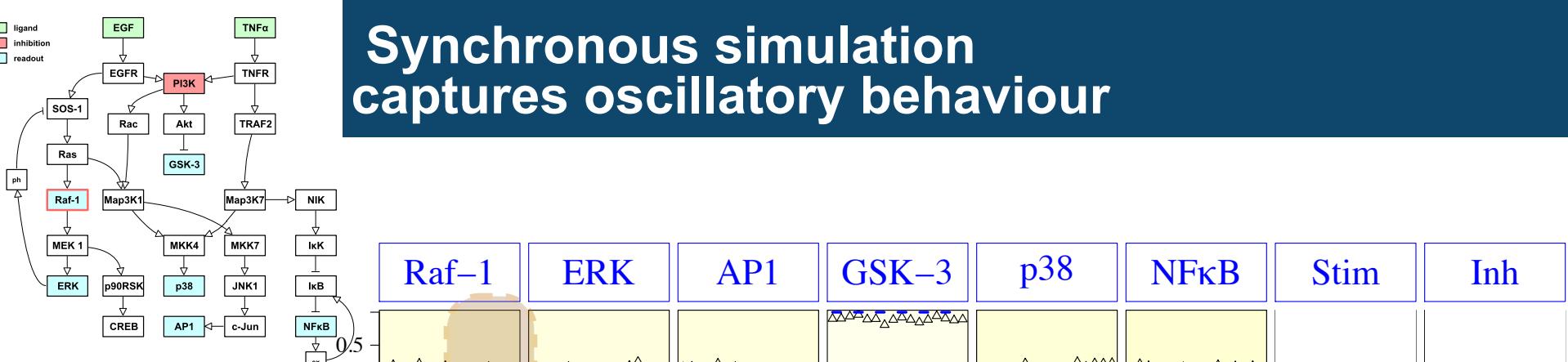


# Approximation of dynamics using synchronous simulation & multiple time-scales



More recently: update to link to MaBoss

Rough approximation of dynamics,  
still computationally efficient

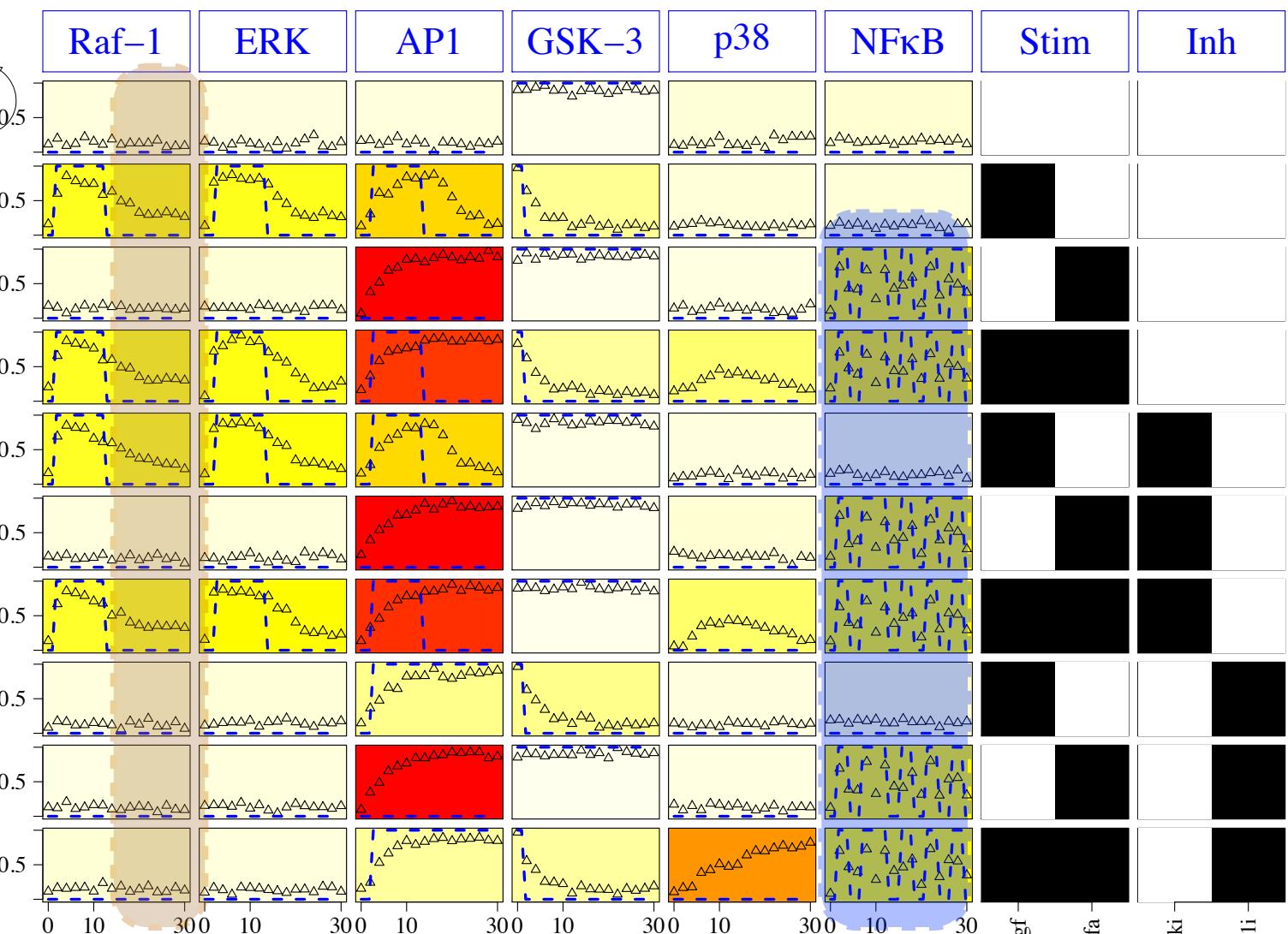


# Synchronous simulation captures oscillatory behaviour

Can not explain data due to missing links

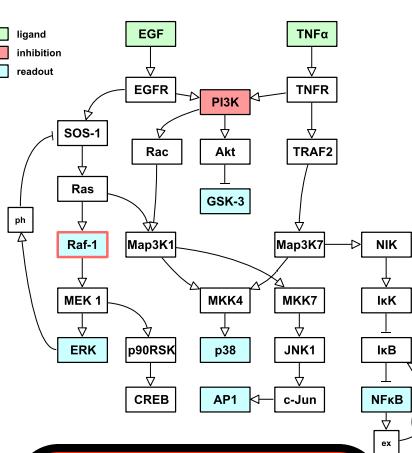
Identifies strong active links

Captures strong feedbacks





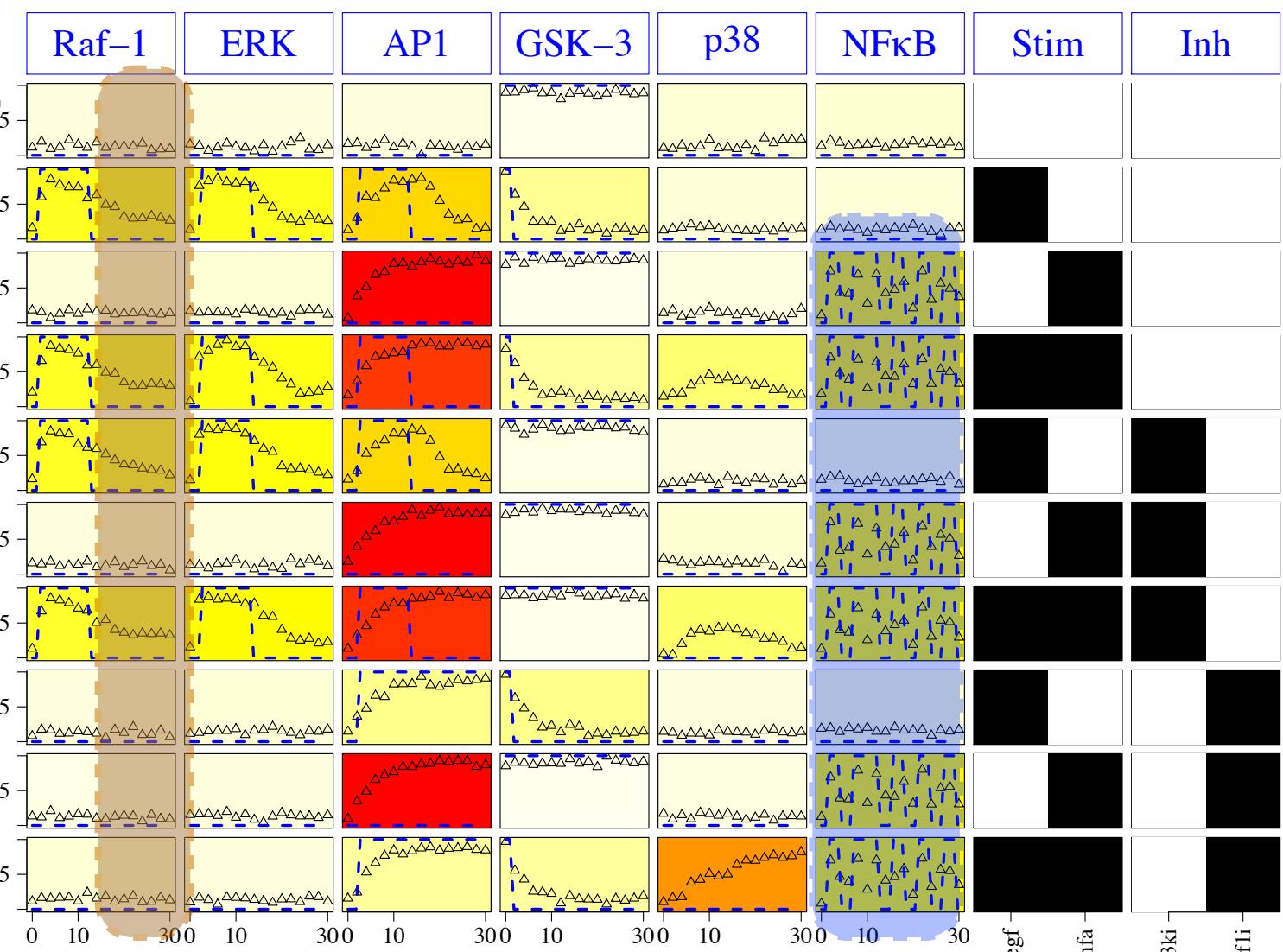
# Synchronous simulation captures oscillatory behaviour



Can not explain data due to missing links

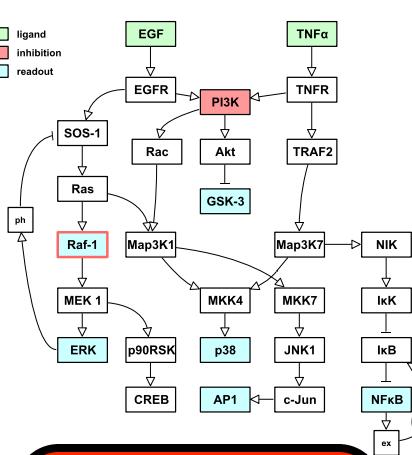
Identifies strong active links

Captures strong feedbacks





# Synchronous simulation captures oscillatory behaviour

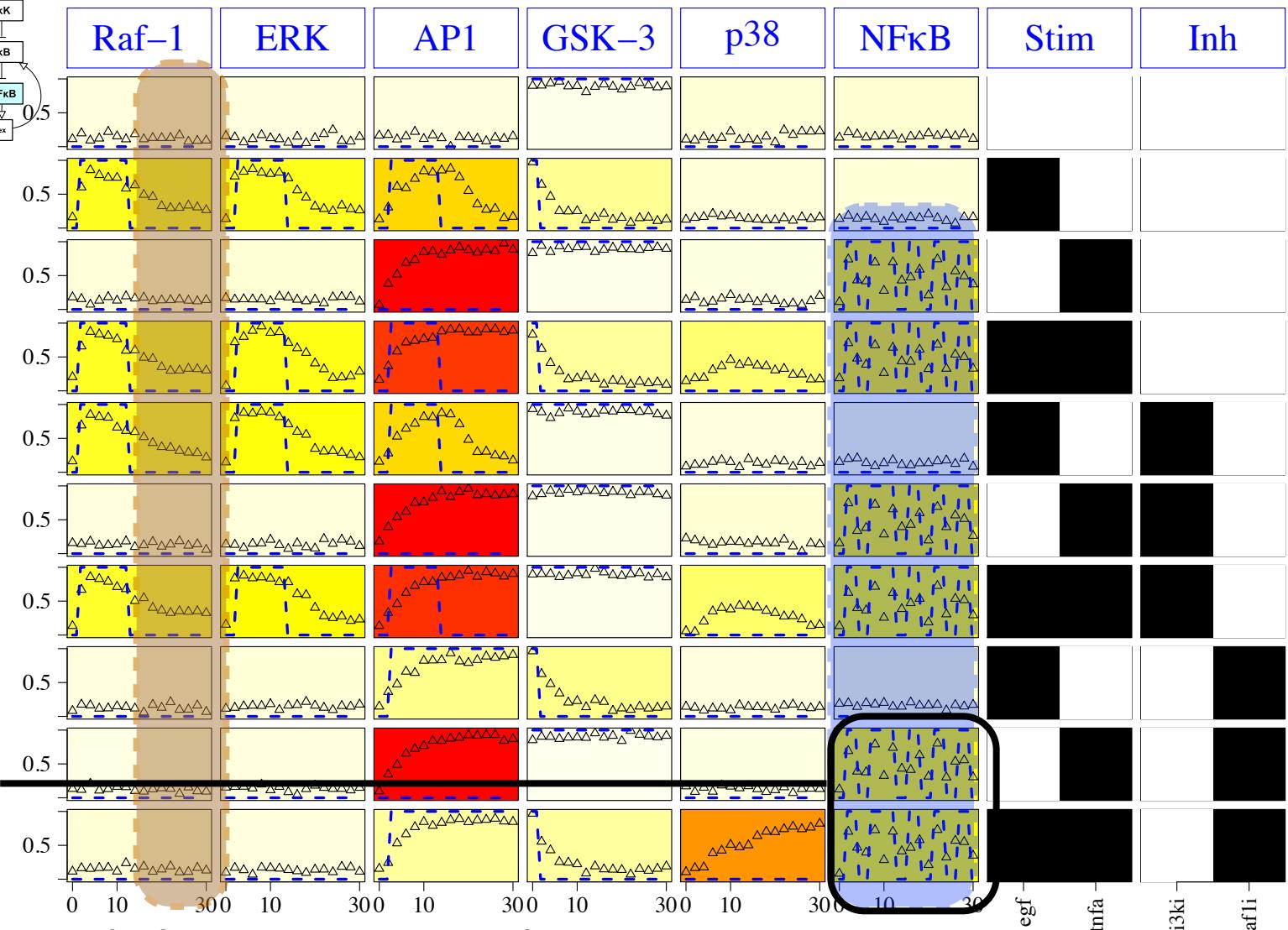


Can not explain data due to missing links

Identifies strong active links

Captures strong feedbacks

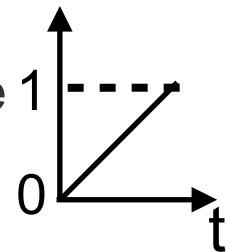
Captures 'weak' feedbacks



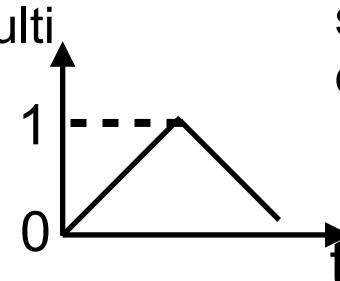


# From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state

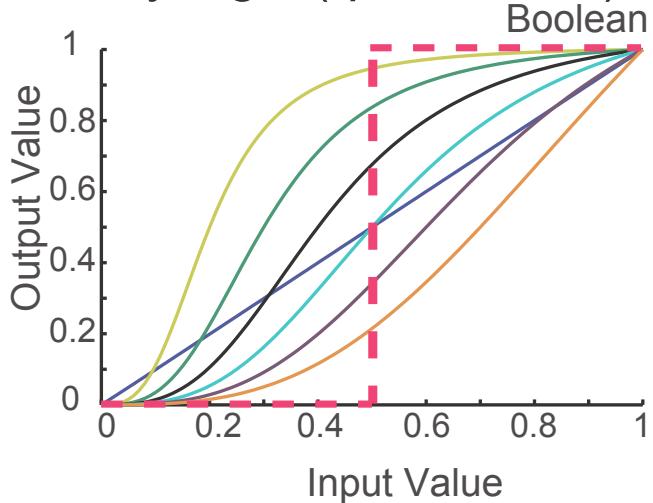


Boolean multi time-scale



sync.  
dynamics

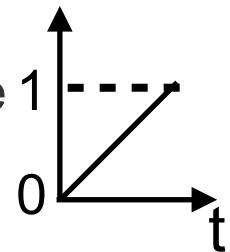
Fuzzy logic (quantitative)





# From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state

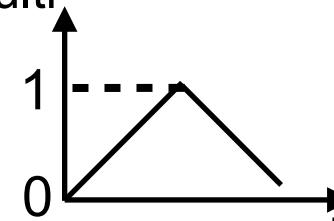


Boolean multi time-scale

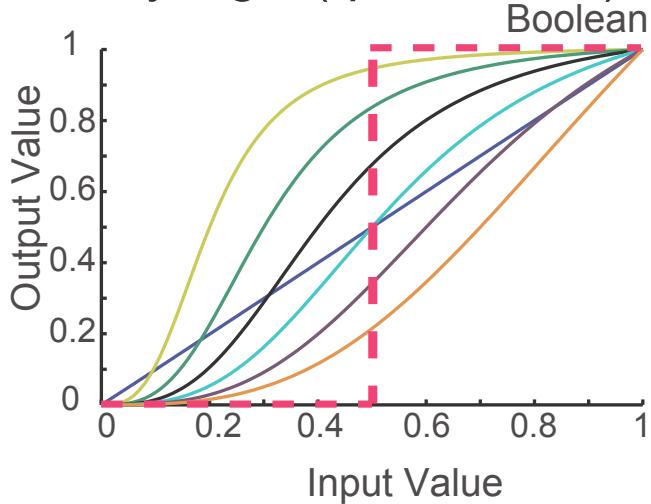
Camille Terfve

sync. dynamics

Aidan MacNamara



Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011



# Logic-based ODEs

- Convert Boolean update function  $B_i$  into a *continuous homologue*  $\bar{B}_i$  using multivariate polynomial interpolation
  - **Accuracy** (same behavior as  $B_i$  for 0/1  
→ same monotony & steady state behavior)
  - Good **analytical** properties (smoothness)
  - **Minimal and unique**
- Make non linear replacing variable with Hill function
- Transform into differential equation

$$f(\bar{x}_i) = \frac{\bar{x}_i^n}{(\bar{x}_i^n + k^n)}$$

$$\boxed{\bar{x}_i(t+1) = \bar{B}_i(\bar{x}_{i1}(t), \bar{x}_{i2}(t), \dots, \bar{x}_{iN_i}(t))} \rightarrow \boxed{\dot{\bar{x}}_i = \frac{1}{\tau_i} \cdot (\bar{B}_i(\bar{x}_{i1}, \bar{x}_{i2}, \dots, \bar{x}_{iN}) - \bar{x}_i)}$$

- E.g. a AND b inactivate C

$$\begin{aligned}\frac{dc}{dt} = \frac{1}{\tau} \left( \frac{a^{n_a} * (1 + k_a^{n_a}) * (1 - b^{n_b}) * (1 + k_b^{n_b})}{(a^{n_a} + k_a^{n_a}) * (b^{n_b} + k_b^{n_b})} + \frac{(1 - a^{n_a}) * (1 + k_a^{n_a}) * b^{n_b} * (1 + k_b^{n_b})}{(a^{n_a} + k_a^{n_a}) * (b^{n_b} + k_b^{n_b})} \right. \\ \left. + \frac{a^{n_a} * (1 + k_a^{n_a}) * b^{n_b} * (1 + k_b^{n_b})}{(a^{n_a} + k_a^{n_a}) * (b^{n_b} + k_b^{n_b})} - c \right)\end{aligned}$$



# Logic-based ODEs

- Convert Boolean update function  $B_i$  into a continuous homologue  $\bar{B}_i$  using multivariate polynomial interpolation
  - Accuracy (same behavior as  $B_i$  for 0/1  
→ same monotony & steady state behavior)
  - Good analytical properties (smoothness)
  - Minimal and unique
- Make non linear replacing variable with Hill function
- Transform into differential equation that matches the Boolean model
  - IDEA:  
ODE model  
mathematically ‘well-behaved’  
when states are 0 or 1
- E.g. a AND b inactivate C

$$f(\bar{x}_i) = \frac{\bar{x}_i^n}{(\bar{x}_i^n + k^n)}$$

$$\dot{x}_i = \frac{1}{\tau_i} (\bar{B}_i(\bar{x}_{i1}, \bar{x}_{i2}, \dots, \bar{x}_{iN}) - \bar{x}_i)$$

$$\begin{aligned}\frac{d}{dt}c = \frac{1}{\tau} &\left( \frac{a^{n_a} * (1 + k_a^{-n_a}) * (1 - b^{n_b}) * (1 + k_b^{-n_b})}{(a^{n_a} + k_a^{-n_a}) * (b^{n_b} + k_b^{-n_b})} + \frac{(1 - a^{n_a}) * (1 + k_a^{-n_a}) * b^{n_b} * (1 + k_b^{-n_b})}{(a^{n_a} + k_a^{-n_a}) * (b^{n_b} + k_b^{-n_b})} \right. \\ &\left. + \frac{a^{n_a} * (1 + k_a^{-n_a}) * b^{n_b} * (1 + k_b^{-n_b})}{(a^{n_a} + k_a^{-n_a}) * (b^{n_b} + k_b^{-n_b})} - c \right)\end{aligned}$$



# ODEs can be automatically generated from Boolean model (Odefy)

```
d/dt(tnfa) = 0*(1-tnfa_inh) %Note that this implies a continuous stimulus
```

```
d/dt(tgfa) = 0*(1-tgfa_inh) % Note that this implies a continuous stimulus
```

```
d/dt(raf) = ((egfr^raf_n_egfr/(egfr^raf_n_egfr+raf_k_egfr^raf_n_egfr)*(1+raf_k_egfr^raf_n_egfr)-raf) * raf_tauinv)*(1-raf_inh)
```

```
d/dt(pi3k) = ((egfr^pi3k_n_egfr/(egfr^pi3k_n_egfr+pi3k_k_egfr^pi3k_n_egfr)*(1+pi3k_k_egfr^pi3k_n_egfr)-pi3k) * pi3k_tauinv)*(1-pi3k_inh)
```

```
d/dt(ikb) = ((tnfa^ikb_n_tnfa/(tnfa^ikb_n_tnfa+ikb_k_tnfa^ikb_n_tnfa)*(1+ikb_k_tnfa^ikb_n_tnfa)*(1-pi3k^ikb_n_pi3k/(pi3k^ikb_n_pi3k+ikb_k_pi3k^ikb_n_pi3k)*(1+ikb_k_pi3k^ikb_n_pi3k))+(1-tnfa^ikb_n_tnfa/(tnfa^ikb_n_tnfa+ikb_k_tnfa^ikb_n_tnfa)*(1+ikb_k_tnfa^ikb_n_tnfa))*pi3k^ikb_n_pi3k/(pi3k^ikb_n_pi3k+ikb_k_pi3k^ikb_n_pi3k)*(1+ikb_k_pi3k^ikb_n_pi3k)+tnfa^ikb_n_tnfa/(tnfa^ikb_n_tnfa+ikb_k_tnfa^ikb_n_tnfa)*(1+ikb_k_tnfa^ikb_n_tnfa)*pi3k^ikb_n_pi3k/(pi3k^ikb_n_pi3k+ikb_k_pi3k^ikb_n_pi3k)*(1+ikb_k_pi3k^ikb_n_pi3k)-ikb) * ikb_tauinv)*(1-ikb_inh)
```

```
d/dt(gsk3) = (((1-akt^gsk3_n_akt/(akt^gsk3_n_akt+gsk3_k_akt^gsk3_n_akt)*(1+gsk3_k_akt^gsk3_n_akt))-gsk3) * gsk3_tauinv)*(1-gsk3_inh)
```

```
d/dt(erk12) = (((1-raf^erk12_n Raf/(raf^erk12_n Raf+erk12_k Raf^erk12_n Raf)*(1+erk12_k Raf^erk12_n Raf))*(1-ikb^erk12_n Ikb/(ikb^erk12_n Ikb+erk12_k Ikb^erk12_n Ikb)*(1+erk12_k Ikb^erk12_n Ikb))+raf^erk12_n Raf/(raf^erk12_n Raf+erk12_k Raf^erk12_n Raf)*(1-ikb^erk12_n Ikb/(ikb^erk12_n Ikb+erk12_k Ikb^erk12_n Ikb)*(1+erk12_k Ikb^erk12_n Ikb))+raf^erk12_n Raf/(raf^erk12_n Raf+erk12_k Raf^erk12_n Raf)*(1+erk12_k Raf^erk12_n Raf)*ikb^erk12_n Ikb/(ikb^erk12_n Ikb+erk12_k Ikb^erk12_n Ikb)*(1+erk12_k Ikb^erk12_n Ikb)-erk12) * erk12_tauinv)*(1-erk12_inh)
```

```
d/dt(egfr) = ((tgfa^egfr_n_tgfa/(tgfa^egfr_n_tgfa+egfr_k_tgfa^egfr_n_tgfa)*(1+egfr_k_tgfa^egfr_n_tgfa)-egfr) * egfr_tauinv)*(1-egfr_inh)
```

```
d/dt(casp8) = ((tnfa^casp8_n_tnfa/(tnfa^casp8_n_tnfa+casp8_k_tnfa^casp8_n_tnfa)*(1+casp8_k_tnfa^casp8_n_tnfa)-casp8) * casp8_tauinv)*(1-casp8_inh)
```

```
d/dt(akt) = ((pi3k^akt_n_pi3k/(pi3k^akt_n_pi3k+akt_k_pi3k^akt_n_pi3k)*(1+akt_k_pi3k^akt_n_pi3k)-akt) * akt_tauinv)*(1-akt_inh)
```



# ODEs can be automatically generated from Boolean model (Odefy)

```
d/dt(tnfa) = 0*(1-tnfa_inh) %Note that this implies a continuous stimulus
```

```
d/dt(tgfa) = 0*
```

```
d/dt(raf) = ((egfr_n_tgfa^egfr_k_tgfa + egfr_n_tgfa^egfr_k_tgfa)*((1+egfr_k_tgfa^egfr_n_tgfa)-egfr) * egfr_tauinv)*(1-egfr_inh)
```

```
d/dt(pi3k) = ((egfr_n_tgfa^egfr_k_tgfa + egfr_n_tgfa^egfr_k_tgfa)*((1+egfr_k_tgfa^egfr_n_tgfa)-egfr) * egfr_tauinv)*(1-egfr_inh)
```

```
d/dt(ikb) = ((tnfa_n_ikb^tnfa_k_ikb + tnfa_n_ikb^tnfa_k_ikb)*((1+tnfa_k_ikb^tnfa_n_ikb)-tnfa) * tnfa_tauinv)*(1-tnfa_inh)
```

```
d/dt(gsk3) = ((tnfa_n_ikb^tnfa_k_ikb + tnfa_n_ikb^tnfa_k_ikb)*((1+tnfa_k_ikb^tnfa_n_ikb)-tnfa) * tnfa_tauinv)*(1-tnfa_inh)
```

```
d/dt(erk12) = ((egfr_n_tgfa^egfr_k_tgfa + egfr_n_tgfa^egfr_k_tgfa)*((1+egfr_k_tgfa^egfr_n_tgfa)-egfr) * egfr_tauinv)*(1-egfr_inh)
```

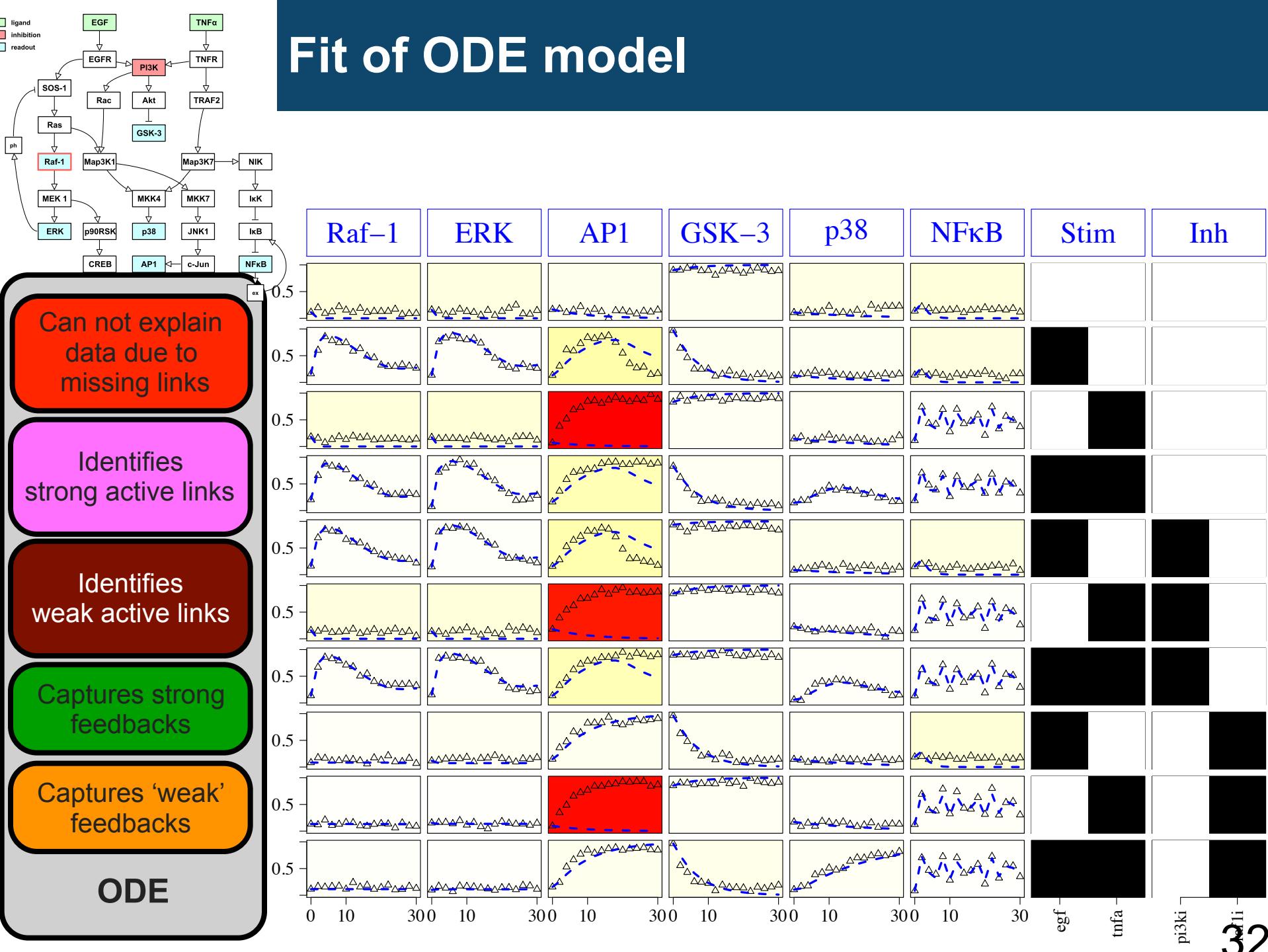
```
d/dt(egfr) = ((tgfa_n_egfr^tgfa_k_egfr + tgfa_n_egfr^tgfa_k_egfr)*((1+tgfa_k_egfr^tgfa_n_egfr)-tgfa) * tgfa_tauinv)*(1-tgfa_inh)
```

```
d/dt(casp8) = ((tnfa_n_casp8^tnfa_k_casp8 + tnfak_n_casp8^tnfak_k_casp8)*((1+tnfak_k_casp8^tnfak_n_casp8)-tnfak) * tnfak_tauinv)*(1-tnfak_inh)
```

```
d/dt(akt) = ((pi3k_n_akt^pi3k_k_akt + pi3k_n_akt^pi3k_k_akt)*((1+akt_k_pi3k^akt_n_pi3k)-akt) * akt_tauinv)*(1-akt_inh)
```

**Even if structure is known need to identify parameters, difficult optimisation problem (similar to biochemical ODEs)**

$$\frac{d}{dt}c = \frac{1}{\tau} \left( \frac{a^{na} * (1 + k_a^{na}) * (1 - k_b^{nb}) * (1 + k_b^{nb})}{(a^{na} + k_a^{na}) * (b^{nb} + k_b^{nb})} + \frac{(1 - a^{na}) * (1 + k_a^{na}) * b^{nb} * (1 + k_b^{nb})}{(a^{na} + k_a^{na}) * (b^{nb} + k_b^{nb})} \right. \\ \left. + \frac{a^{na} * (1 + k_a^{na}) * b^{nb} * (1 + k_b^{nb})}{(a^{na} + k_a^{na}) * (b^{nb} + k_b^{nb})} - c \right)$$





# CellNOpt-MaBOSS Fits can approximate dynamics



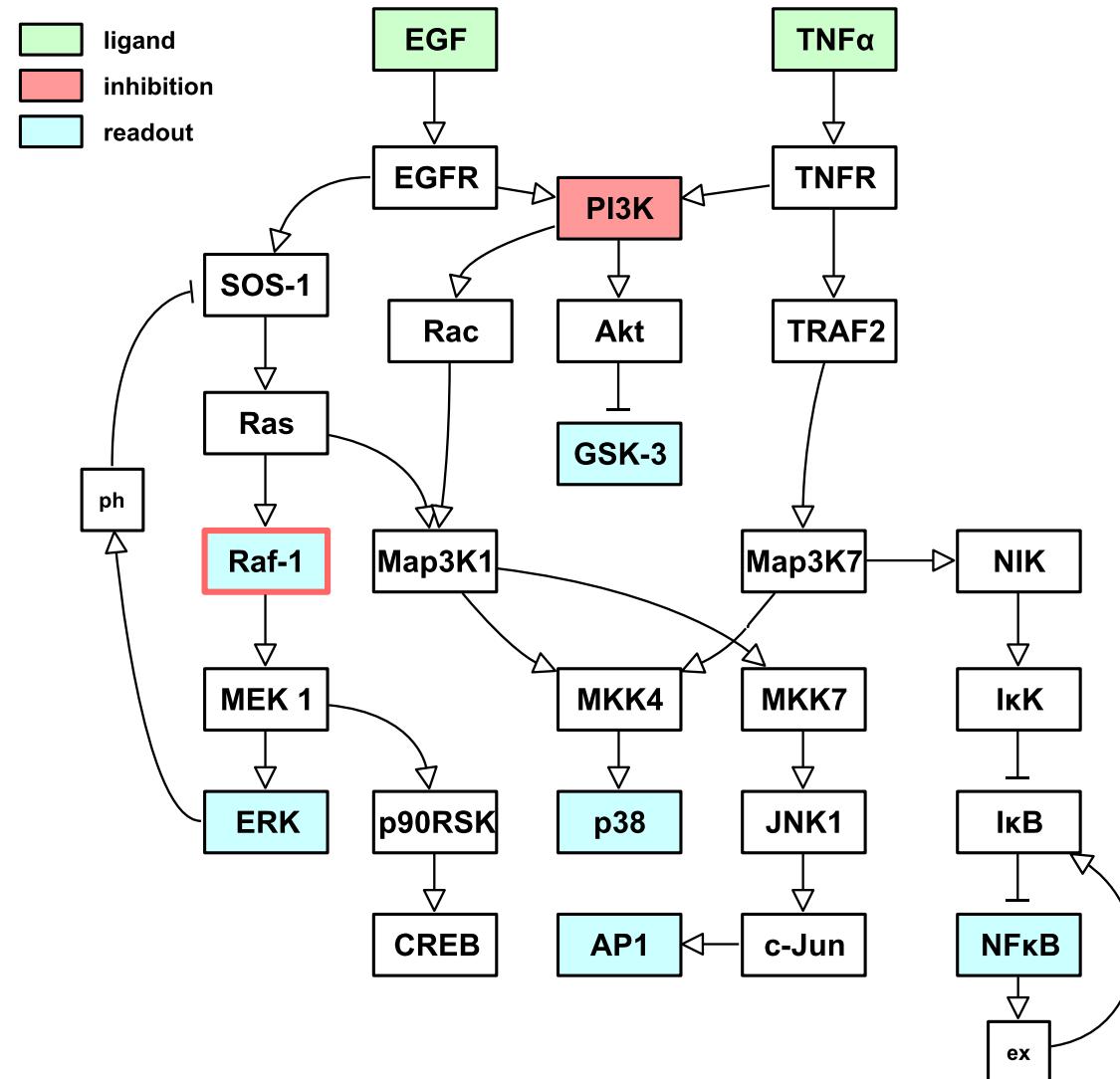
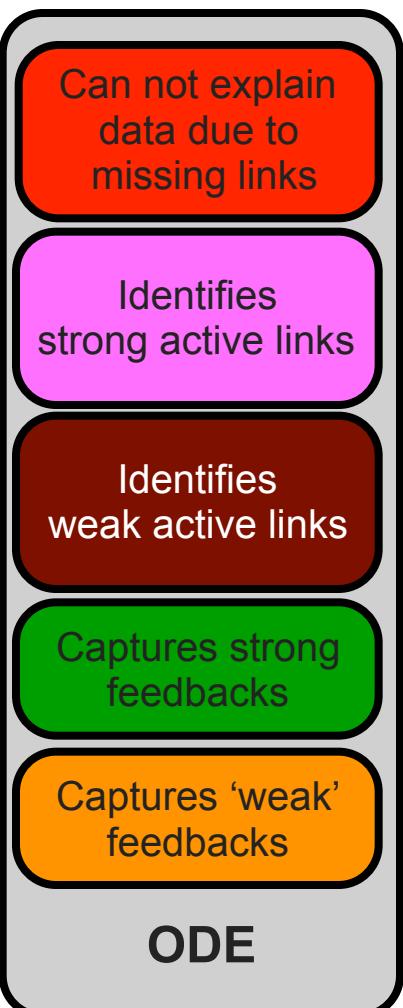
# CellNOpt-MaBOSS Fits can approximate dynamics

Integration nearly ready:

fit of models simulated with MaBoss  
(asynchronous, time-continuous)  
- fits time dynamics, still Boolean  
- worse fit but faster than logic-ODES

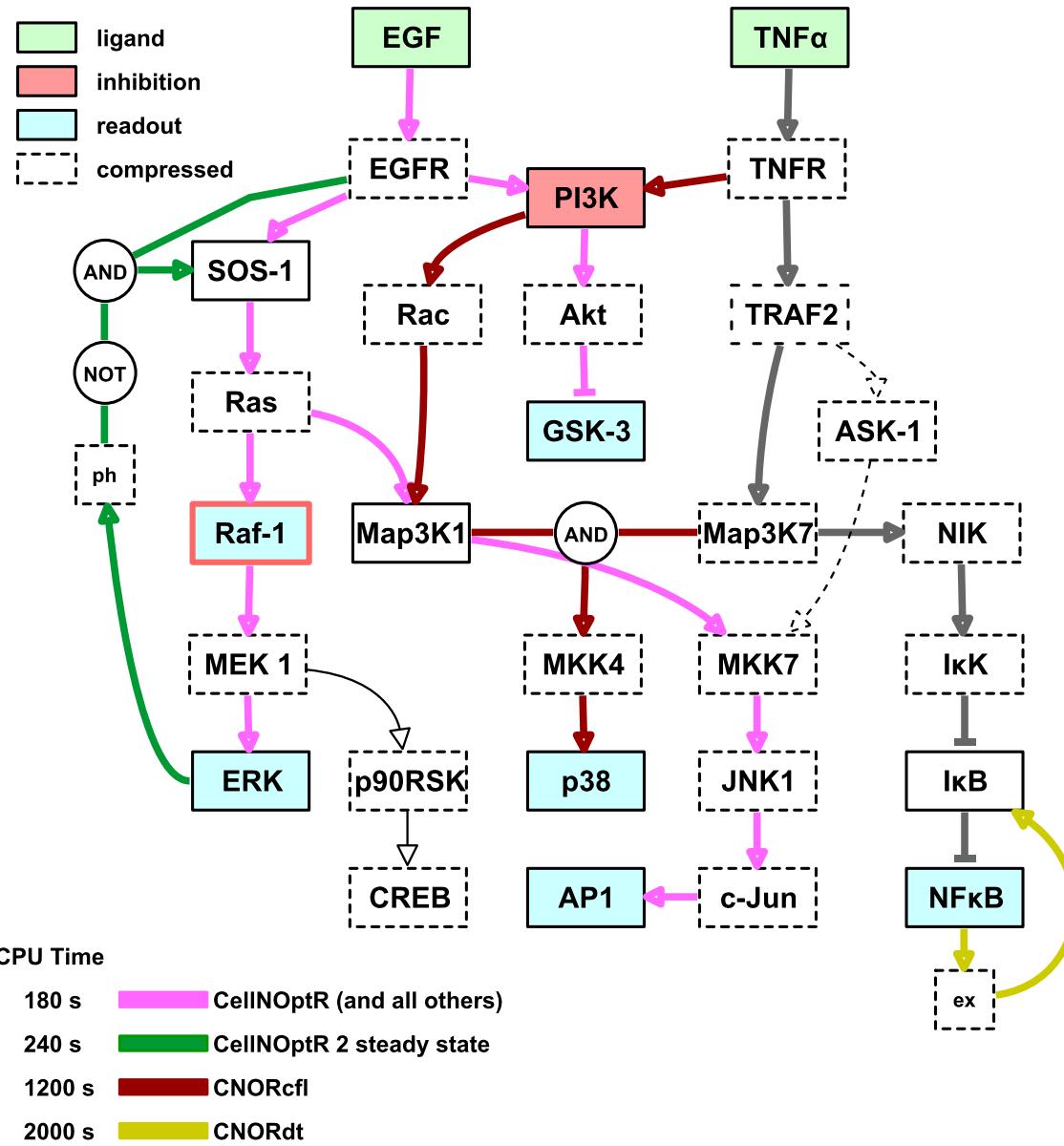
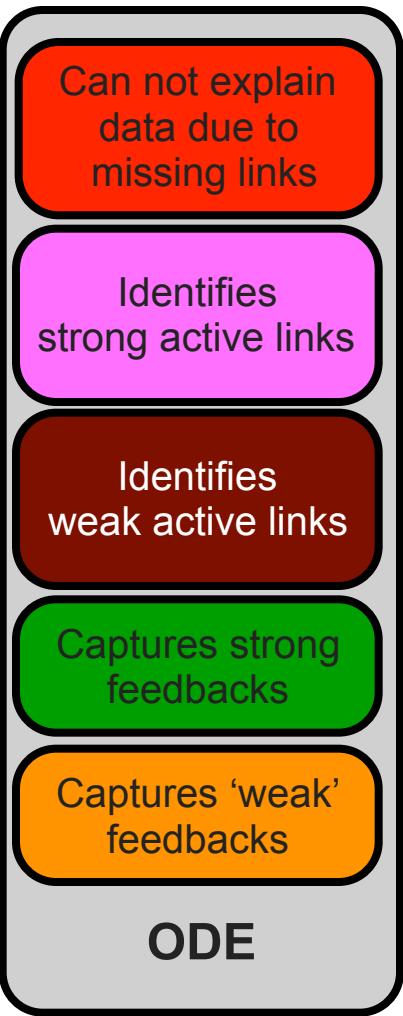


# Different methods capture different aspects





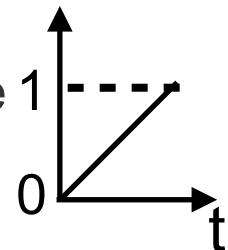
# Different methods capture different aspects





# From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state

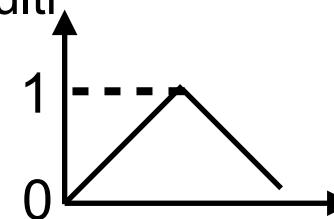


Boolean multi time-scale

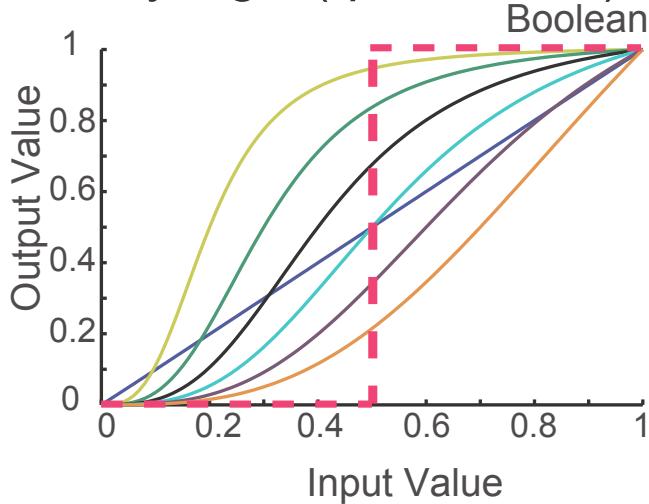
Camille Terfve

sync. dynamics

Aidan MacNamara



Fuzzy logic (quantitative)

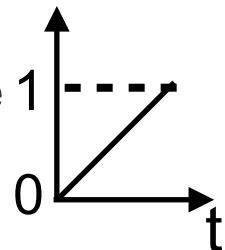


Morris et al., PloS Comp Bio 2011



# From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state



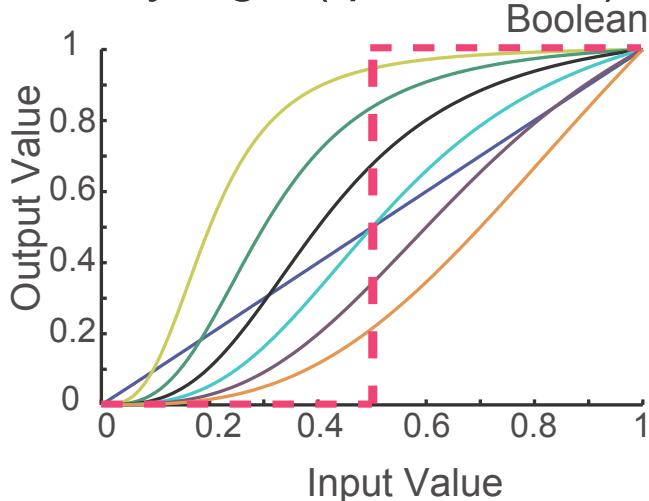
Boolean multi time-scale

Camille Terfve

sync. dynamics

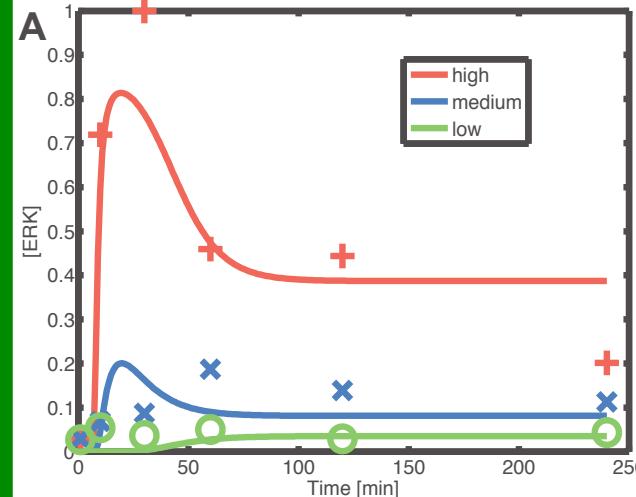
Aidan MacNamara

Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

Logic ODEs (dynamic)



w. J Banga & J. Egea,

David Henriques



# From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary logic steady-state)

Boolean multi-time-scale

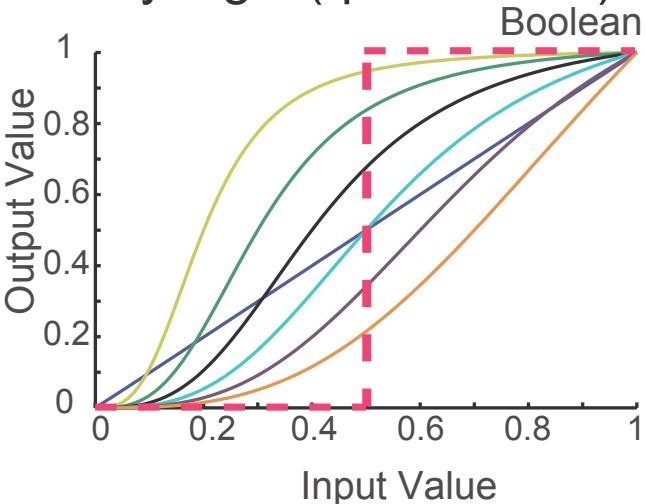
sync. dynamics

Camille Terfve

Aidan MacNamara

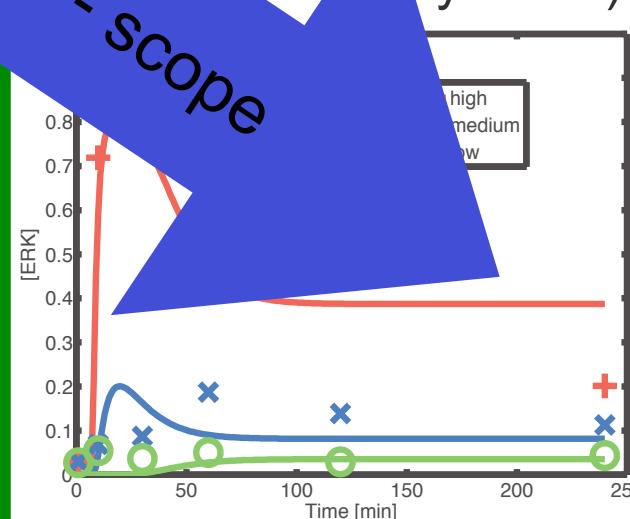
+ detail

Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

ODE (dynamic)



w. J Banga & J. Egea,

MEIGO:  
Global  
optimization  
in R/Matlab  
Egea et al.  
BMC Bioinf  
2014

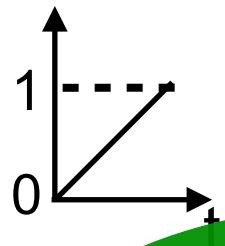
Identify  
structure  
+ parameters  
Henriques et al.  
Bioinformatics  
2015

David  
Henriques



# From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state



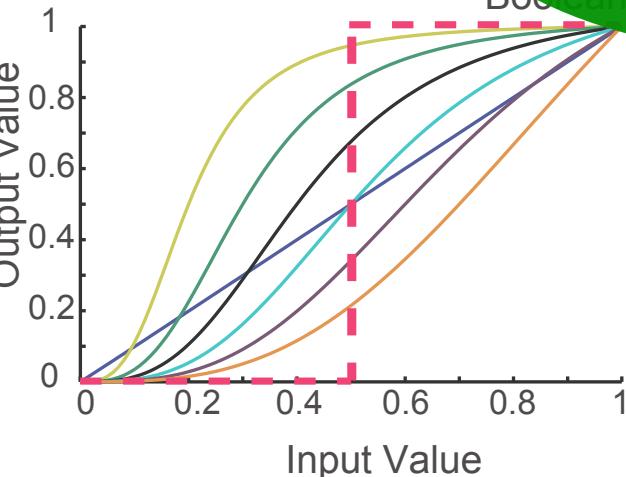
Boolean multi time-scale

Camille  
Terfve

sync.  
dynamics

Aidan  
MacNamara

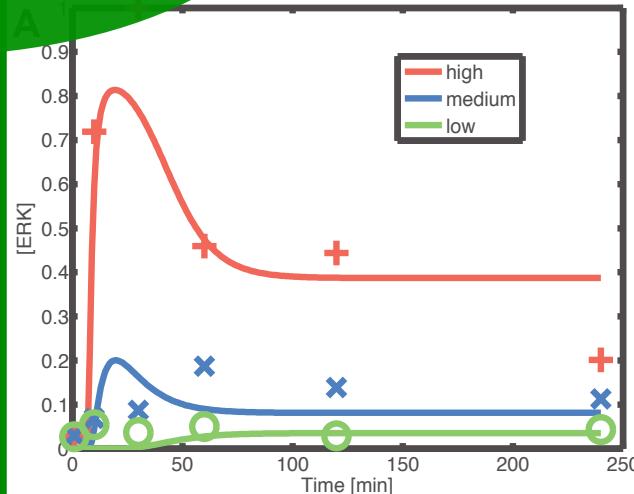
Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

*CellNOpt*

Logic ODEs (dynamic)

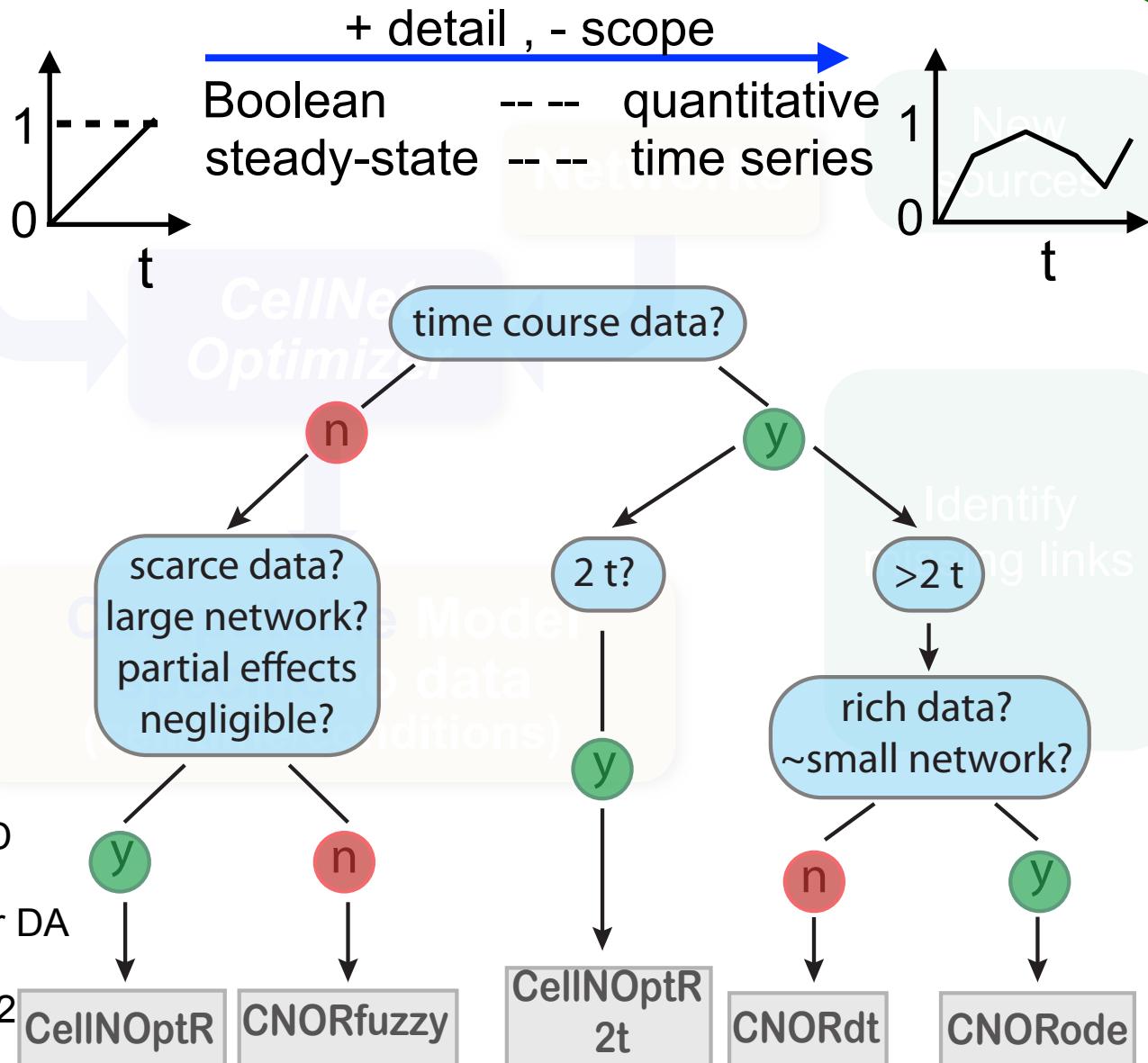


w. J Banga & J. Egea,

David  
Henriques



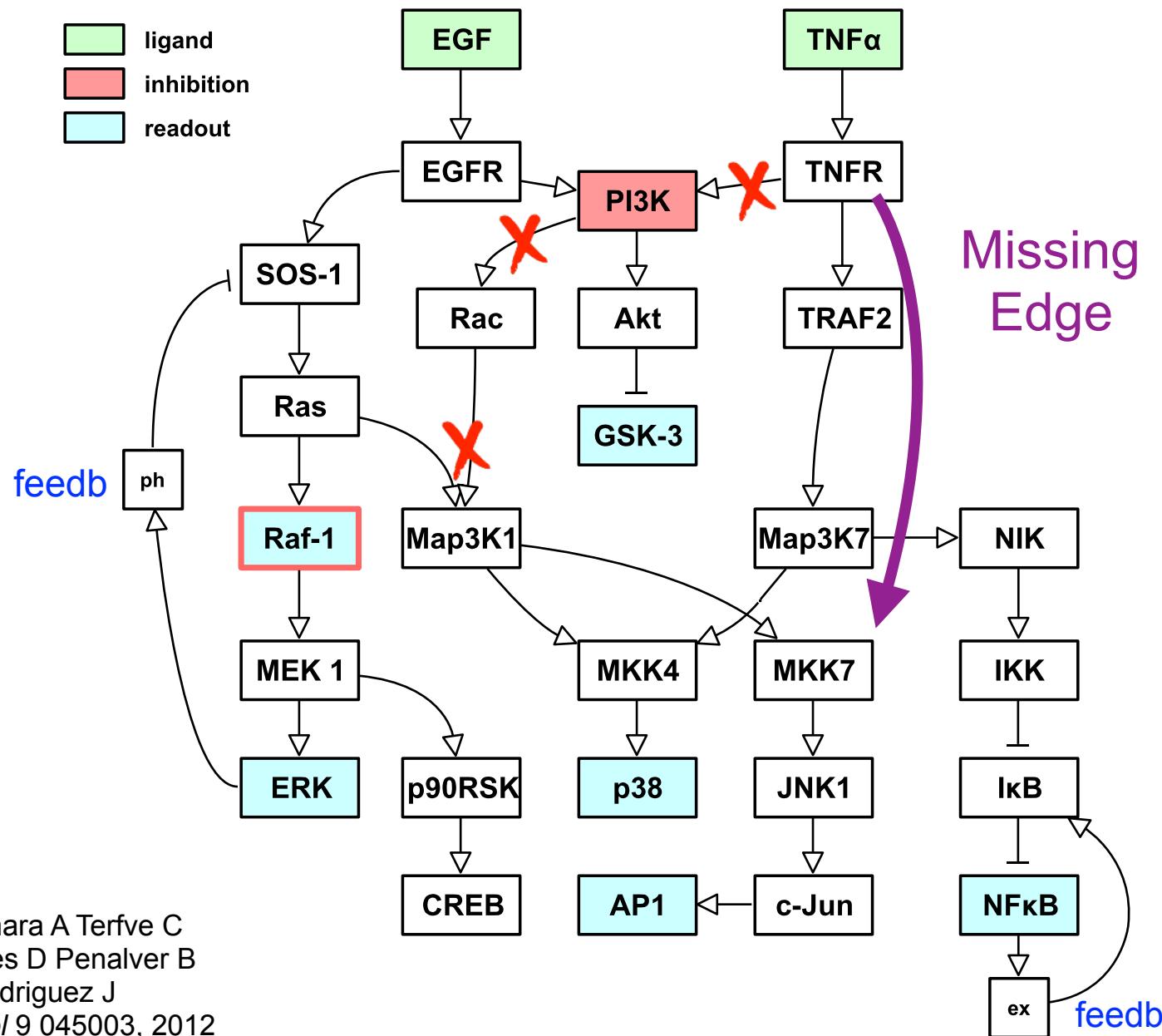
# Broad spectrum of modelling formalism with different level of detail



Terfve C Cokelaer T  
MacNamara A Henriques D  
Gonçalves E Morris MK  
van Iersel M Lauffenburger DA  
Saez-Rodriguez J  
*BMC Syst Biol*, 6:133, 2012



# How to deal with incomplete prior knowledge?





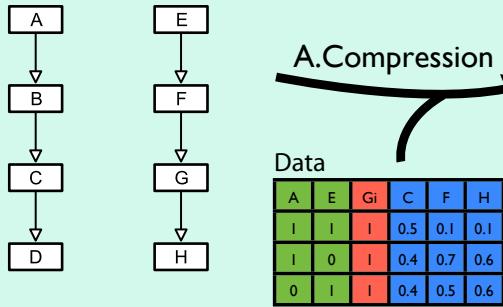
# How to deal with incomplete prior knowledge?

CNOFeed: Link CellNOpt to methods  
to infer new links

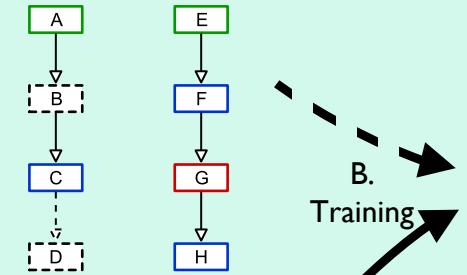
Federica  
Eduati

More types

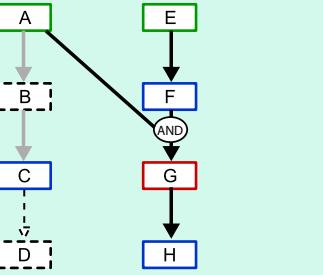
Prior Knowledge Network (PKN)



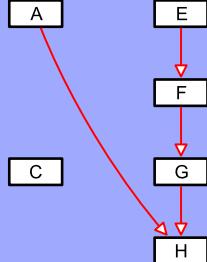
Compressed Network



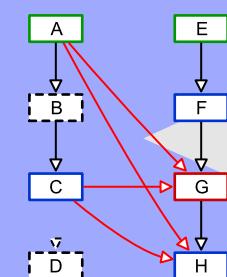
Trained Model



Data-driven  
Network (DDN)

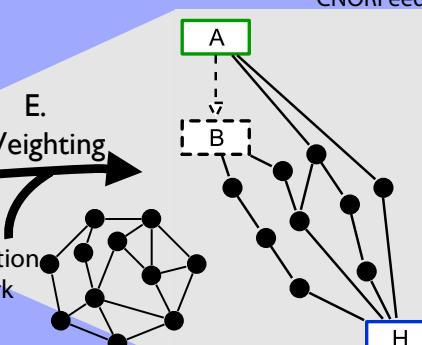


Integrated Network



E. Weighting

Protein  
Interaction  
Network  
(PIN)



Eduati F, de las Rivas J, di Camilo B, Toffolo G, Saez-Rodriguez J  
*Bioinformatics* 10.1093/bts363, 2012



# Steps in building (and using) a model

- Set up experiments to extract most information
- Process data efficiently
- Choose type of mathematical model  
(given data, question, etc)
- Train models to experimental data
- Use models to gain insight



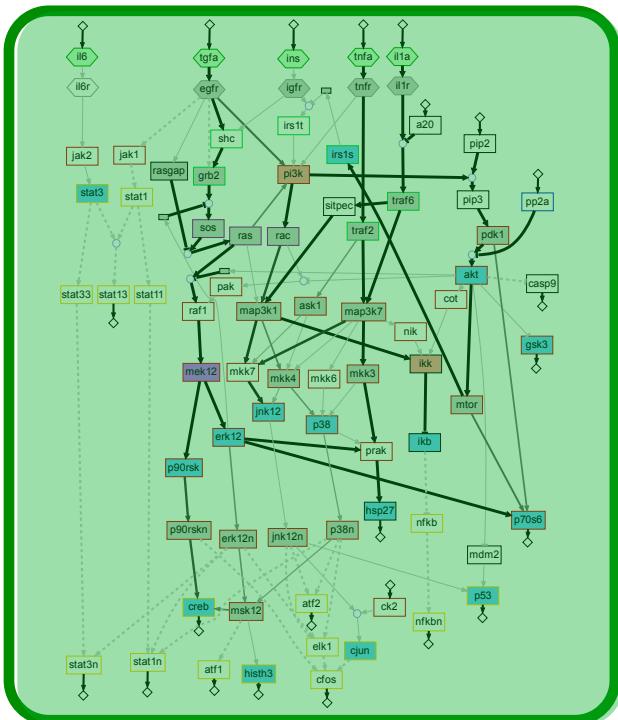
# Steps in building (and using) a model

- Set up experiments to extract most information
- Process data efficiently
- Choose type of mathematical model  
(given data, question, etc)
- Train models to experimental data
- Use models to gain insight



# How is signal processing altered in disease?

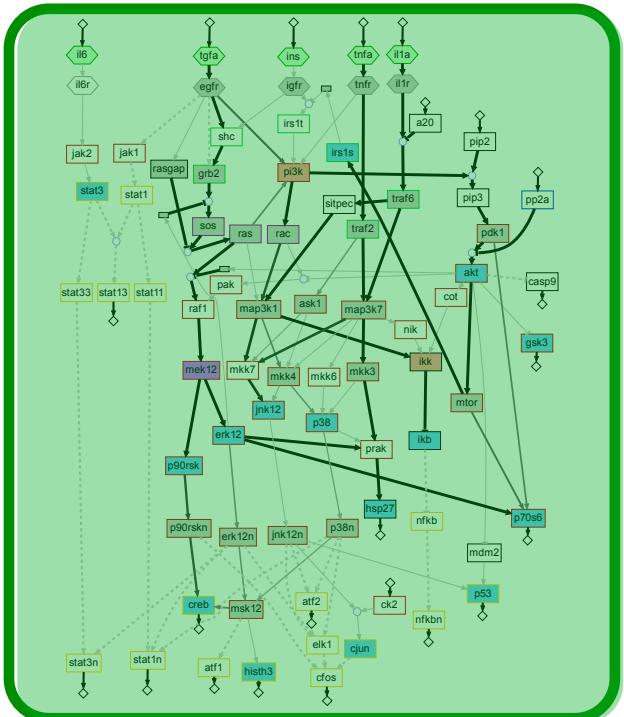
## Health



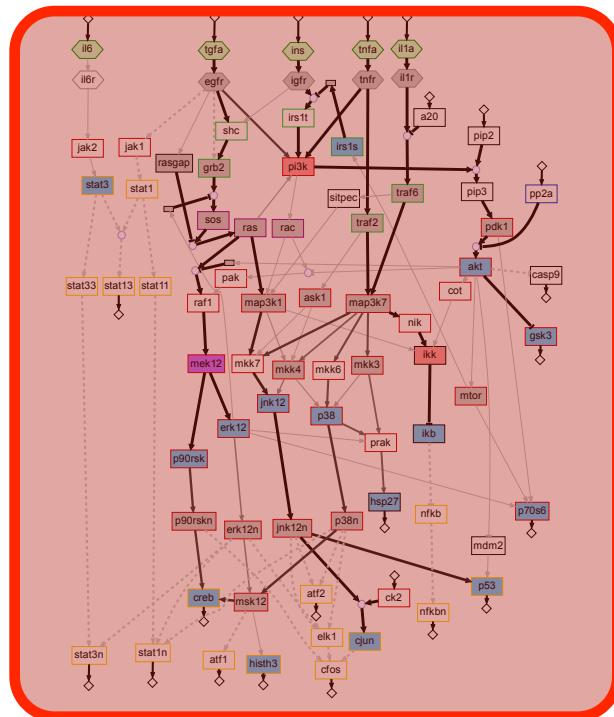


# How is signal processing altered in disease?

## Health



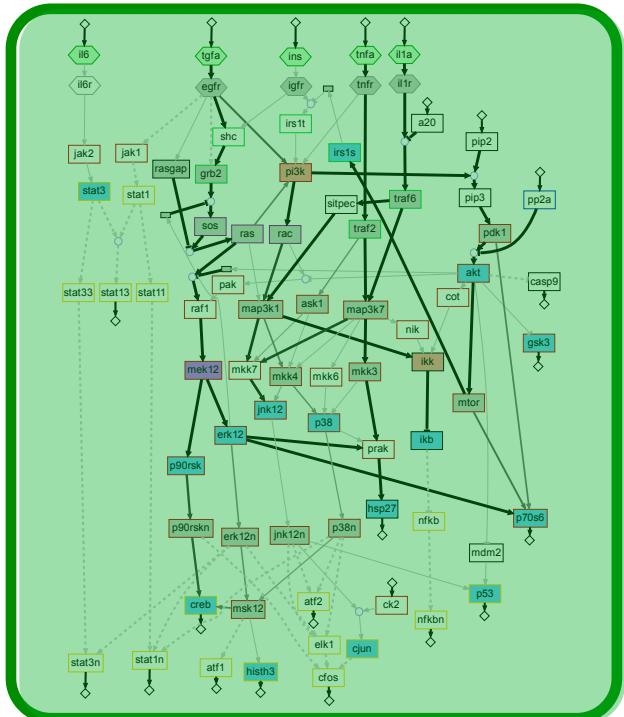
## Disease



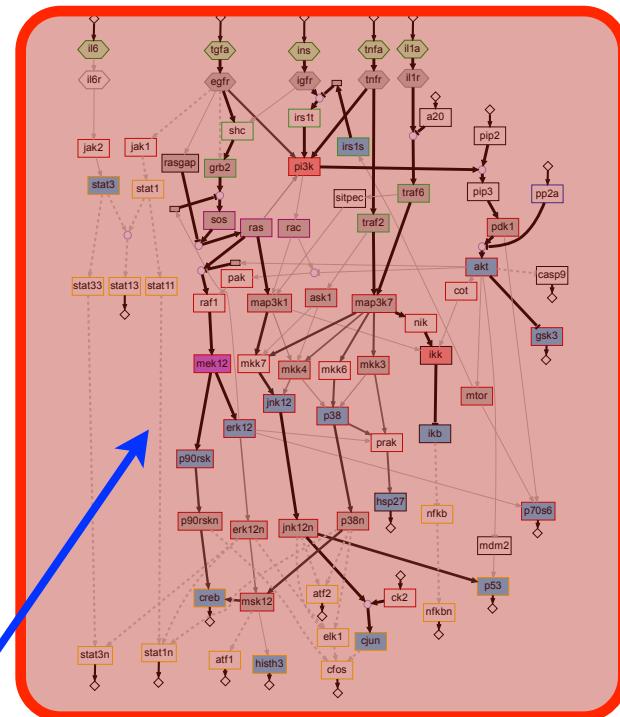


# How is signal processing altered in disease?

## Health



## Disease



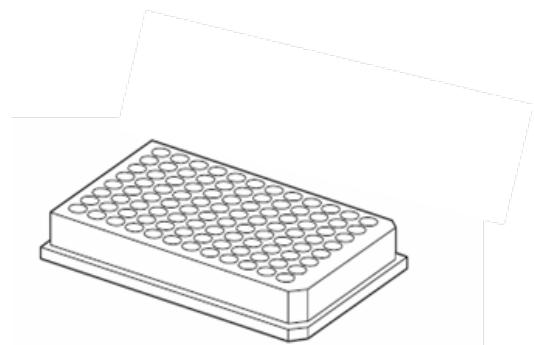
... and how can we target this with therapeutics?



# An example of a perturbation-based high-throughput data sets



# An example of a perturbation-based high-throughput data sets



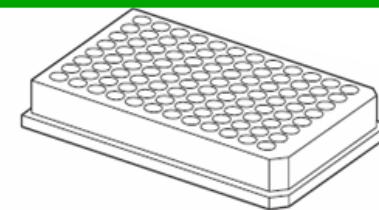


# An example of a perturbation-based high-throughput data sets

Cue

→ 7 extracellular ligands

→ 7 specific chemical inhibitors (drugs)





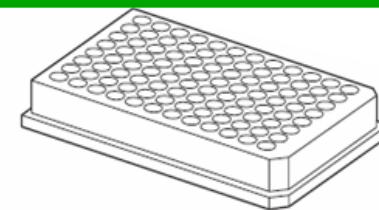
# An example of a perturbation-based high-throughput data sets

Cue

→ 7 extracellular ligands

→ 7 specific **chemical inhibitors** (drugs)

at different times  
after stimulation



Signal

→ **Phosphorylation** of 17 key proteins (30 min, 3h)



# An example of a perturbation-based high-throughput data sets

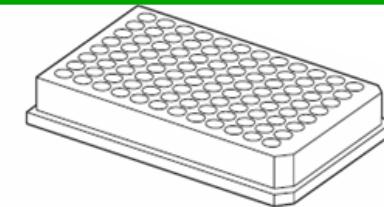
Primary human hepatocytes & HCC cell lines (HepG2, Hep3B, Huh7, Focus)

Cue

→ 7 extracellular ligands

→ 7 specific **chemical inhibitors** (drugs)

at different times  
after stimulation



Signal

→ **Phosphorylation** of 17 key proteins (30 min, 3h)

Response

→ **Release** of 20 cytokines (3h, 24h)

using Luminex/xMAP  
(bead-based ELISA)



# Comparison of primary hepatocytes to 4 HCC cell lines

Primary

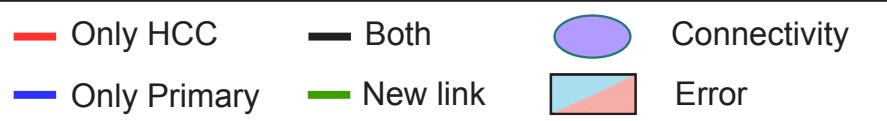
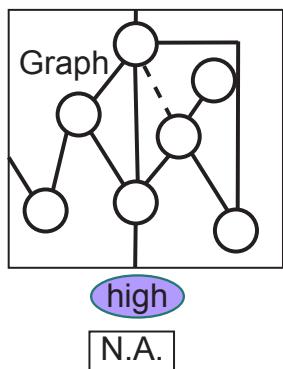
HepG2

Hep3

Huh7

Focus

Generic  
network





# Comparison of primary hepatocytes to 4 HCC cell lines

Primary

HepG2

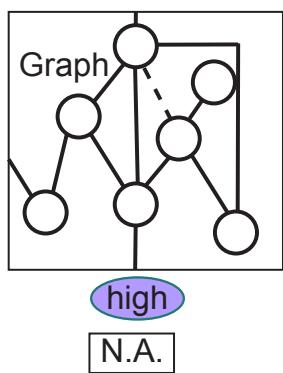
Hep3

Huh7

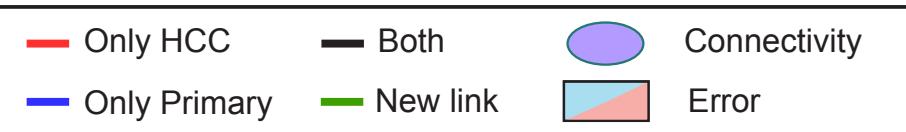
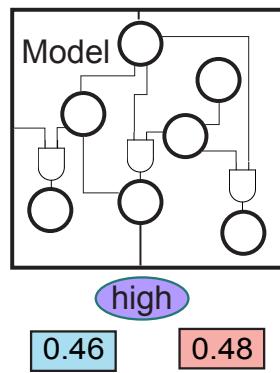
Focus

Generic  
network

Scaffold of  
logical models



Process  
CNO





# Comparison of primary hepatocytes to 4 HCC cell lines

Primary

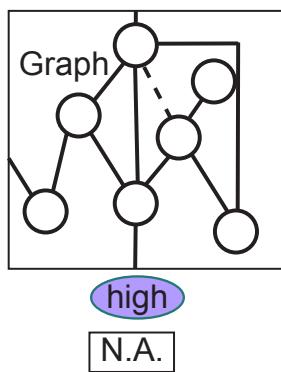
HepG2

Hep3

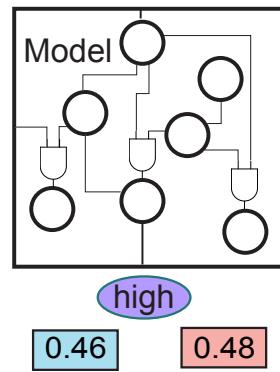
Huh7

Focus

Generic  
network



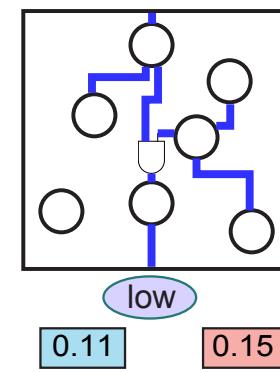
Scaffold of  
logical models



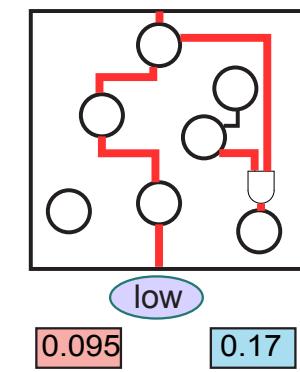
Process  
CNO

Train to  
Cell Response  
Data

Primary

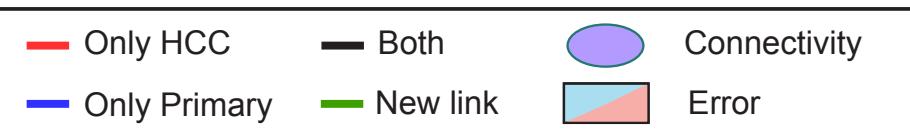


HCC



0.15

0.17





# Comparison of primary hepatocytes to 4 HCC cell lines

Primary

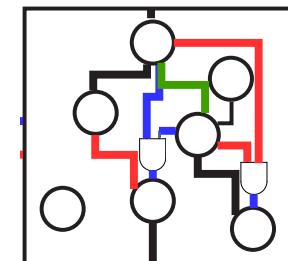
HepG2

Hep3

Huh7

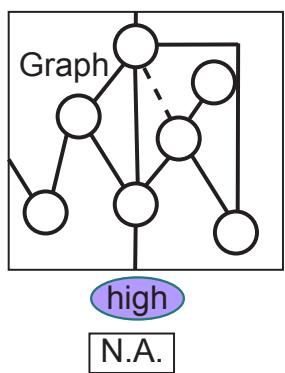
Focus

Specific Networks

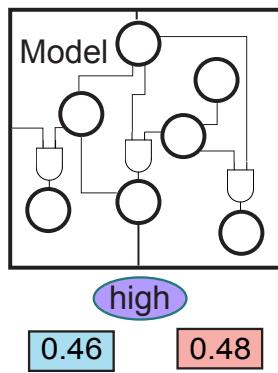


Generic  
network

Scaffold of  
logical models

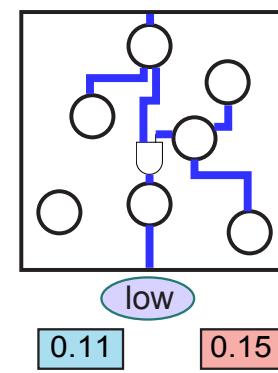


Process  
CNO

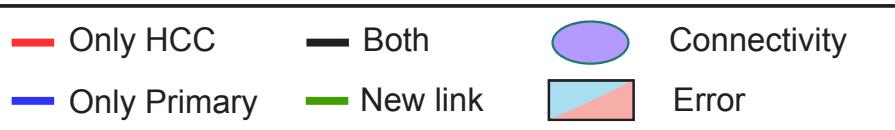
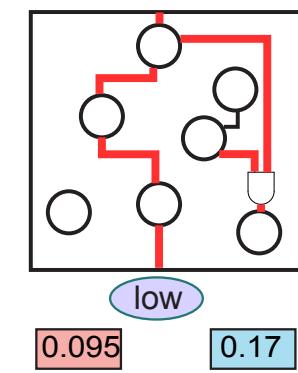


Train to  
Cell Response  
Data

Primary



HCC



# Primary



Stimulus

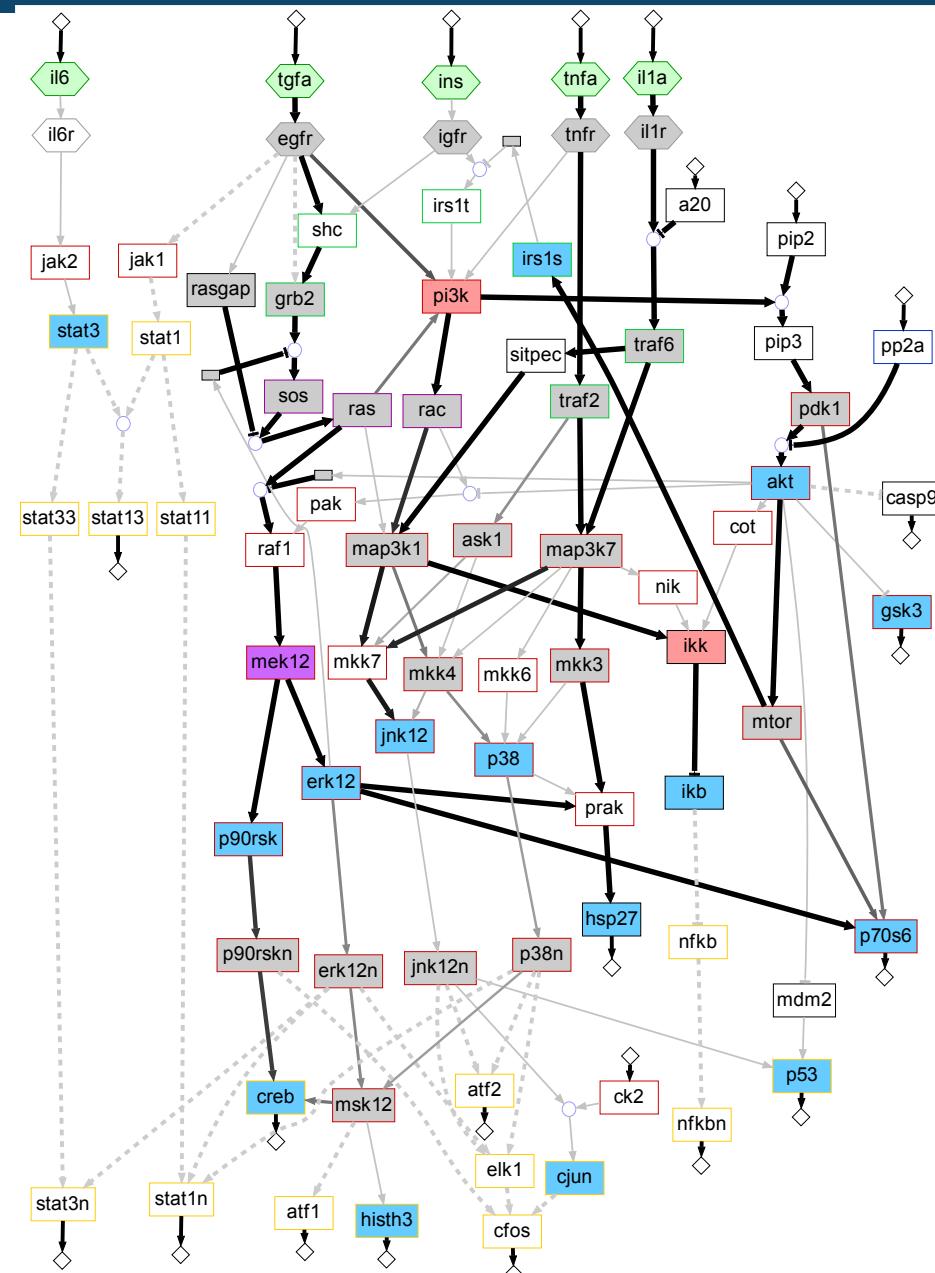
Perturbation

Readout

Perturb&Read

Kept

Removed  
No effect

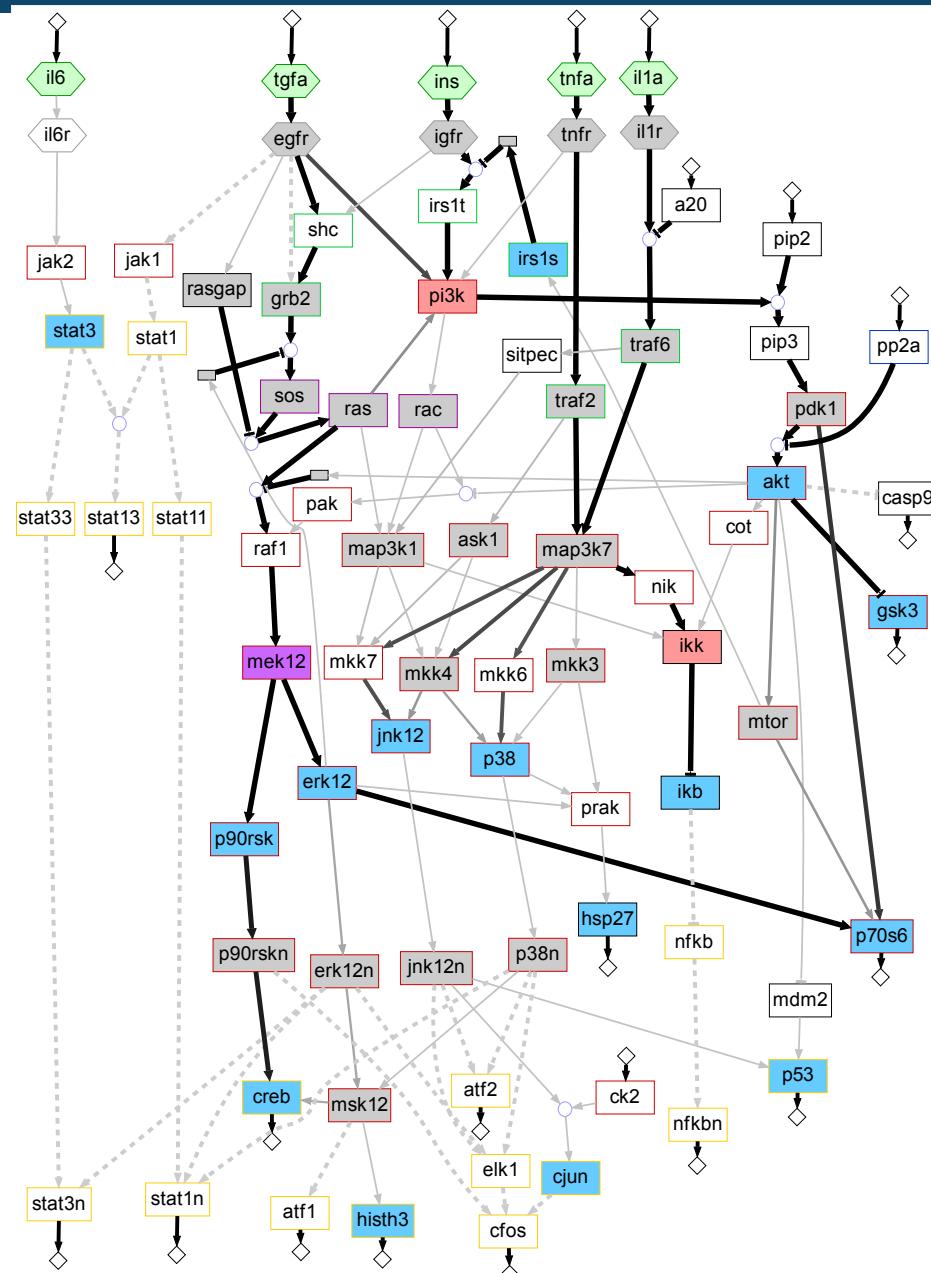




# HepG2

Stimulus  
Perturbation  
Readout  
Perturb&Read

Kept →  
Removed →  
No effect →





# Hep3B

Stimulus

Perturbation

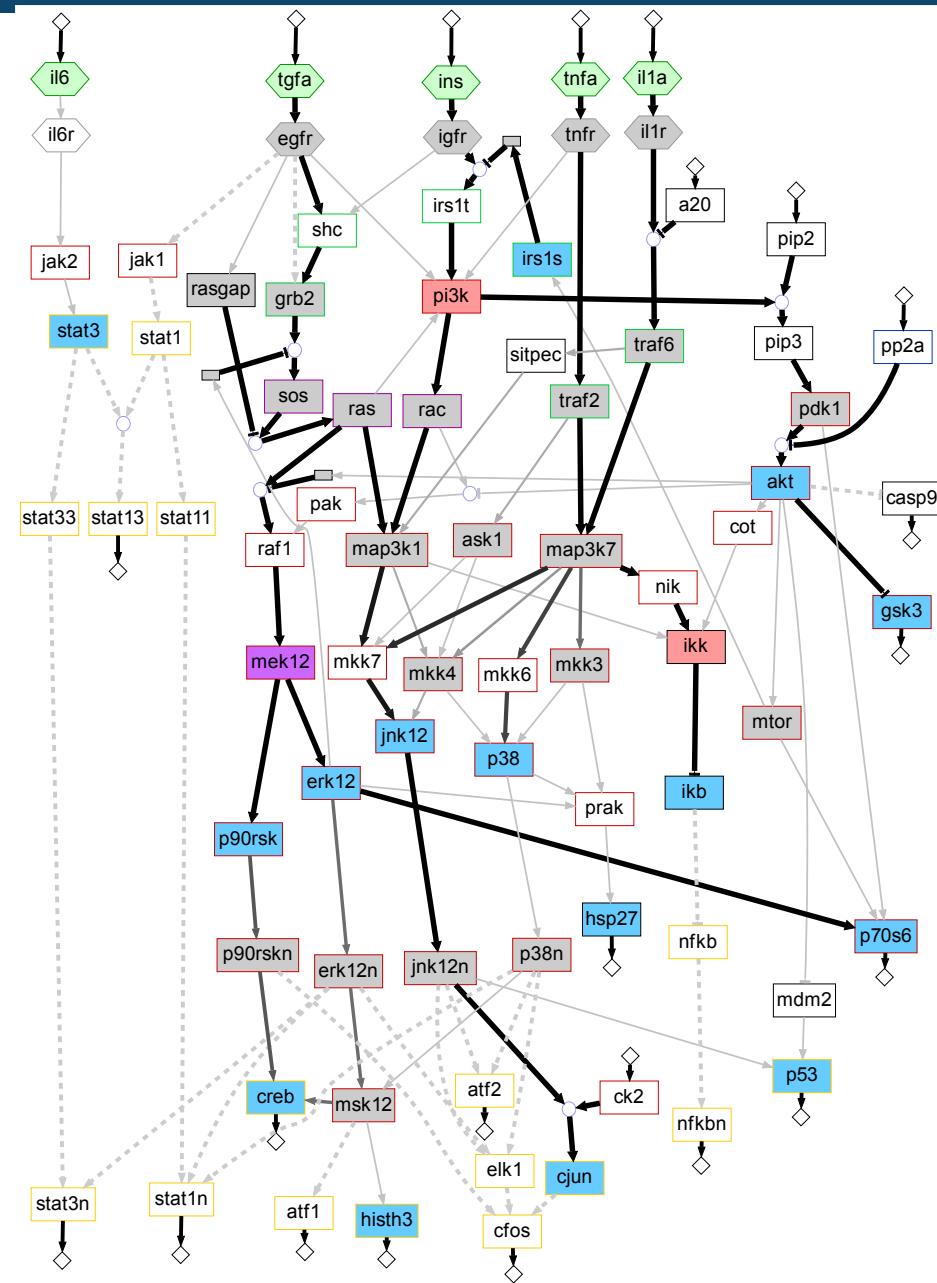
Readout

Perturb&Read

Kept

Removed

No effect





# Huh7

Stimulus

Perturbation

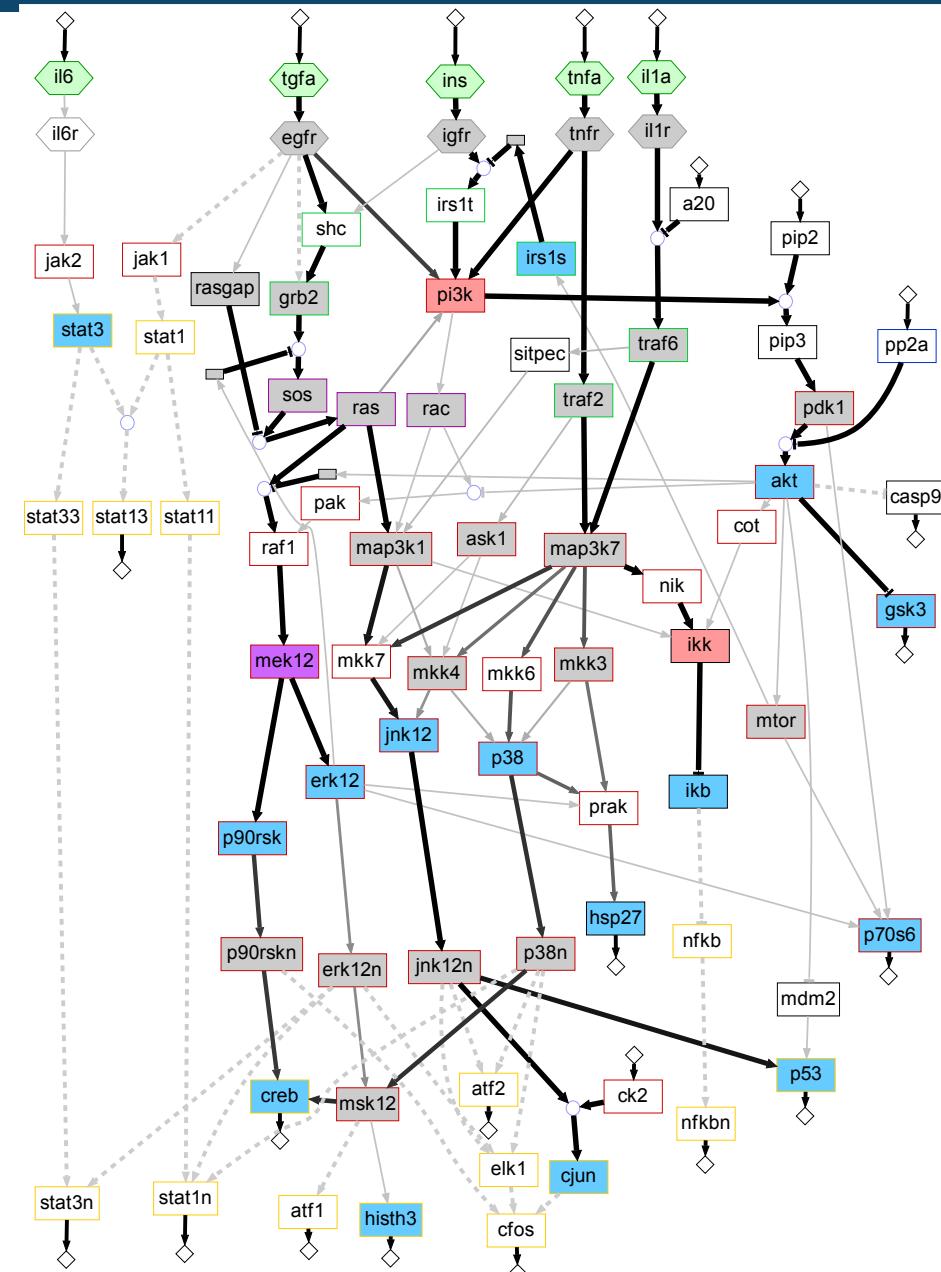
Readout

Perturb&Read

Kept

Removed

No effect





Stimulus

Perturbation

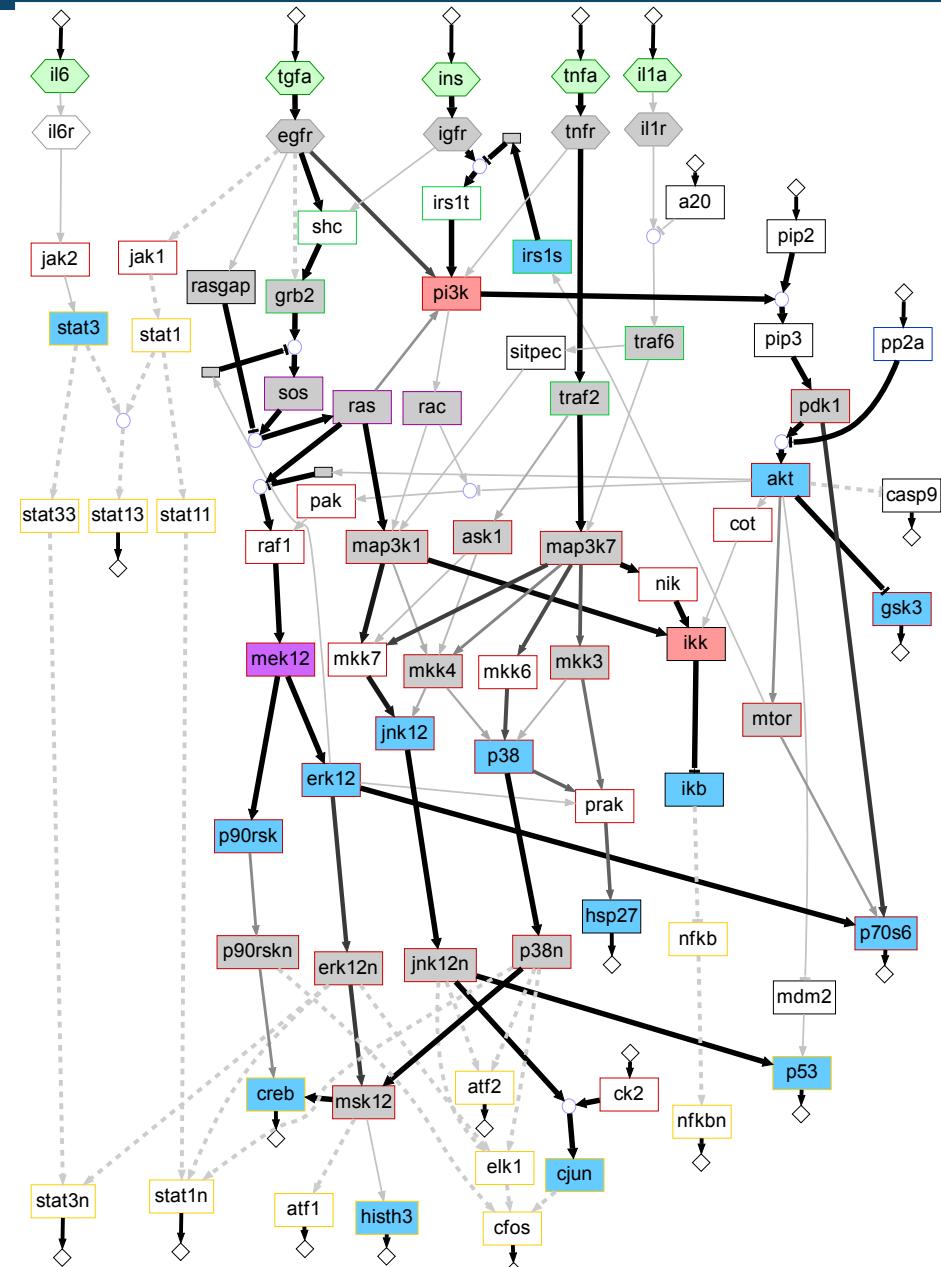
Readout

Perturb&Read

Kept

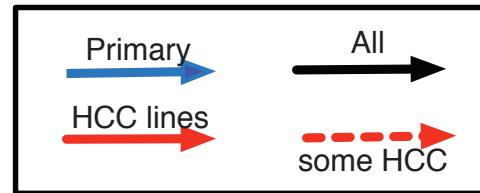
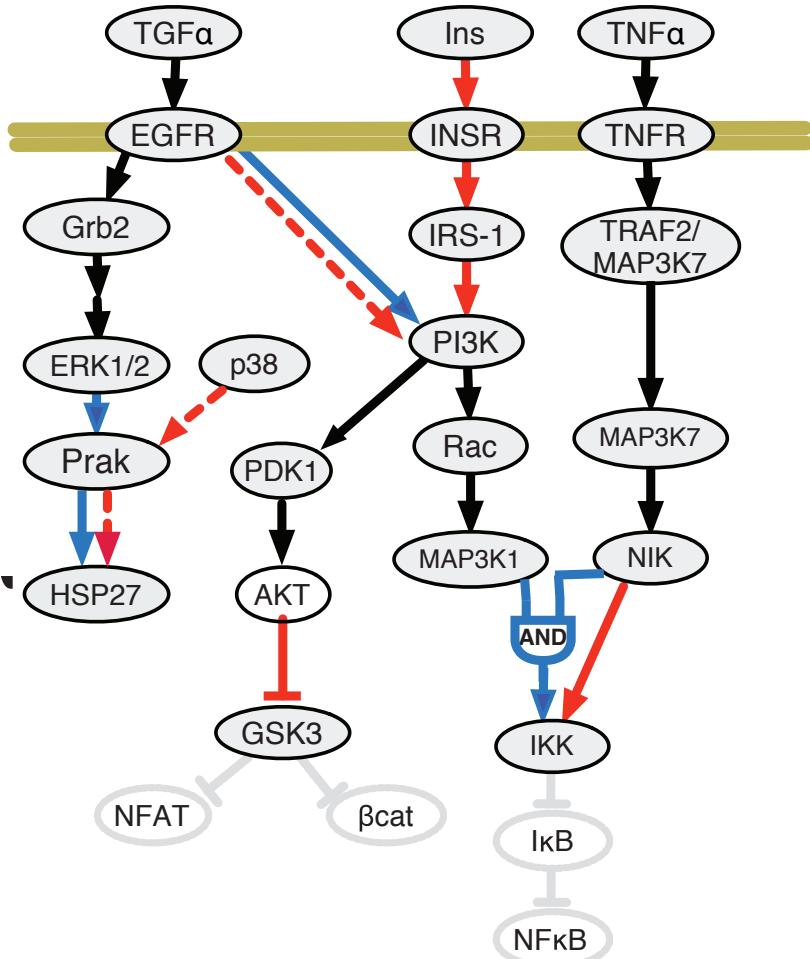
Removed

No effect



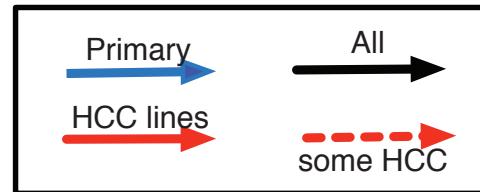
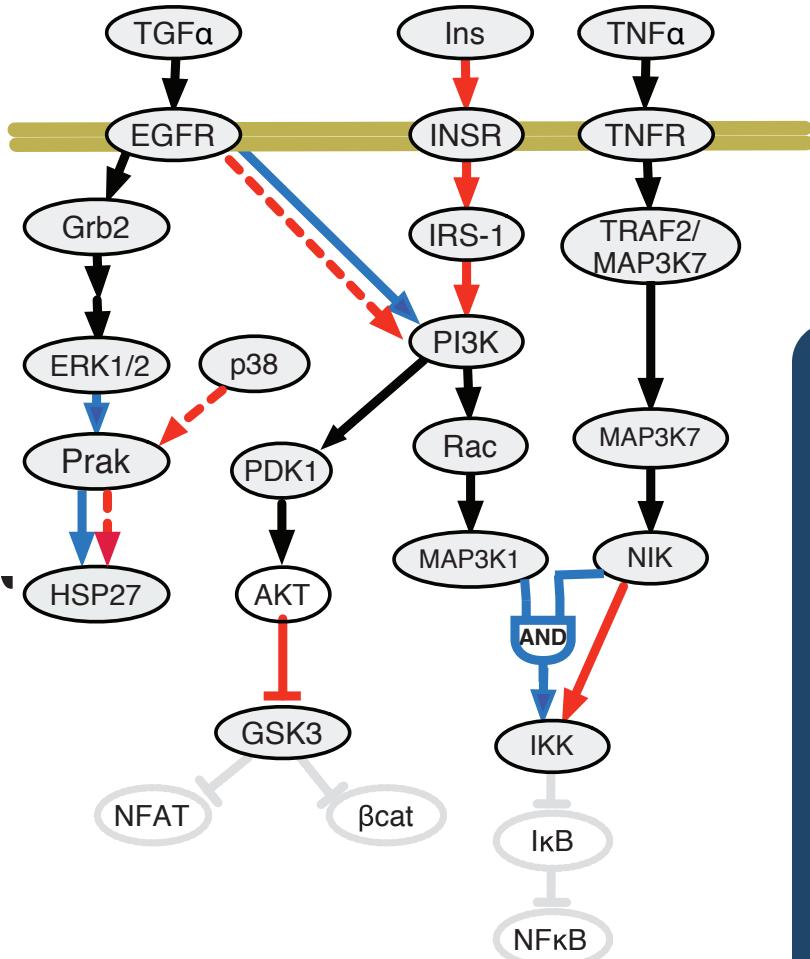


# Logic models provide mechanistic insight in signal deregulation





# Logic models provide mechanistic insight in signal deregulation



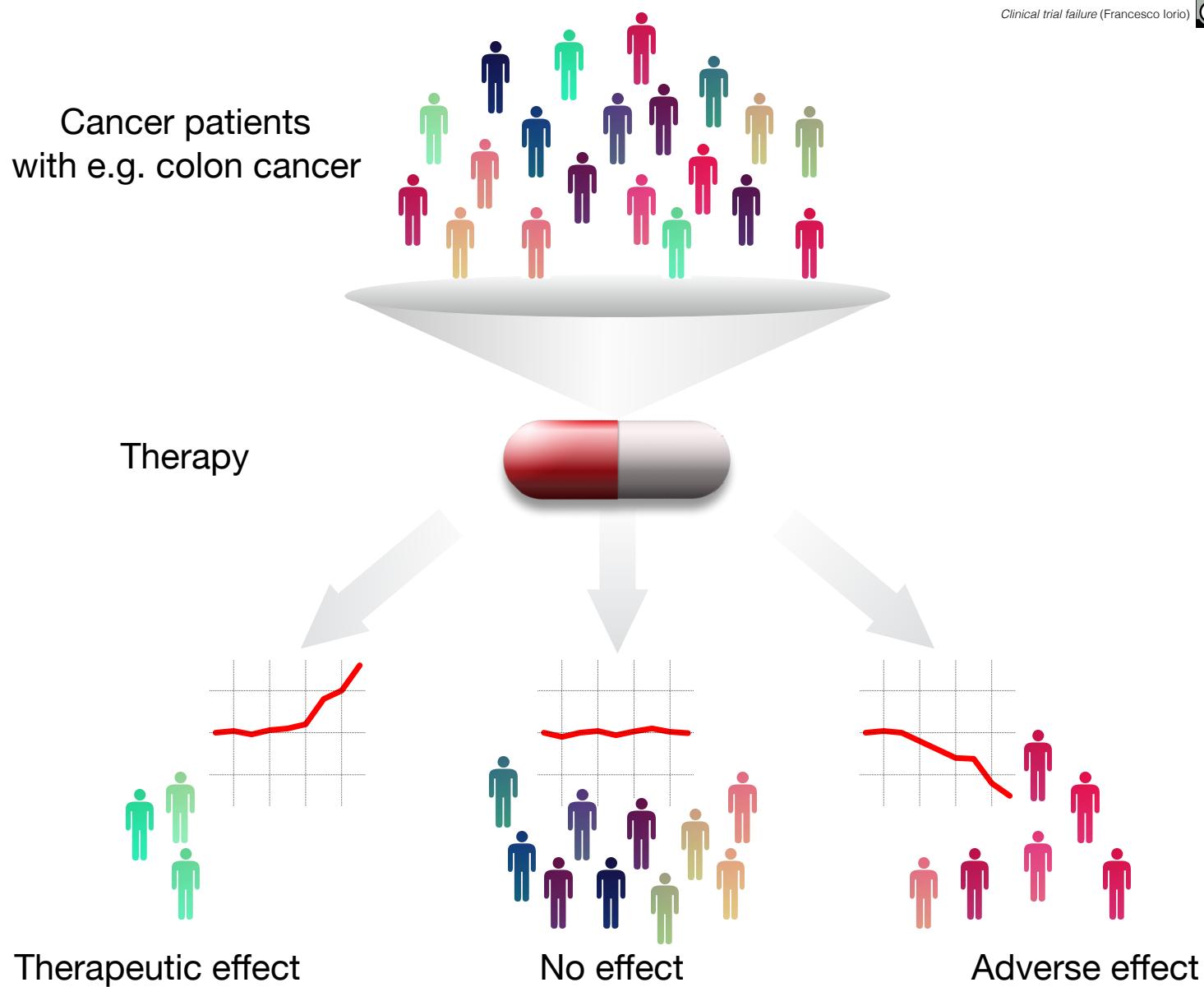
These models can:

- Identify functional differences between cell types (e.g. health vs disease)  
→ therapeutic targets
- Predict outcome of new perturbations (single or combination)
- Characterize targets and mode of action of drugs (Mitsos et al. *PLoS C.B.* 2009)



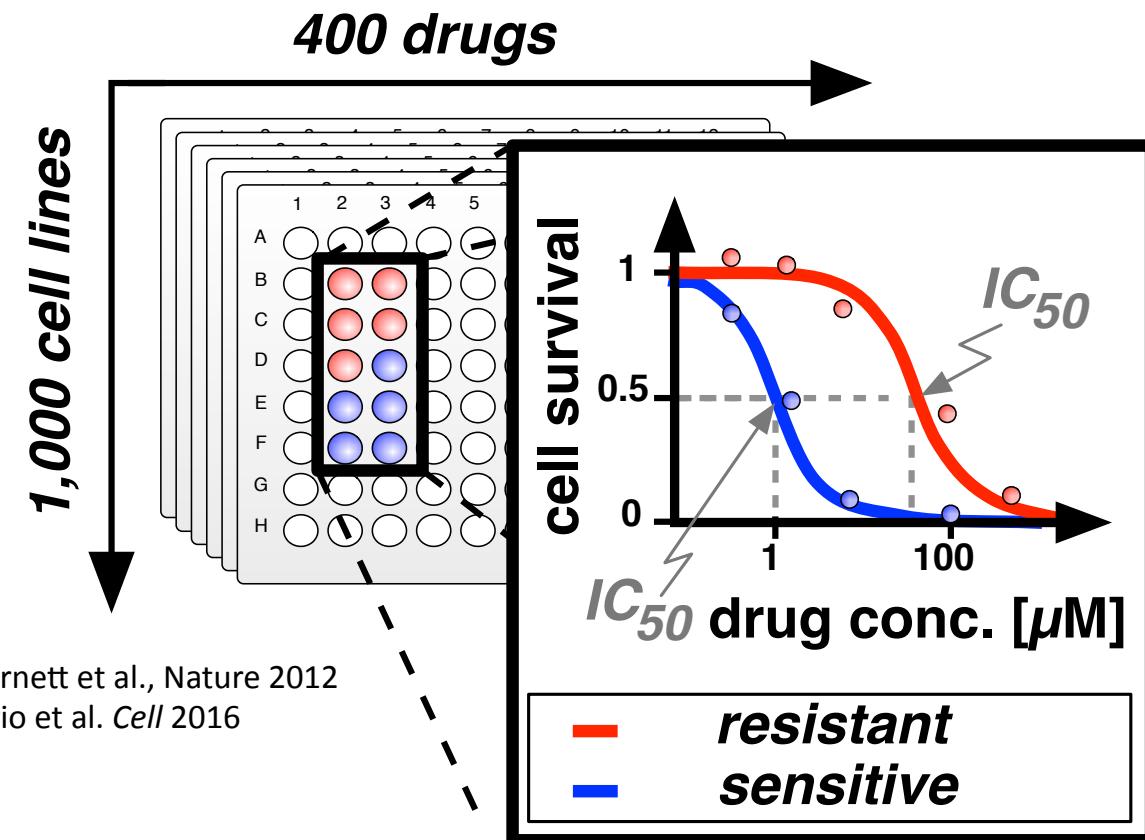
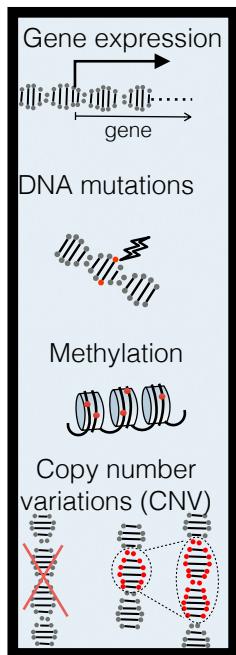
# Can we use logic models to understand drug efficacy in cancer?

Clinical trial failure (Francesco Iorio)



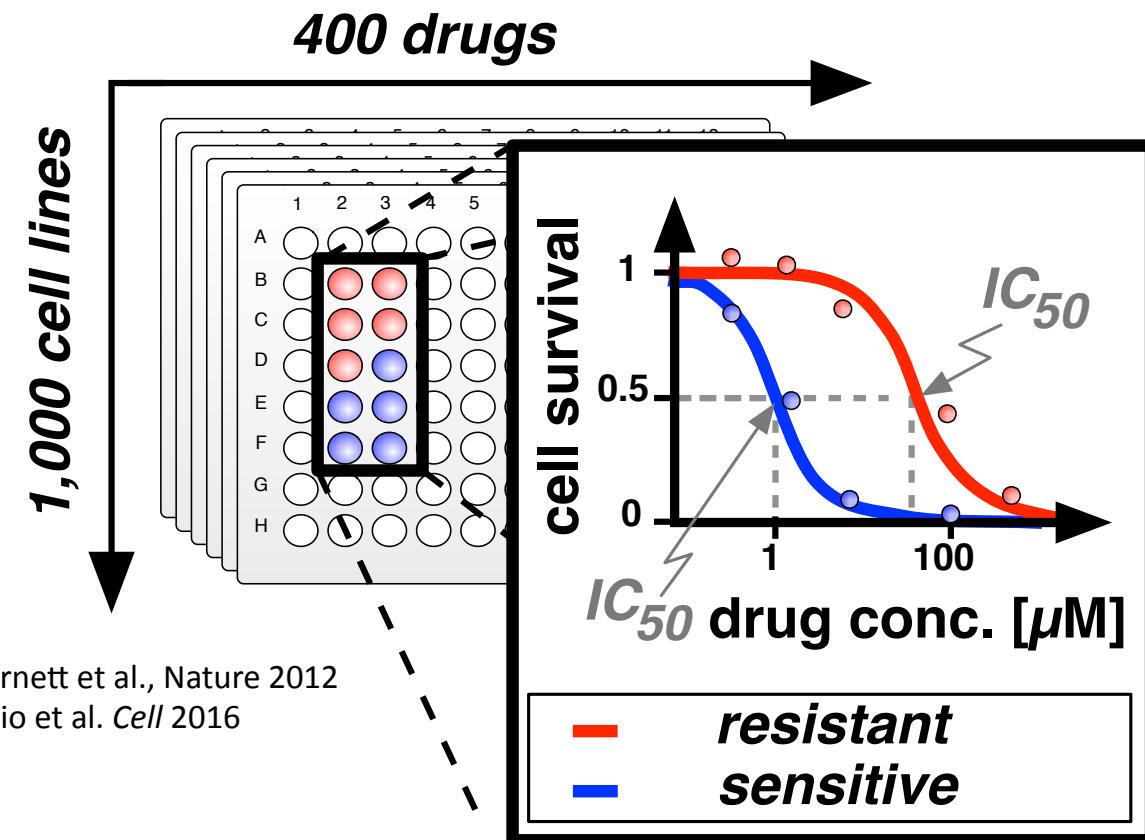
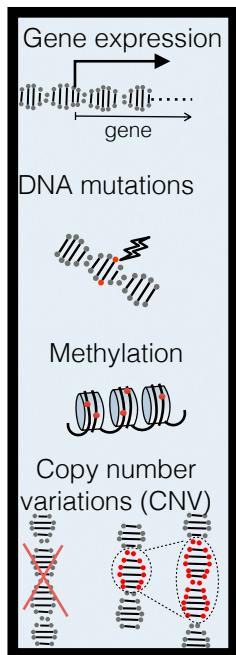


# Can we use logic models to understand drug efficacy in cancer?



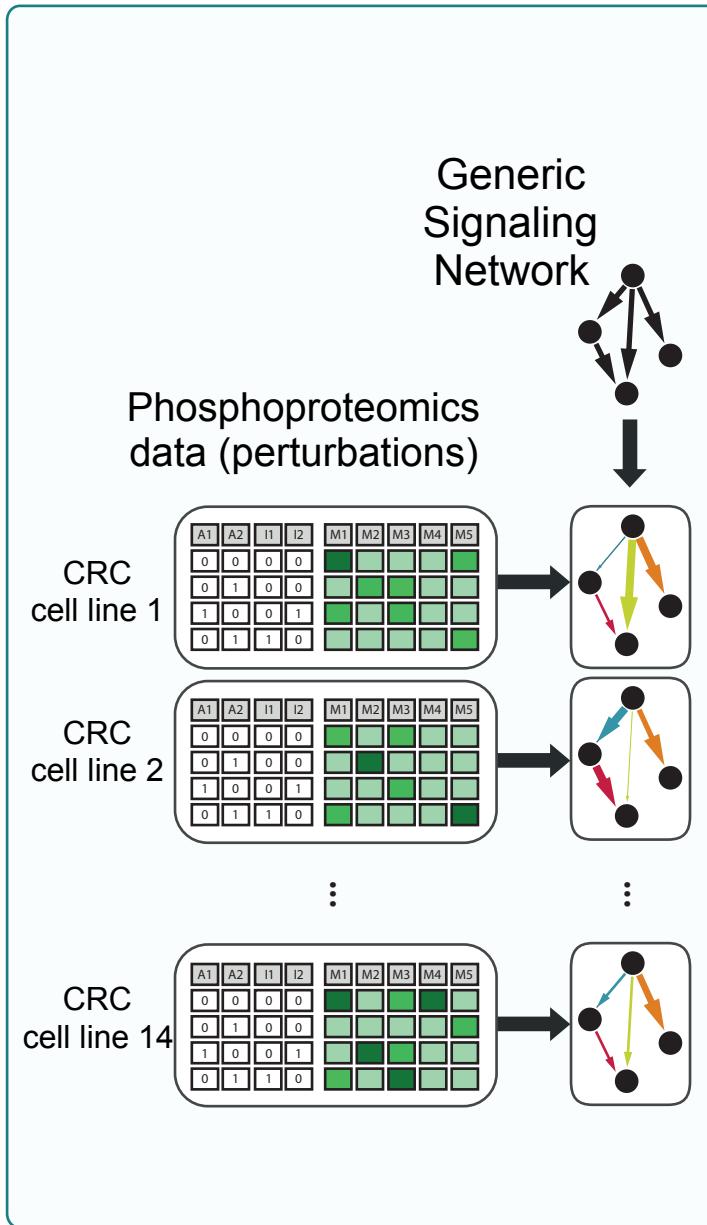


# Can we use logic models to understand drug efficacy in cancer?



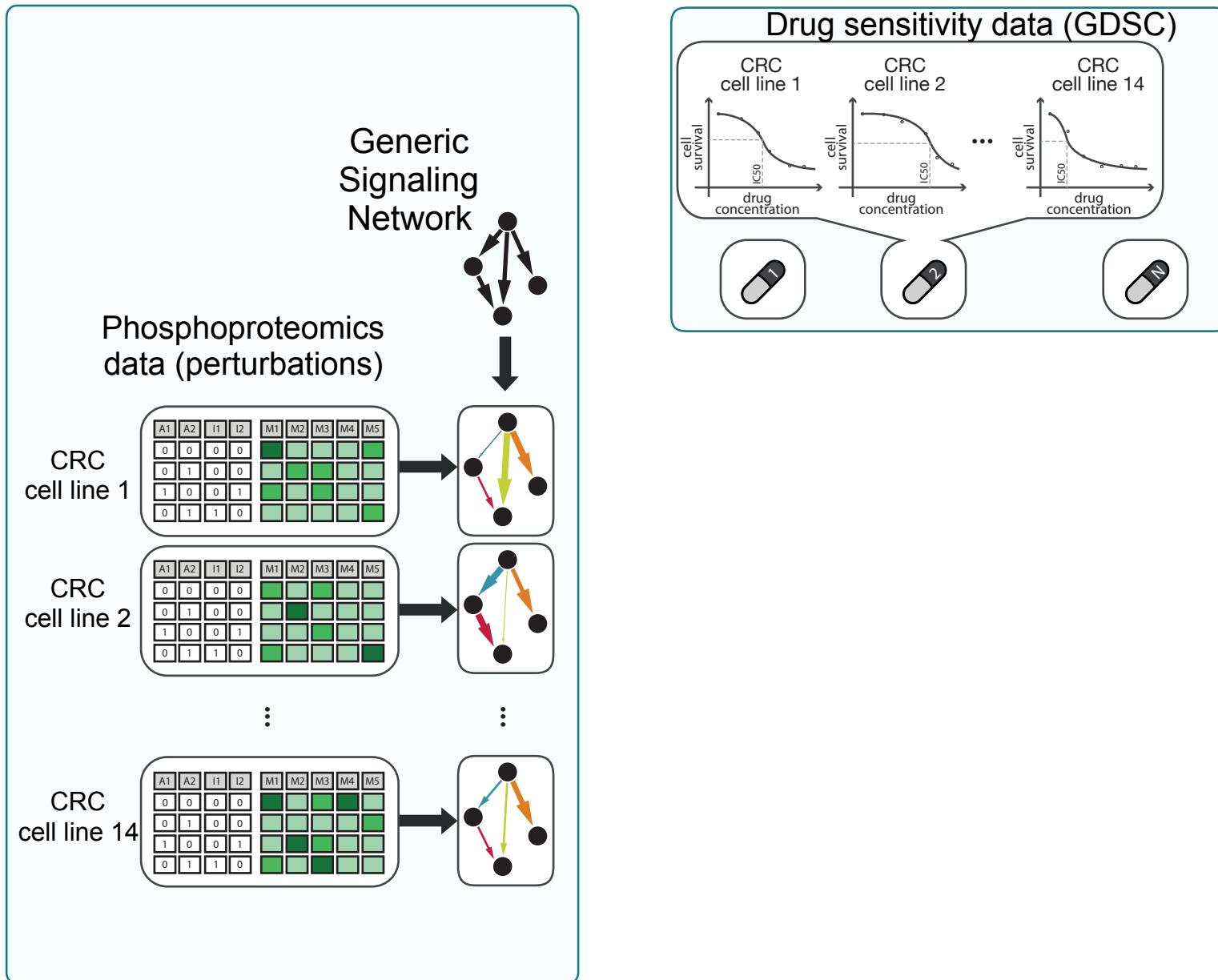


# Looking for model-based biomarkers of drug sensitivity



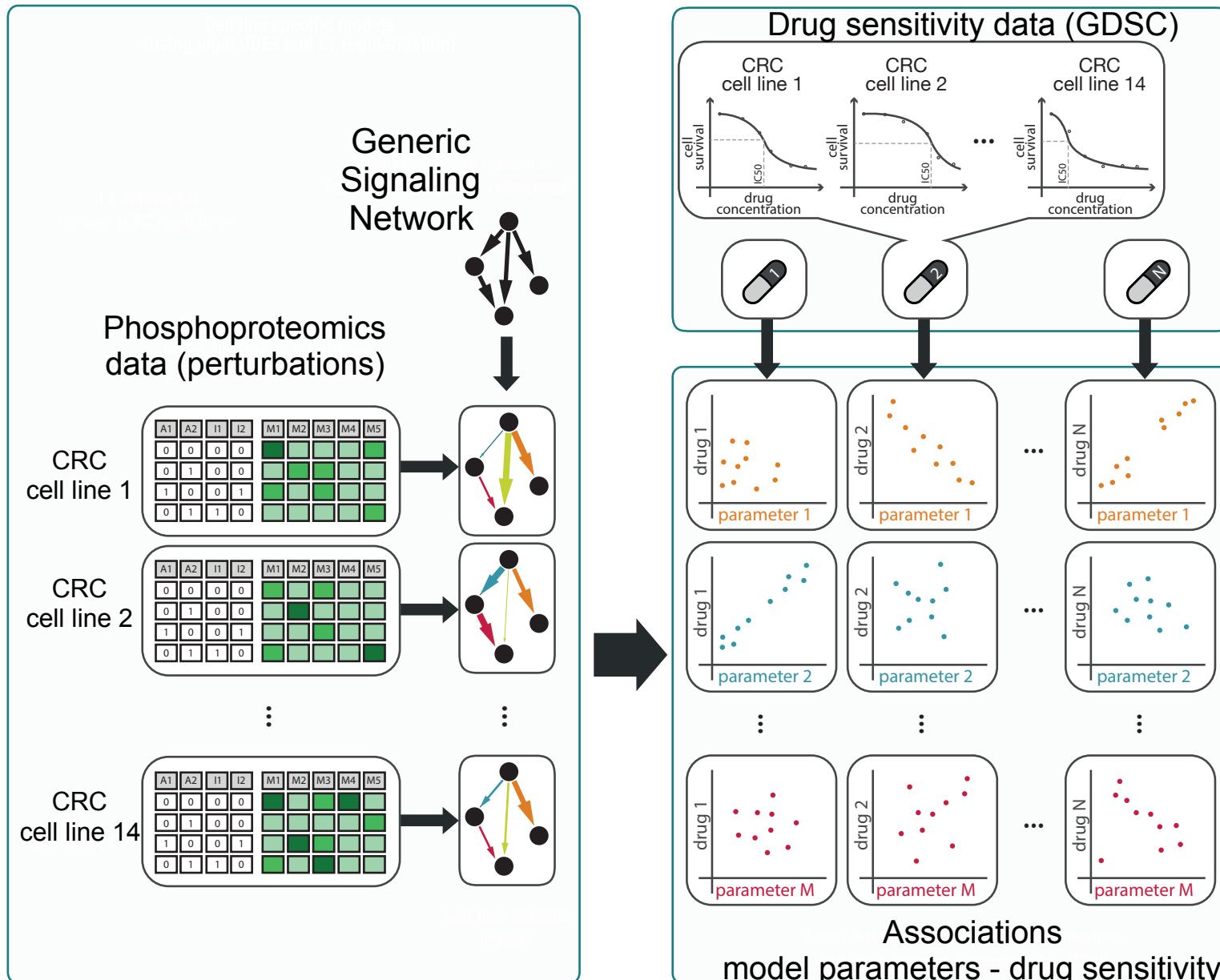


# Looking for model-based biomarkers of drug sensitivity





# Looking for model-based biomarkers of drug sensitivity





# Case study: understanding drug resistance in colorectal cancer

## 14 colorectal cancer cell lines

From GDSC (genomic & drug response data available)

## Luminex phospho data:

- 14 measured phospho-proteins
- 7 targeted drugs + 4 ligands (42 conditions)

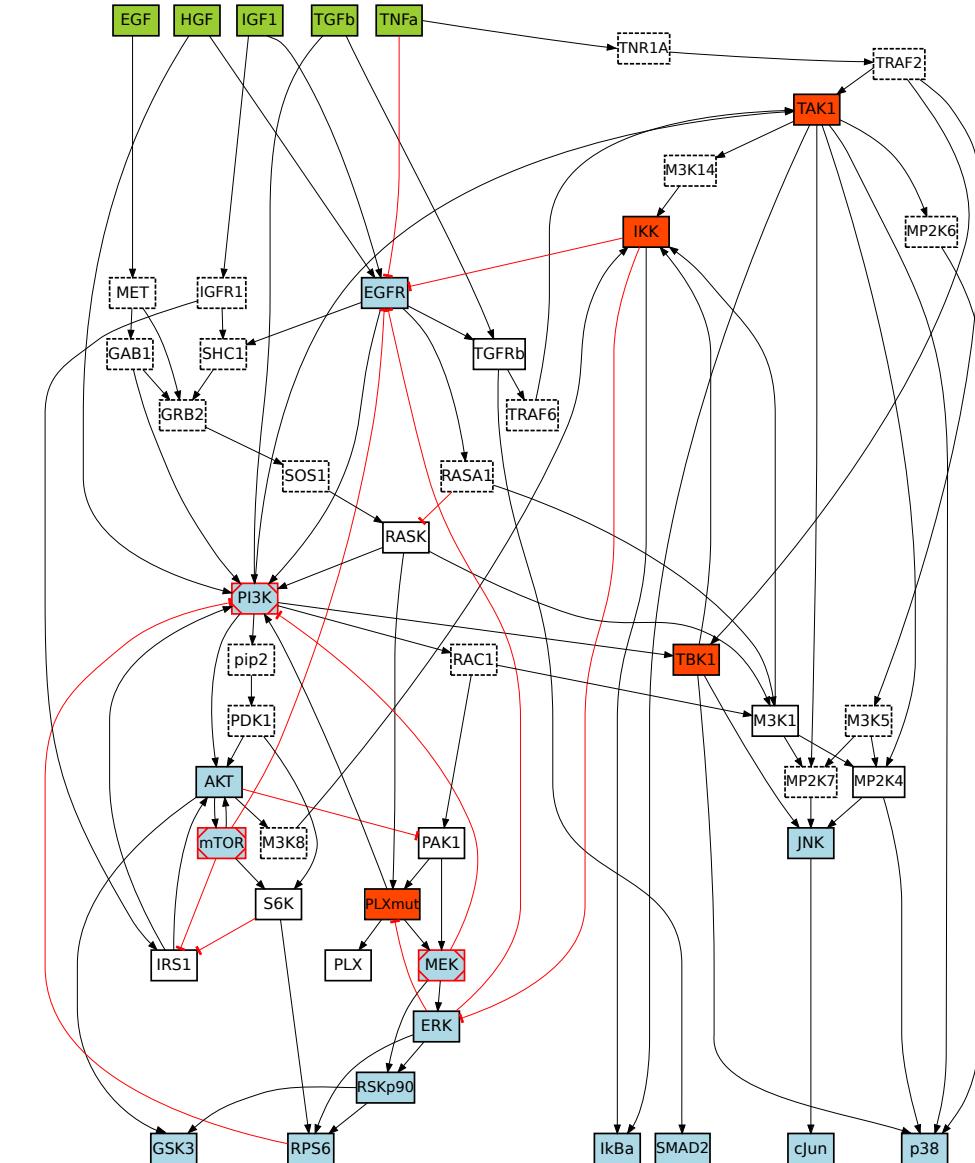
## Prior knowledge network

- derived from literature (~ 50 references)

Federica Eduati

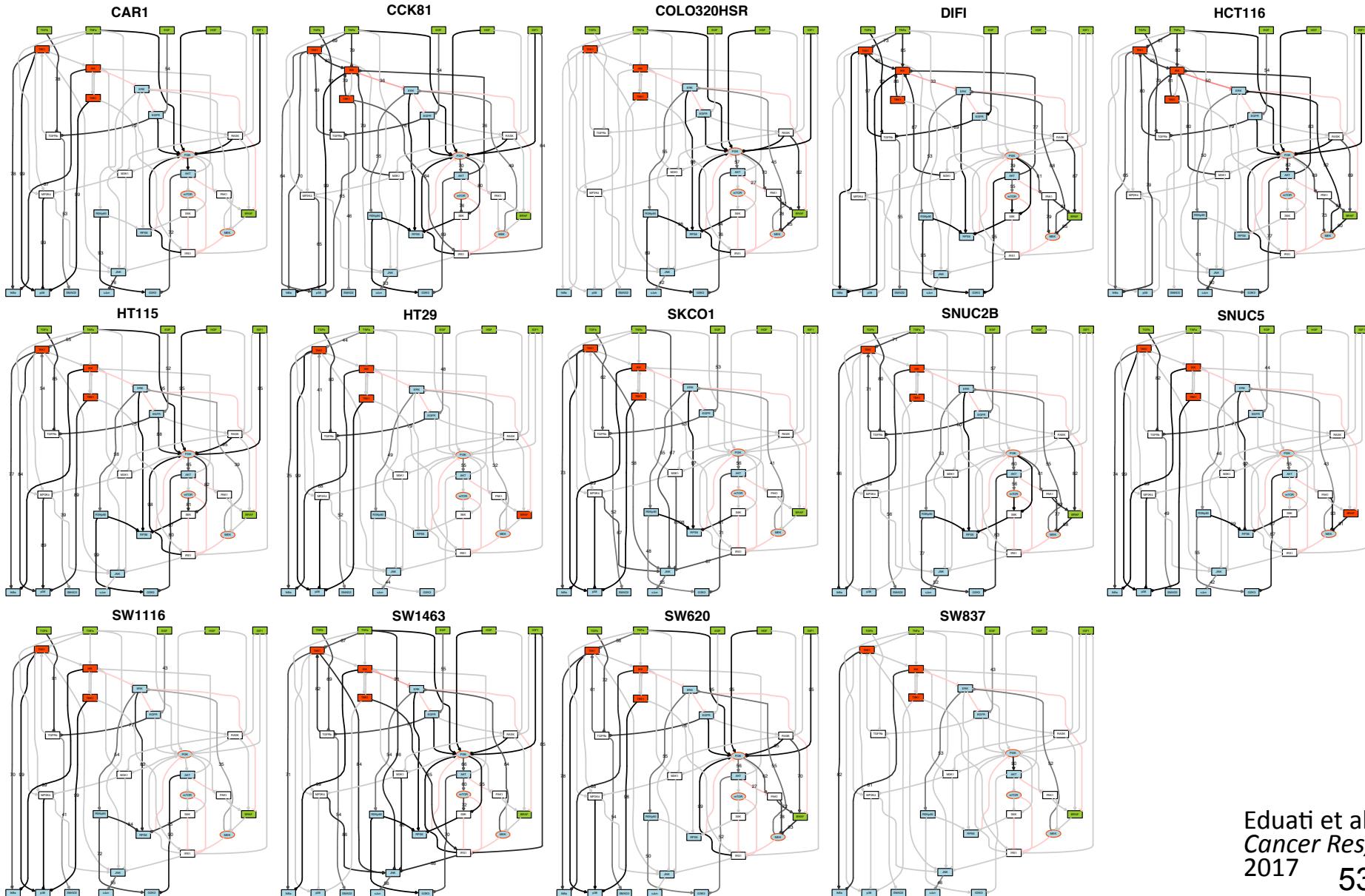
w. Nils Bluetghen (Charite) & Mathew Garnett (Sanger)

Eduati et al. submitted



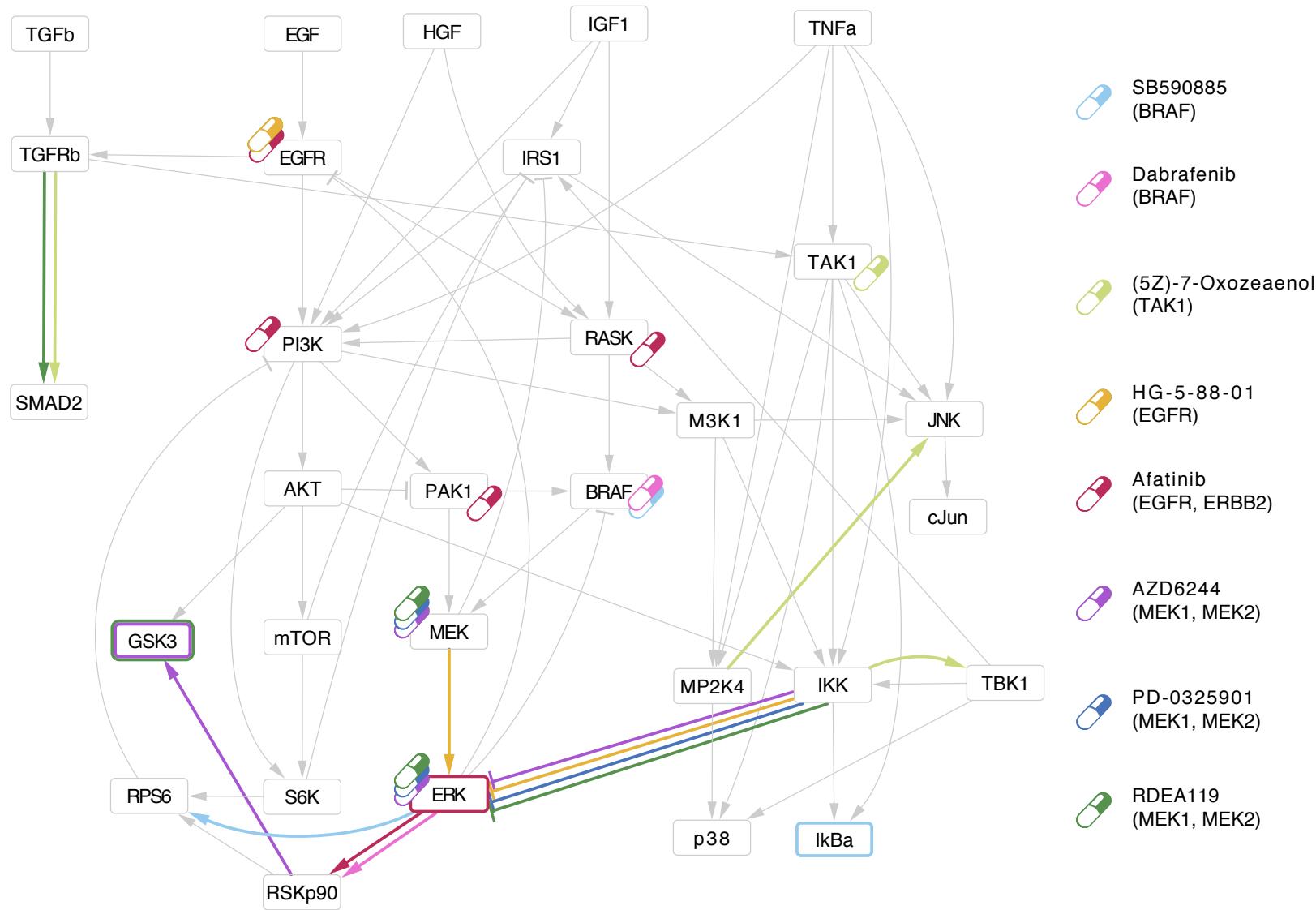


# CRC Cell line specific models





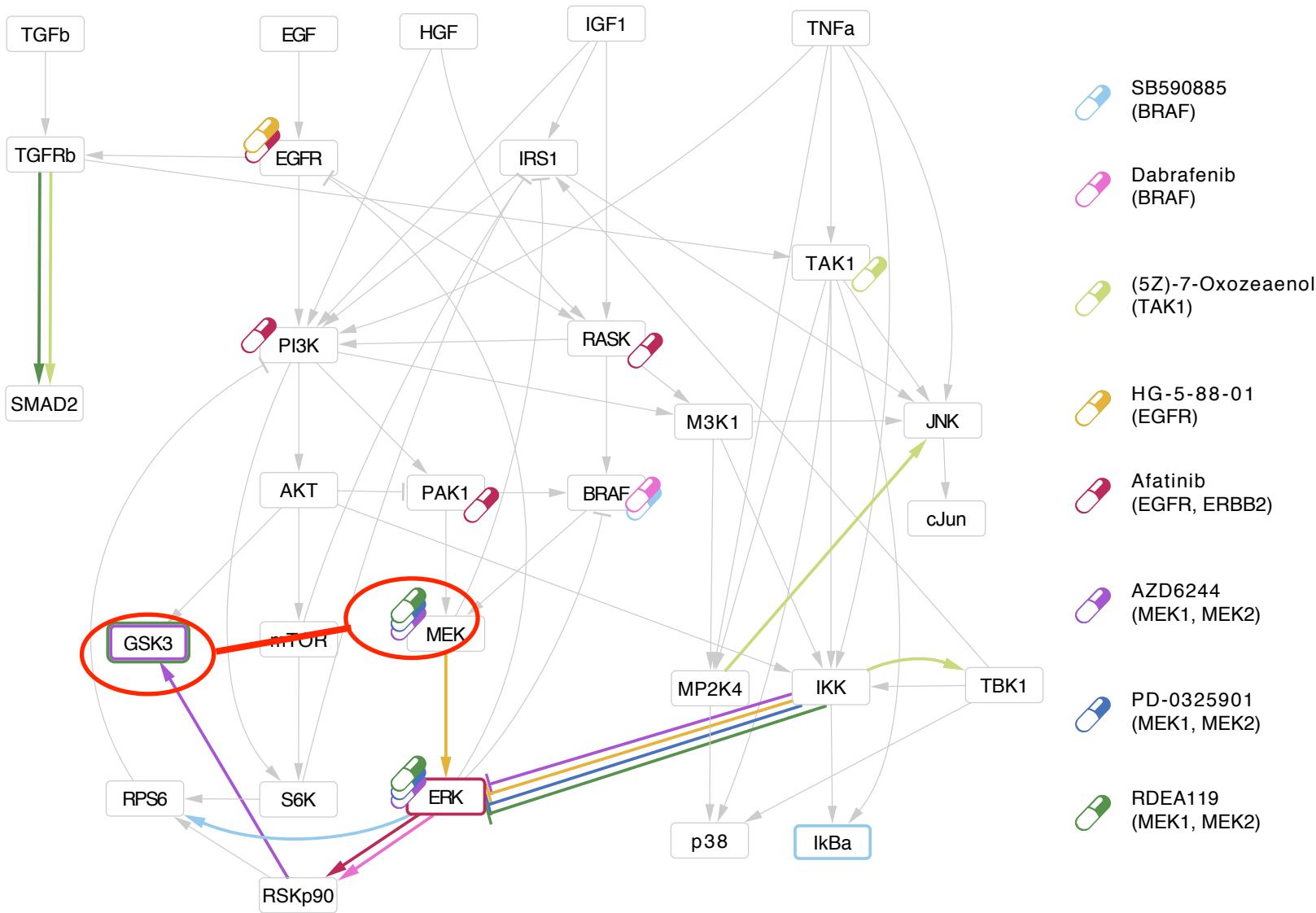
# Model-based biomarkers of drug efficacy and resistance



No genetic biomarker



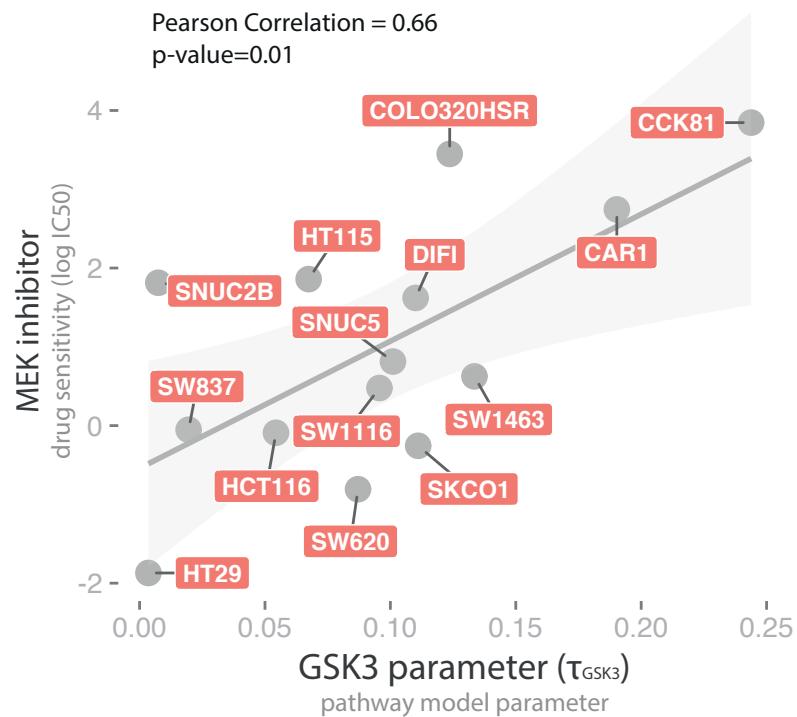
# Model-based biomarkers of drug efficacy and resistance



No genetic biomarker

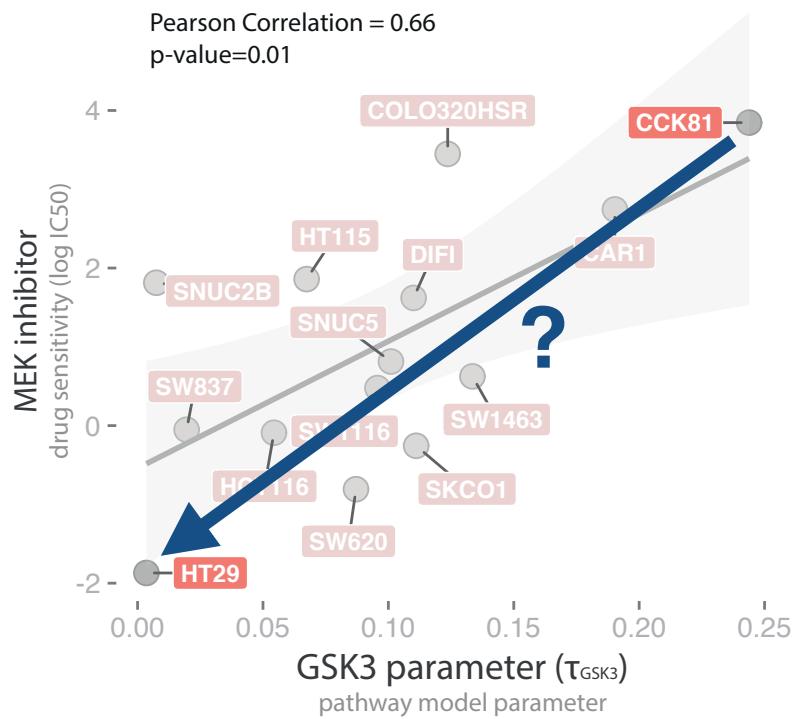


# Association between GSK3 functionality and MEK inhibitor efficacy suggests combination



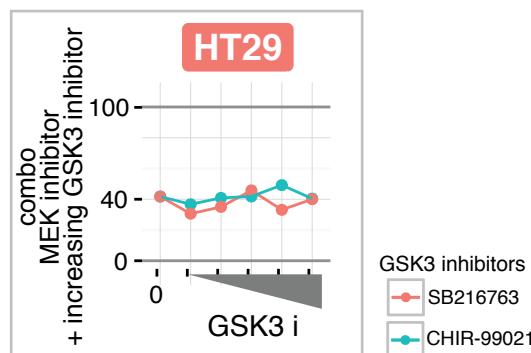
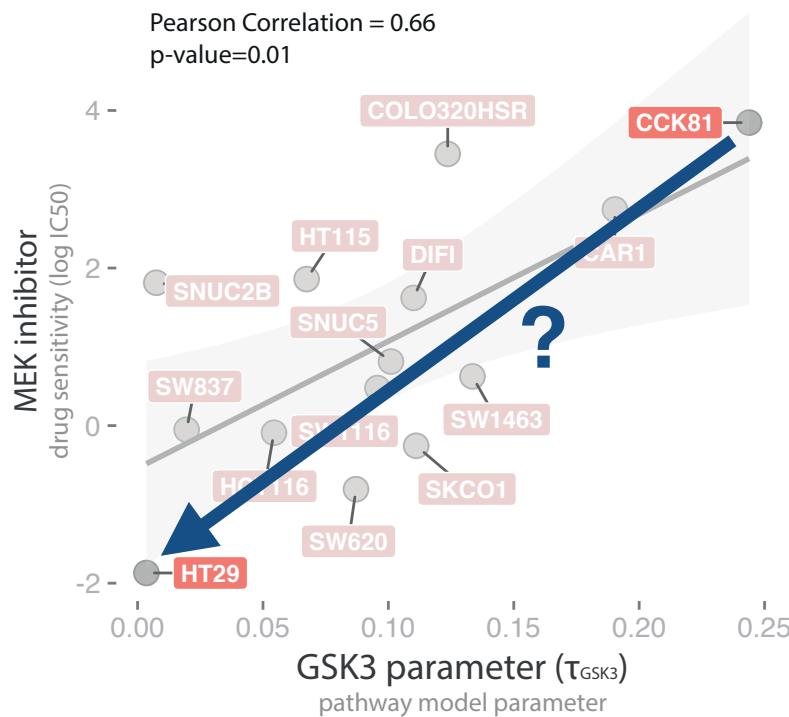


# Association between GSK3 functionality and MEK inhibitor efficacy suggests combination





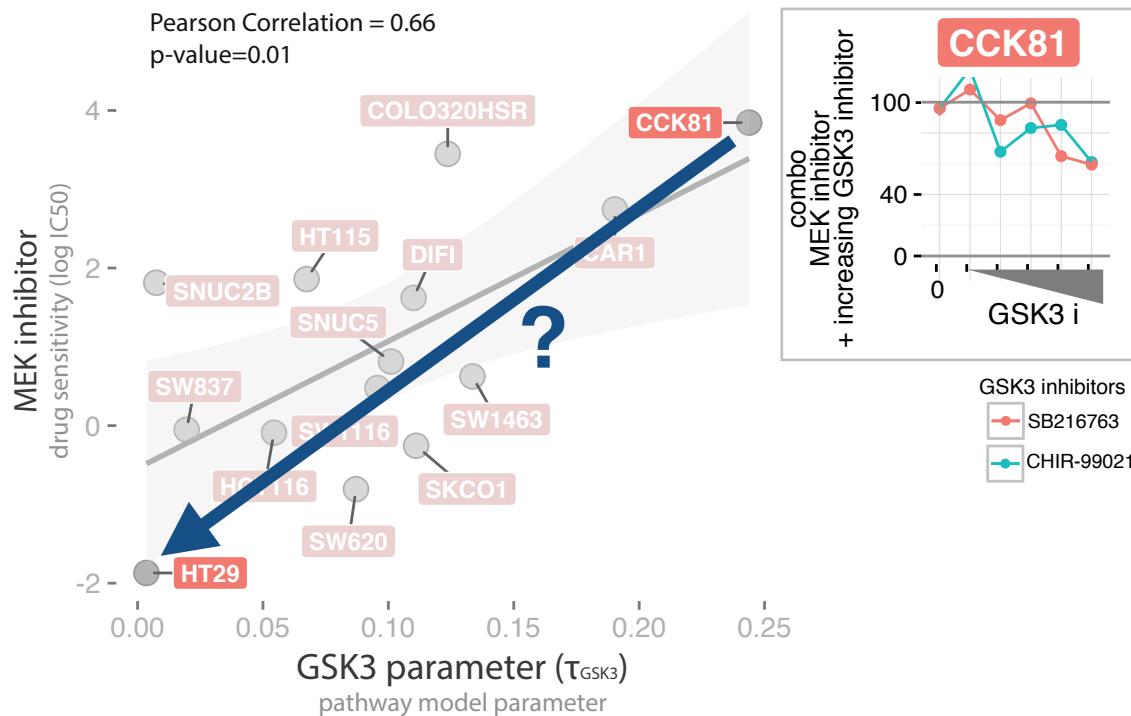
# Association between GSK3 functionality and MEK inhibitor efficacy suggests combination



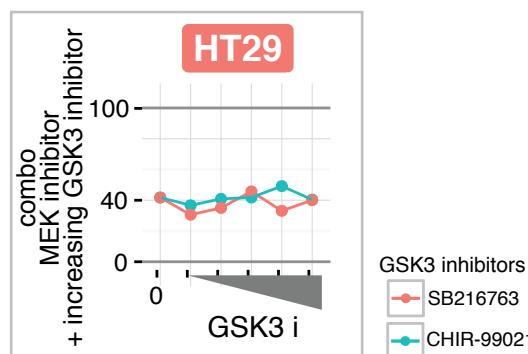
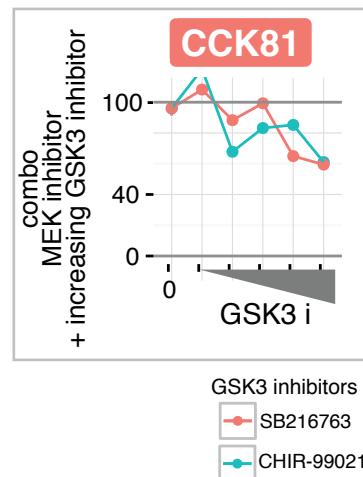
no improved sensitivity when GSK3 is not functional



# Association between GSK3 functionality and MEK inhibitor efficacy suggests combination



synergistic combo when GSK3 is functional



no improved sensitivity when GSK3 is not functional



# How to...

- Set up experiments to extract most information
- Process data efficiently
- Choose type of mathematical model  
(given data, question, etc)
- Train models to experimental data
- Use models to gain insight



# How to...

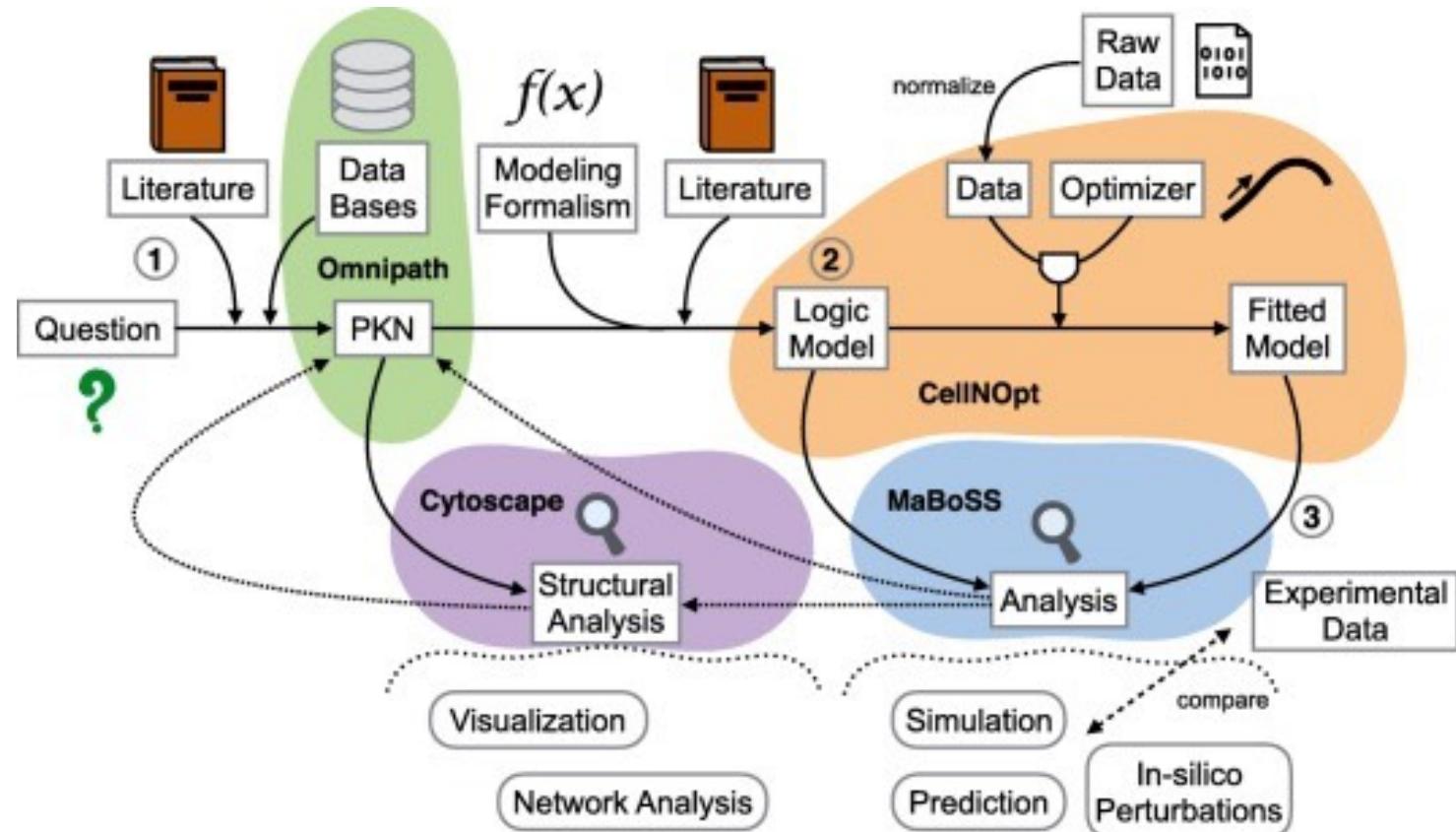
- Set up experiments to extract most information
- Process data efficiently
- Choose type of mathematical model (given data, question, etc)
- Train models to experimental data
- Use models to gain insight

Used logic modelling & applications to signalling, but general principles hold for other modelling approaches & applications



# A detailed tutorial

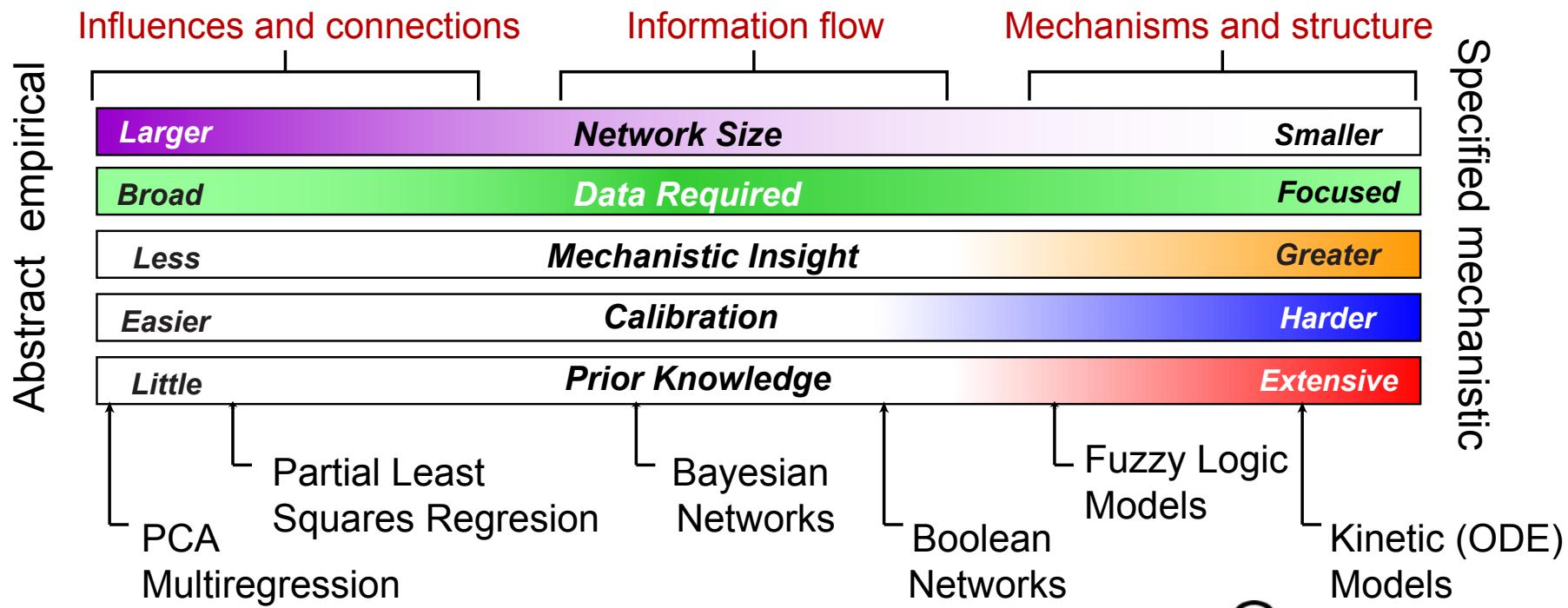
- Application of OmniPath, CellNOpt, MaBoSS and Cytoscape to a prostate cancer example





# Spectrum of modeling approaches

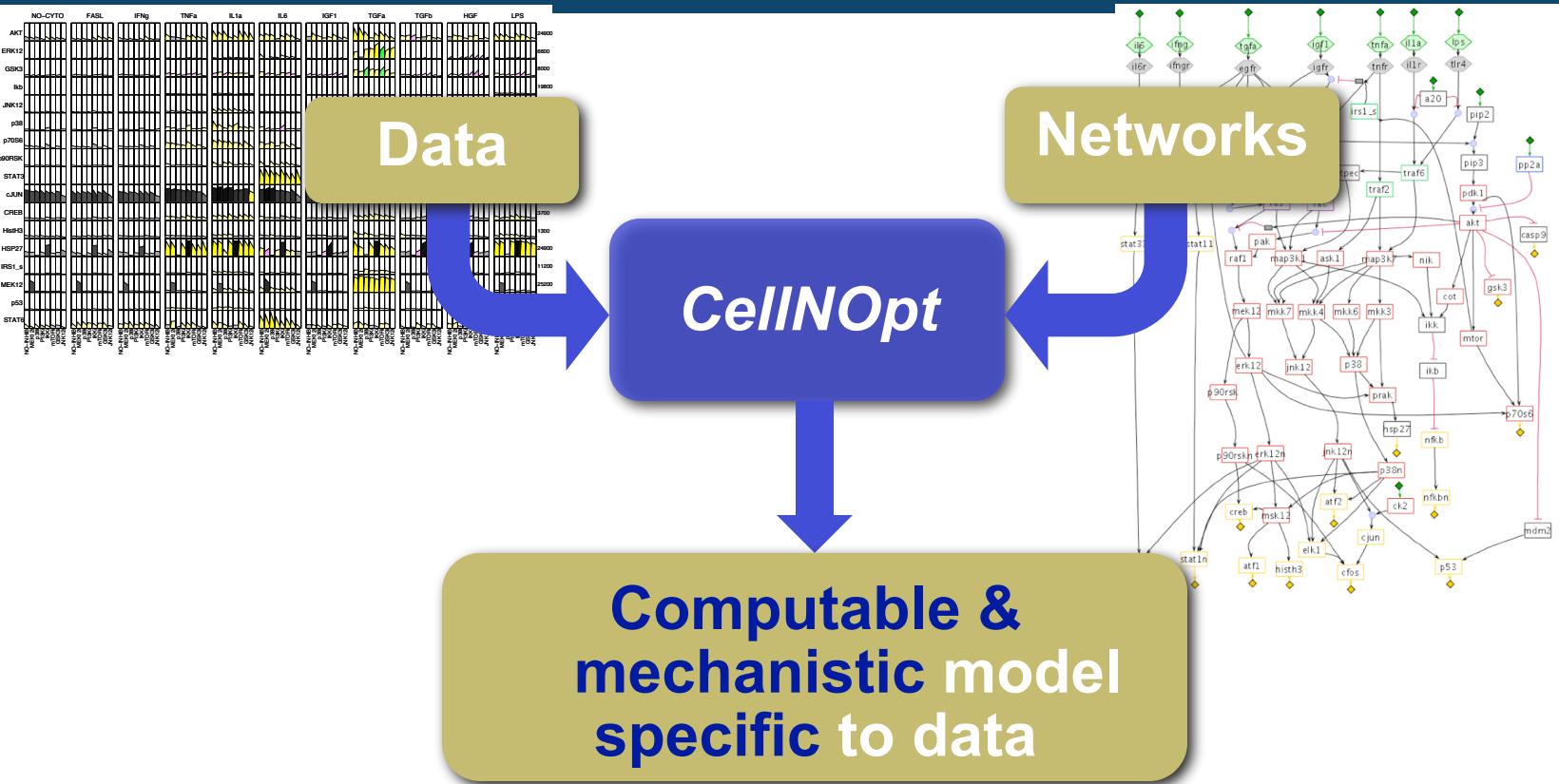
- Choice of method depends on:
  - Question, prior knowledge, data, ... (+ modeler's expertise)
  - Art more than science





# All models are wrong, but some are useful

G. Box



- Logic models: intermediate between data-driven & biochemical models
- Flexible and scalable framework
- Suitable to integrate large-scale data + networks



# Acknowledgements

[www.saezlab.org](http://www.saezlab.org)  [sysbiomed](https://sysbiomed.org)  
[www.github.com/saezlab](https://github.com/saezlab)

Current members:

Denes Turei	Angeliki Kalamara
Damien Arnol	Aurelien Dugourd
Melanie Rinas	Luis Tobalina
Charlie Pieterman	
Vignesh Subramanian	
Mi Yang	Attila Gabor
Hyojin Kim	Nicolas Palacios
Christian Holland	Enio Gjerga
Panuwat Trairatphisan	Anika Liu



Mathew Garnett Ultan McDermott (Sanger)  
DREAM Challenges, sp. Gustavo Stolovitzky (IBM)  
Thorsten Cramer Ulf Neumann (RWTH Aachen)  
Ruedi Aebersold Uwe Sauer (ETH Zurich)  
Laurence Calzone (Institut Curie)  
Bernd Bodenmiller (U Zurich)  
Leonidas Alexopoulos (NTUA)  
Rafael Kramann (Uniklinik Aachen)  
Christian Frezza (Cambridge)

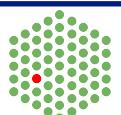
Lodewyk Wessels (NKI)  
Oliver Stegle Pedro Beltrao (EMBL-EBI)  
Christoph Merten (EMBL)  
CoLoMoTo consortium  
Tamas Korcsmaros (EI)  
Nils Bluethgen (Charite)  
Julio Banga (CSIC)  
Anne Claude Gavin (EMBL)  
Jesper Olsen (Copenhagen)



Open Targets



Sys4MS





# Acknowledgements

[www.saezlab.org](http://www.saezlab.org)  [sysbiomed](https://sysbiomed.org)  
[www.github.com/saezlab](https://github.com/saezlab)

Current members:

Denes Turei	Angeliki Kalamara
Damien Arnol	Aurelien Dugourd
Melanie Rinas	Luis Tobalina
Charlie Pieterman	
Vignesh Subramanian	
Mi Yang	Attila Gabor
Hyojin Kim	Nicolas Palacios
Christian Holland	Enio Gjerga
Panuwat Trairatphisan	Anika Liu



Mathew Garnett Ultan McDermott (Sanger)  
DREAM Challenges, sp. Gustavo Stolovitzky (IBM)  
Thorsten Cramer Ulf Neumann (RWTH Aachen)  
Ruedi Aebersold Uwe Sauer (ETH Zurich)  
Laurence Calzone (Institut Curie)  
Bernd Bodenmiller (U Zurich)  
Leonidas Alexopoulos (NTUA)  
Rafael Kramann (Uniklinik Aachen)  
Christian Frezza (Cambridge)

Lodewyk Wessels (NKI)  
Oliver Stegle Pedro Beltrao (EMBL-EBI)  
Christoph Merten (EMBL)  
CoLoMoTo consortium  
Tamas Korcsmaros (EI)  
Nils Bluethgen (Charite)  
Julio Banga (CSIC)  
Anne Claude Gavin (EMBL)  
Jesper Olsen (Copenhagen)



Open Targets



Sys4MS

