Logic models of signalling networks with CellNOpt:

from Boolean logic to differential equations

Julio Saez-Rodriguez

Joint Research Center for Computational Biomedicine

RWTH AACHEN UNIVERSITY

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

Plato’s allegory of the cave

http://www.zlw-ima.rwth-aachen.de/forschung/

Plato's allegory of the cave

http://www.zlw-ima.rwth-aachen.de/forschung/

- **Cues** are **lights**

Plato’s allegory of the cave

http://www.zlw-ima.rwth-aachen.de/forschung/

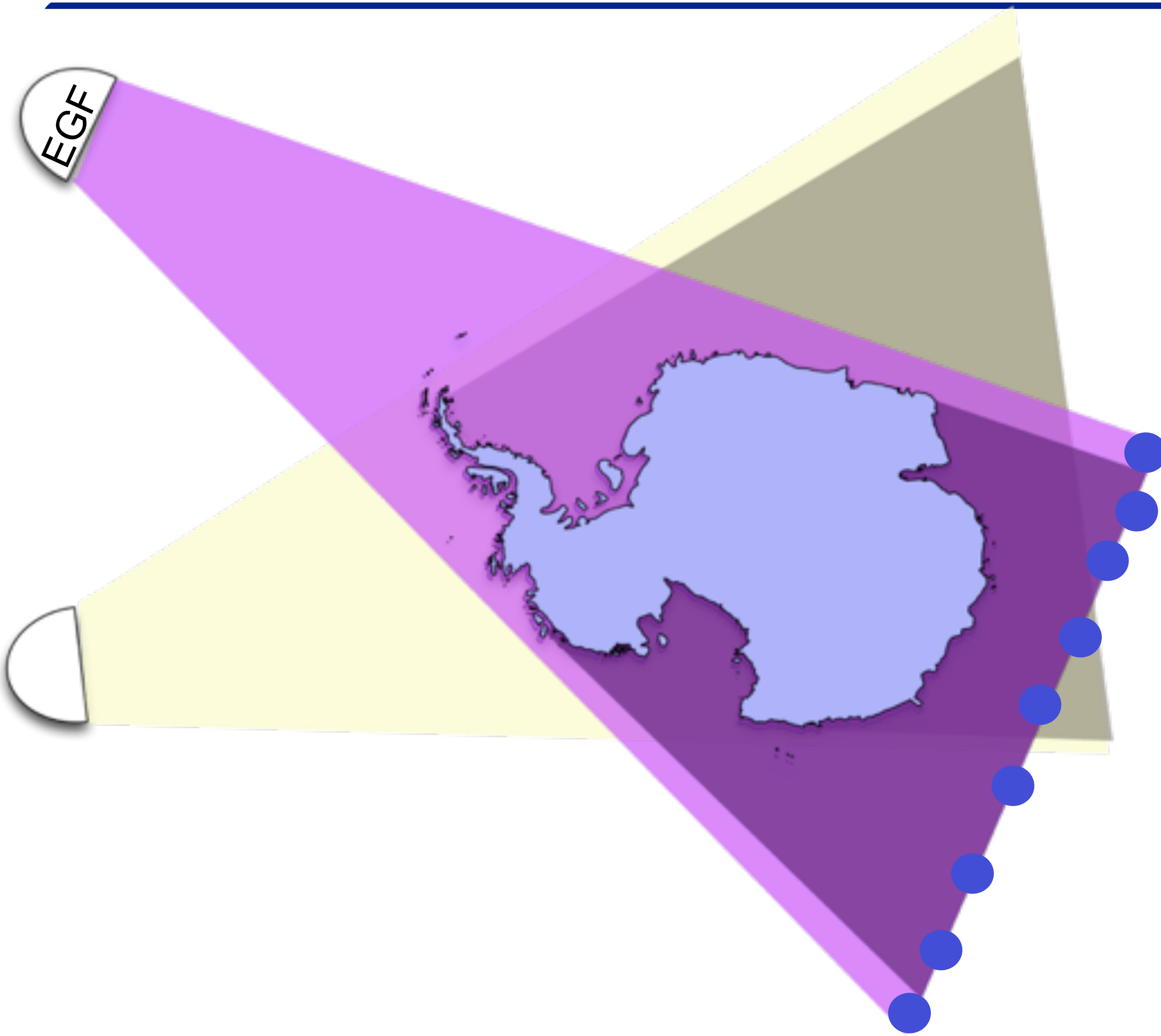
- **Cues** are **lights**
  - **Measurements** are **shadows**:
    - Phosphorylation = Activation? Which site? How does it affect the regulation of the protein?

Plato's allegory of the cave

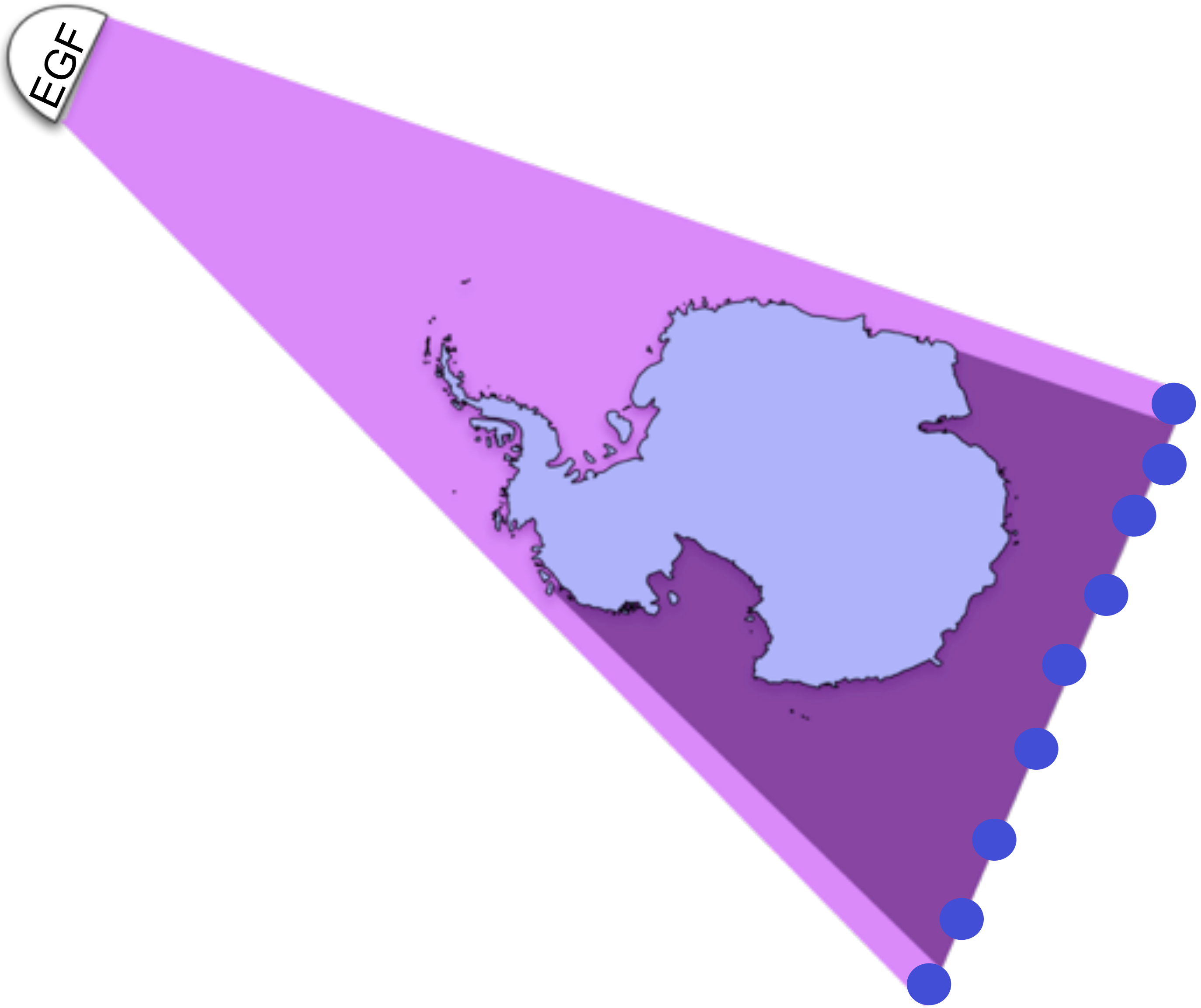
- **Cues** are **lights**
  - **Measurements** are **shadows**:
    - Phosphorylation = Activation? Which site? How does it affect the regulation of the protein?
  - Fluorescence=phosphorylation? Signal saturated? Below detection level?

http://www.zlw-ima.rwth-aachen.de/forschung/

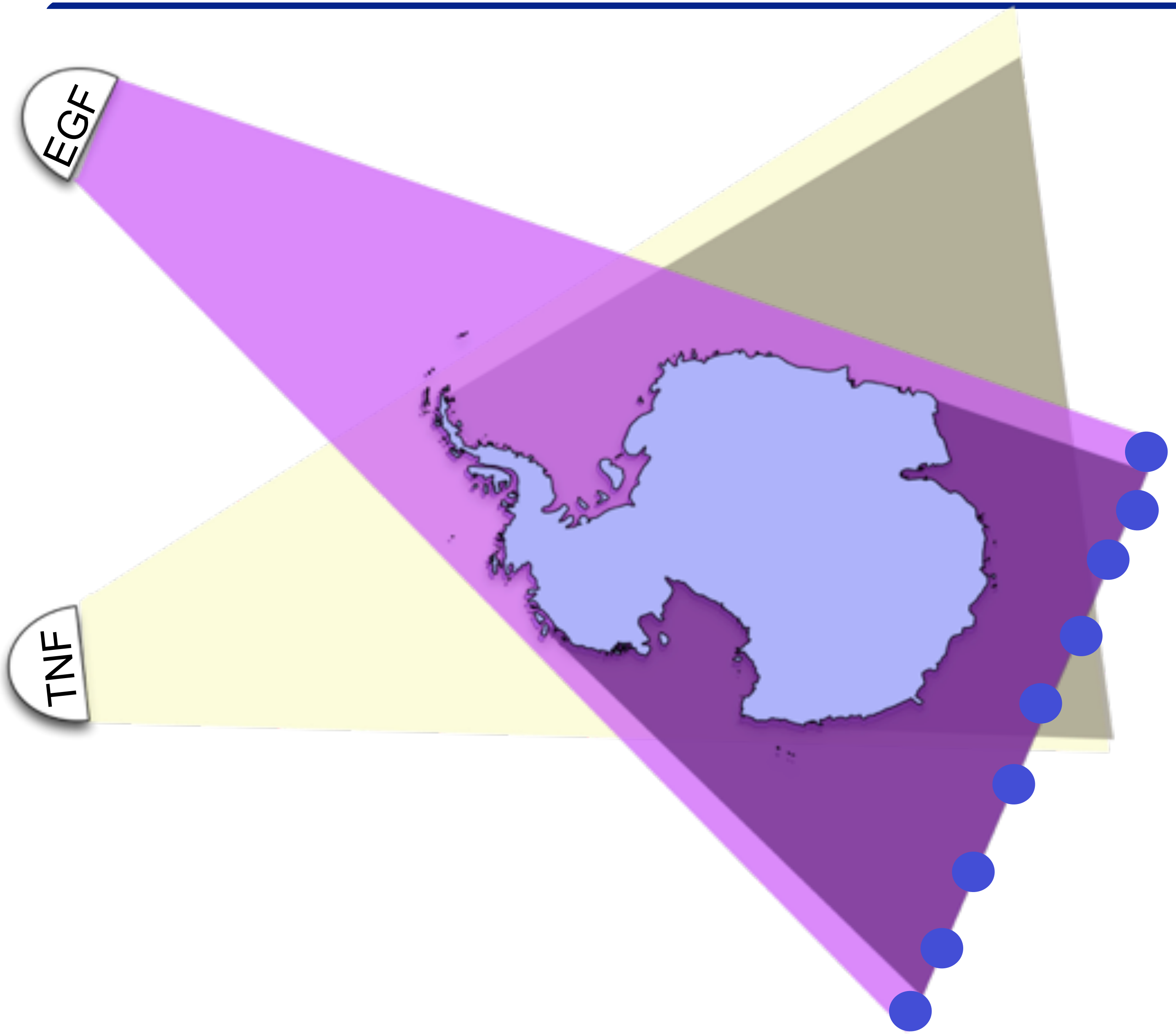
Choosing cues to maximize information



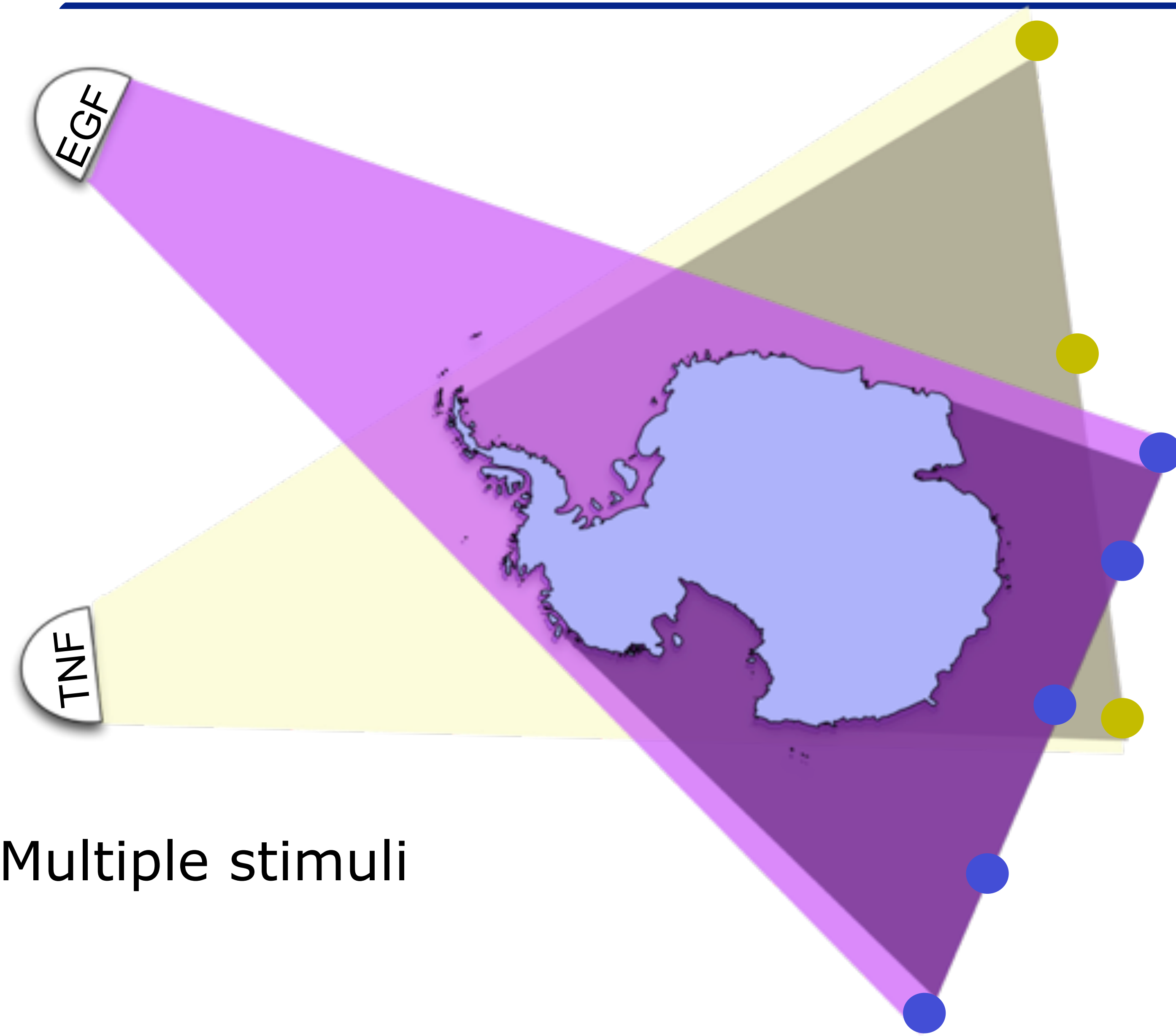
Choosing cues to maximize information



Choosing cues to maximize information

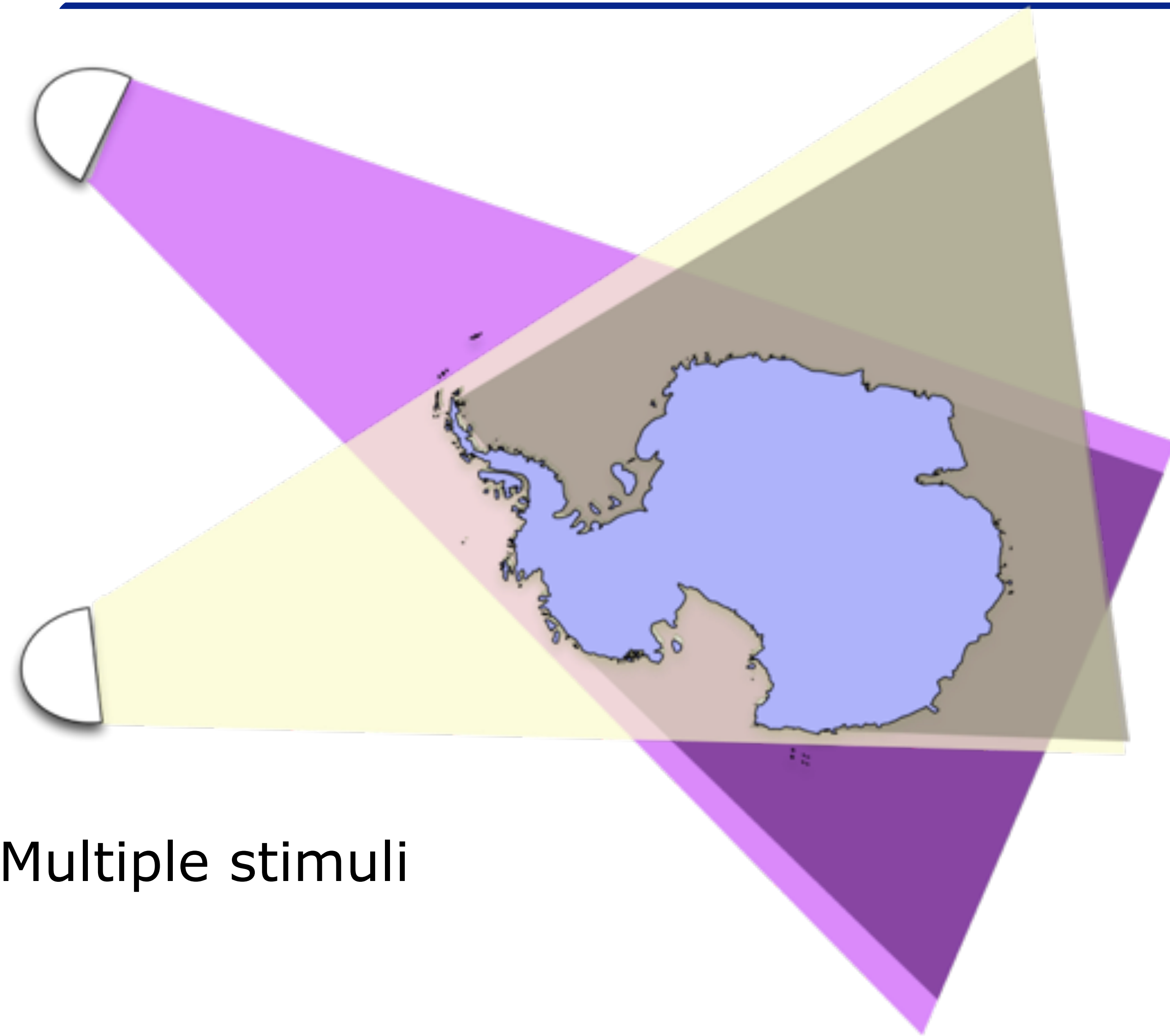


Choosing cues to maximize information



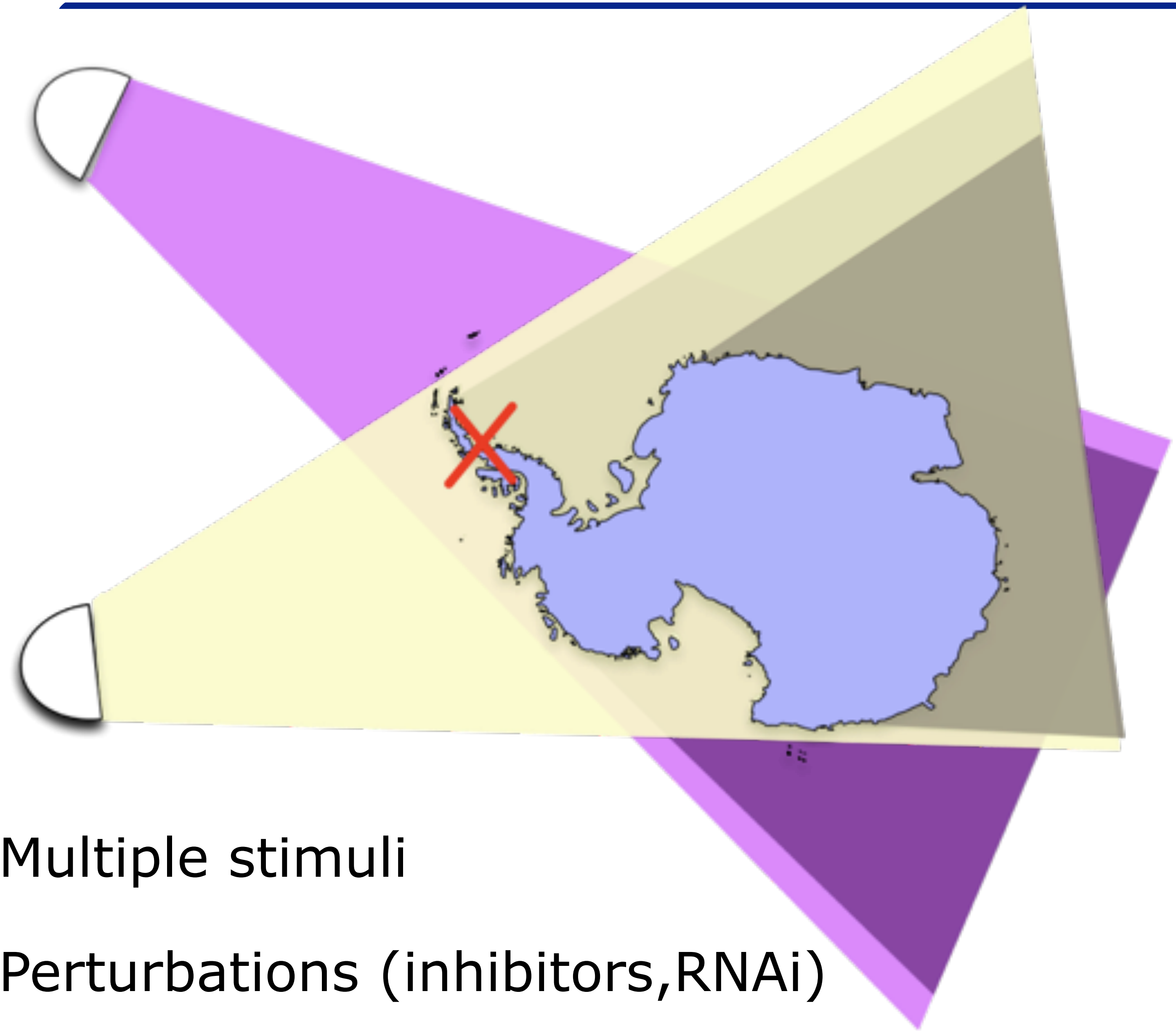
- Multiple stimuli

Choosing cues to maximize information



- Multiple stimuli

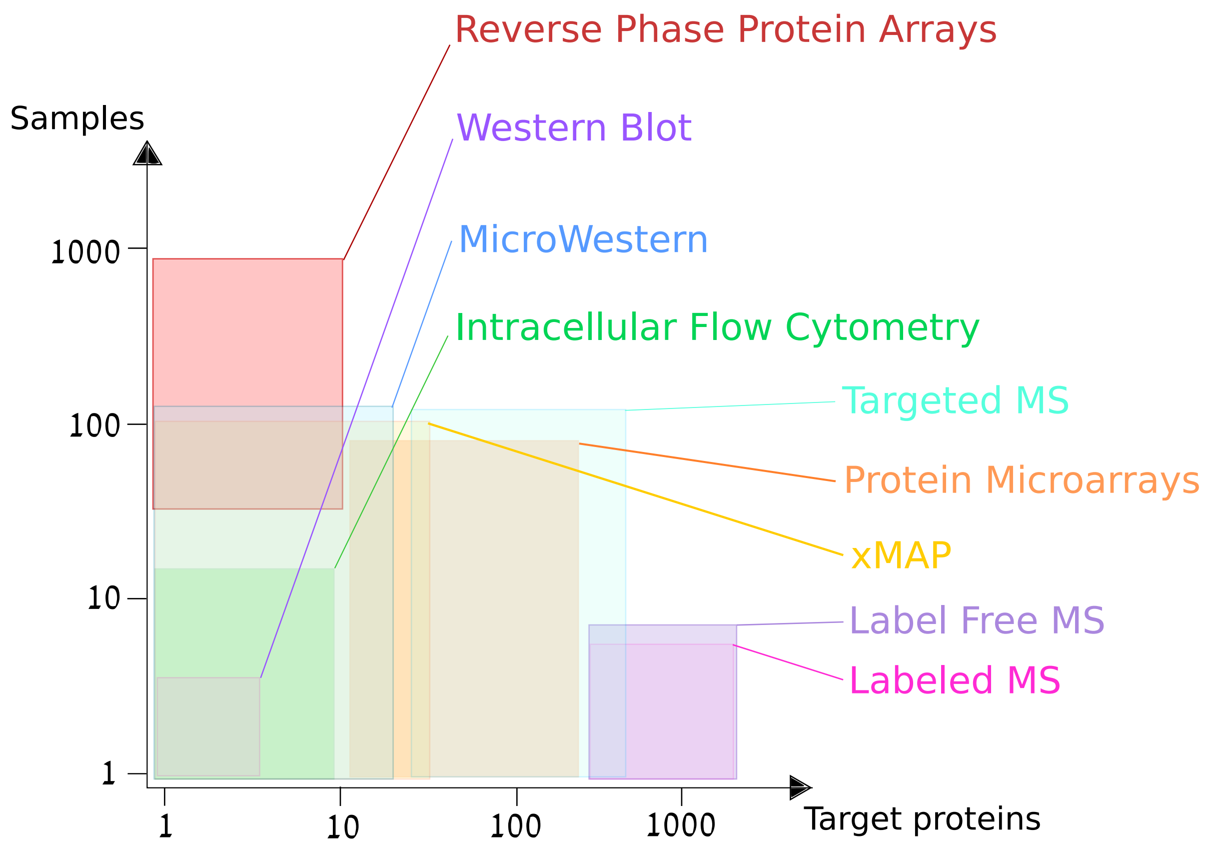
Choosing cues to maximize information



- Multiple stimuli
- Perturbations (inhibitors,RNAi)

To look at signal transduction, proteomics is the natural platform

Different platforms with pros and cons



Terfve C, Saez-Rodriguez J, *Adv. Syst. Biol.*, 2012
Saez-Rodriguez J, MacNamara A, Cook, S, *Annual Rev Biomed Eng*, 2015

Steps in building (and using) a model

- Set up experiments to extract most information
- Process data efficiently
- Choose type of mathematical model
- Train models to experimental data
- Use models to gain insight

Steps in building (and using) a model

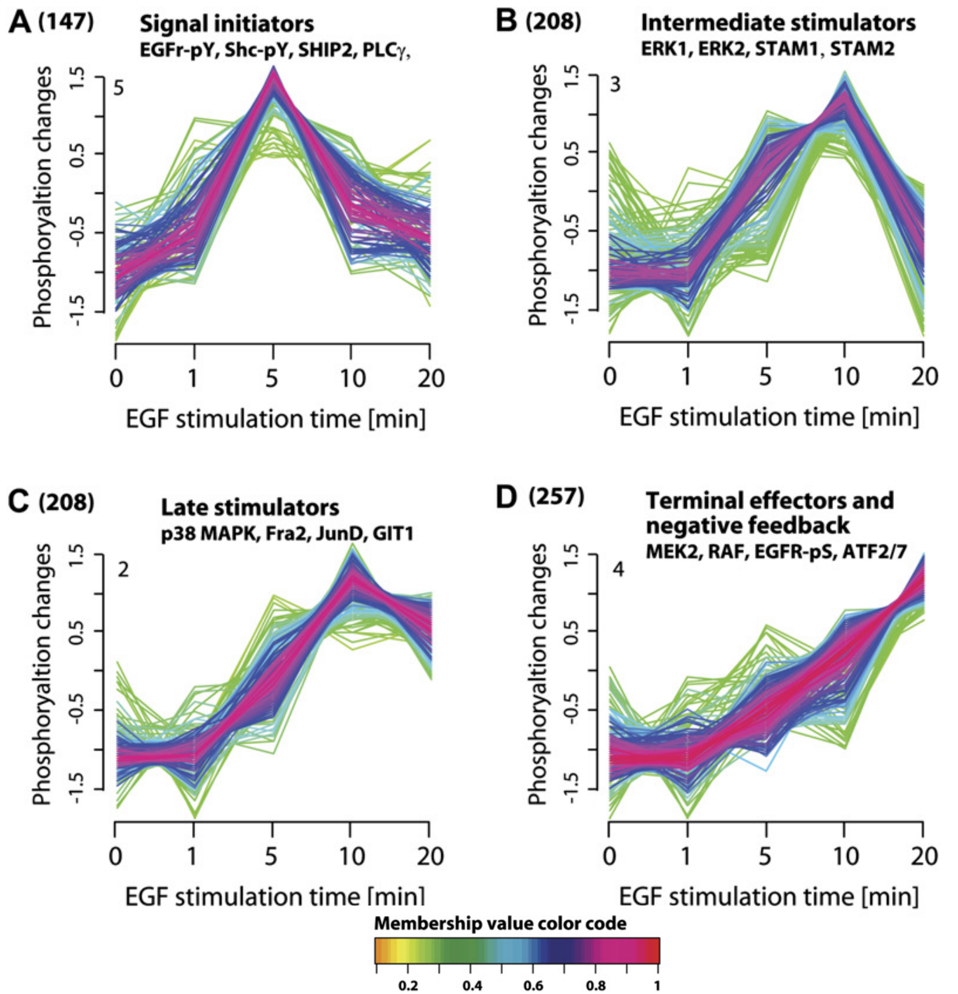
- Set up experiments to extract most information
- Process data efficiently
- Choose type of mathematical model
- Train models to experimental data
- Use models to gain insight

Data mining only requires data and very useful to classify data, generate hypothesis, ...

Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

Data mining only requires data and very useful to classify data, generate hypothesis, ...

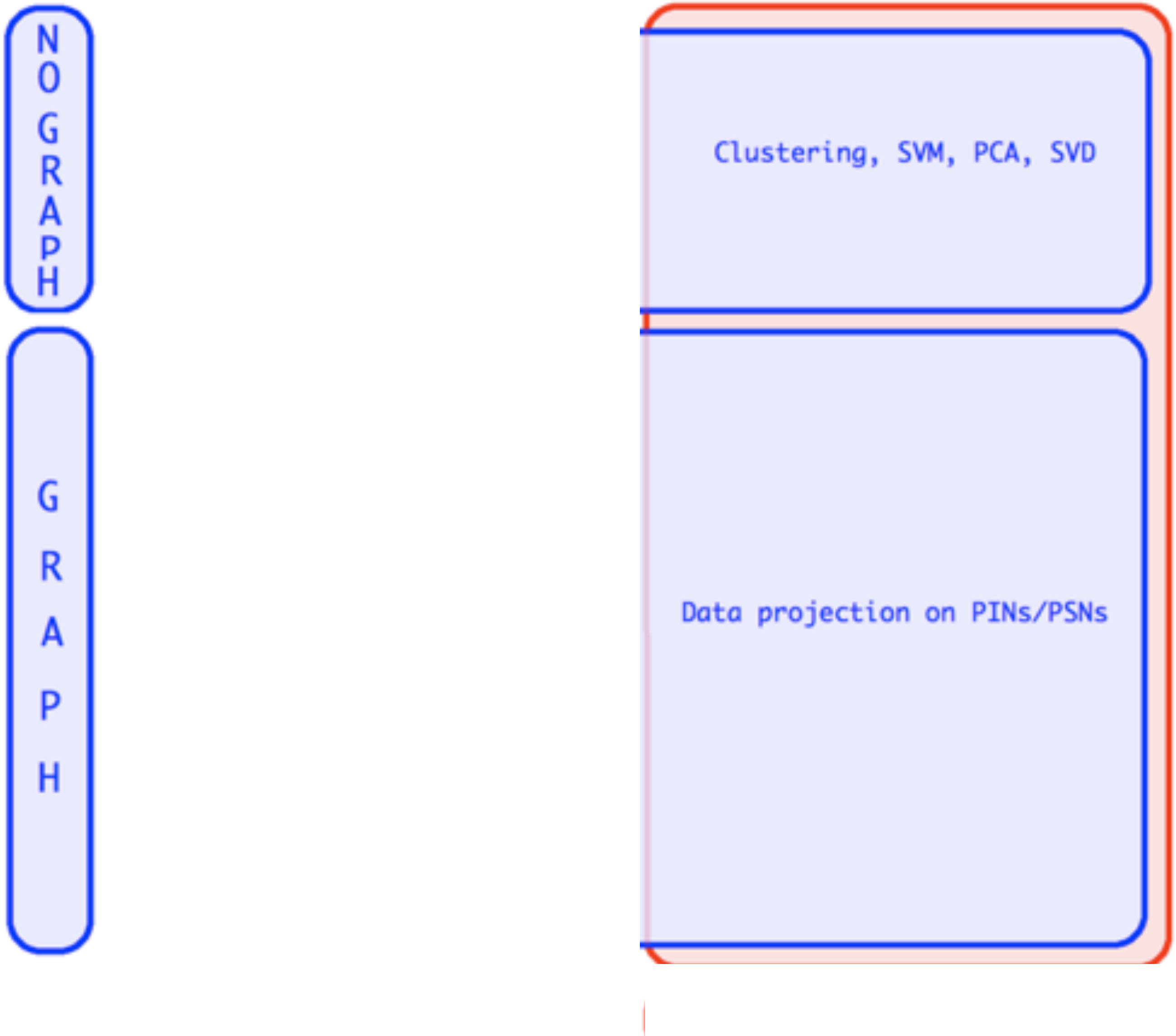
E.g. cluster time series of phosphorylation of peptides upon EGF stimulation to classify them (early/late responders, etc.)



Olsen, J.V. et al., “Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks”, *Cell* 127:635-648, (2006).

Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

Projection on protein networks helps to put data in molecular context



Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

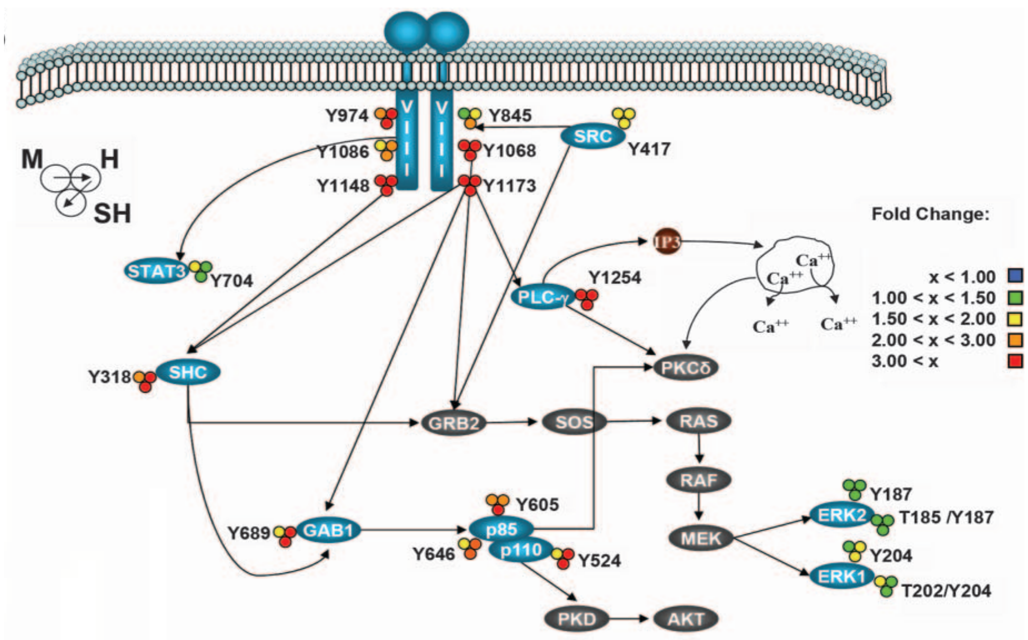
Projection on protein networks helps to put data in molecular context

COMBINE

RWTH AACHEN

N O G R

E.g. visualize fold increase in phosphorylation for different Gliobastoma cell lines (with different EGFR expression)



Huang, P. et al., "Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma", PNAS, (2007).

Clustering, SVM, PCA, SVD

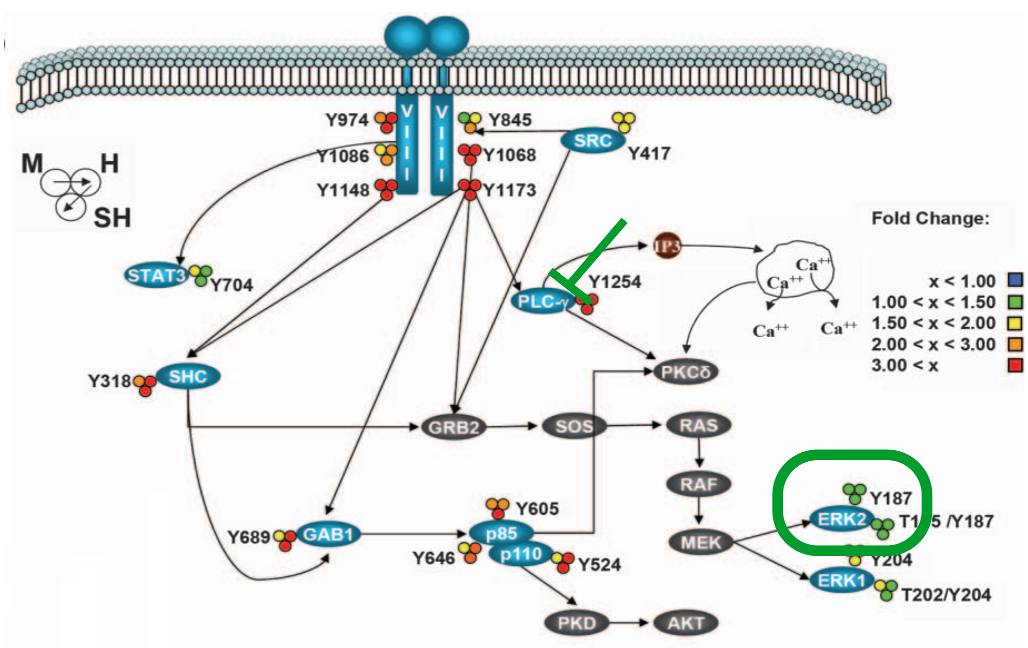
Data projection on PINs/PSNs

Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

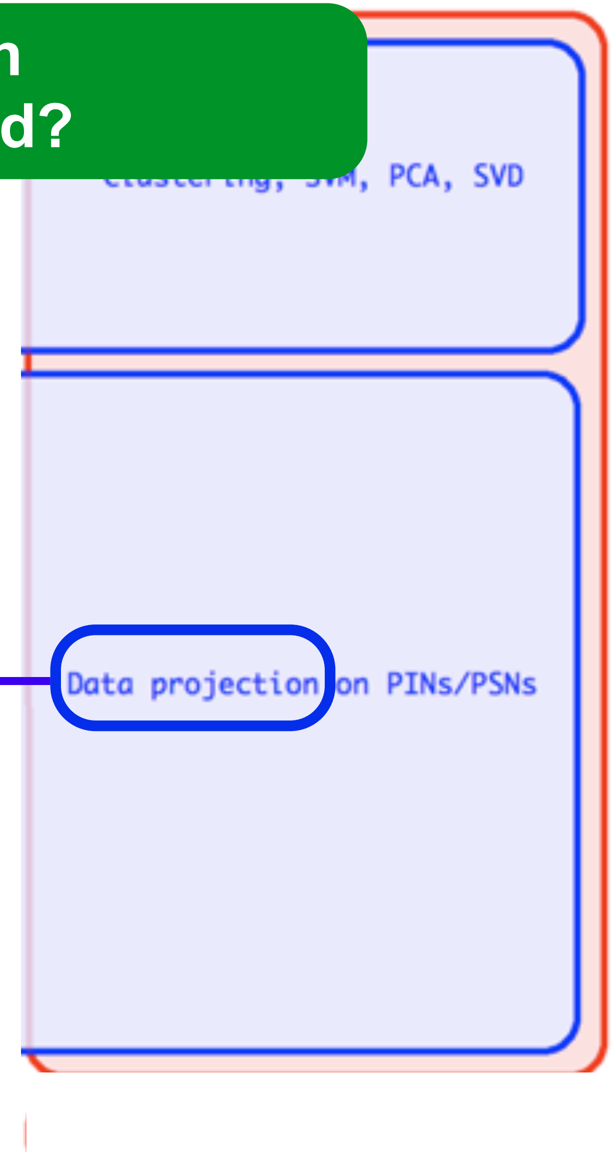
Projection on protein networks helps to put data in molecular context

What would happen on ERK2 if PLCγ is inhibited?

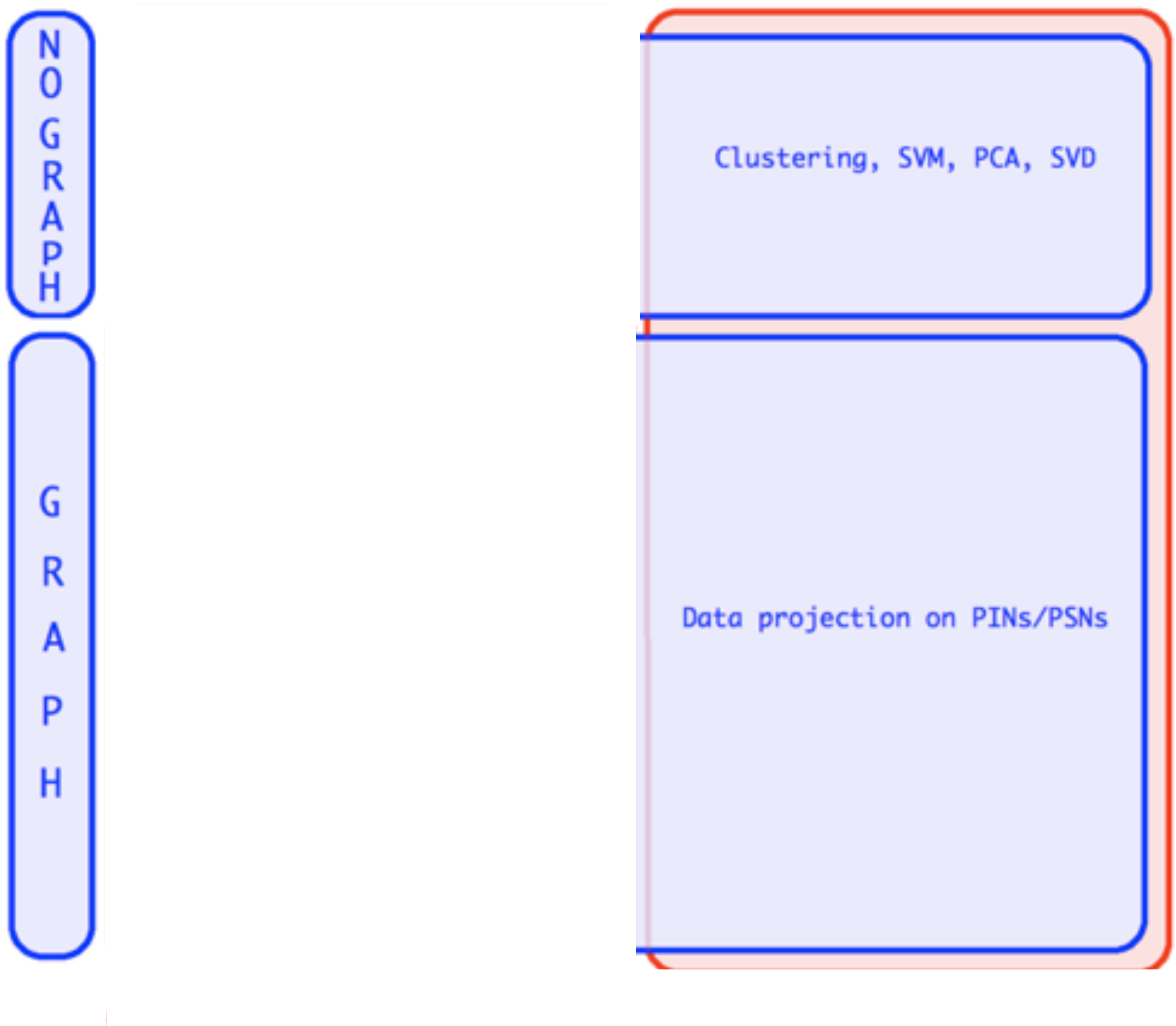
E.g. visualize fold increase in phosphorylation for different Gliobastoma cell lines (with different EGFR expression)



Huang, P. et al., “Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma”, PNAS, (2007).

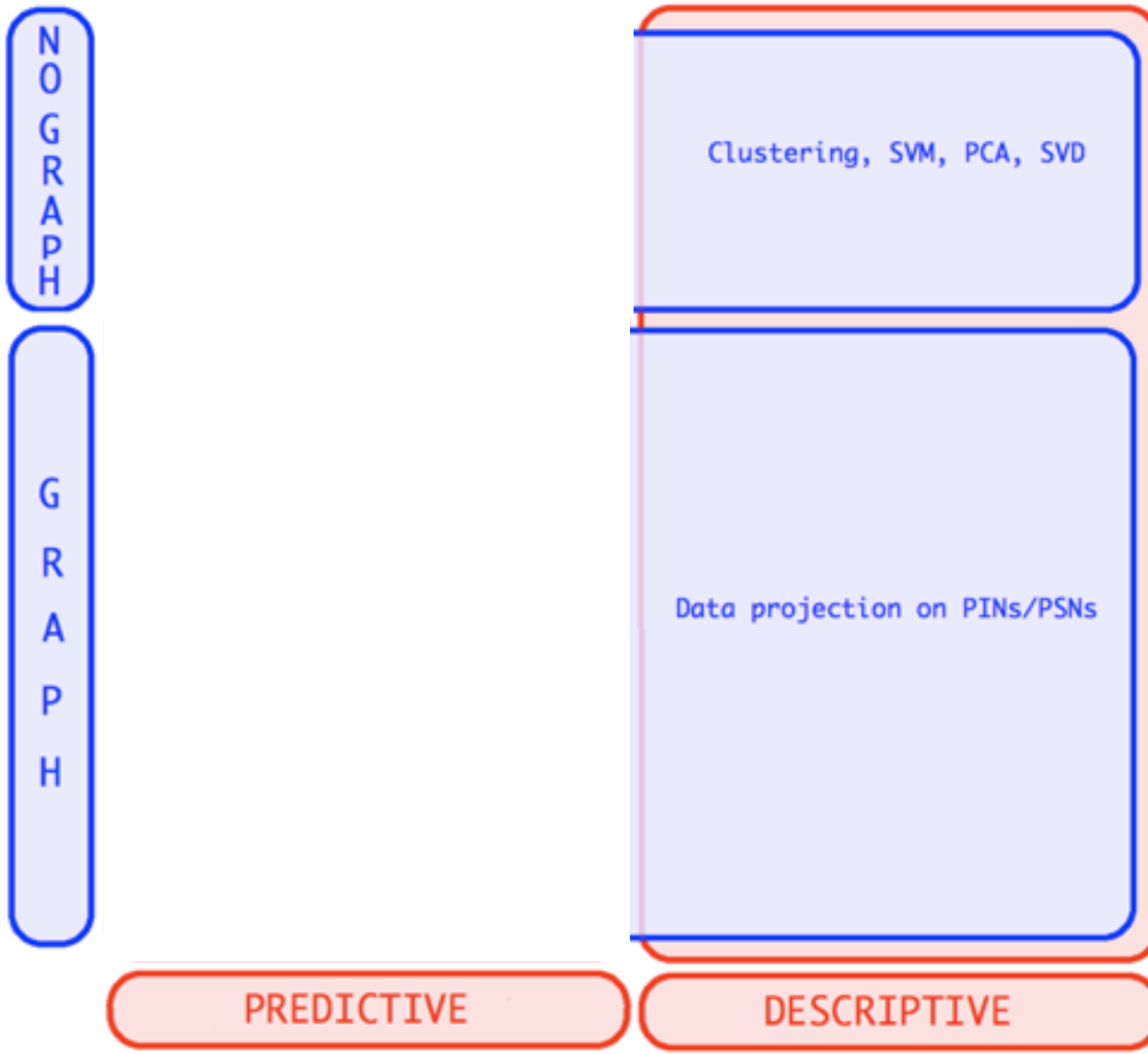


Data-driven predictive models



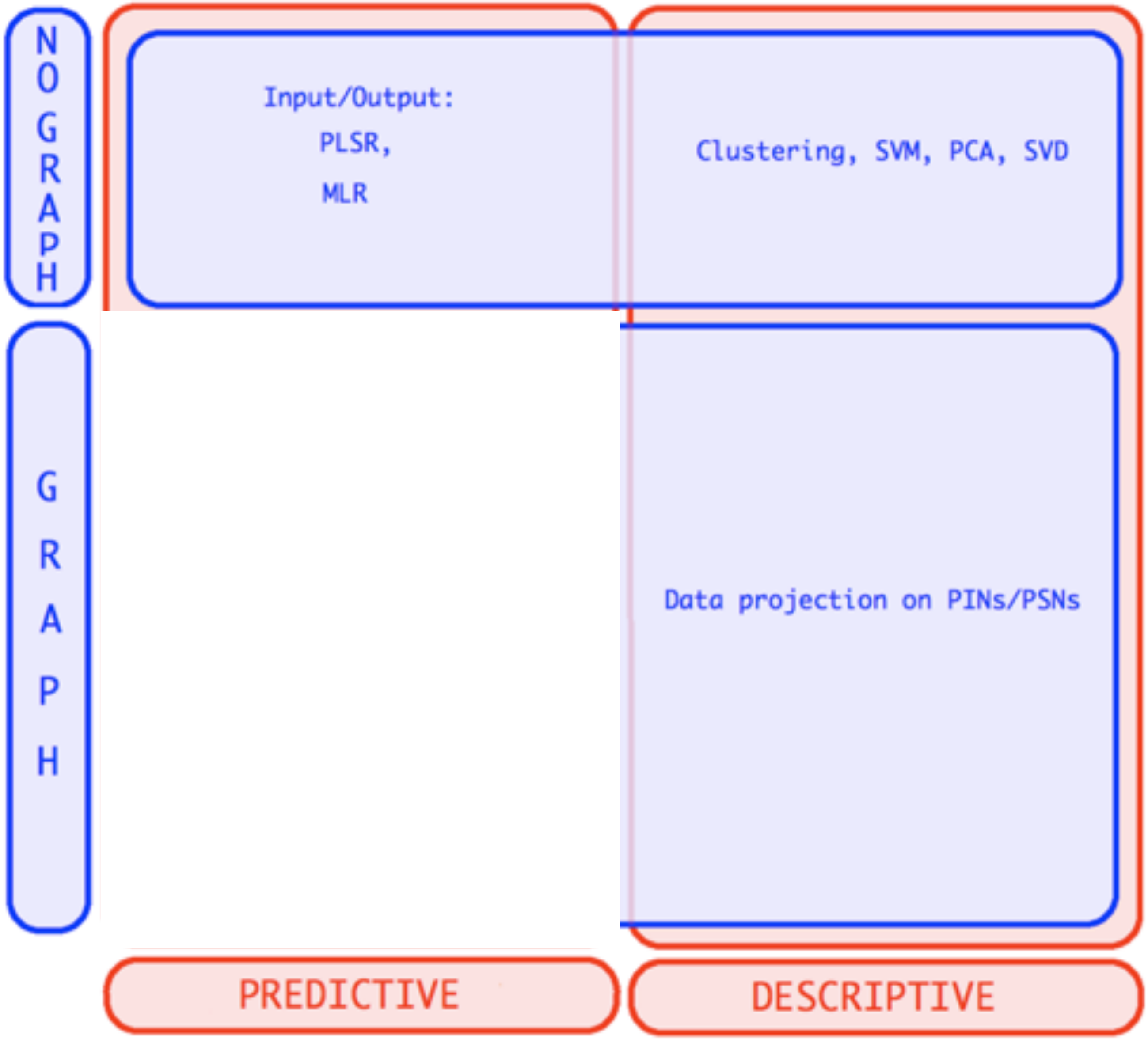
Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

Data-driven predictive models



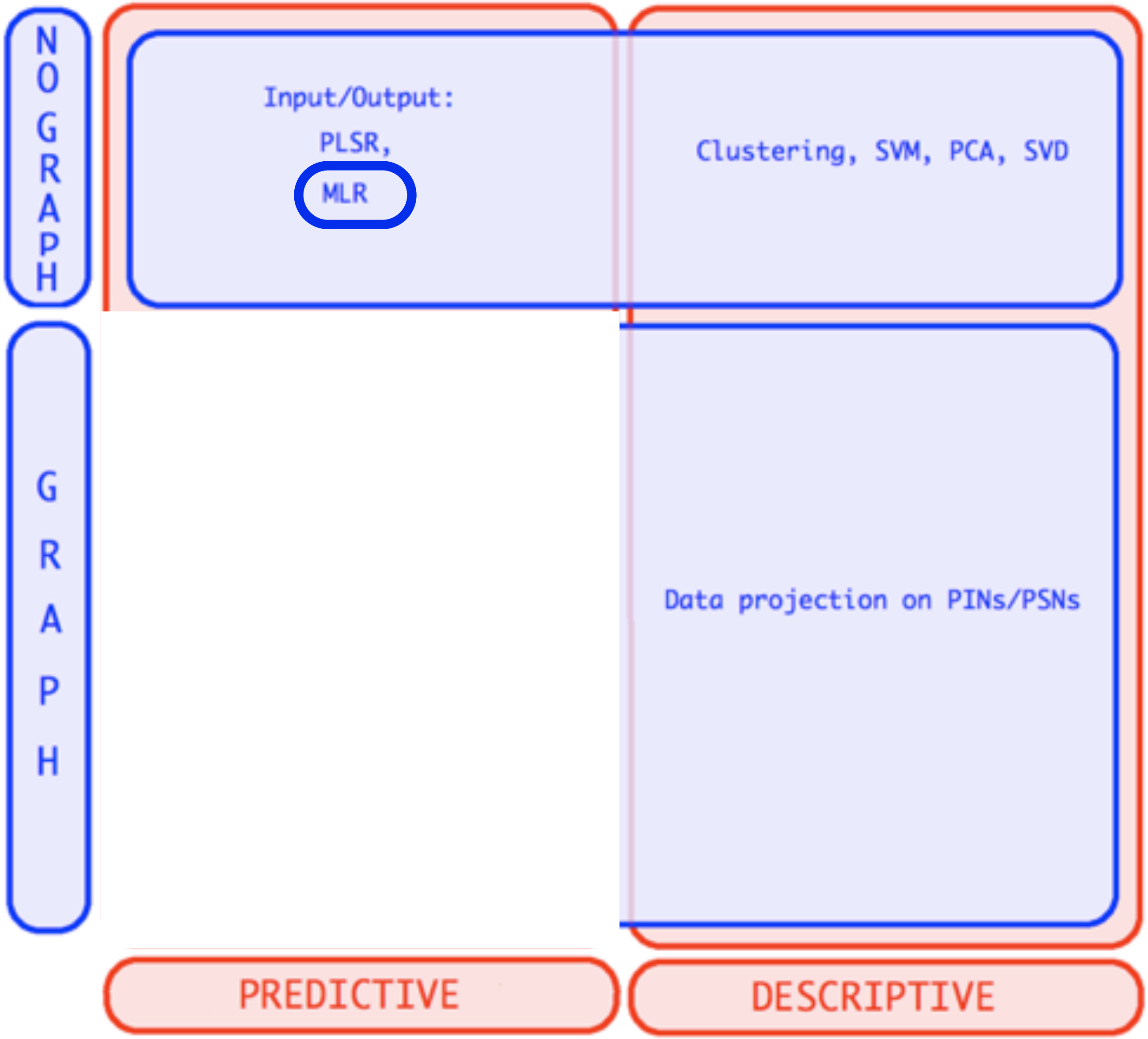
Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

Data-driven predictive models



Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

Data-driven predictive models

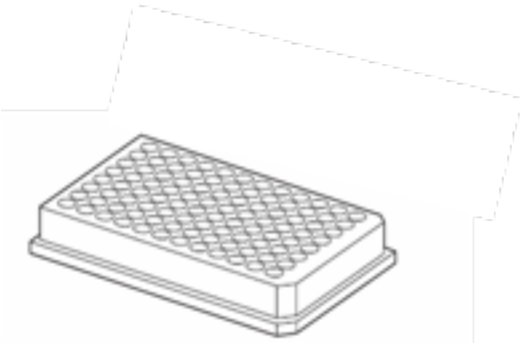


Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

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

Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.,* 9, 1849, 2010

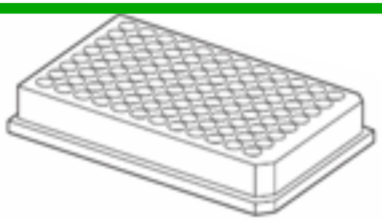
An example of a perturbation-based high-throughput data sets



Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.,* 9, 1849, 2010

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

Cue	
	➜ 7 **extracellular ligands**
	➜ 7 specific **chemical inhibitors** (drugs)



Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

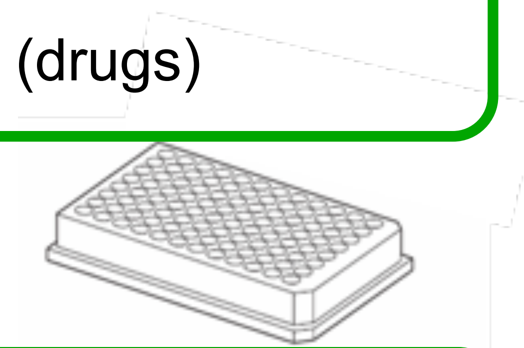
An example of a perturbation-based high-throughput data sets

Cue

➜ 7 **extracellular ligands**

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

at different times after stimulation



Signal

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

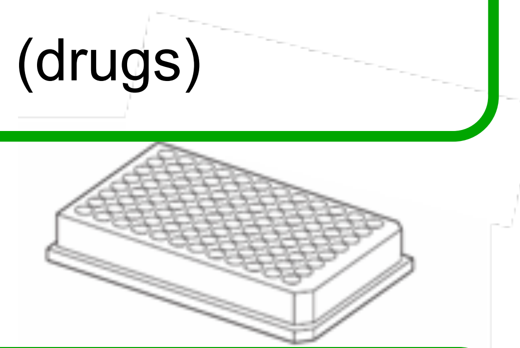
Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

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



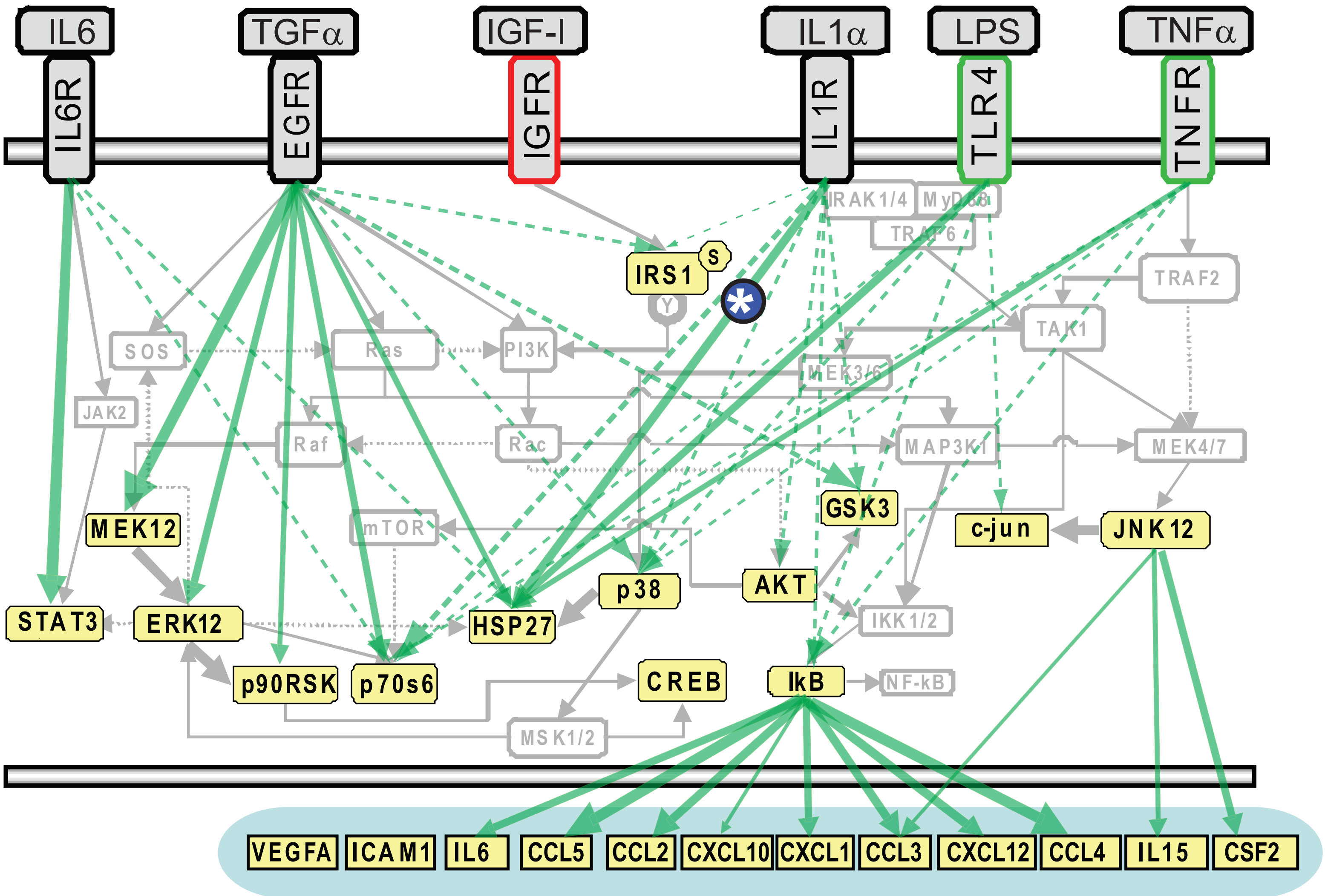
Signal
- ➜ **Phosphorylation** of 17 key proteins (30 min, 3h)

using Luminex/xMAP (bead-based ELISA)

Response
- ➜ **Release** of 20 cytokines (3h, 24h)

Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

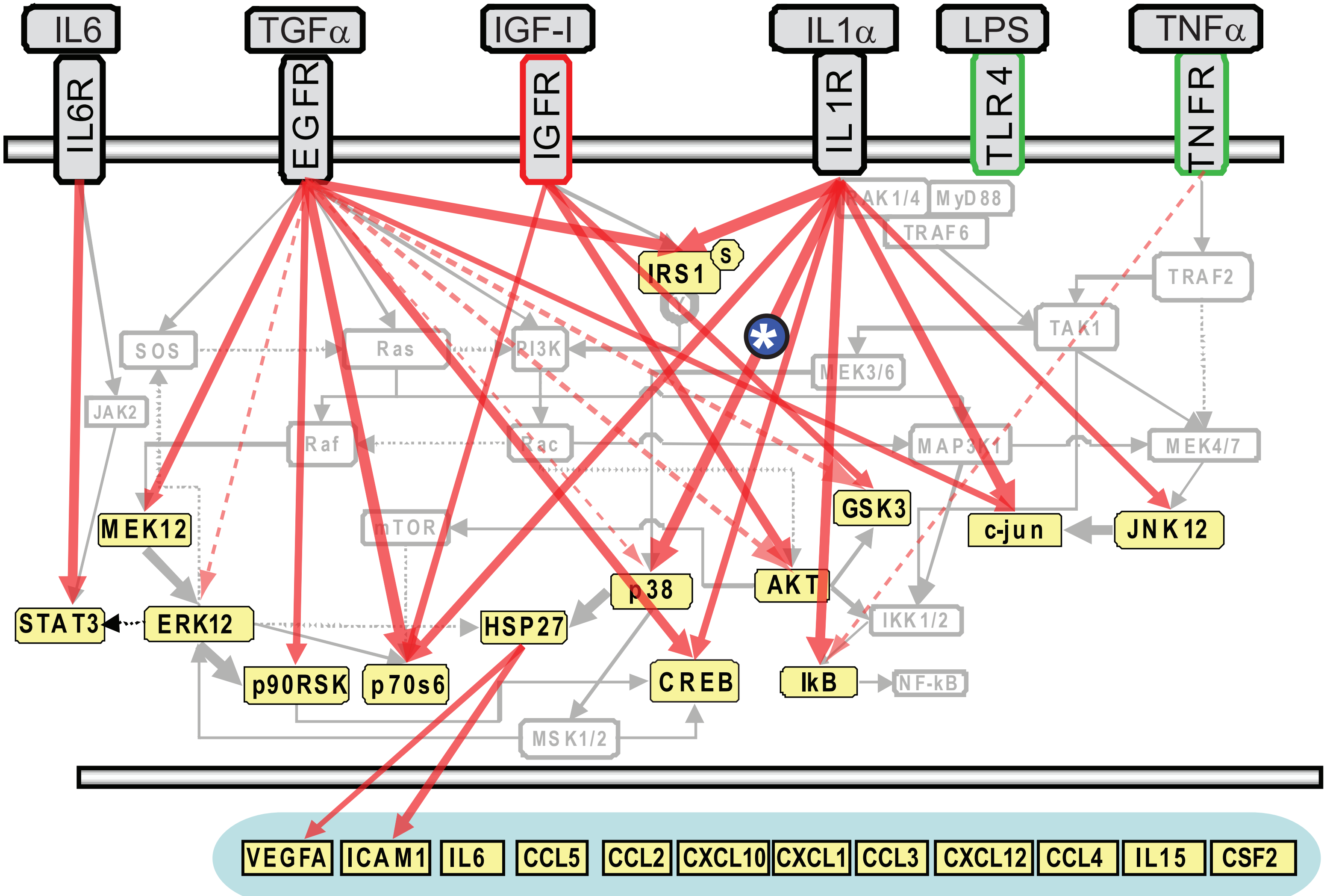
Linear regression uncovers signaling differences between primary & transformed hepatocytes



Draw 30 % strongest weights (correlations) obtained by Multiple Linear Regression

Alexopoulos L, Saez-Rodriguez J, Cosgove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

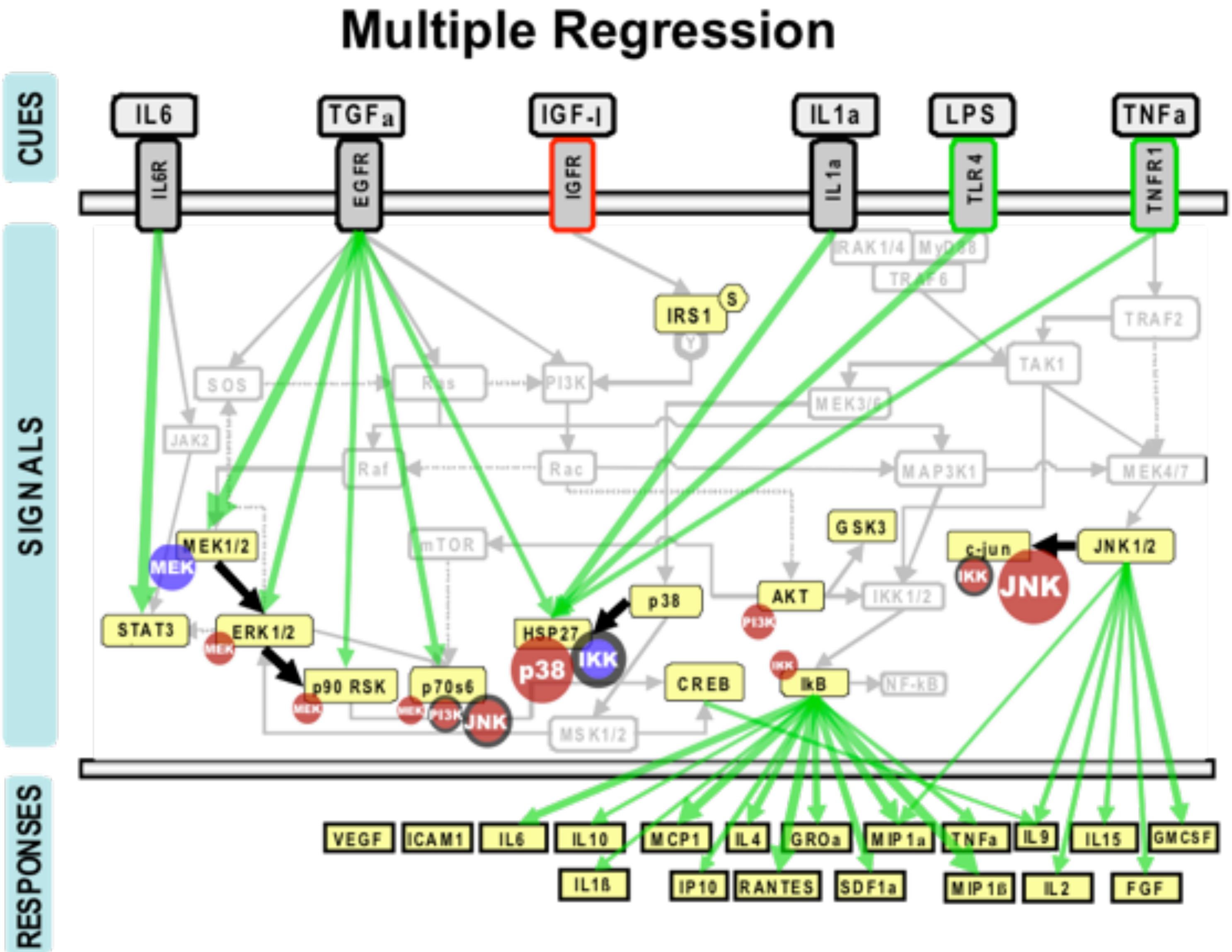
Linear regression uncovers signaling differences between primary & transformed hepatocytes



Draw 30 % strongest weights (correlations) obtained by Multiple Linear Regression

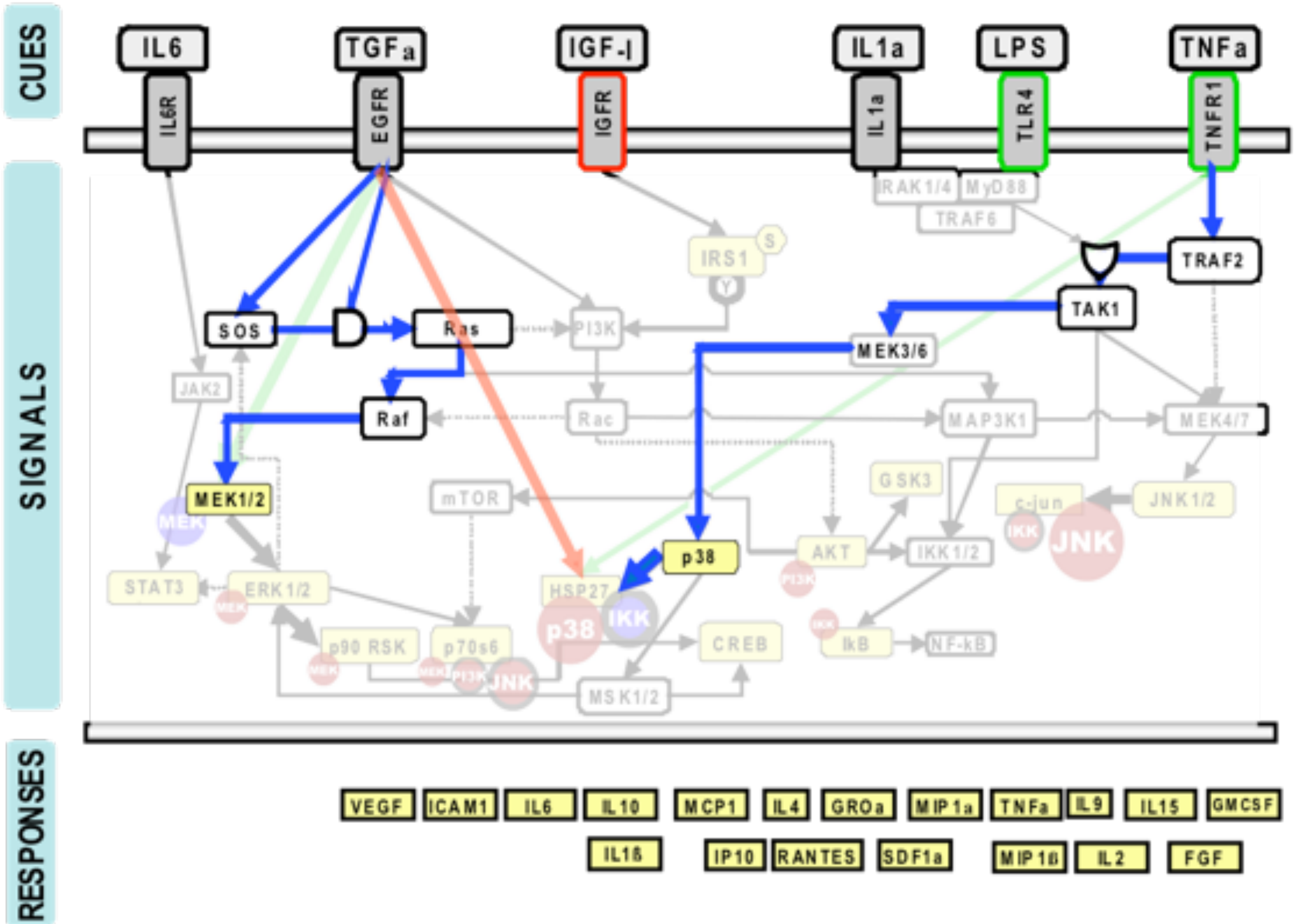
Alexopoulos L, Saez-Rodriguez J, Cosgove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

Data-driven approaches useful but provide limited mechanistic insight



Alexopoulos L, Saez-Rodriguez J, Cosgove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

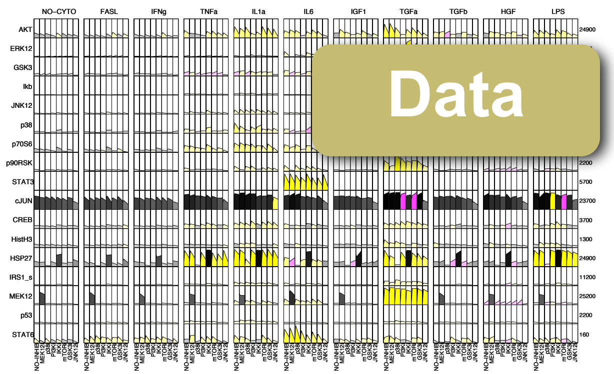
Data-driven approaches useful but provide limited mechanistic insight

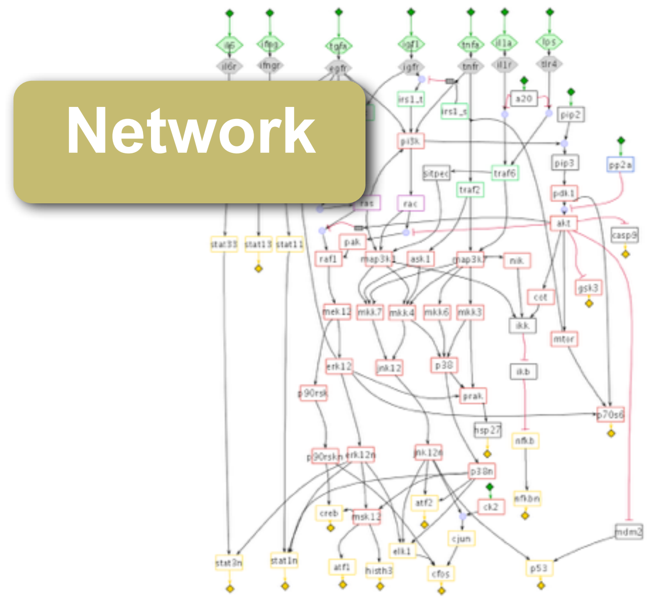


Alexopoulos L, Saez-Rodriguez J, Cosgove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

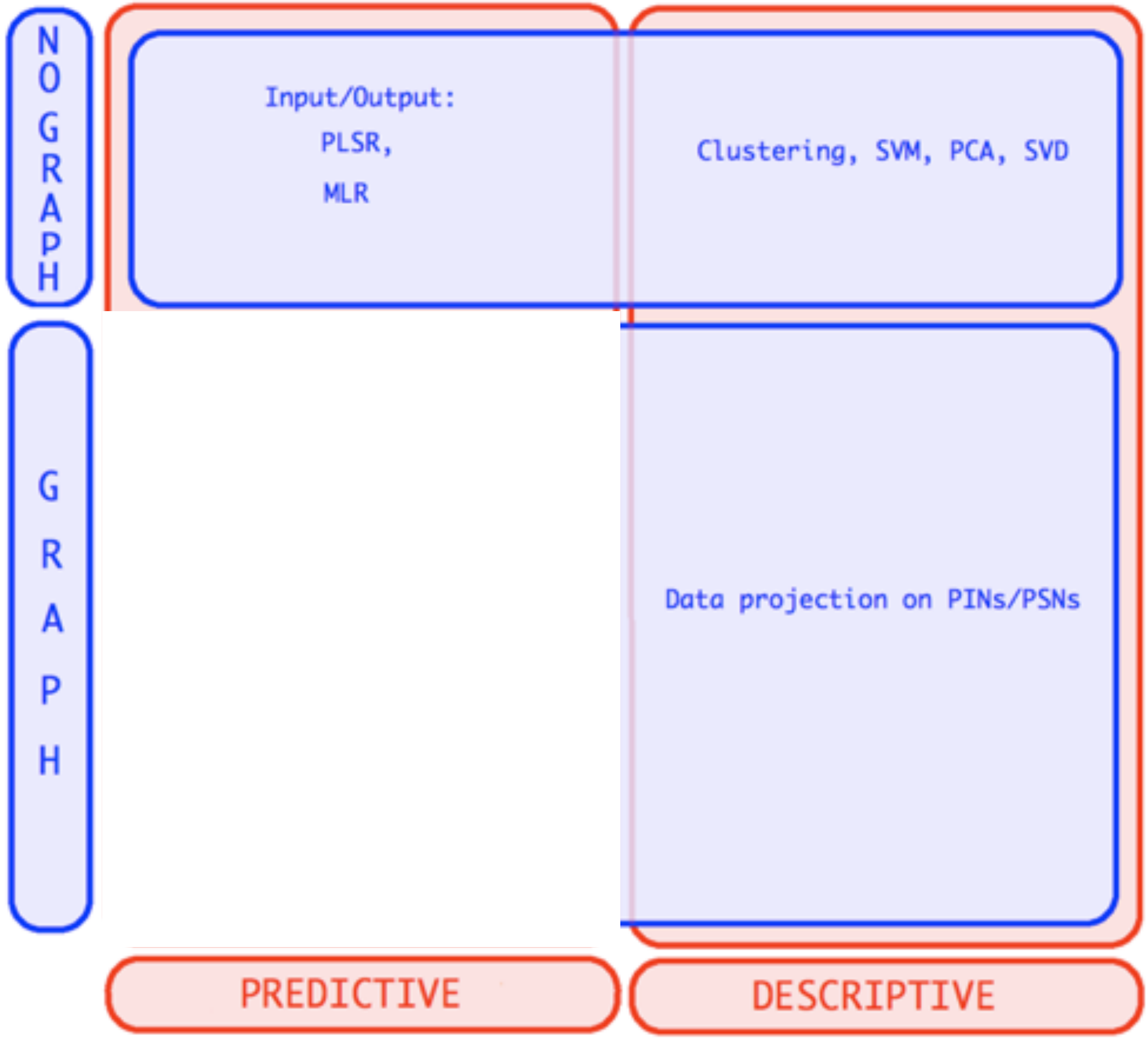
Mechanistic modelling: Integrating network knowledge with functional data

Mechanistic modelling: Integrating network knowledge with functional data



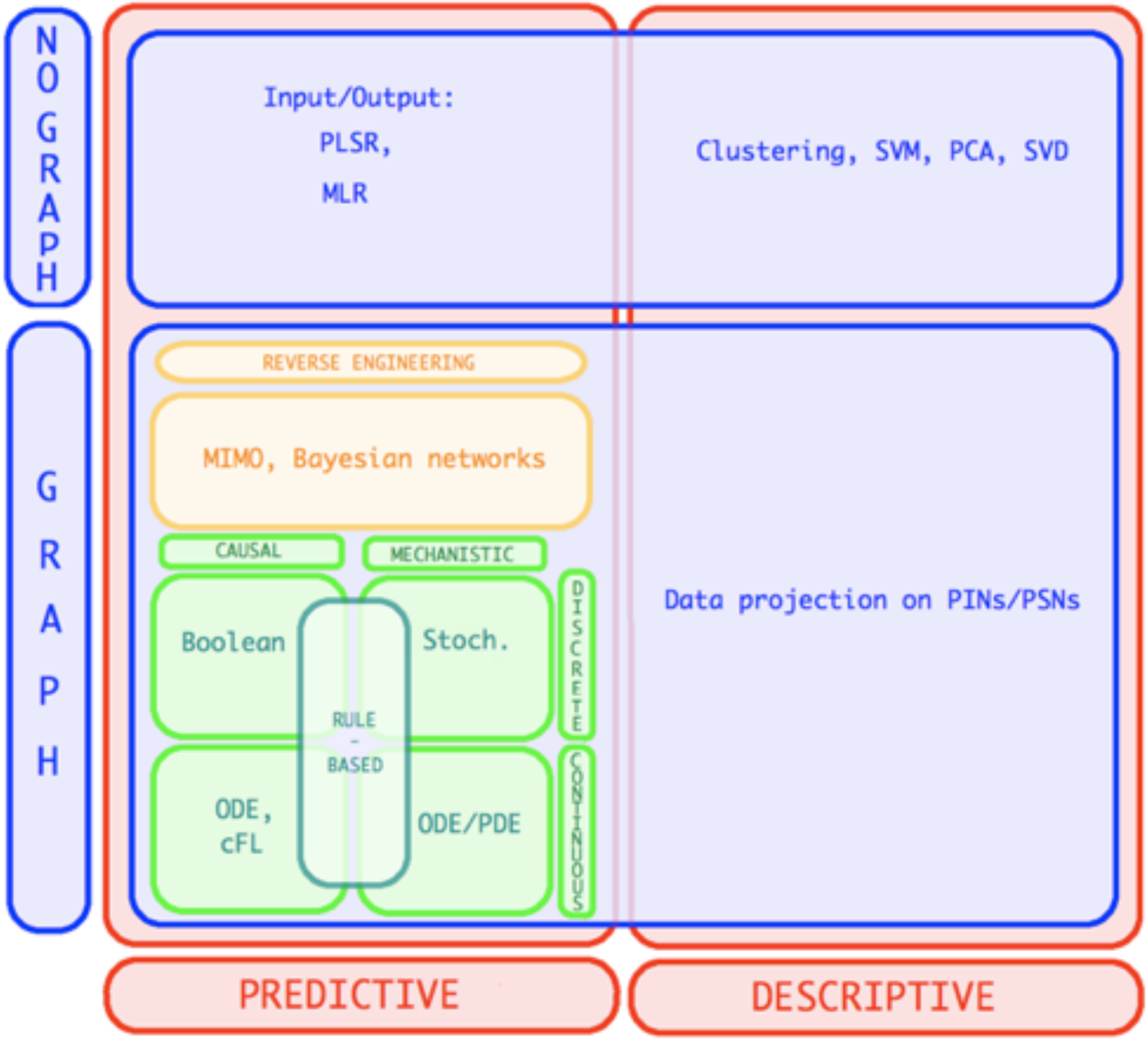


Multiple mechanistic approaches



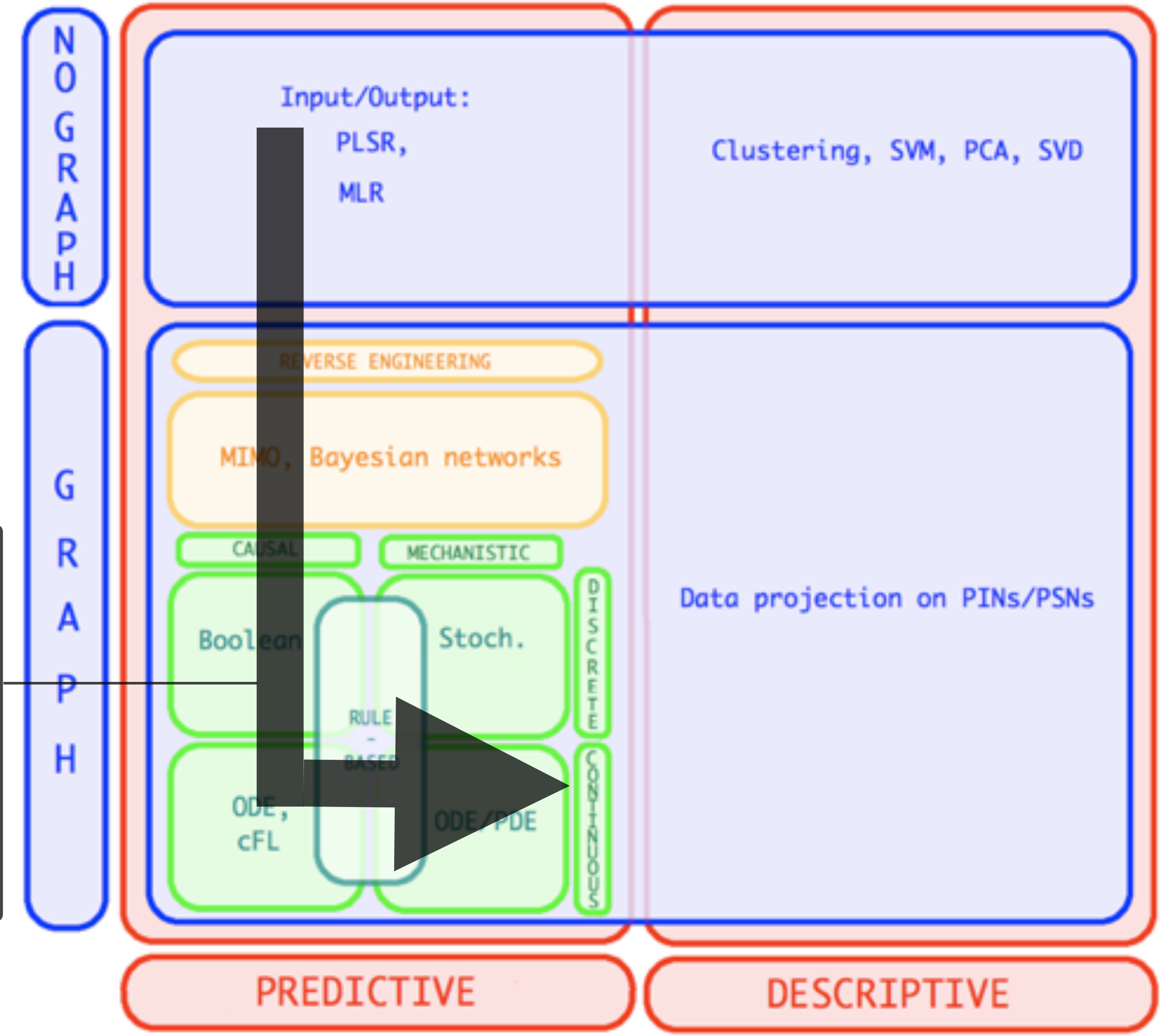
Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57*, 2012

Multiple mechanistic approaches



Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

Multiple mechanistic approaches



+ Mechanistic insight

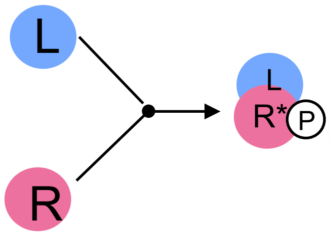
+ Prior knowledge & data required

- Model size

+ Harder calibration

Terfve C, Saez-Rodriguez J, *Advances in Systems Biology, Springer, 736:19-57, 2012*

Causal vs. biochemical representation of signal transduction

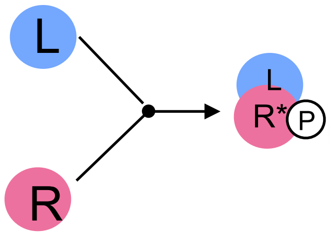


$$R + L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

Causal vs. biochemical representation of signal transduction

- Natural modeling approach is based on **differential equations** describing underlying biochemistry

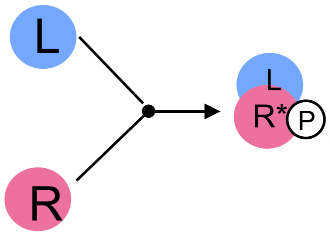


$$R + L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

Causal vs. biochemical representation of signal transduction

- Natural modeling approach is based on **differential equations** describing underlying biochemistry



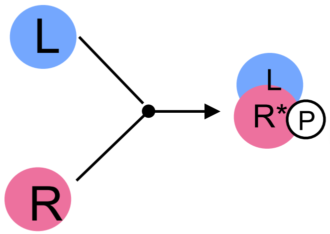
$$R + L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,"ON/OFF") models: **maps + logic gates (AND/OR)**

Causal vs. biochemical representation of signal transduction

- Natural modeling approach is based on **differential equations** describing underlying biochemistry



$$R+L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

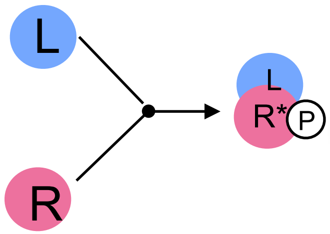
- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,”ON/OFF”) models: **maps + logic gates (AND/OR)**

Map

Boolean
Model

Causal vs. biochemical representation of signal transduction

- Natural modeling approach is based on **differential equations** describing underlying biochemistry

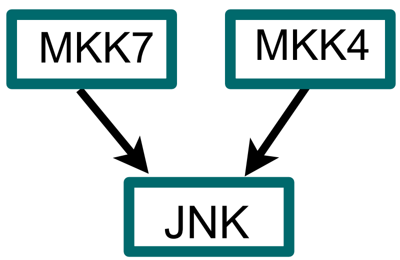


$$R + L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,"ON/OFF") models: **maps + logic gates (AND/OR)**

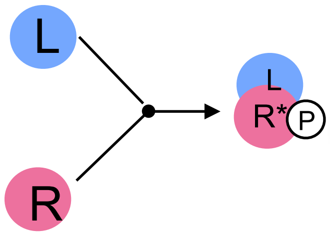
Map



Boolean Model

Causal vs. biochemical representation of signal transduction

- Natural modeling approach is based on **differential equations** describing underlying biochemistry

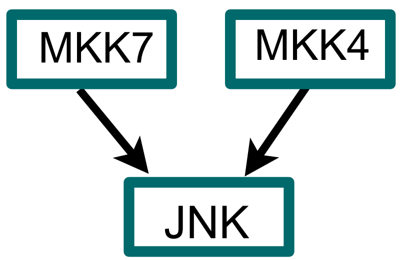


$$R + L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

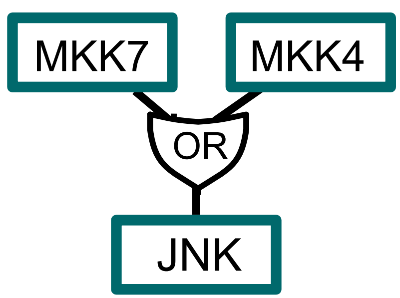
$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,"ON/OFF") models: **maps + logic gates (AND/OR)**

Map



Boolean Model

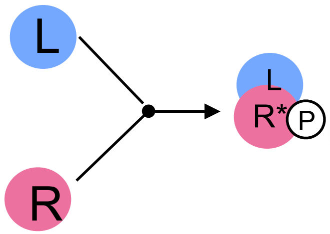


COMBINE

Causal vs. biochemical representation of signal transduction

RWTH AACHEN

- Natural modeling approach is based on **differential equations** describing underlying biochemistry

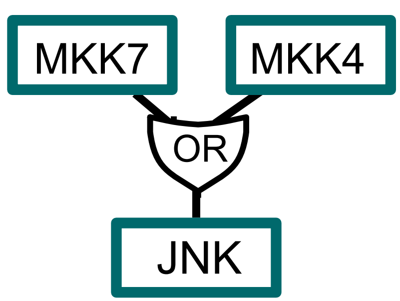


$$R+L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^* \quad \Rightarrow \quad \frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,”ON/OFF”) models: **maps + logic gates (AND/OR)**

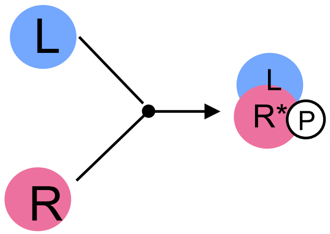
Map

Boolean Model



Causal vs. biochemical representation of signal transduction

- Natural modeling approach is based on **differential equations** describing underlying biochemistry



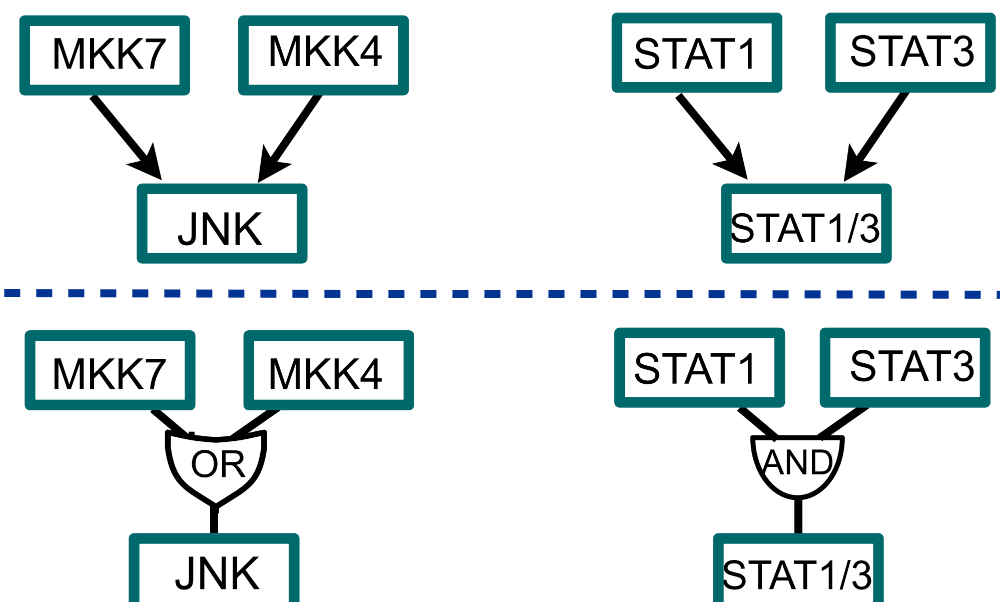
$$R + L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,"ON/OFF") models: **maps + logic gates (AND/OR)**

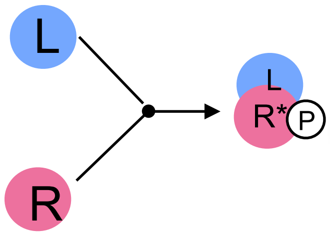
Map

Boolean Model



Causal vs. biochemical representation of signal transduction

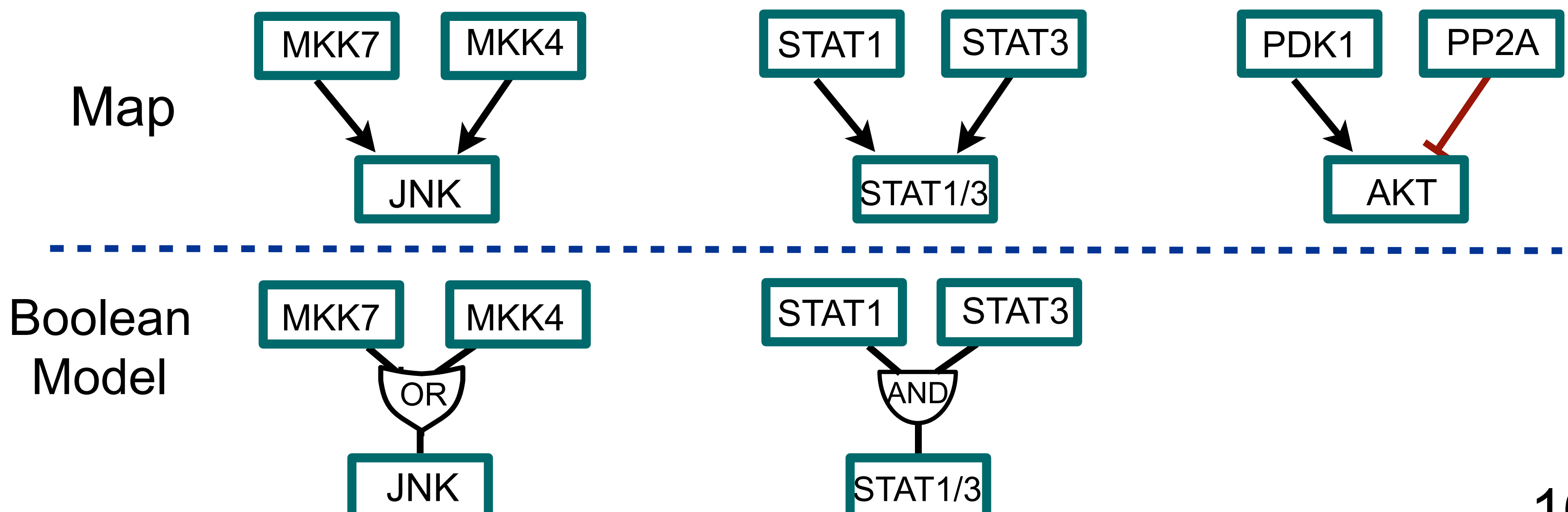
- Natural modeling approach is based on **differential equations** describing underlying biochemistry



$$R+L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

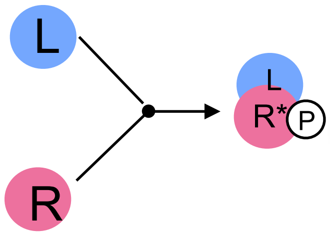
$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,"ON/OFF") models: **maps + logic gates (AND/OR)**



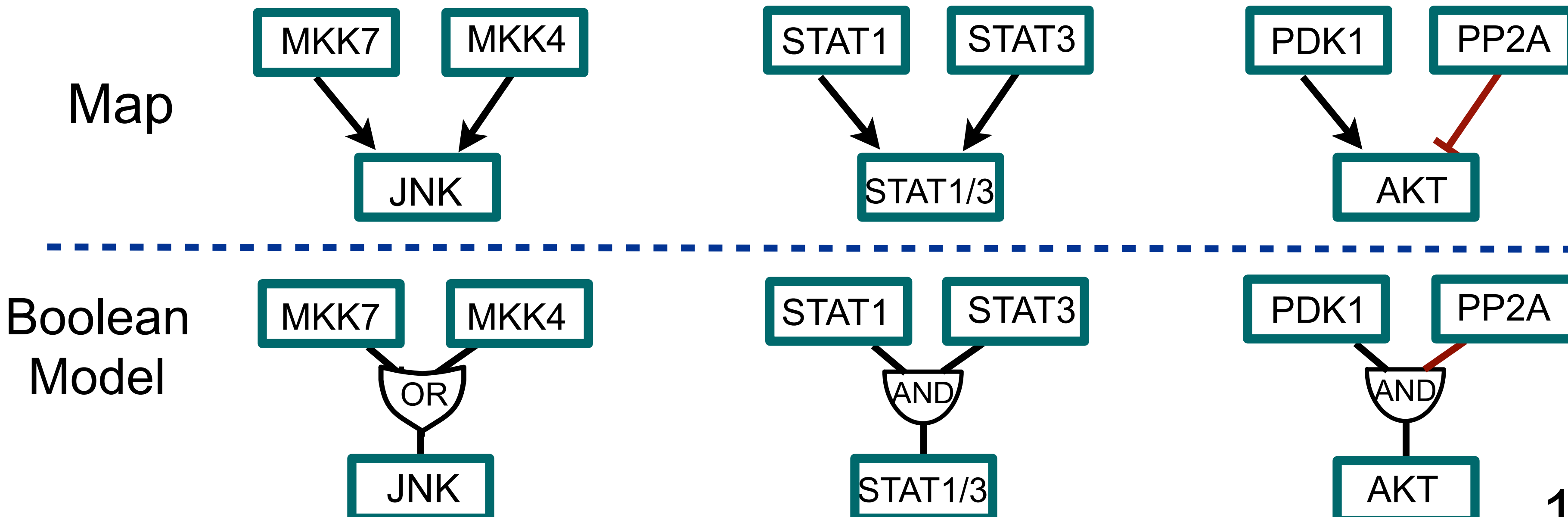
Causal vs. biochemical representation of signal transduction

- Natural modeling approach is based on **differential equations** describing underlying biochemistry

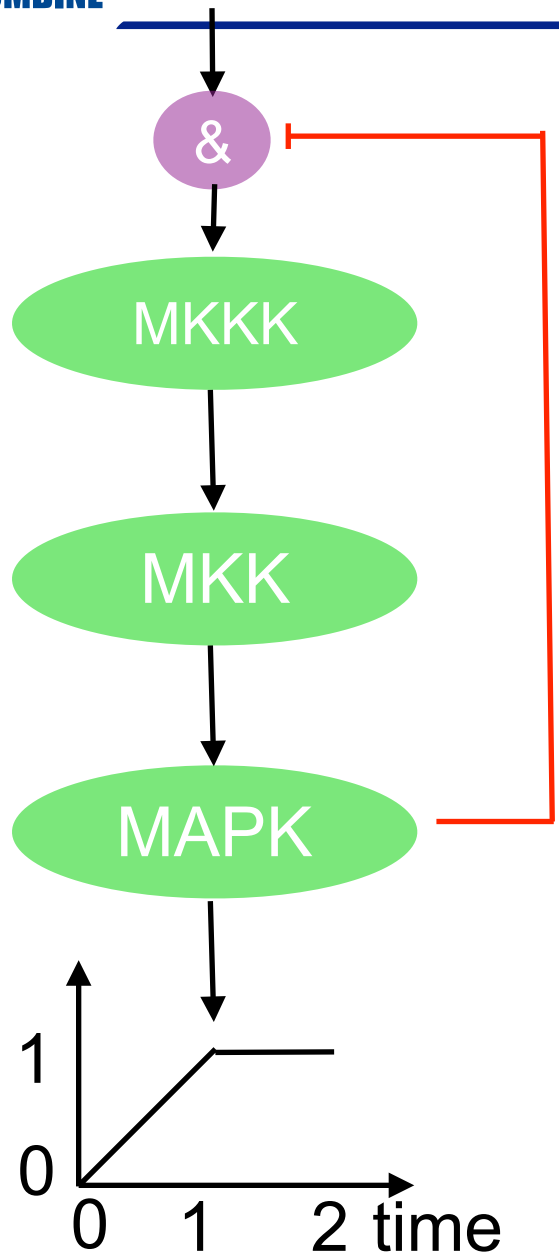


$$R+L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^* \qquad \frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

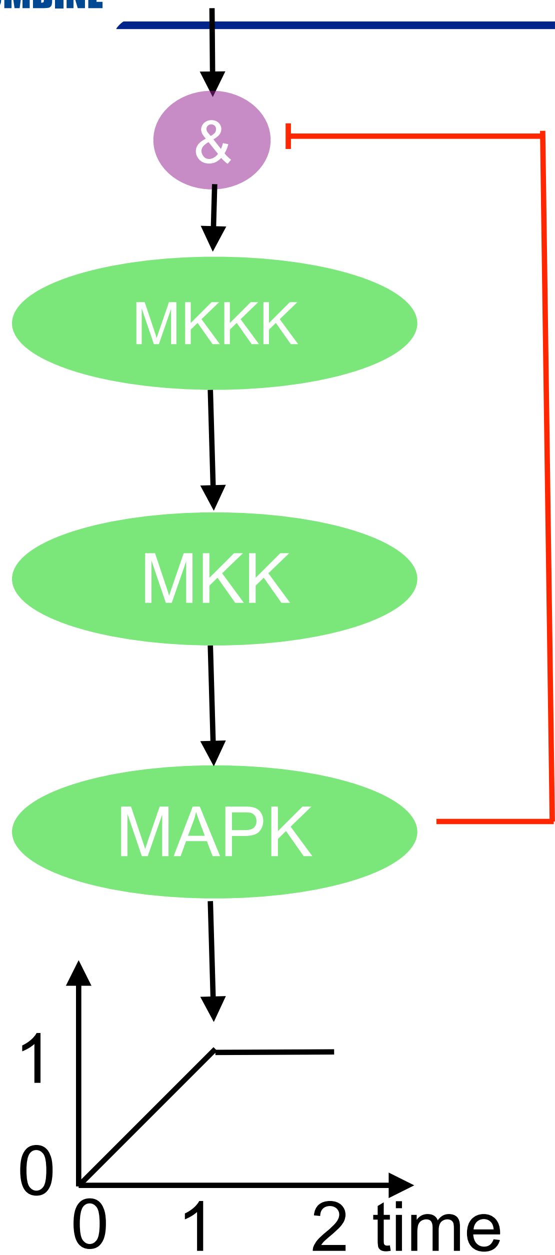
- **If the network is too large** for detailed modeling (parameters to fit, etc.), and not enough molecular information available in pathway maps
  - Boolean (0/1,"ON/OFF") models: **maps + logic gates (AND/OR)**

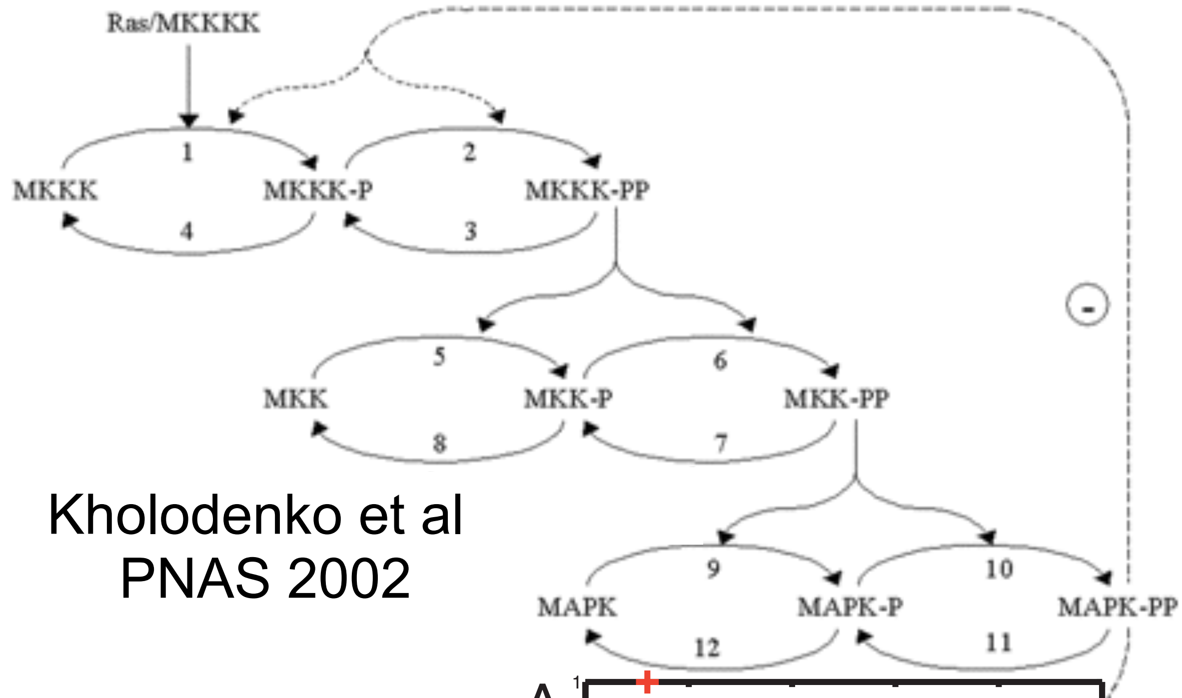


Causal vs. reaction representation

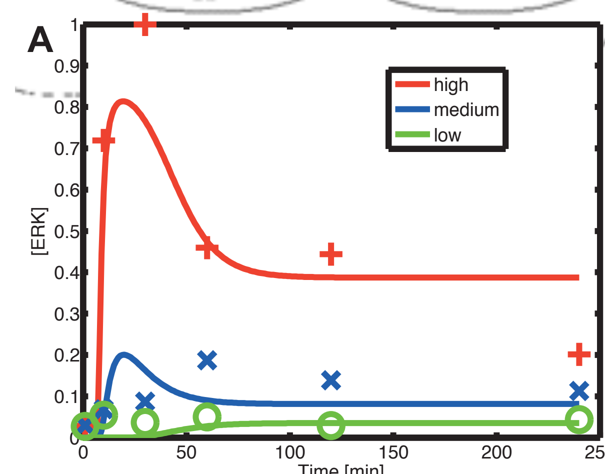


Causal vs. reaction representation

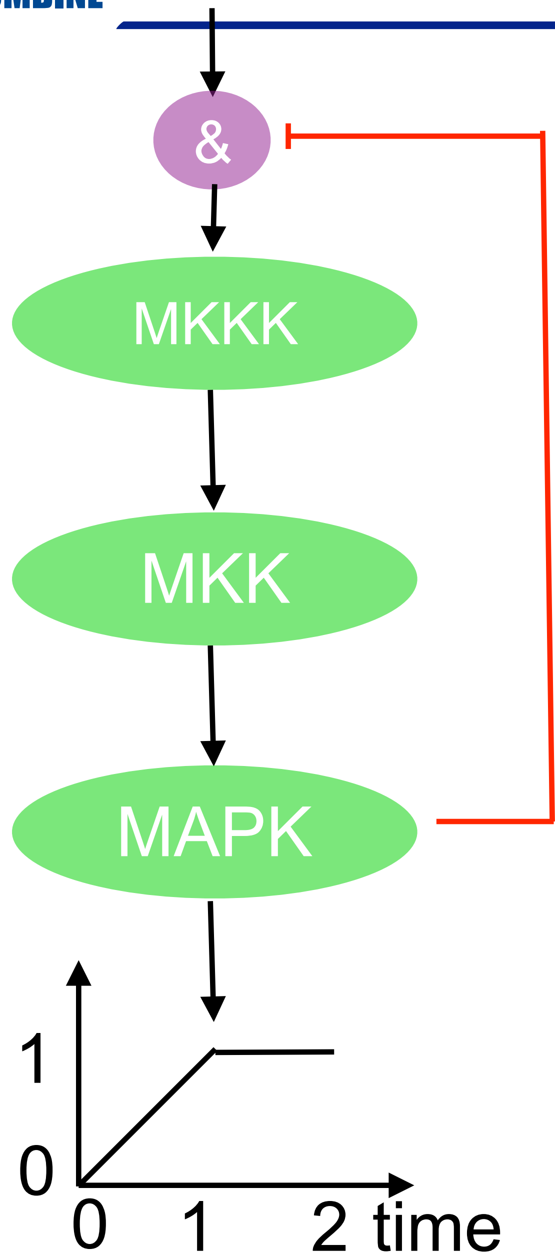




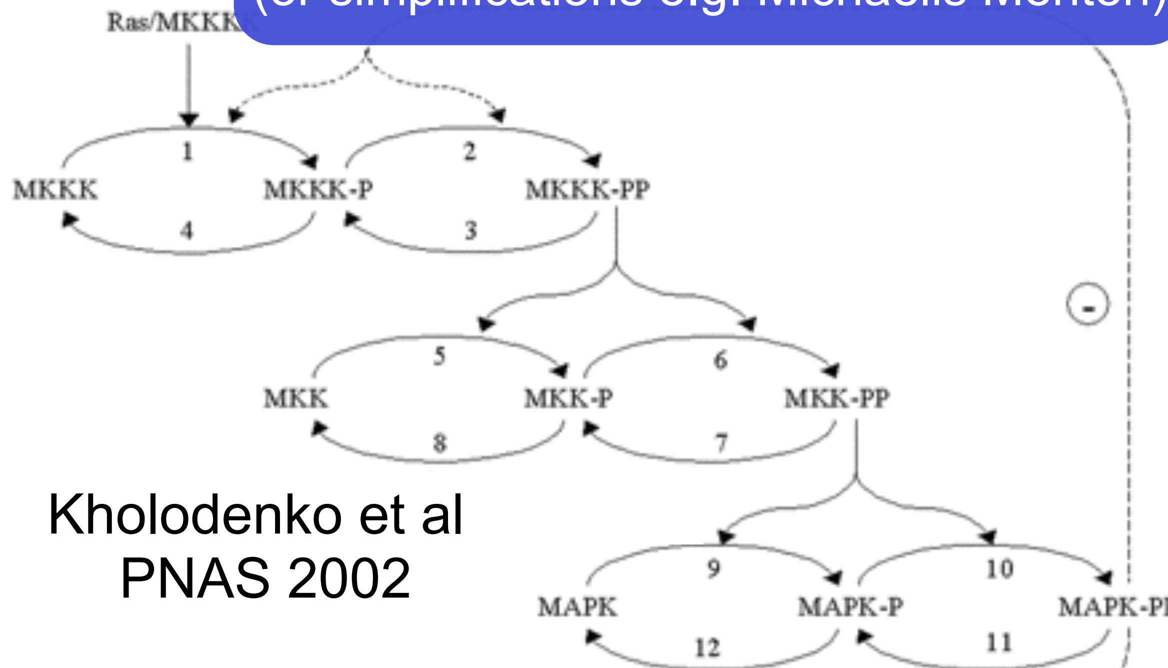
Kholodenko et al
PNAS 2002



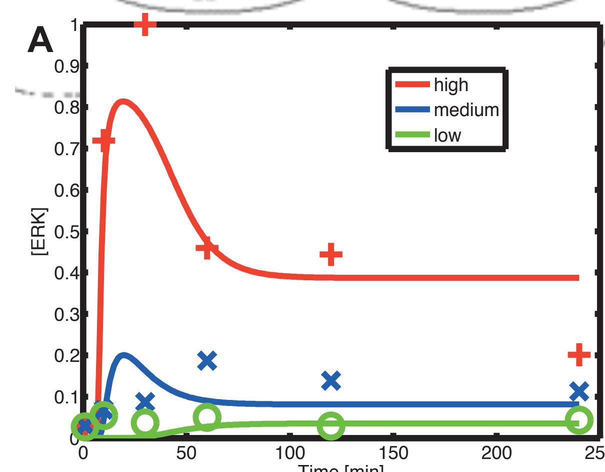
Causal vs. reaction representation



Model reactions with mass action law
(or simplifications e.g. Michaelis Menten)

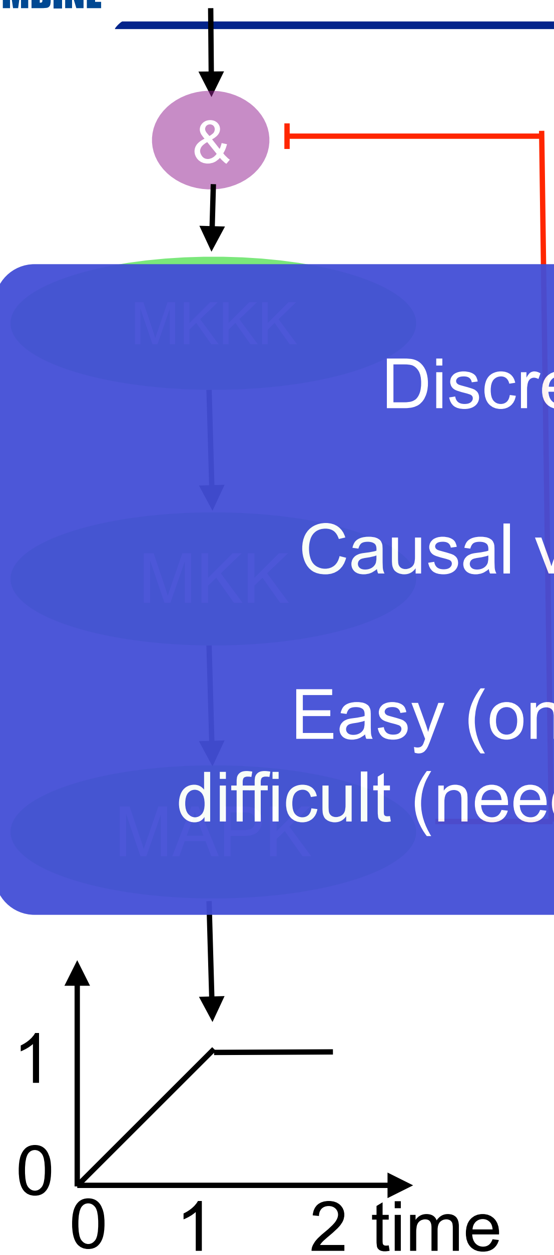


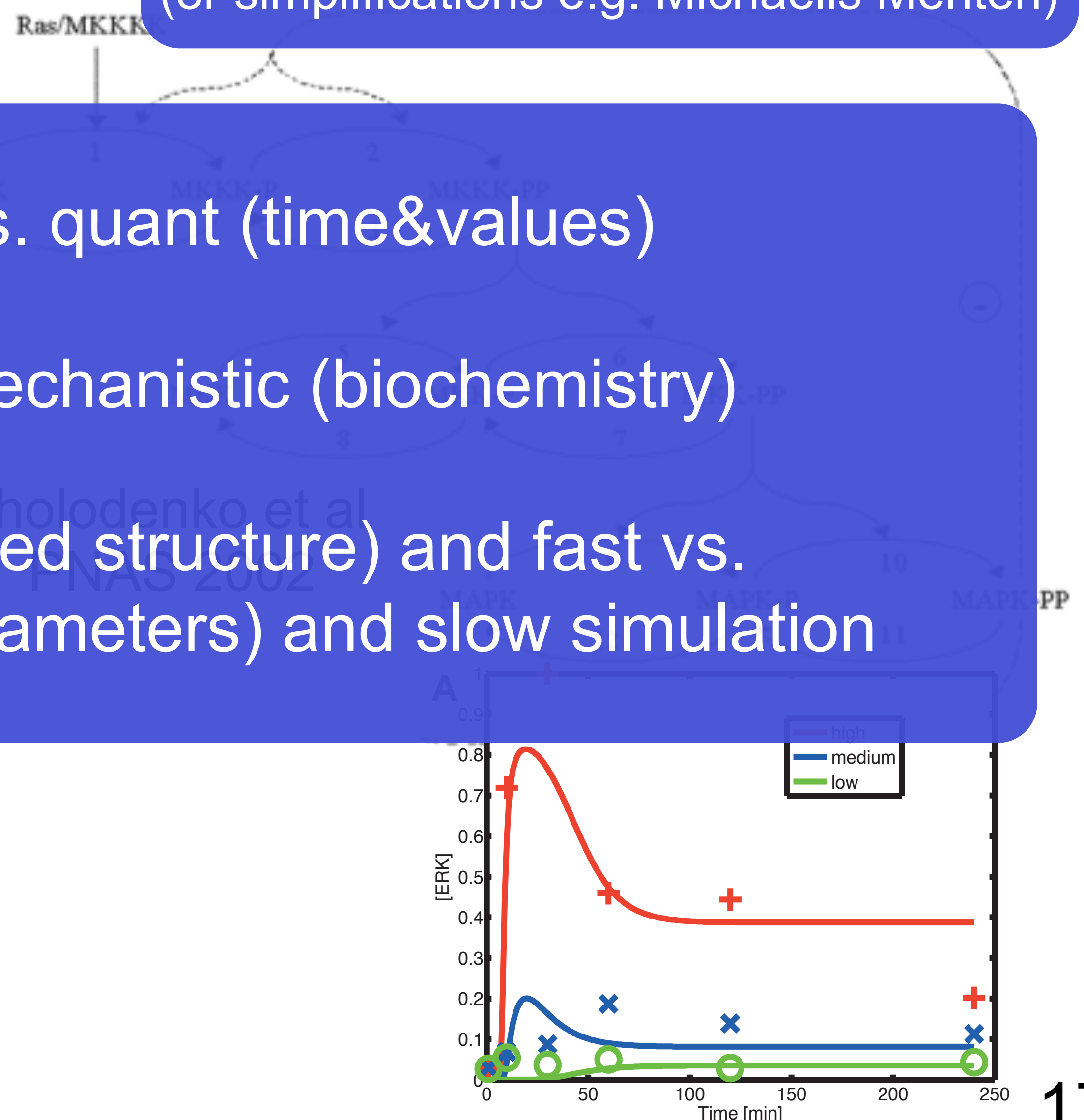
Kholodenko et al
PNAS 2002



Causal vs. reaction representation

Model reactions with mass action law
(or simplifications e.g. Michaelis Menten)





Discrete vs. quant (time&values)

Causal vs. mechanistic (biochemistry)

Easy (only need structure) and fast vs.
difficult (need parameters) and slow simulation

Now Laurence will tell you
all about logic modelling :)

Steps in building (and using) a model

- Set up experiments to extract most information
- Process data efficiently
- Choose type of mathematical model
- 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
- Train models to experimental data
- Use models to gain insight

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

$$\theta_f = \sum_{l=1}^{S} \sum_{K=1}^{M} (\underbrace{Bi_{kl}^{M}}_{\in \{0,1\}} - \underbrace{Bi_{kl}^{E}}_{\in [0,1)})^2$$

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

$$\theta_f = \sum_{l=1}^{S} \sum_{K=1}^{M} (\underbrace{Bi_{kl}^{M}}_{\in \{0,1\}} - \underbrace{Bi_{kl}^{E}}_{\in [0,1)})^2$$

Size of model

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

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

$$\theta_f = \sum_{l=1}^{S} \sum_{K=1}^{M} (\underbrace{Bi_{kl}^M}_{\in \{0,1\}} - \underbrace{Bi_{kl}^E}_{\in [0,1)})^2$$

Relative importance Fit vs. Size

Size of model

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

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

$$\theta_f = \sum_{l=1}^{S} \sum_{K=1}^{M} (\underbrace{Bi_{kl}^{M}}_{\in \{0,1\}} - \underbrace{Bi_{kl}^{E}}_{\in [0,1)})^2$$

Relative importance Fit vs. Size

Size of model

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

In practice: small value (~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

$$\theta_f = \sum_{l=1}^{S} \sum_{K=1}^{M} (\underbrace{Bi_{kl}^{M}}_{\in \{0,1\}} - \underbrace{Bi_{kl}^{E}}_{\in [0,1)})^2$$

Data is normalized between 0 and 1

Relative importance Fit vs. Size

Size of model

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

In practice: small value (~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

$$\theta_f = \sum_{l=1}^{S} \sum_{K=1}^{M} (\underbrace{Bi_{kl}^{M}}_{\in \{0,1\}} - \underbrace{Bi_{kl}^{E}}_{\in [0,1)})^2$$

Data is normalized between 0 and 1

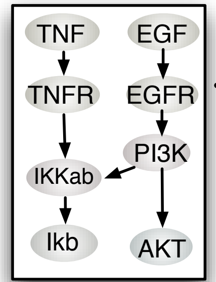
Relative importance Fit vs. Size

Size of model

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

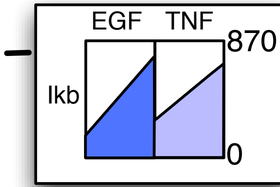
Best model ~ minimum metric
(optimization problem) - can be solved algorithmically

Linking pathway maps to data of signal transduction

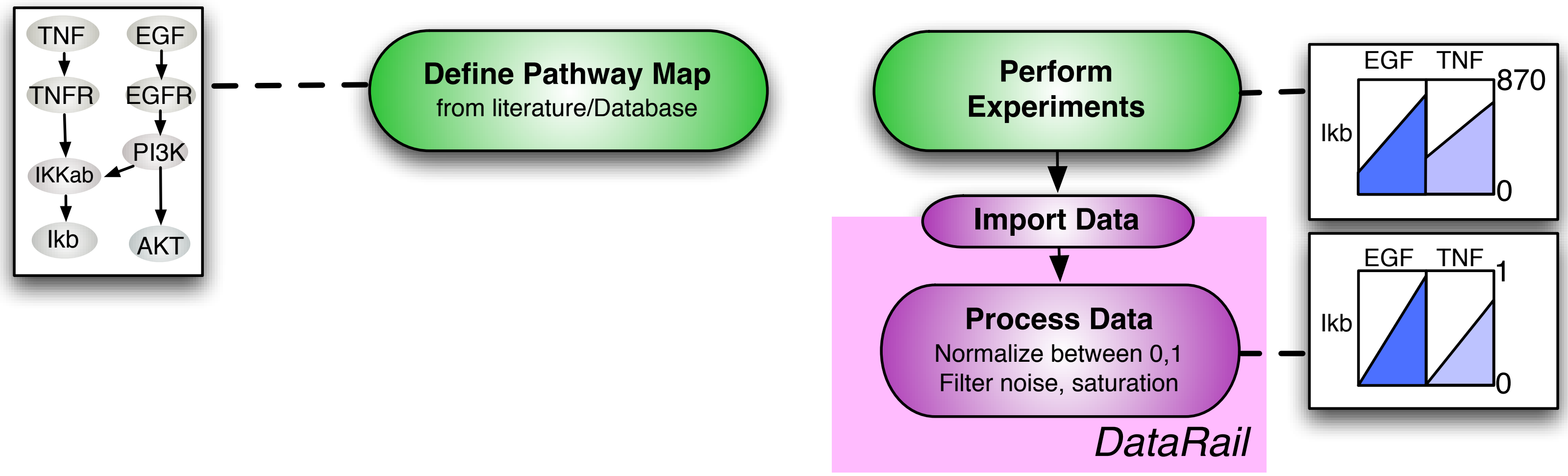


Define Pathway Map
from literature/Database

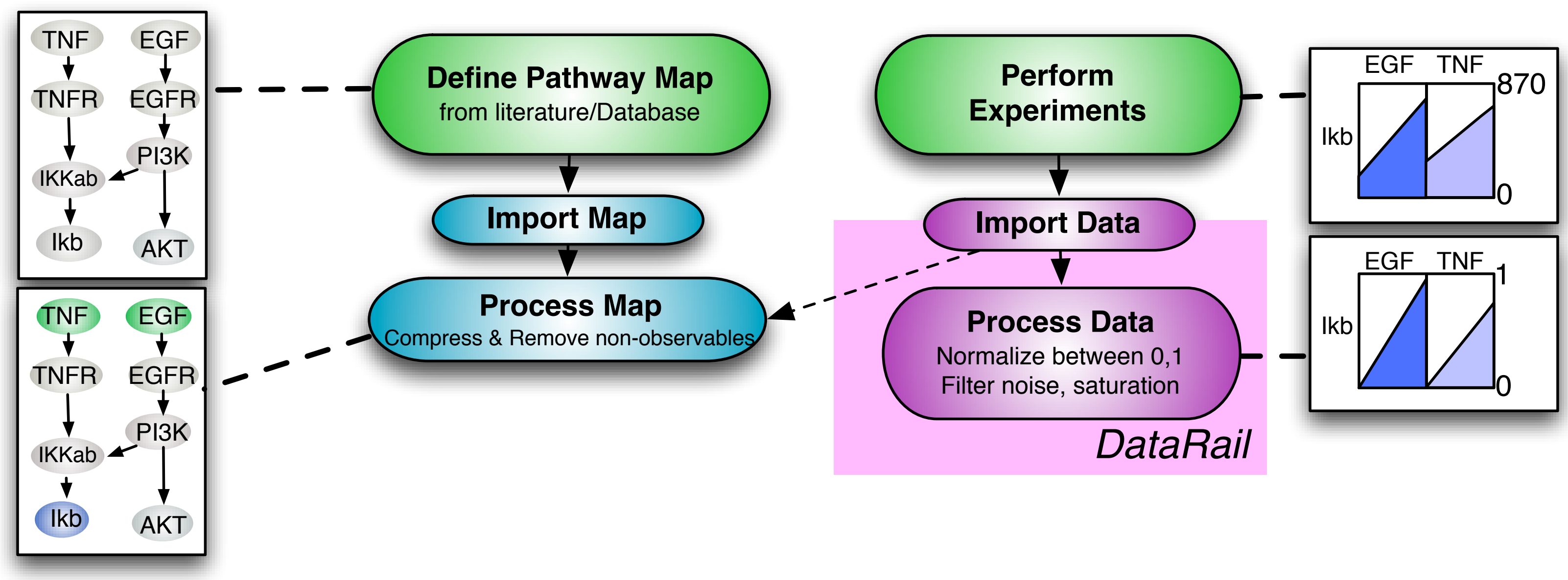
Perform Experiments



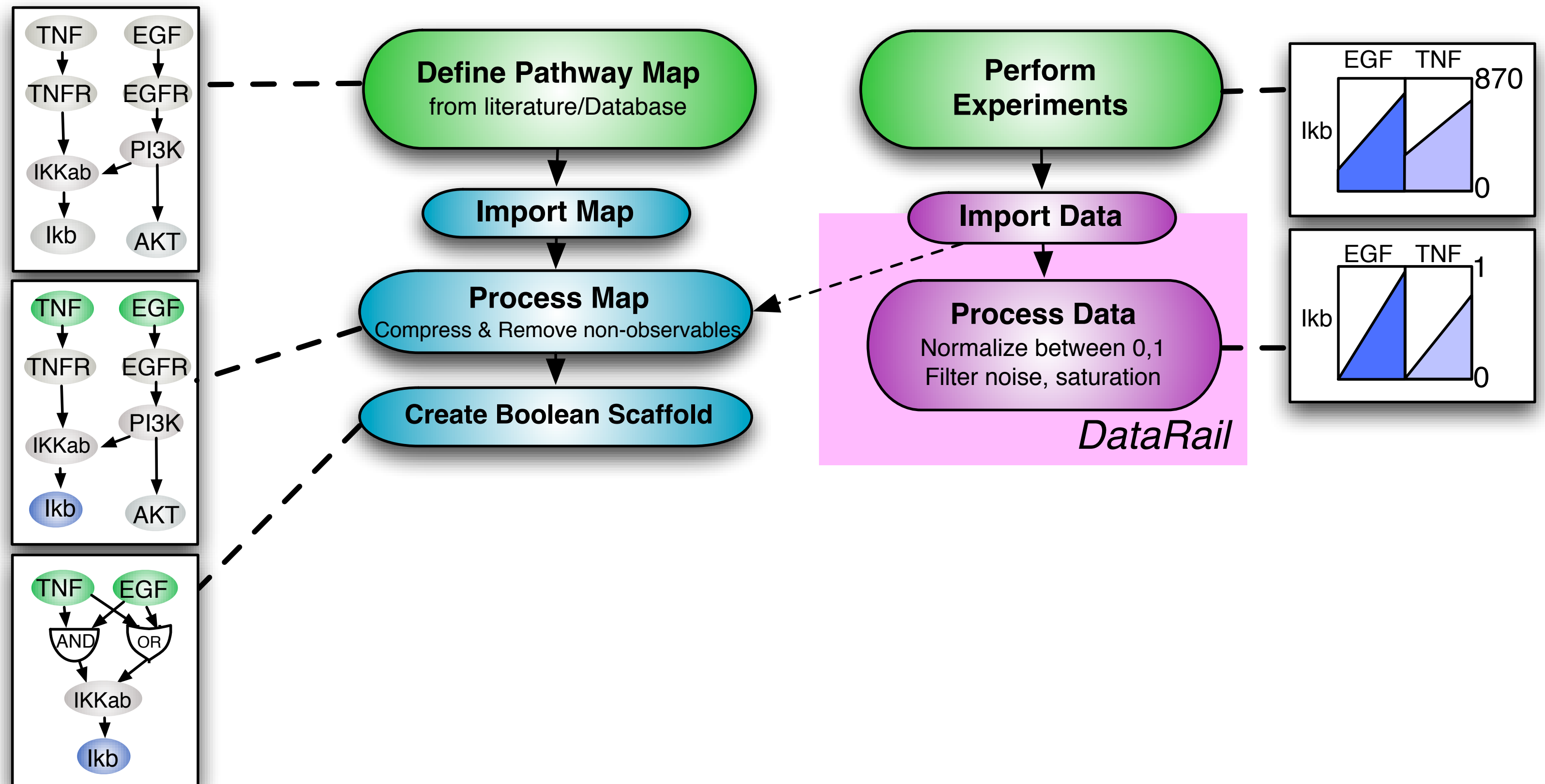
Linking pathway maps to data of signal transduction



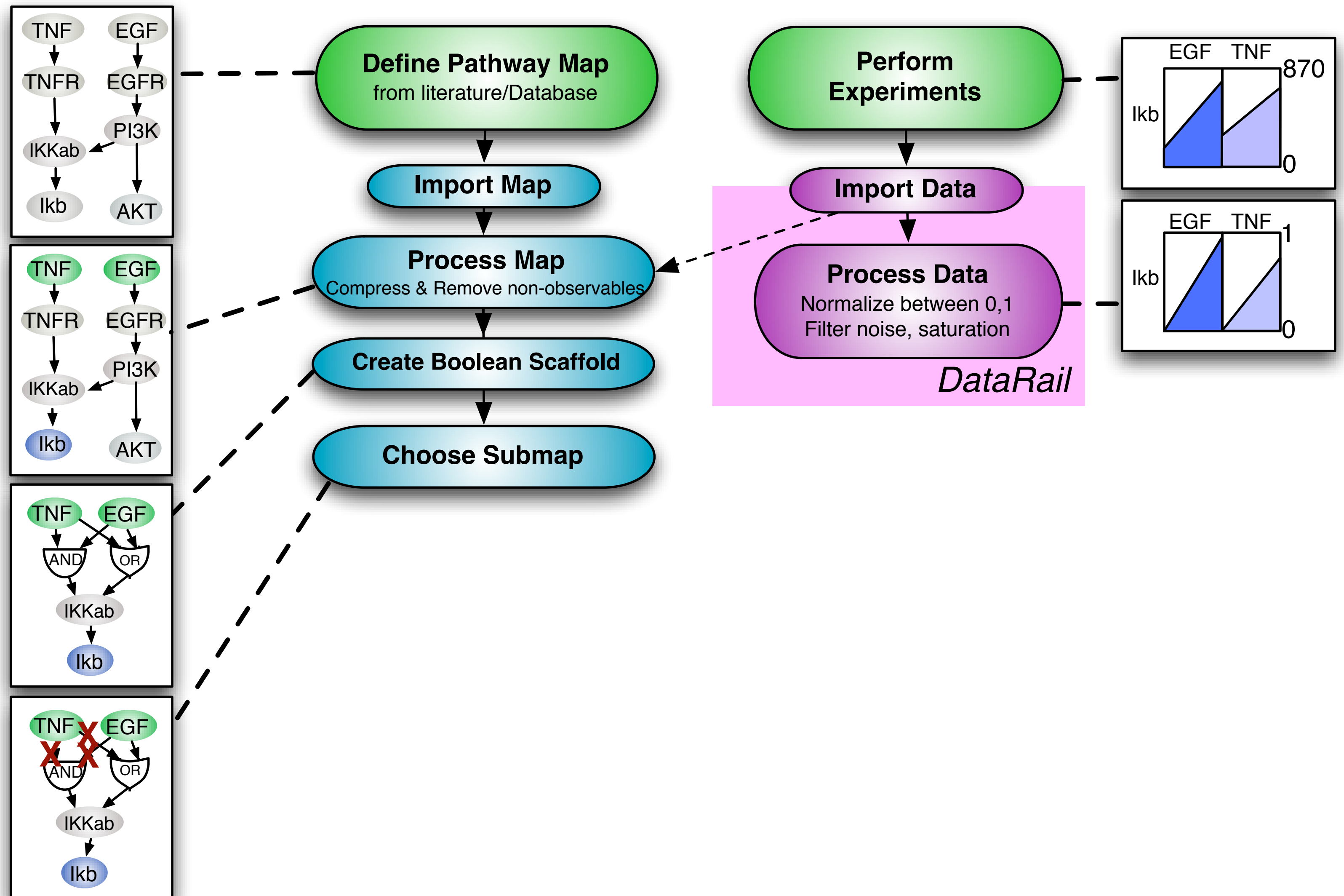
Linking pathway maps to data of signal transduction



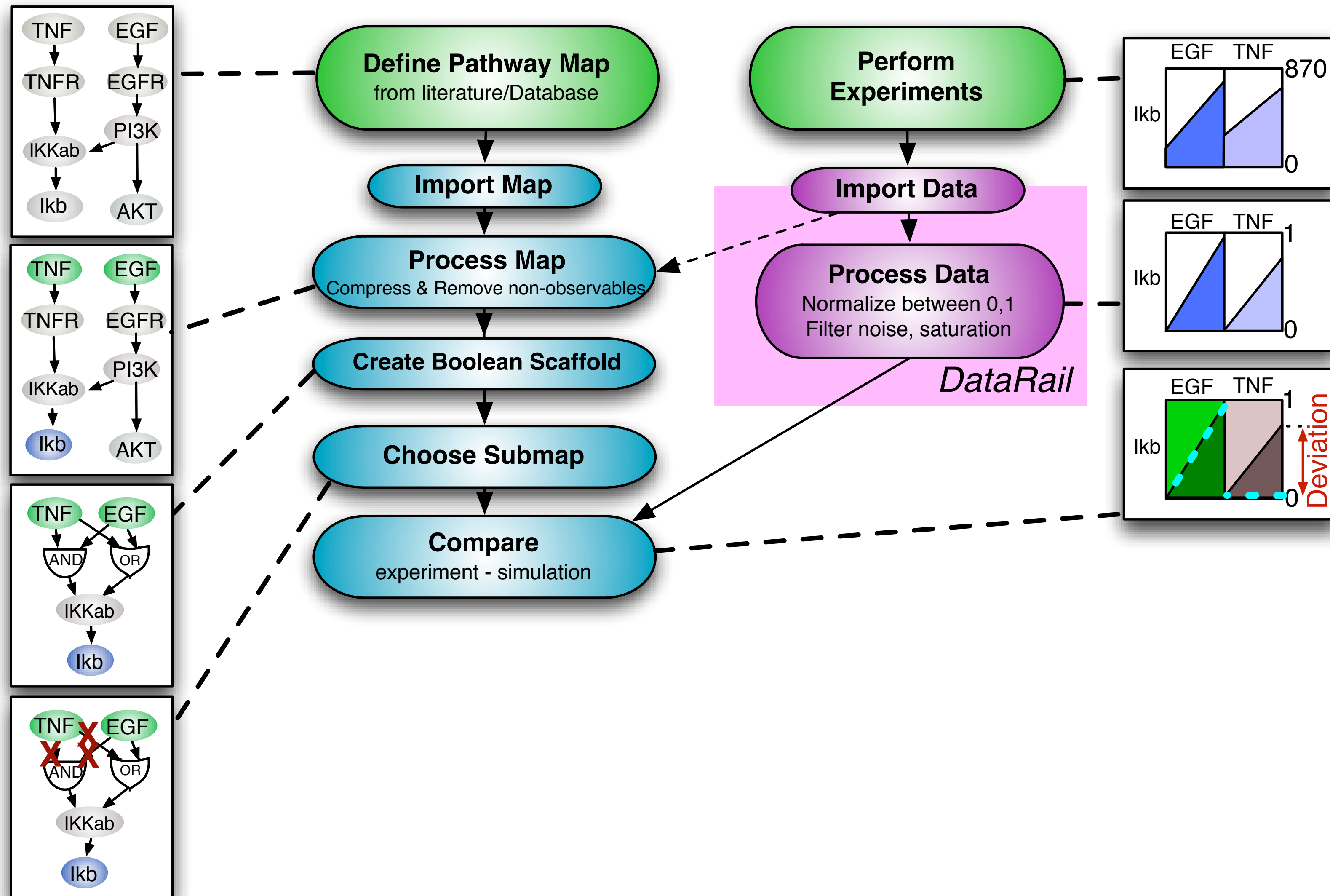
Linking pathway maps to data of signal transduction



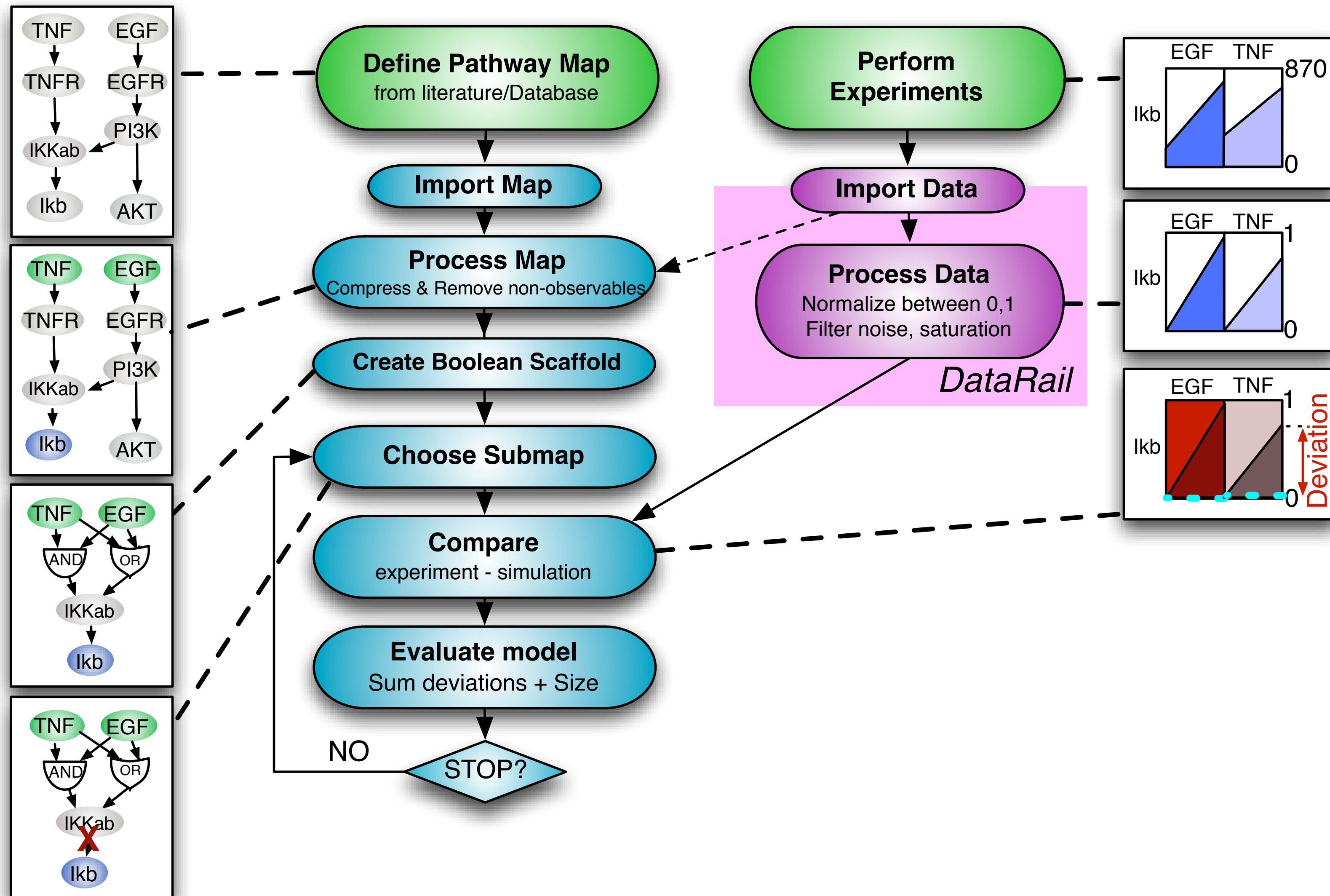
Linking pathway maps to data of signal transduction



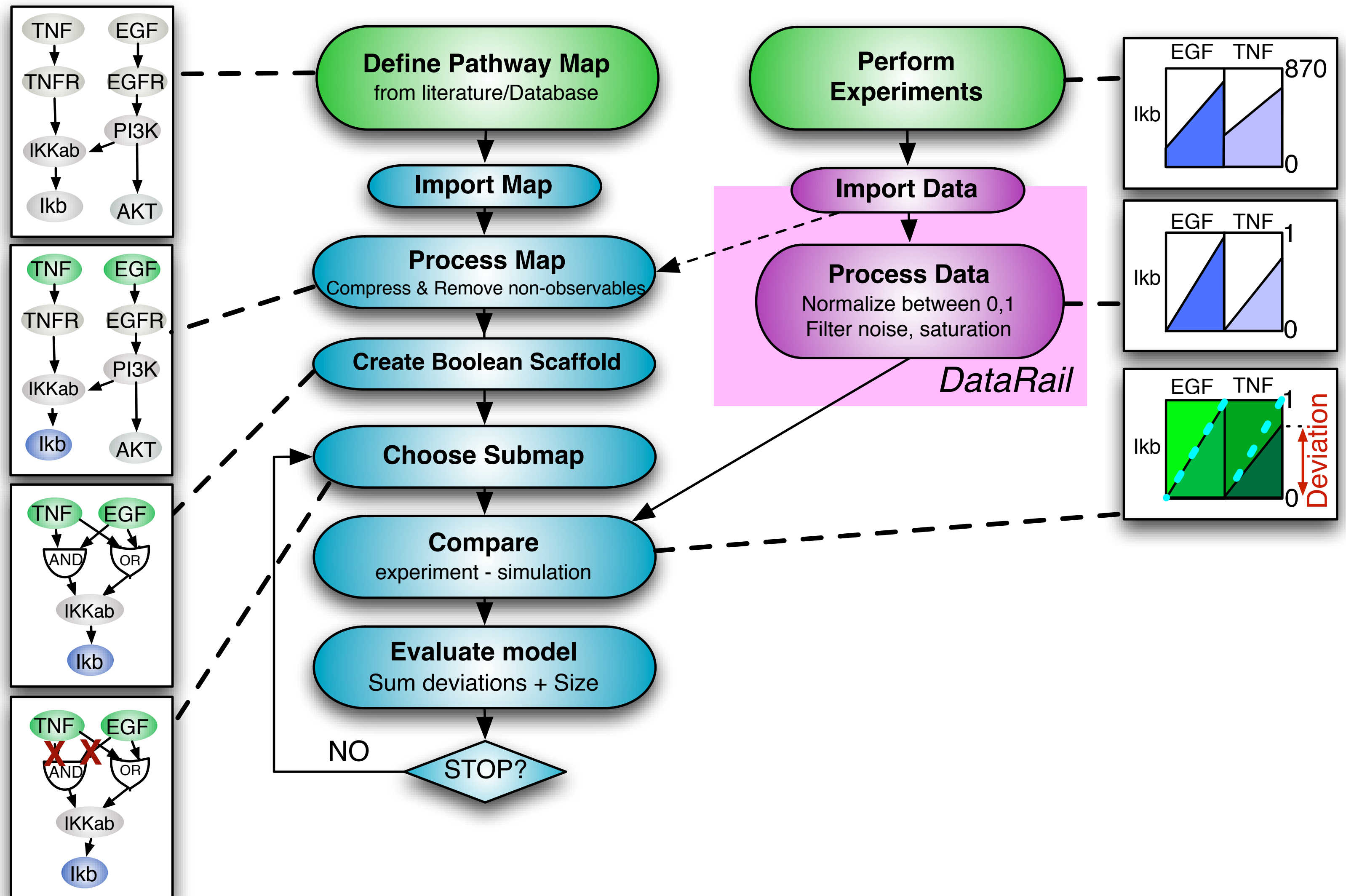
Linking pathway maps to data of signal transduction



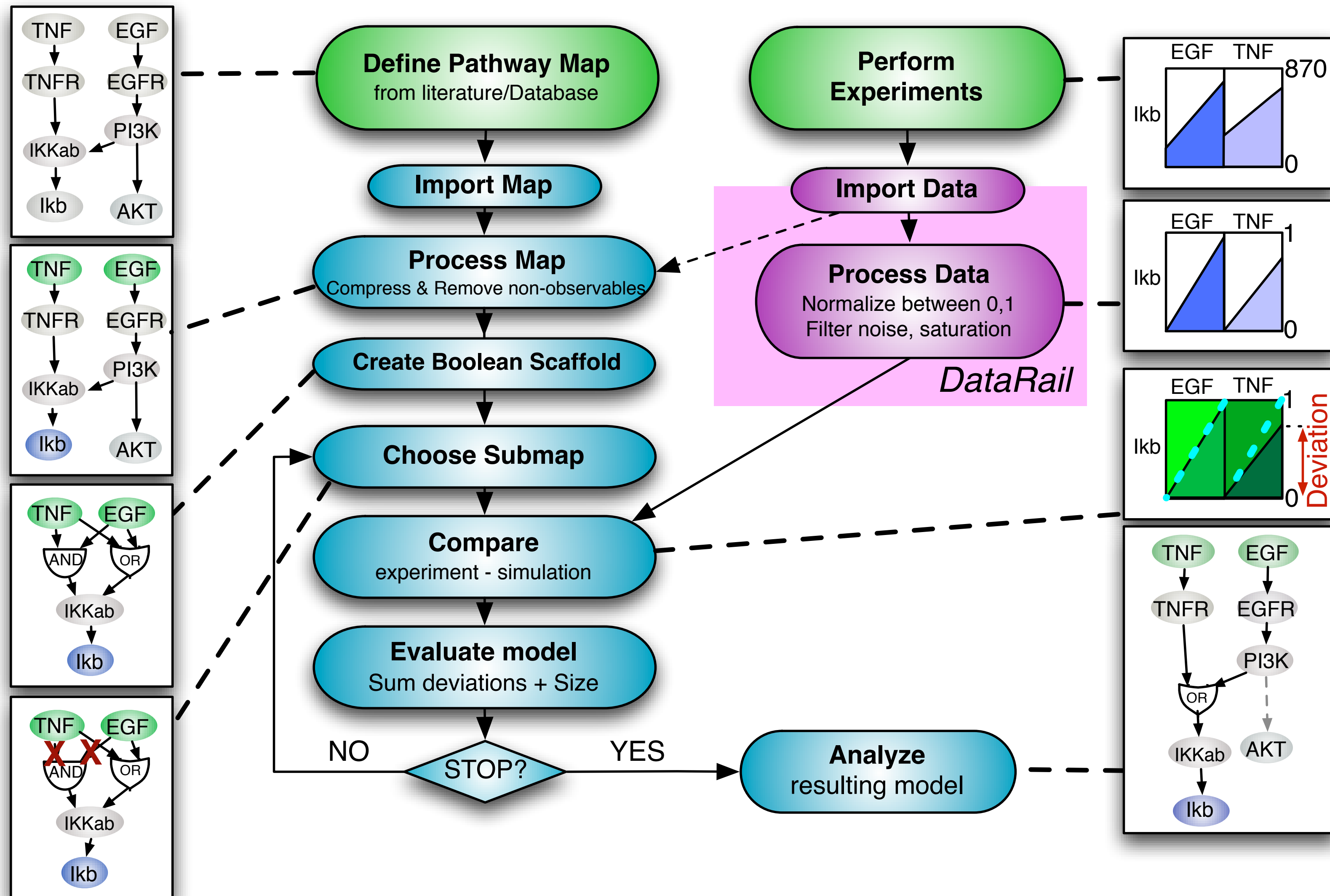
Linking pathway maps to data of signal transduction



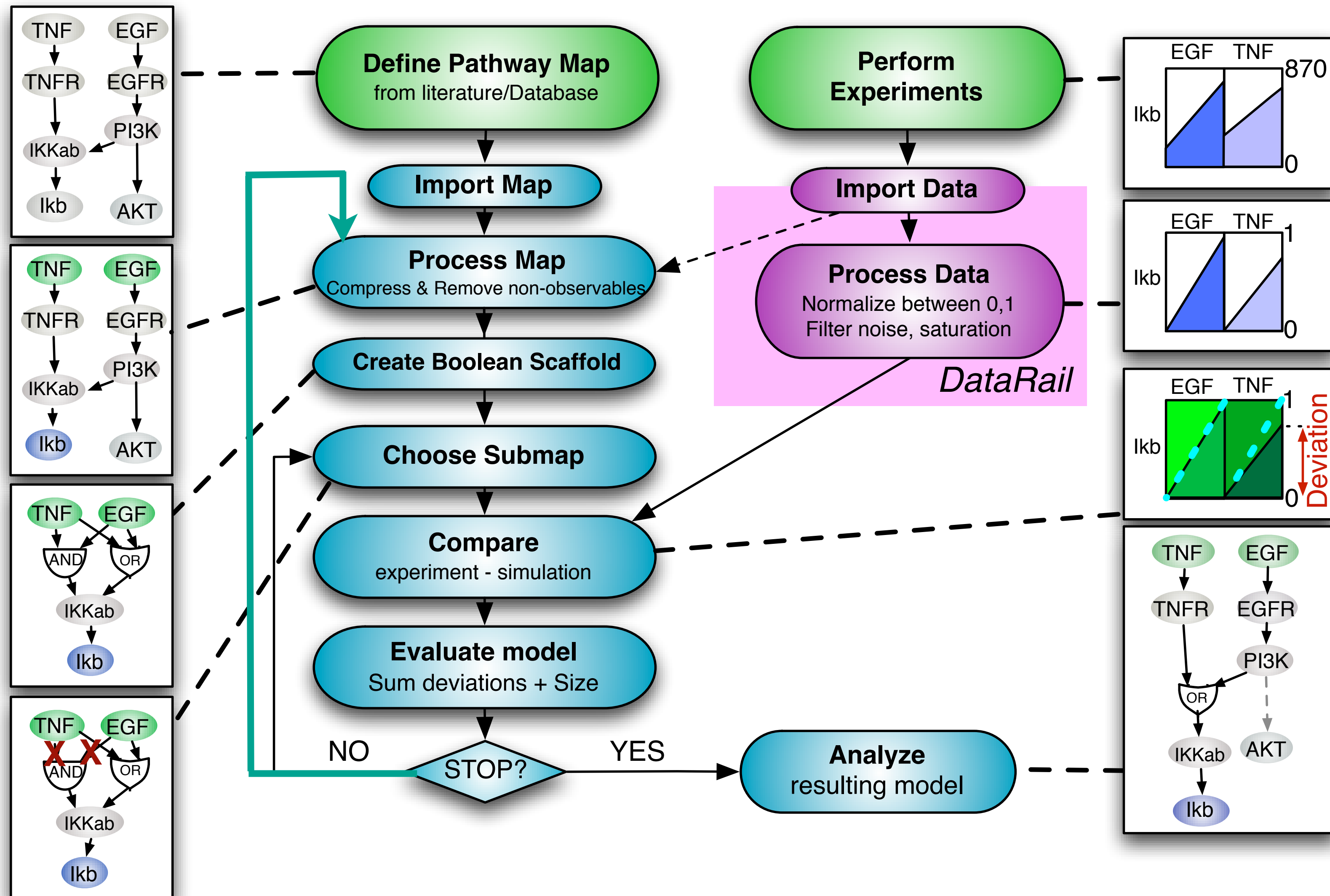
Linking pathway maps to data of signal transduction



Linking pathway maps to data of signal transduction



Linking pathway maps to data of signal transduction



Linking pathway maps to data of signal transduction

Pipeline implemented in

CellNOpt *(for CellNetOptimizer), aka CNO*

a Bioconductor, Python (and Matlab) toolbox

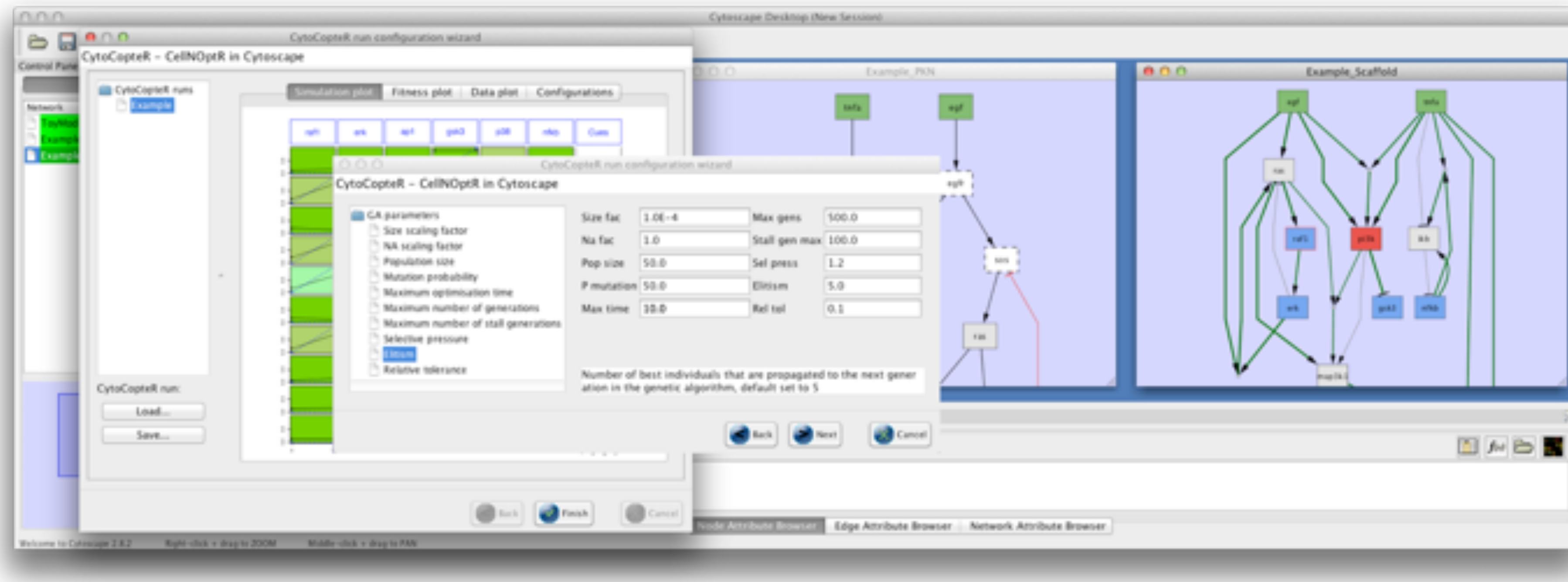
freely available at http:www.cellnopt.org

Bioconductor:
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

Matlab:
Morris MK, Melas I, Saez-Rodriguez J, *Methods Mol. Biol,* 930:179-214, 2013

CytoCopter

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



Emanuel Gonçalves **Martijn van Iersel**

Terfve C *BMC Syst Biol, 6:*133, 2012

There are many modeling tools!

Table 2: Tools Available for the Logic Modeling of Biochemical Signaling Networks

tool	type of logic	functionality	treatment of time
BooleanNet	Boolean, piecewise linear	simulation and visualization	synchronous, asynchronous, or continuous
GinSim	discrete (multistate)	model building, simulation, and analysis	synchronous, asynchronous, or mixed asynchronous
CellNetAnalyzer	Boolean (multistate)	model simulation, visualization, and network properties analysis	logic steady state
CellNetOptimizer	Boolean	model training and simulation	logic steady state
Odefy	Boolean, logic-based ODEs	model simulation and visualization	synchronous, asynchronous, or continuous
Genetic Network Analyzer	piecewise linear	model building and simulation	continuous
ChemChains	Boolean	model simulation, visualization, and analysis	synchronous or asynchronous
MetaReg	discrete (multistate)	model simulation, visualization, and refinement	logic steady state
SQUAD	standardized qualitative dynamical systems	model simulation and analysis	continuous

Morris MK, Saez-Rodriguez J, Sorger PK, Lauffenburger DA, *Biochemistry,* 49(15), 2010.

There are many modeling tools!

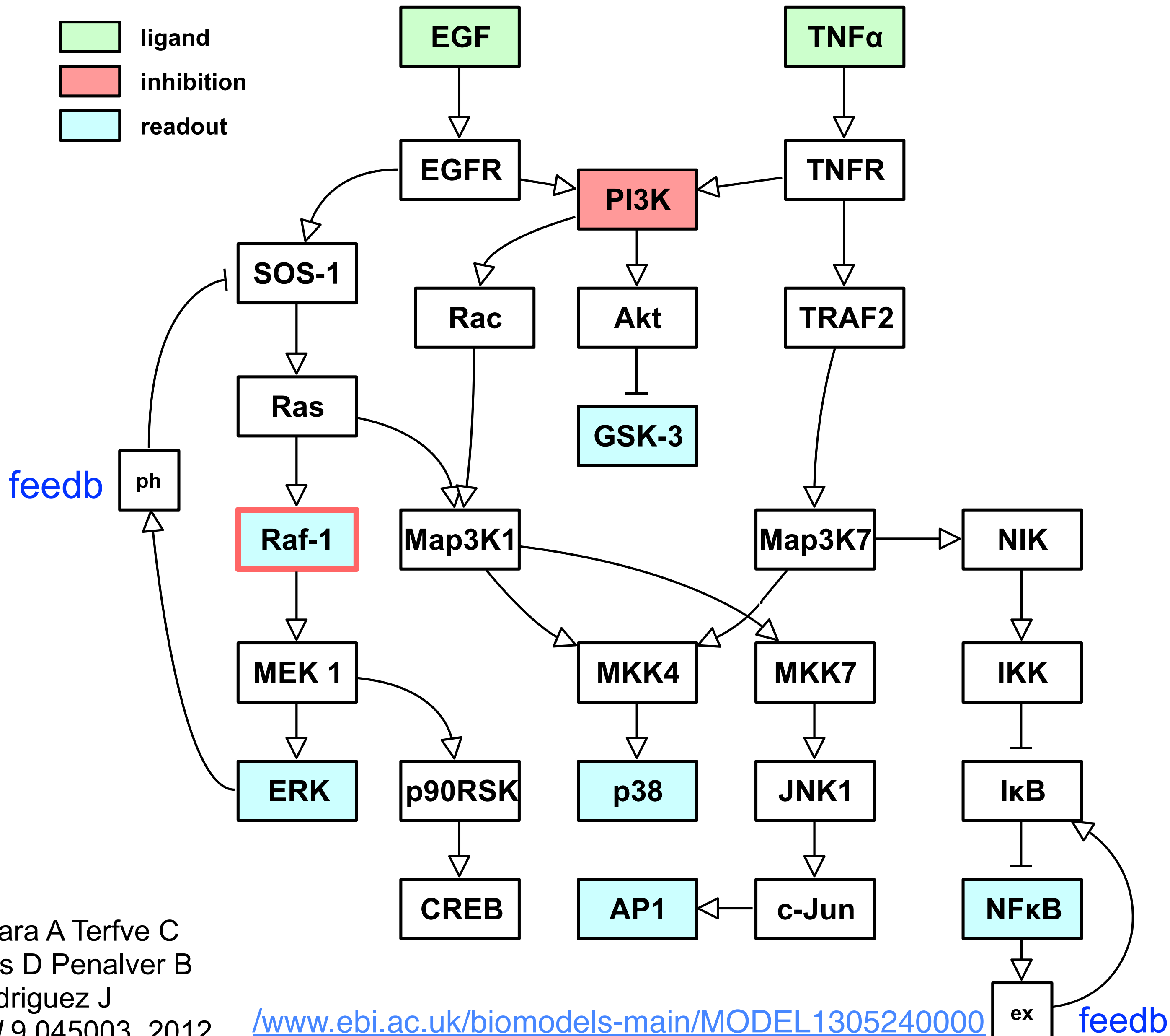
Table 2: Tools Available for the Logic Modeling of Biochemical Signaling Networks

tool	type of logic	functionality	treatment of time
BooleanNet	Boolean, piecewise linear	simulation and visualization	synchronous, asynchronous, or continuous
GinSim	discrete (multistate)	model building, simulation, and analysis	synchronous, asynchronous, or mixed asynchronous
CellNetAnalyzer	Boolean (multistate)	model simulation, visualization, and network properties analysis	logic steady state
CellNetOptimizer	Boolean	model training and simulation	logic steady state
Odefy	Boolean, logic-based ODEs	model simulation and visualization	synchronous, asynchronous, or continuous
Genetic Network Analyzer	piecewise linear	model building and simulation	continuous
ChemChains	Boolean	model simulation, visualization, and analysis	synchronous or asynchronous
MetaReg	discrete (multistate)	model simulation, visualization, and refinement	logic steady state
SQUAD	standardized qualitative dynamical systems	model simulation and analysis	continuous

- SBML-qual to exchange logic models
Chaouiya et al. *BMC Sys Bio* 2013

Morris MK, Saez-Rodriguez J, Sorger PK, Lauffenburger DA, *Biochemistry,* 49(15), 2010.

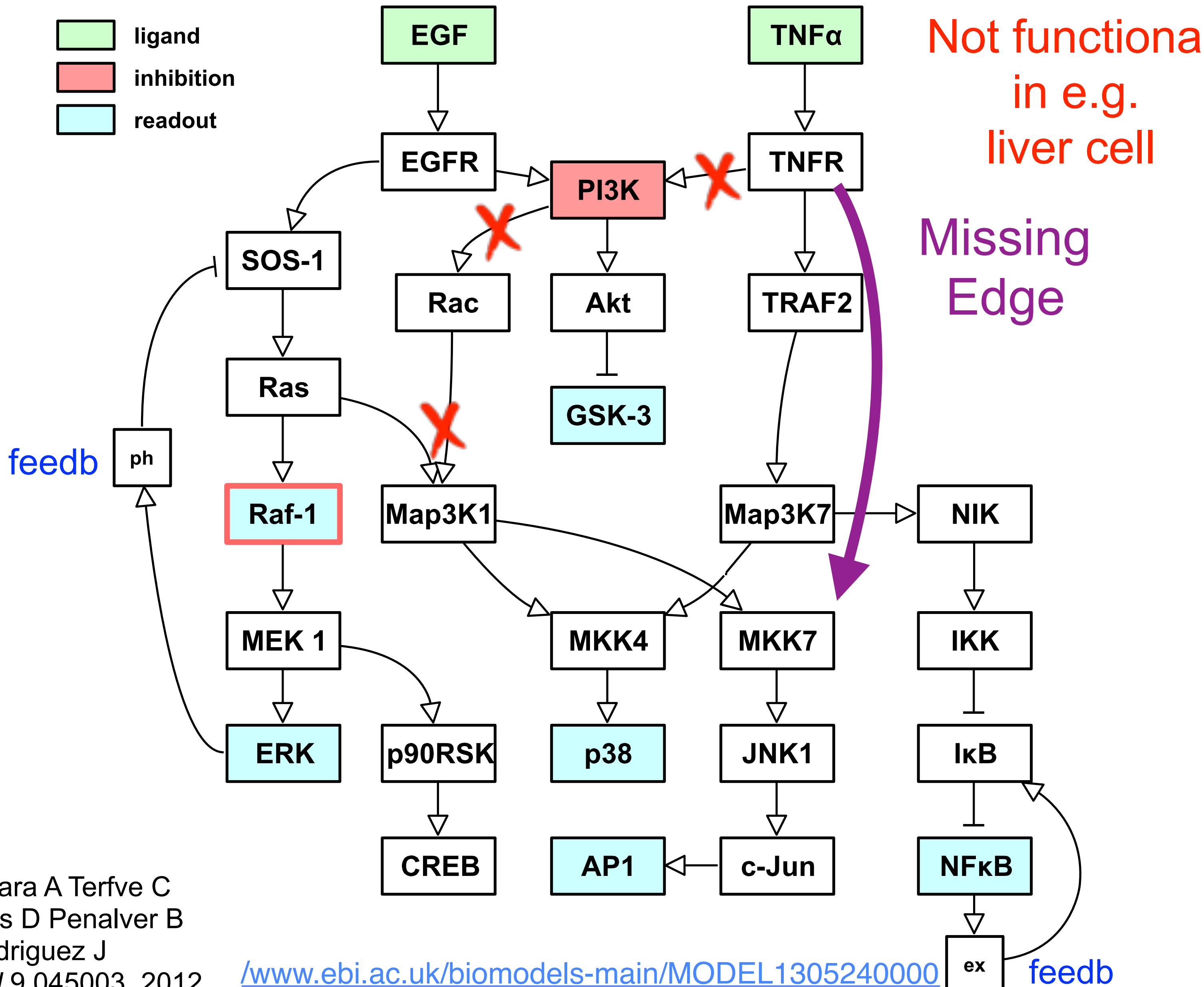
A Toy model



MacNamara A Terfve C
Henriques D Penalver B
Saez-Rodriguez J
Phys Biol 9 045003, 2012

/www.ebi.ac.uk/biomodels-main/MODEL1305240000

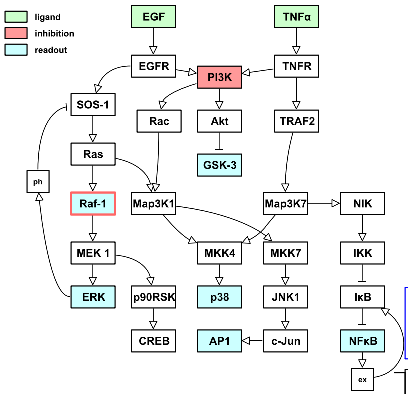
A Toy model

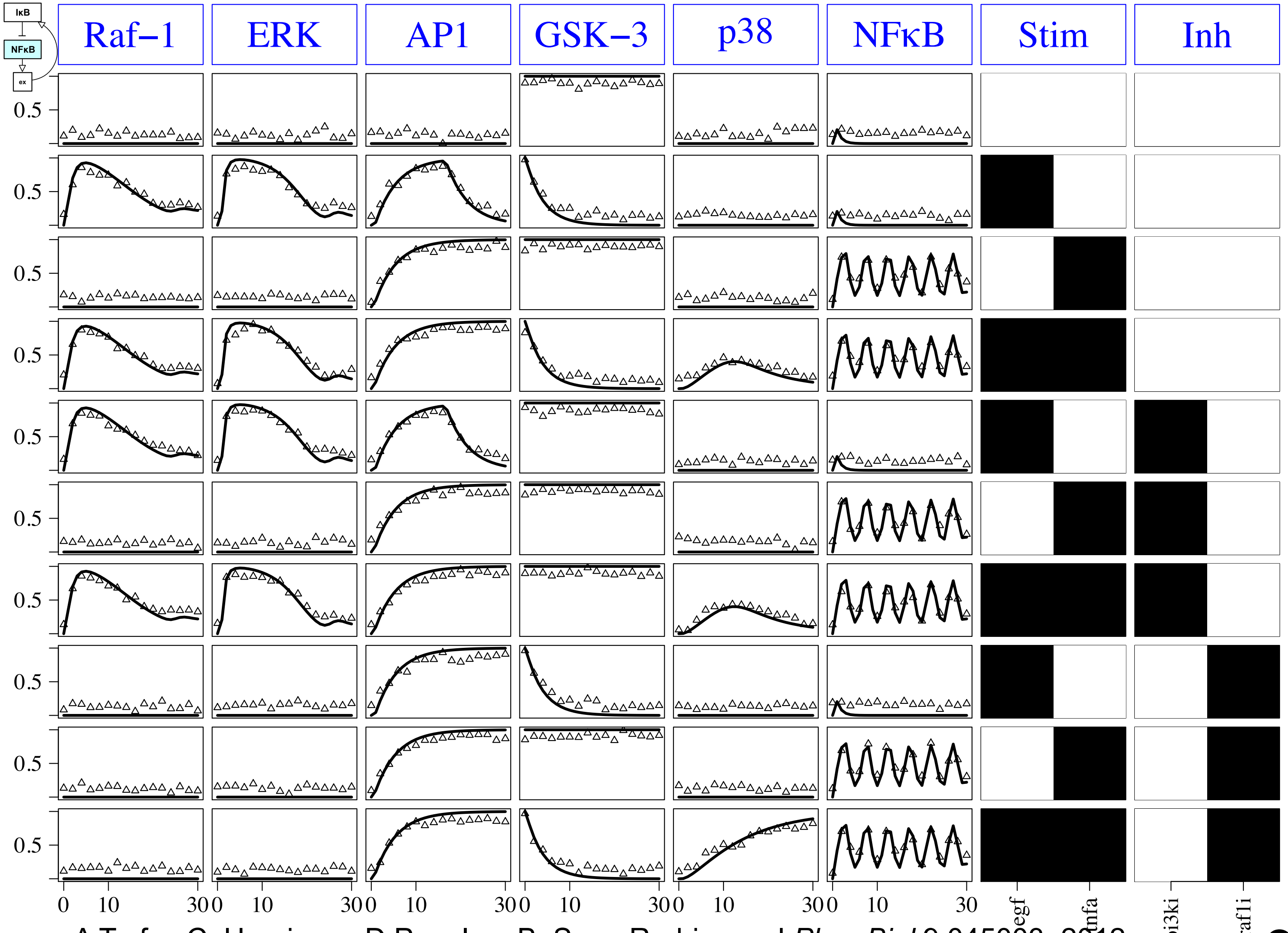


MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

/www.ebi.ac.uk/biomodels-main/MODEL1305240000

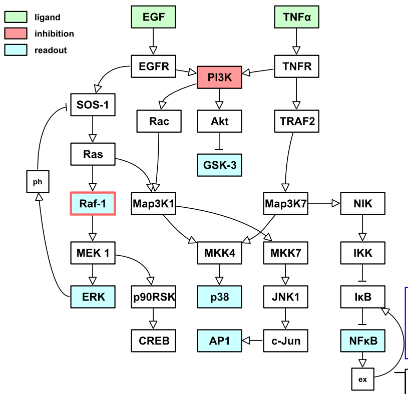
The ‘real’ data



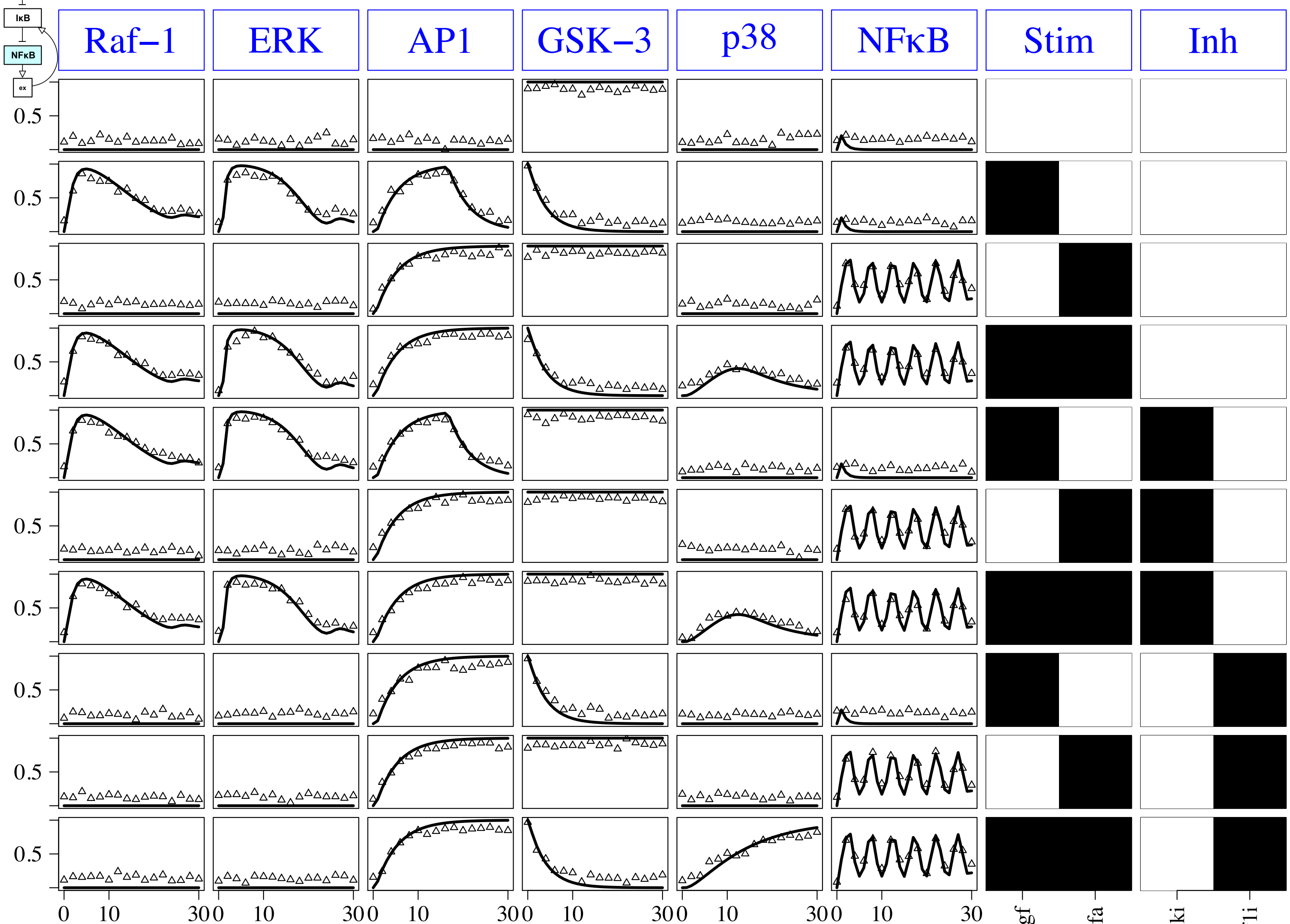


MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

The ‘real’ data

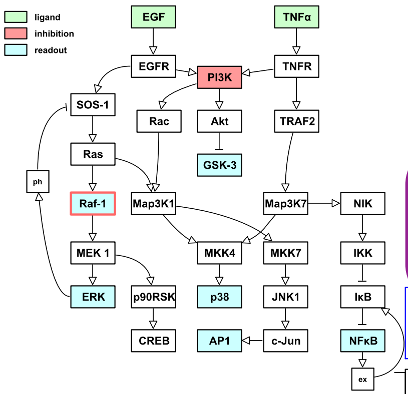


How to pick right time to measure?



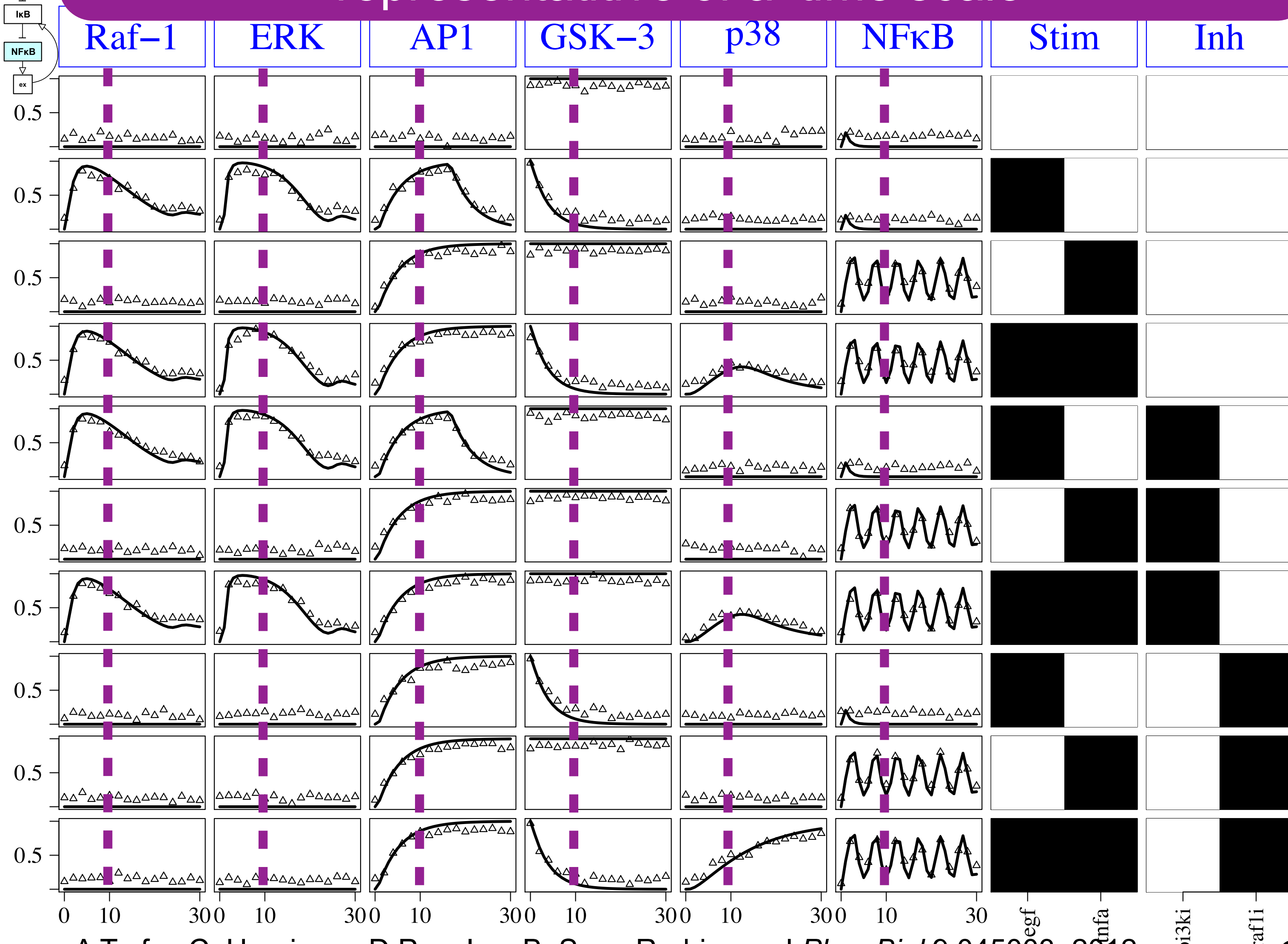
MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

The 'real' data



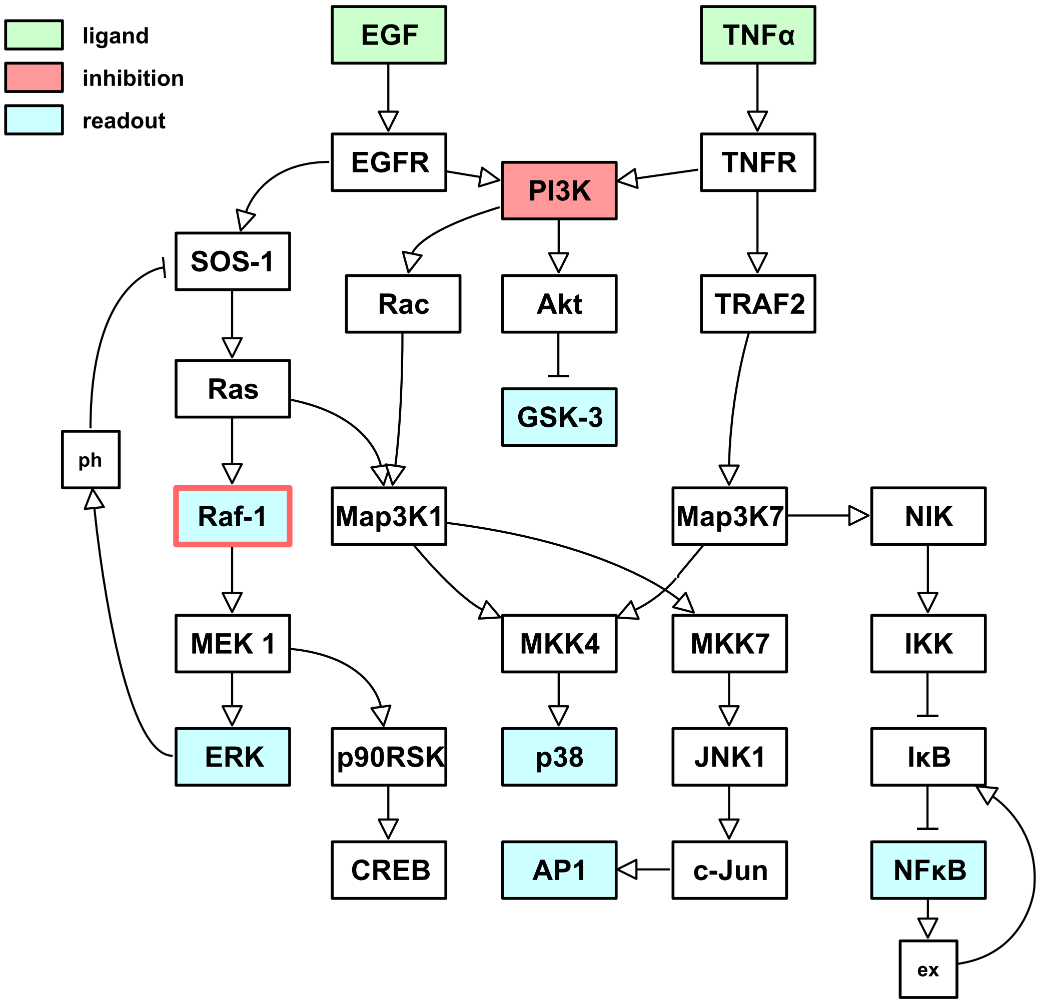
If you can only pick one ($$), choose one representatitive of a 'time scale'

How to pick right time to measure?

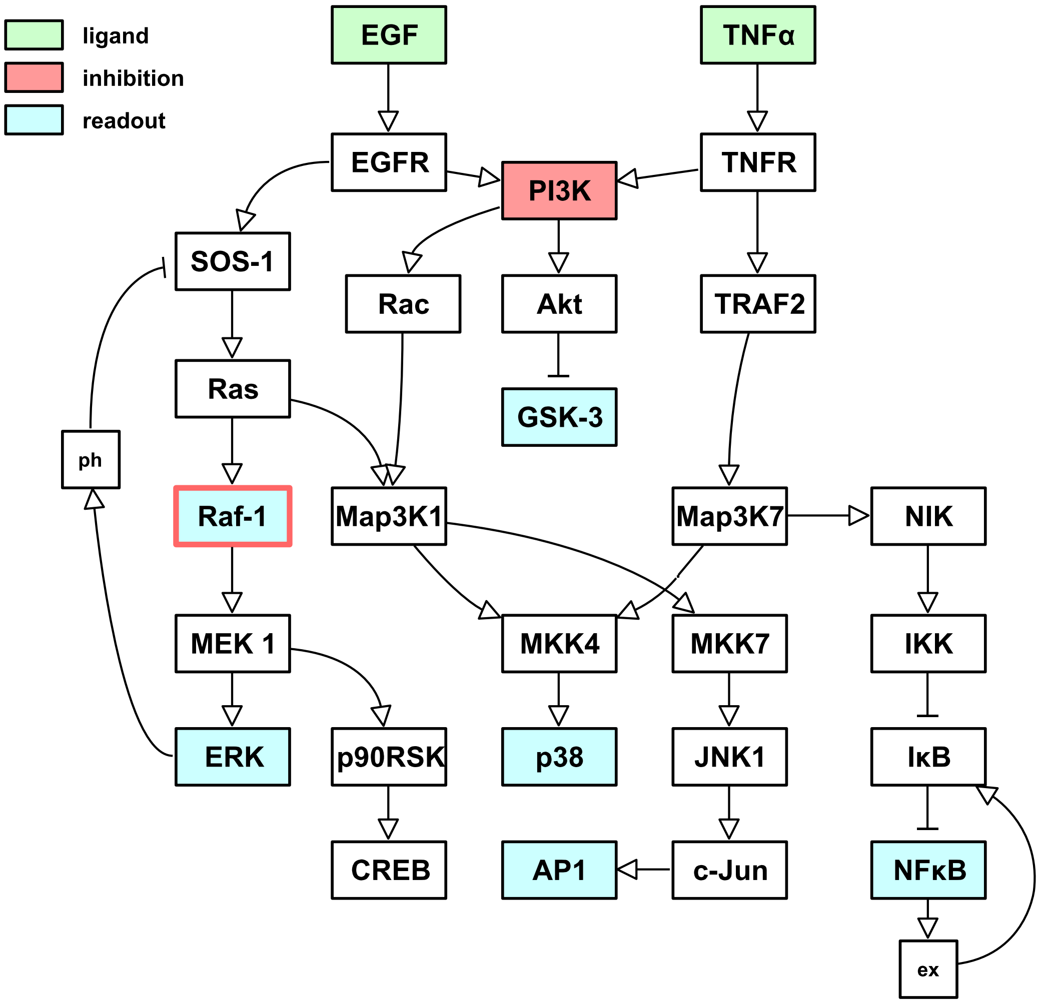


MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

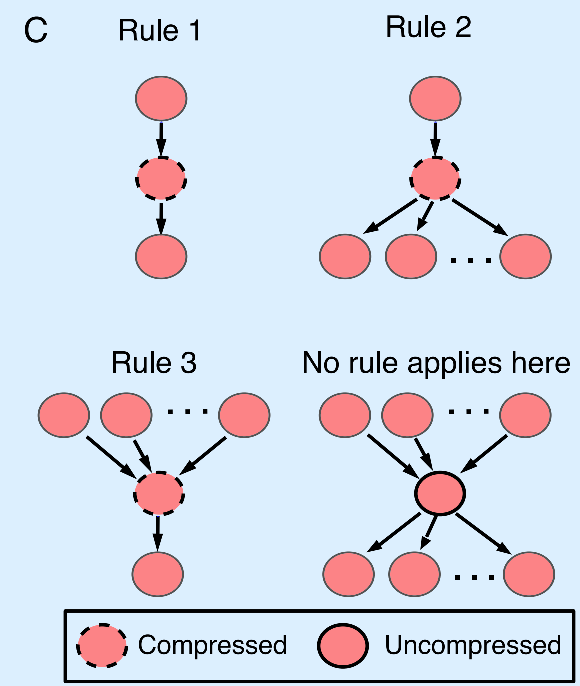
Model preprocessing:



Model preprocessing:

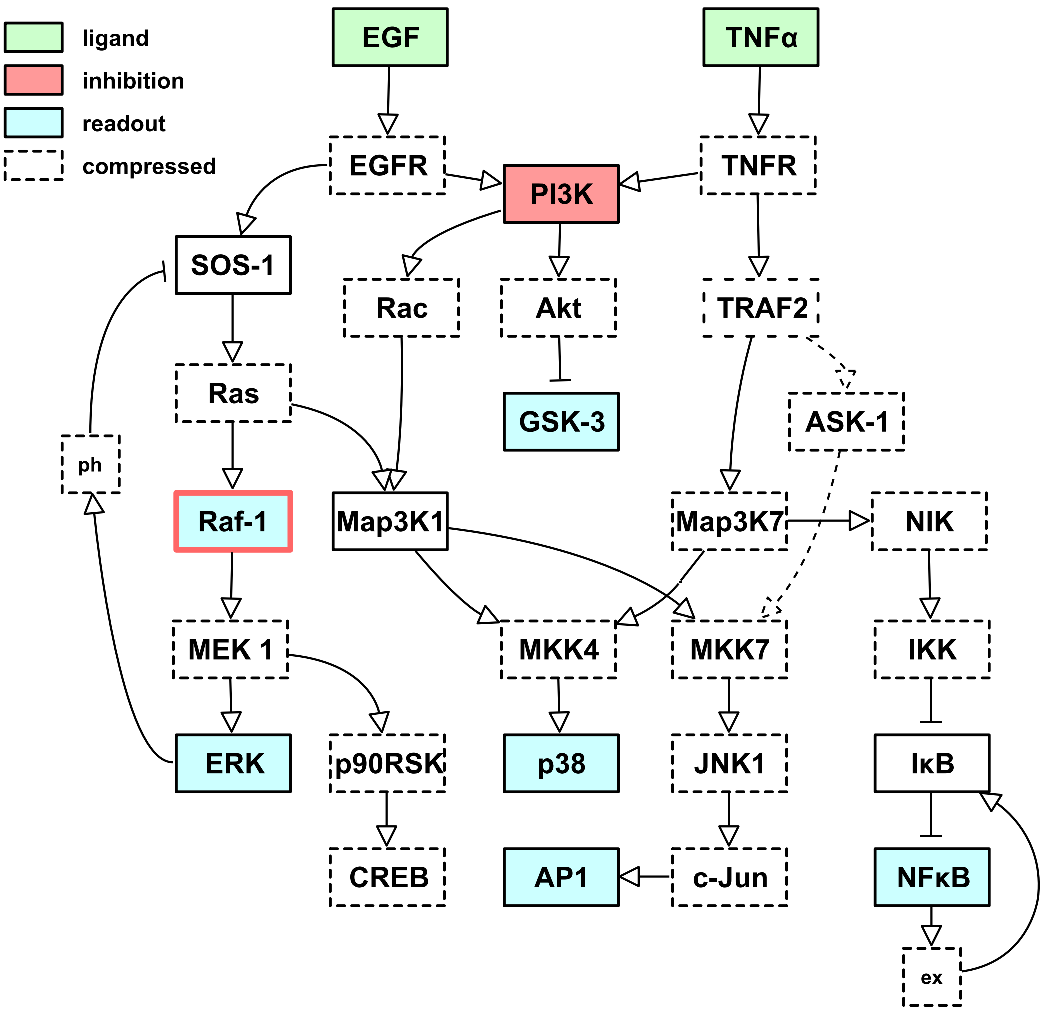


Model compression& removal of non-controlable & non-observable branches

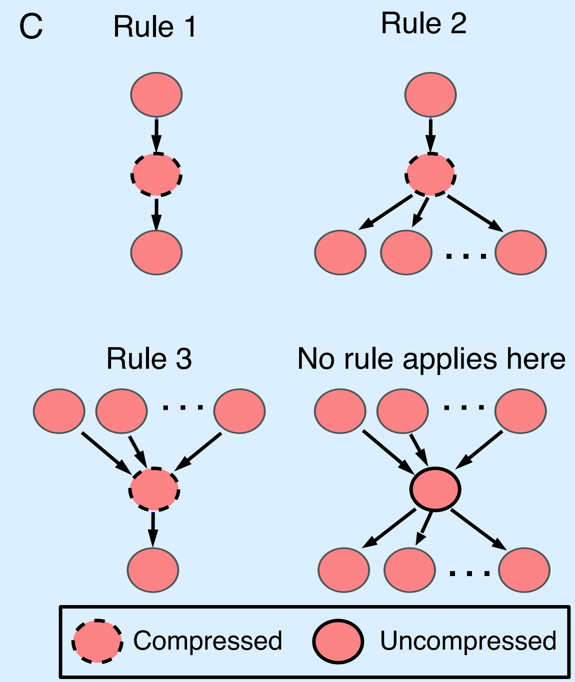


Saez-Rodriguez J, et al., *Mol. Syst. Biol*, 2009

Model preprocessing:

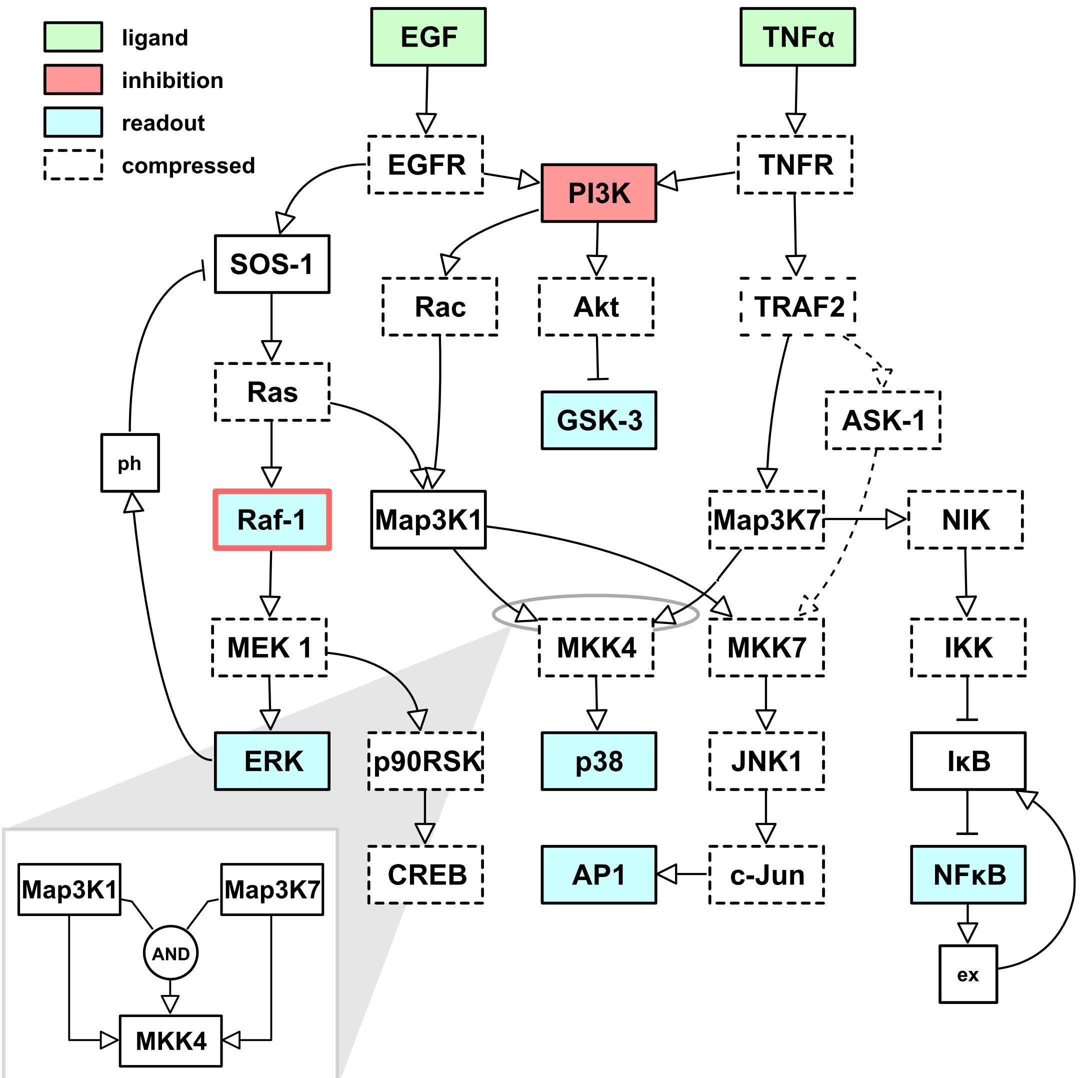


Model compression& removal of non-controlable & non-observable branches

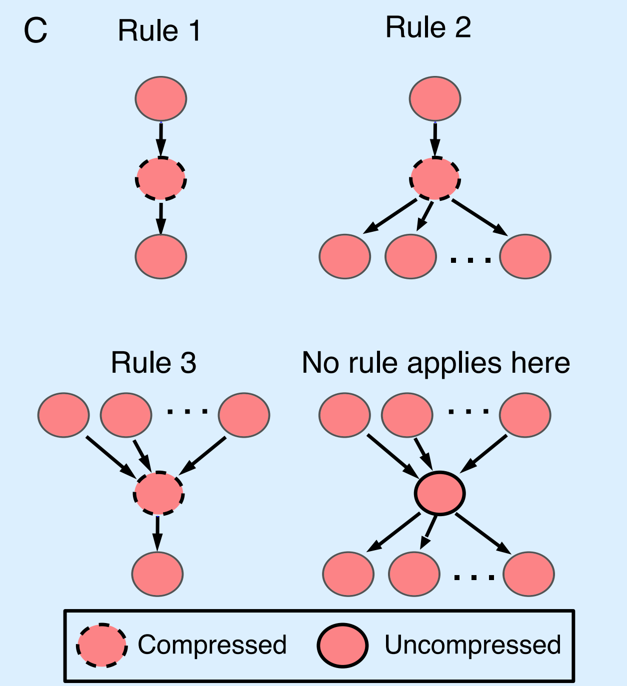


Saez-Rodriguez J, et al., *Mol. Syst. Biol*, 2009

Model preprocessing:

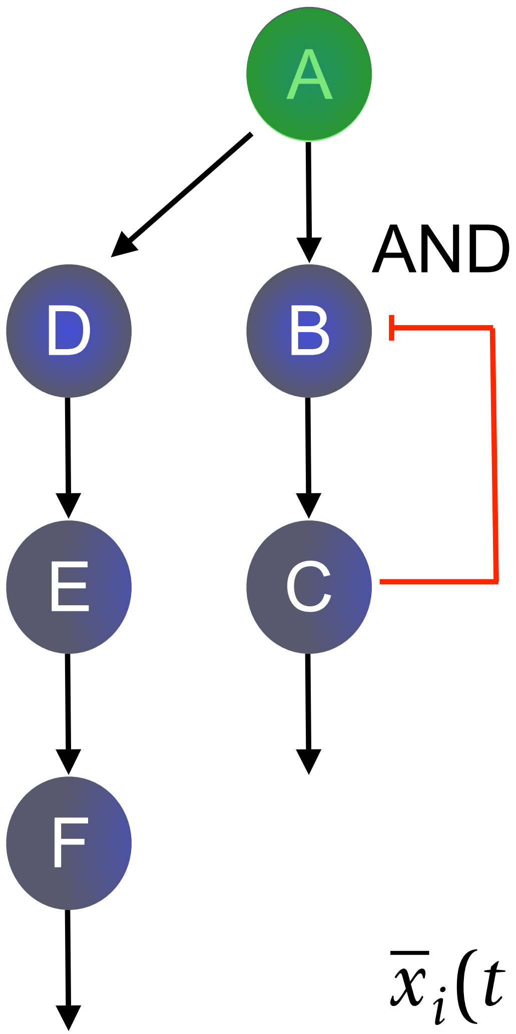


Model compression& removal of non-controlable & non-observable branches



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

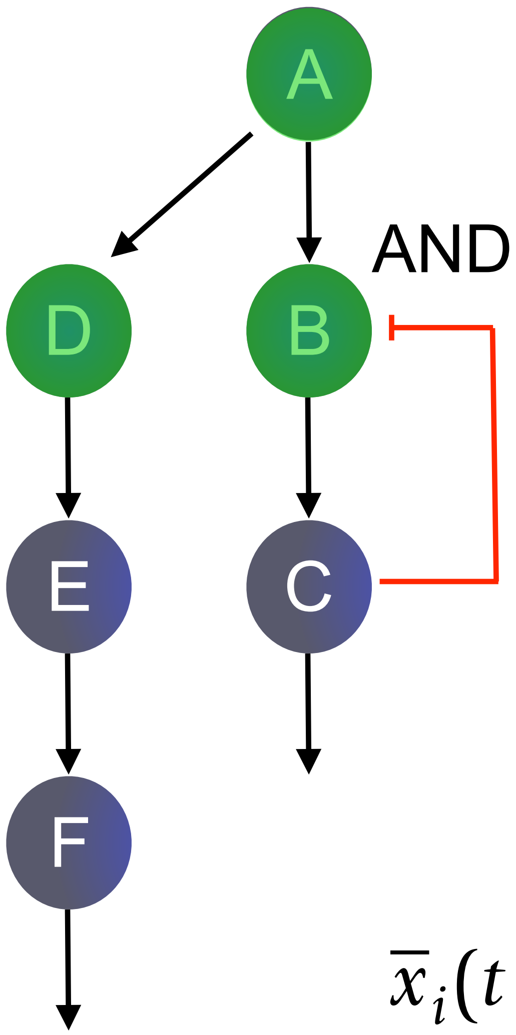
Simplest Boolean simulation: Synchronous updates



System can reach steady state or oscillate

$$\overline{x}_i(t+1) = \overline{B}_i(\overline{x}_{i1}(t), \overline{x}_{i2}(t), \ldots, \overline{x}_{iN_i}(t)).$$

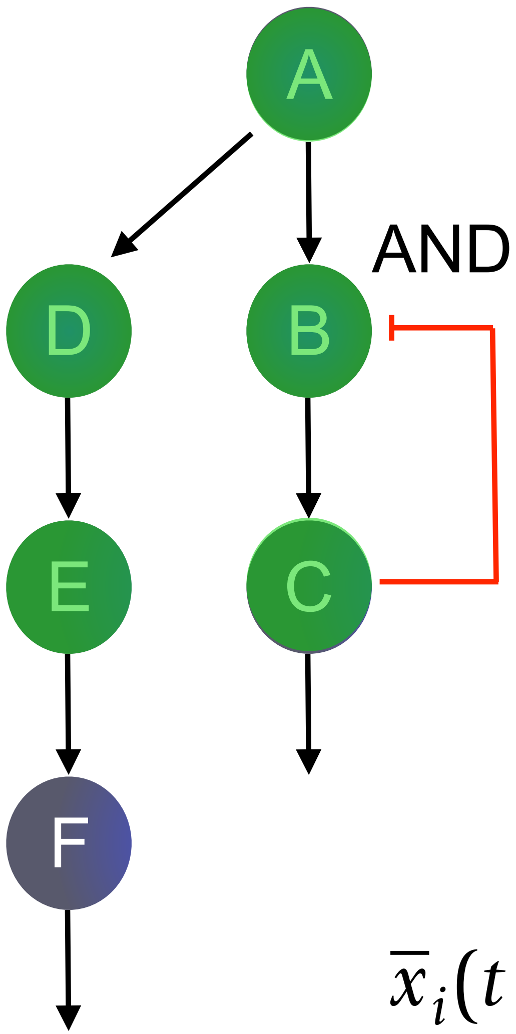
Simplest Boolean simulation: Synchronous updates



System can reach steady state or oscillate

$$\overline{x}_i(t+1) = \overline{B}_i(\overline{x}_{i1}(t), \overline{x}_{i2}(t), \ldots, \overline{x}_{iN_i}(t)).$$

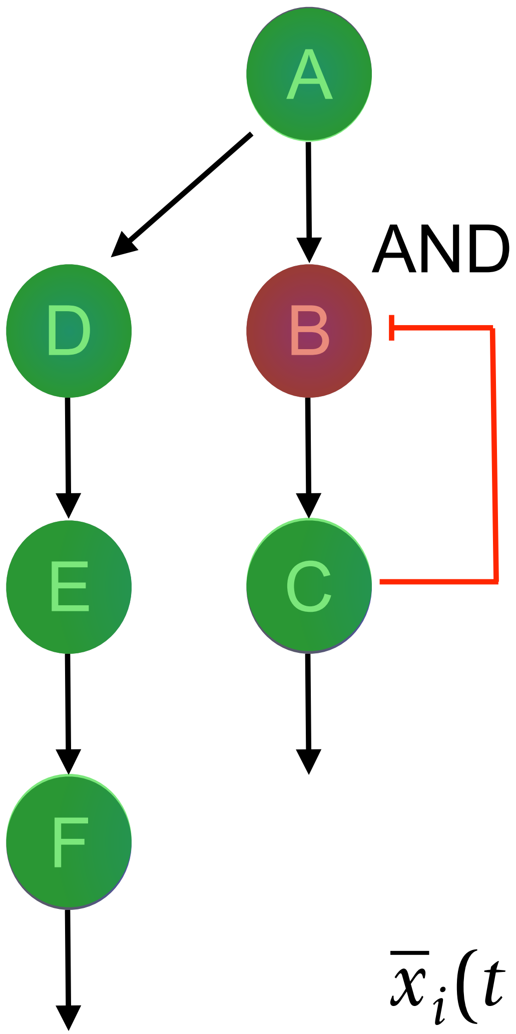
Simplest Boolean simulation: Synchronous updates



System can reach steady state or oscillate

$$\overline{x}_i(t+1) = \overline{B}_i(\overline{x}_{i1}(t), \overline{x}_{i2}(t), \ldots, \overline{x}_{iN_i}(t)).$$

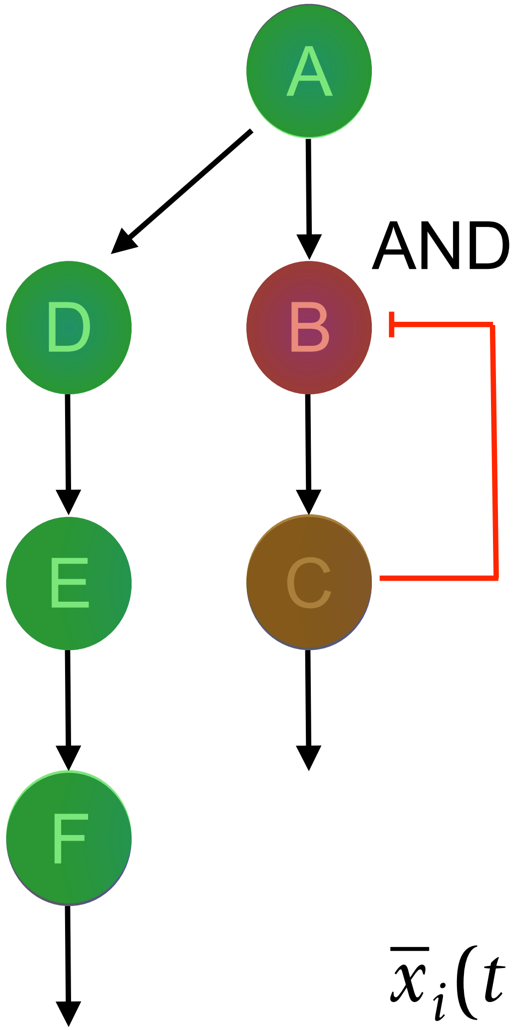
Simplest Boolean simulation: Synchronous updates



System can reach steady state or oscillate

$$\overline{x}_i(t+1) = \overline{B}_i(\overline{x}_{i1}(t), \overline{x}_{i2}(t), \ldots, \overline{x}_{iN_i}(t)).$$

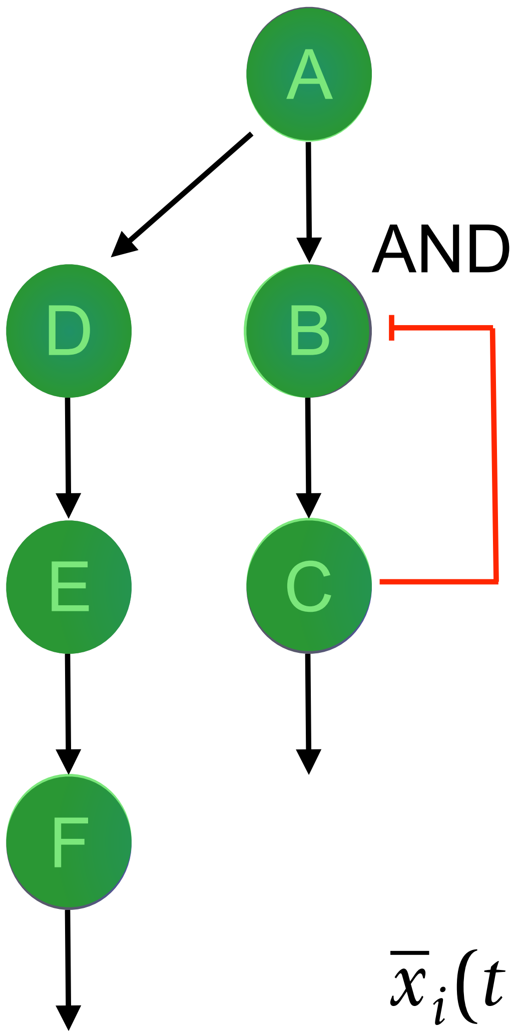
Simplest Boolean simulation: Synchronous updates



System can reach steady state or oscillate

$$\overline{x}_i(t+1) = \overline{B}_i(\overline{x}_{i1}(t), \overline{x}_{i2}(t), \ldots, \overline{x}_{iN_i}(t)).$$

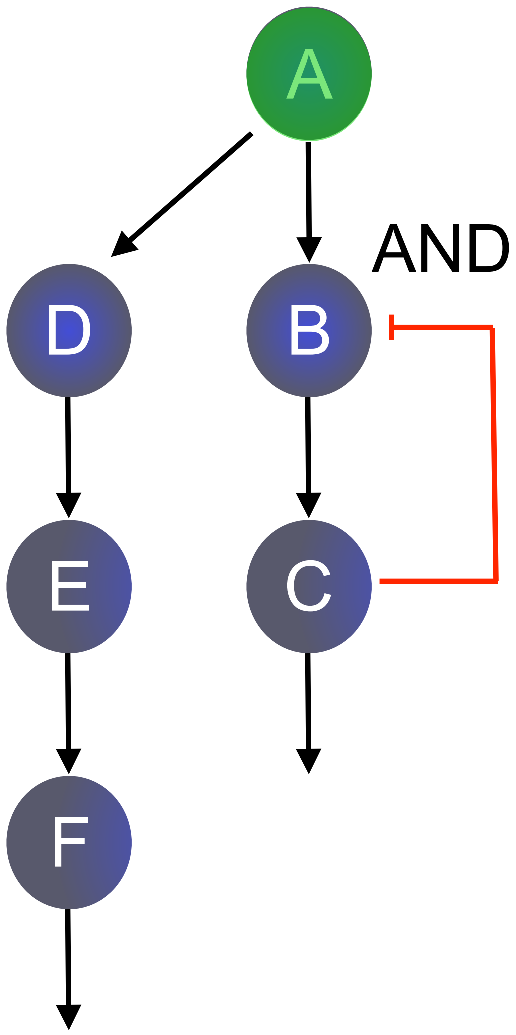
Simplest Boolean simulation: Synchronous updates



System can reach steady state or oscillate

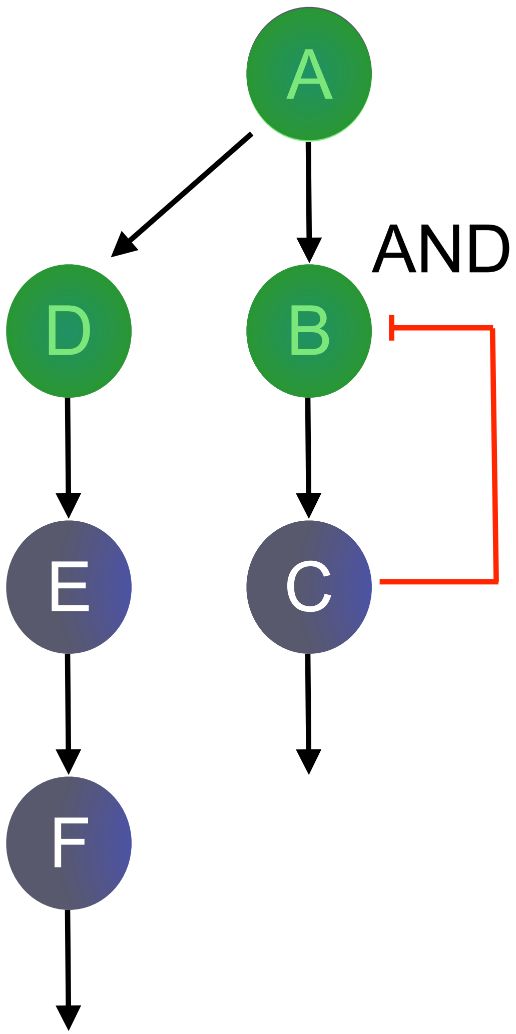
$$\overline{x}_i(t+1) = \overline{B}_i(\overline{x}_{i1}(t), \overline{x}_{i2}(t), \ldots, \overline{x}_{iN_i}(t)).$$

Boolean simulation performed using pseudo-steady state



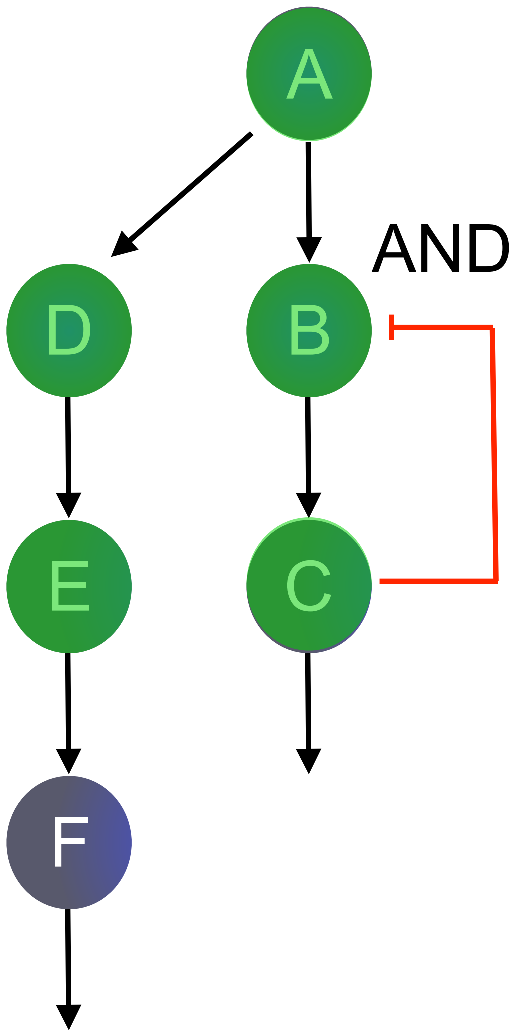
Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

Boolean simulation performed using pseudo-steady state



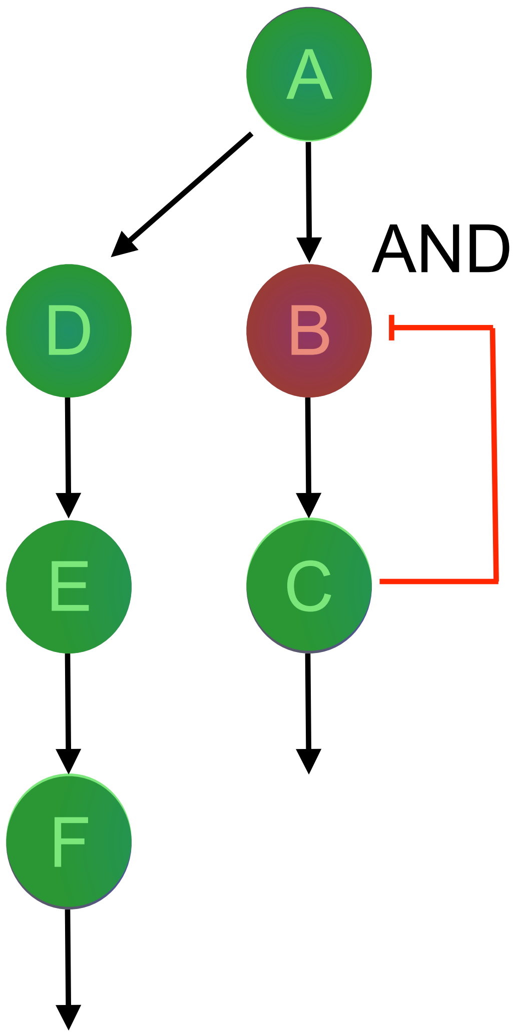
Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

Boolean simulation performed using pseudo-steady state



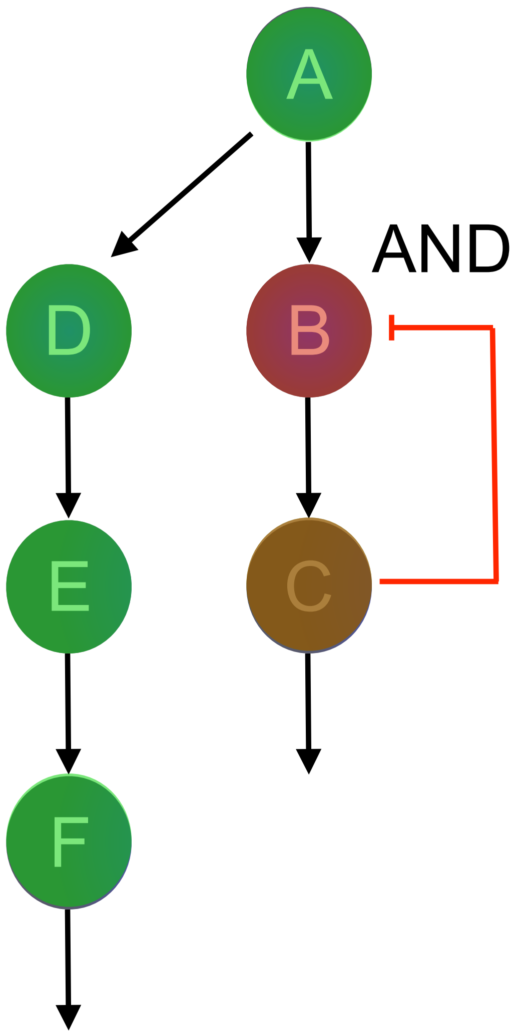
Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

Boolean simulation performed using pseudo-steady state



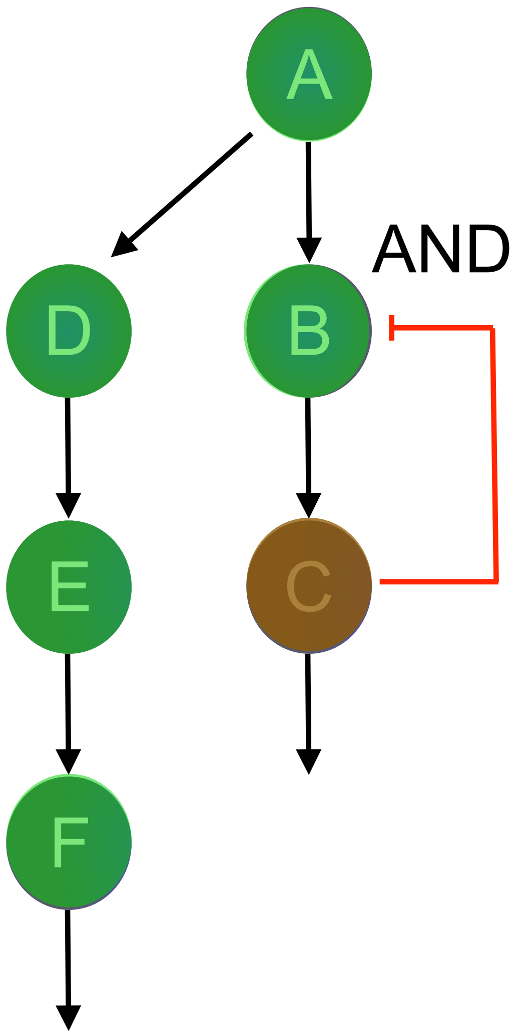
Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

Boolean simulation performed using pseudo-steady state



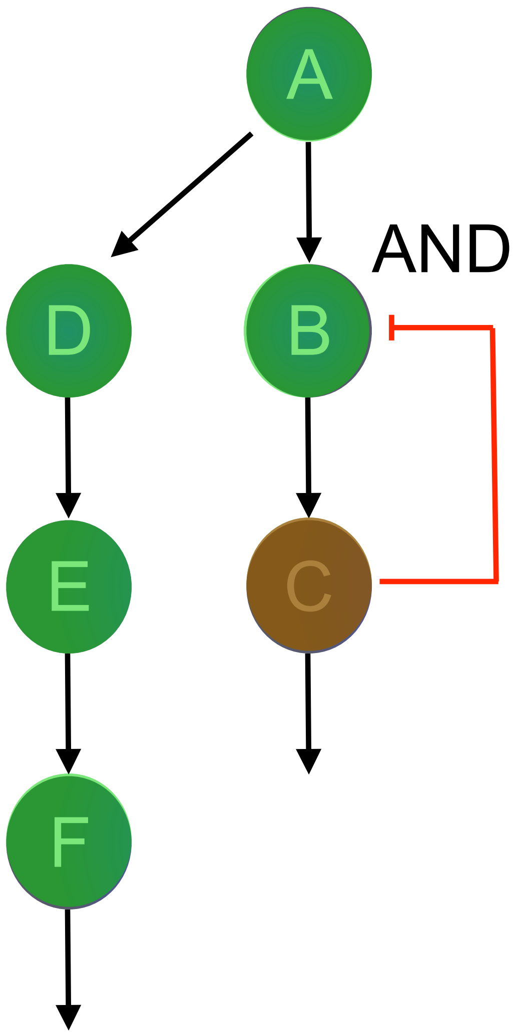
Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

Boolean simulation performed using pseudo-steady state



Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

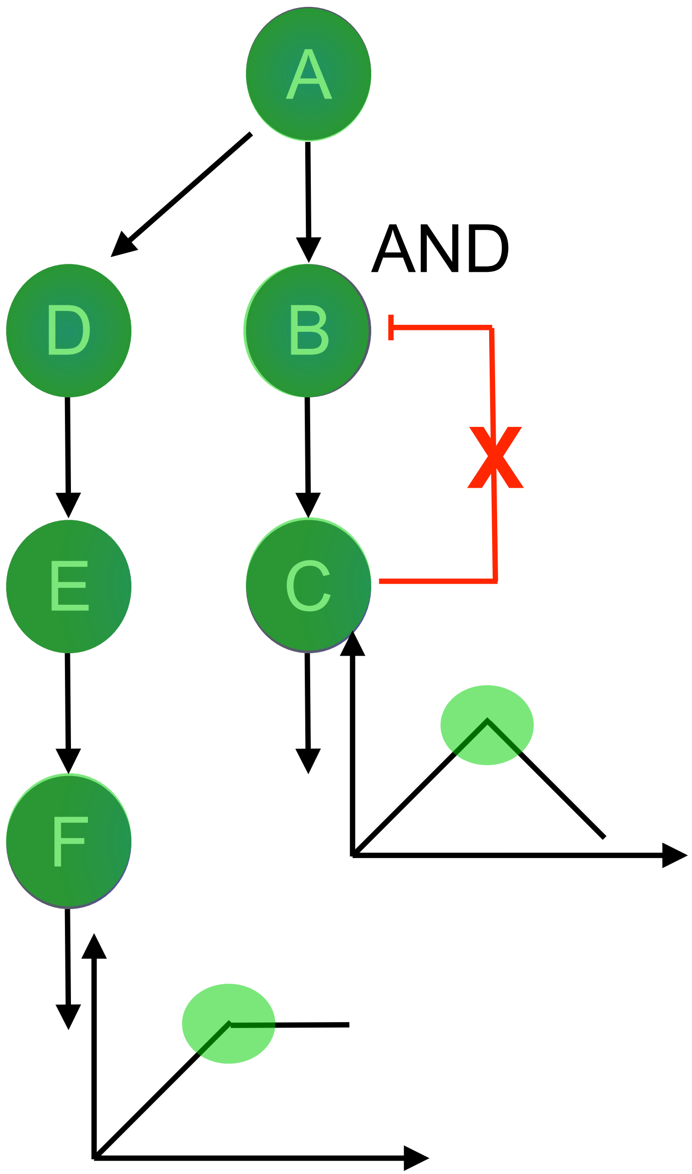
Boolean simulation performed using pseudo-steady state



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

Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

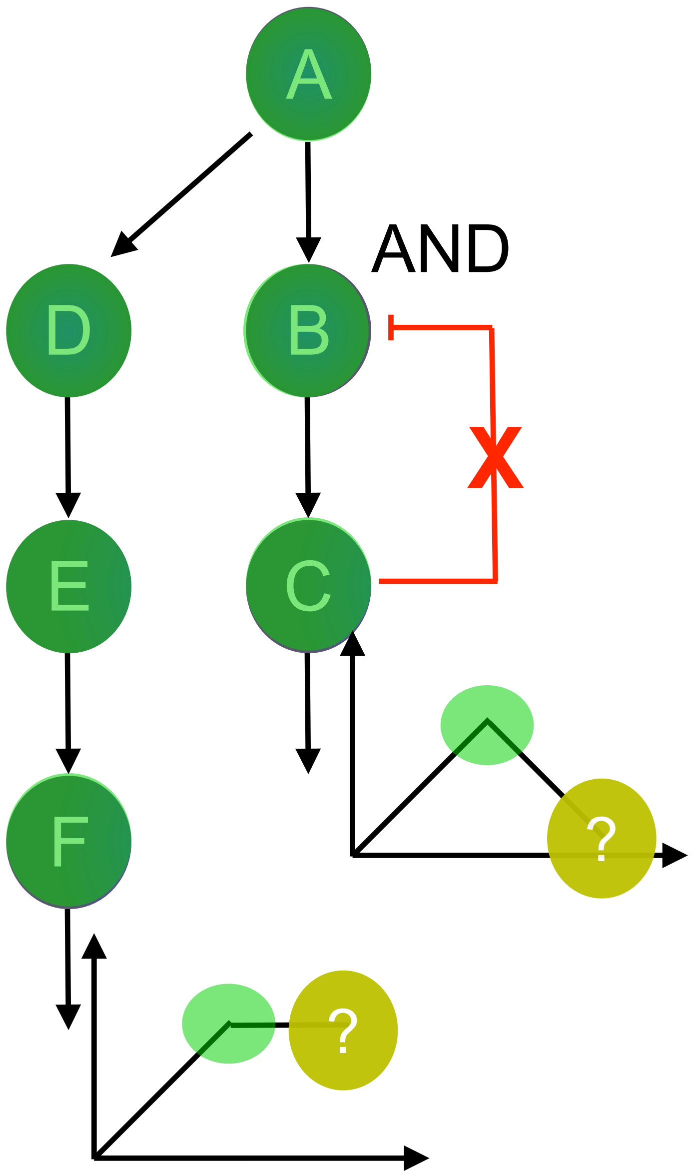
Boolean simulation performed using pseudo-steady state



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

Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

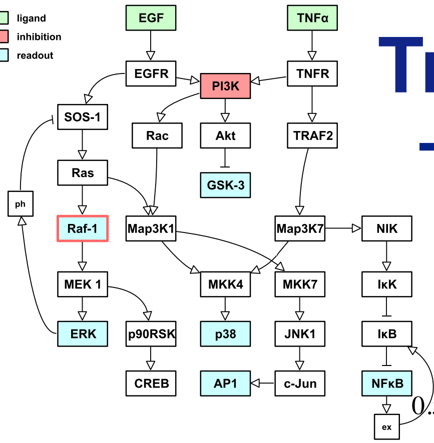
Boolean simulation performed using pseudo-steady state

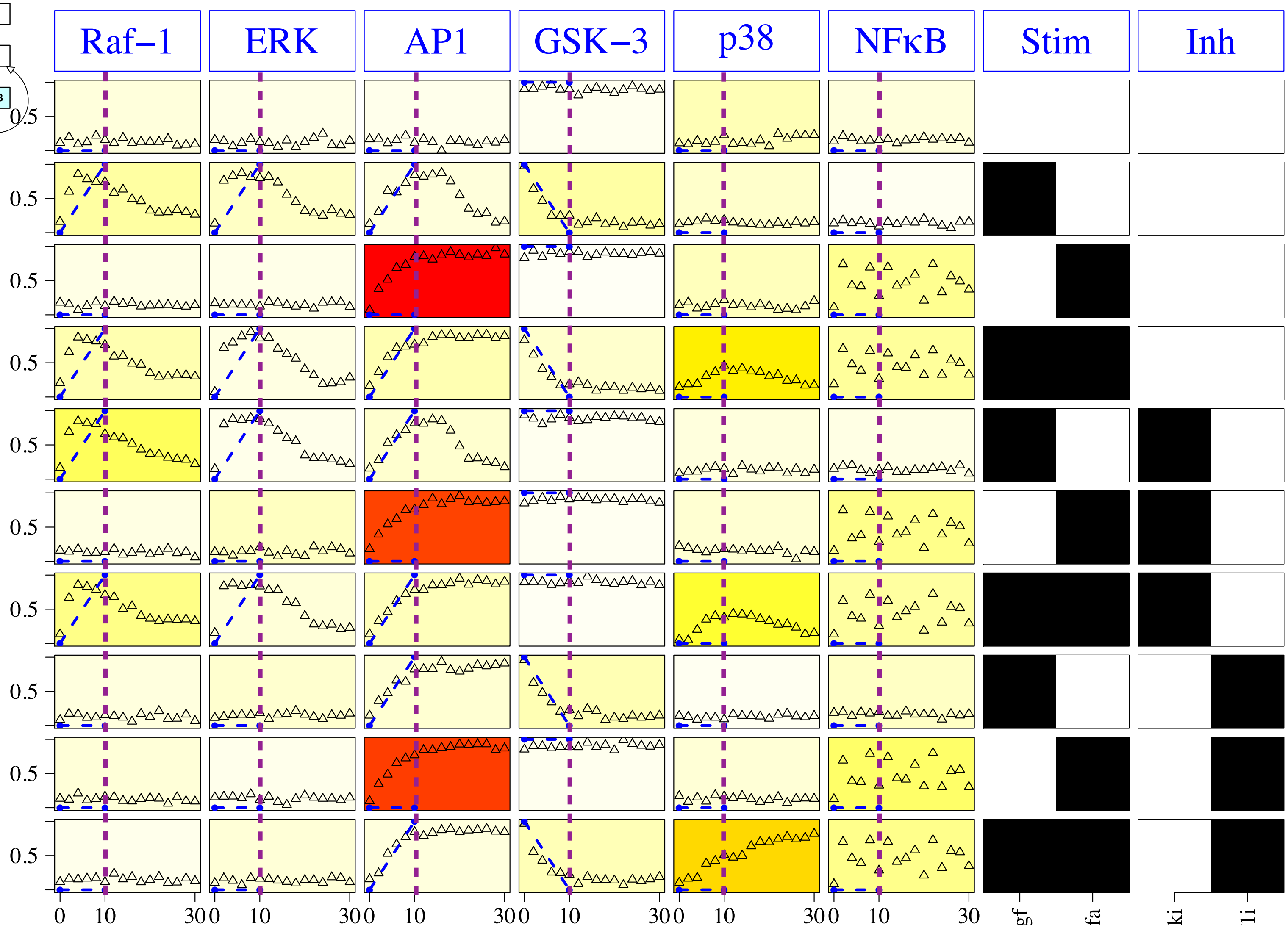


Algorithm penalizes lack of steady state, only effective for one ‘early’ time

Klamt, S., Saez-Rodriguez, J., Lindquist, J., Simeoni, L., Gilles, E. D., *BMC Bioinformatics,* 7:56,2006.

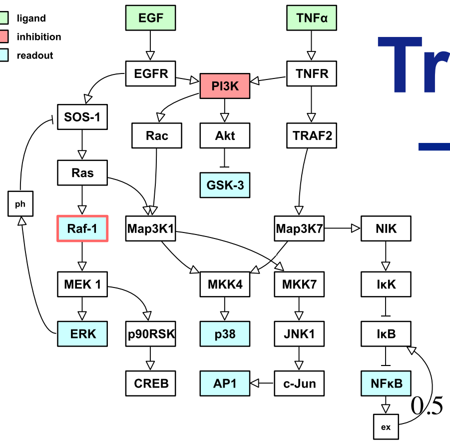
Training to data recovers structure



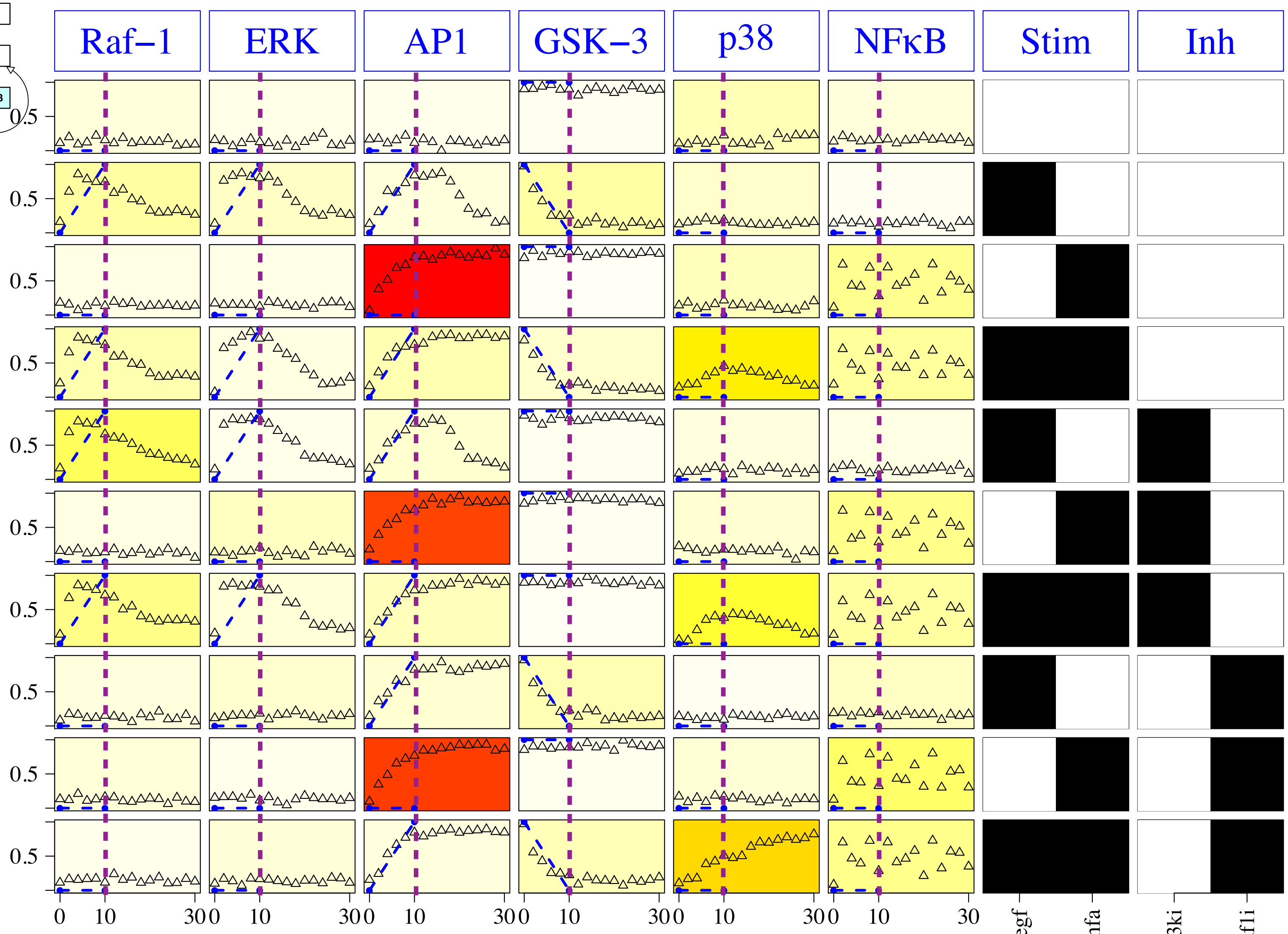


MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Training to data recovers structure

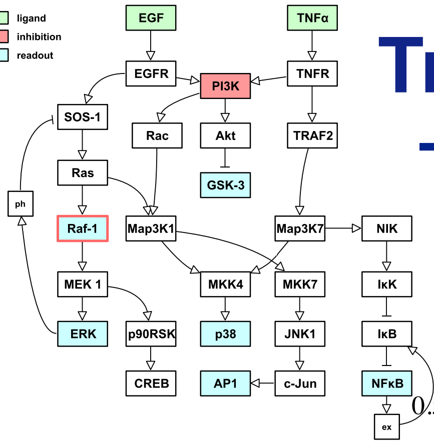


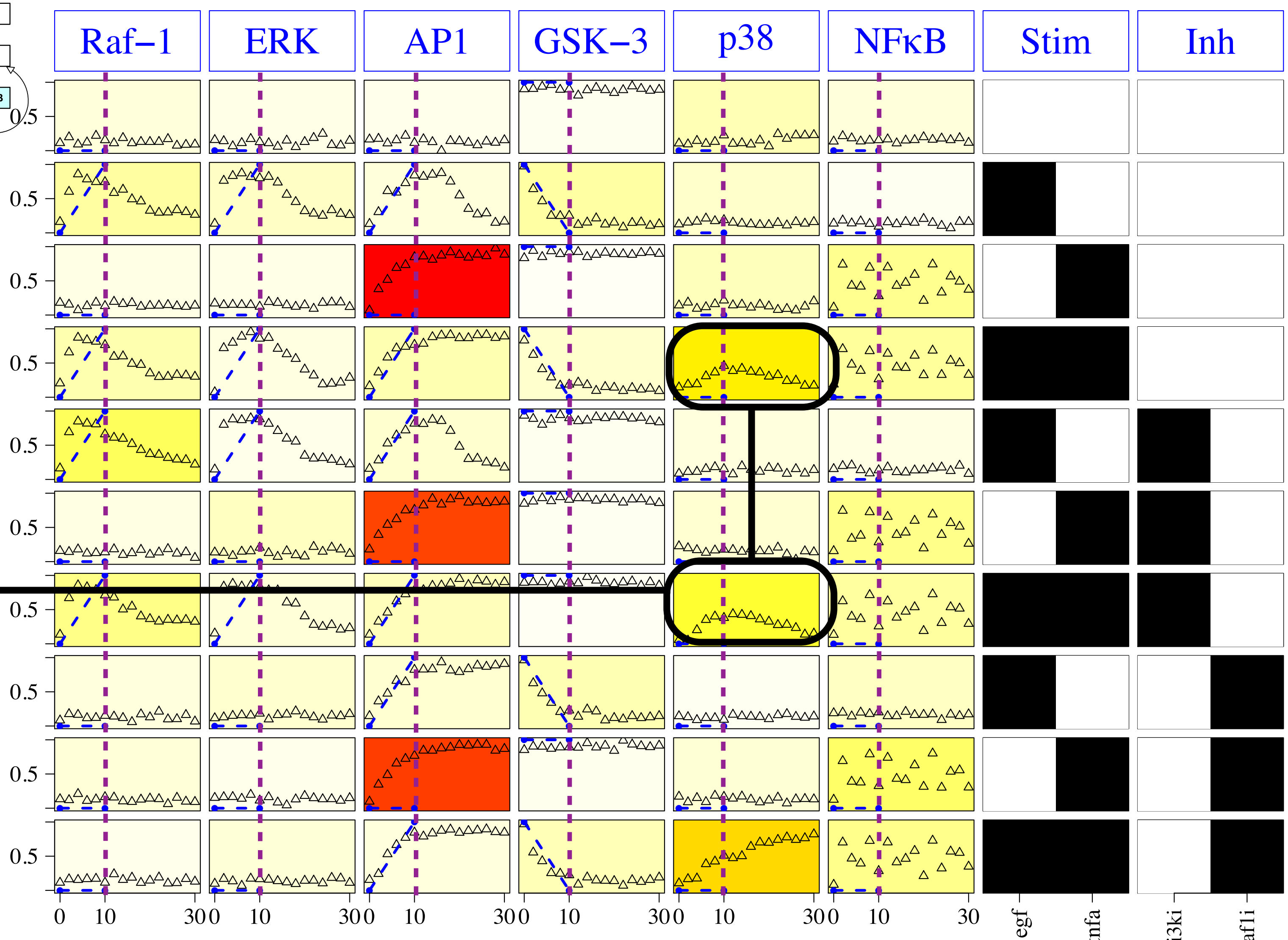
Identifies strong active links (except feedbacks)



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Training to data recovers structure



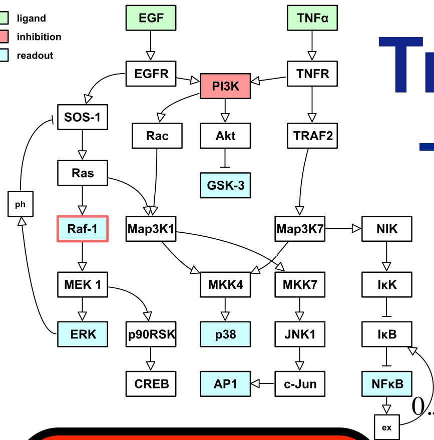


Identifies strong active links (except feedbacks)

Does not identify weak effects

MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

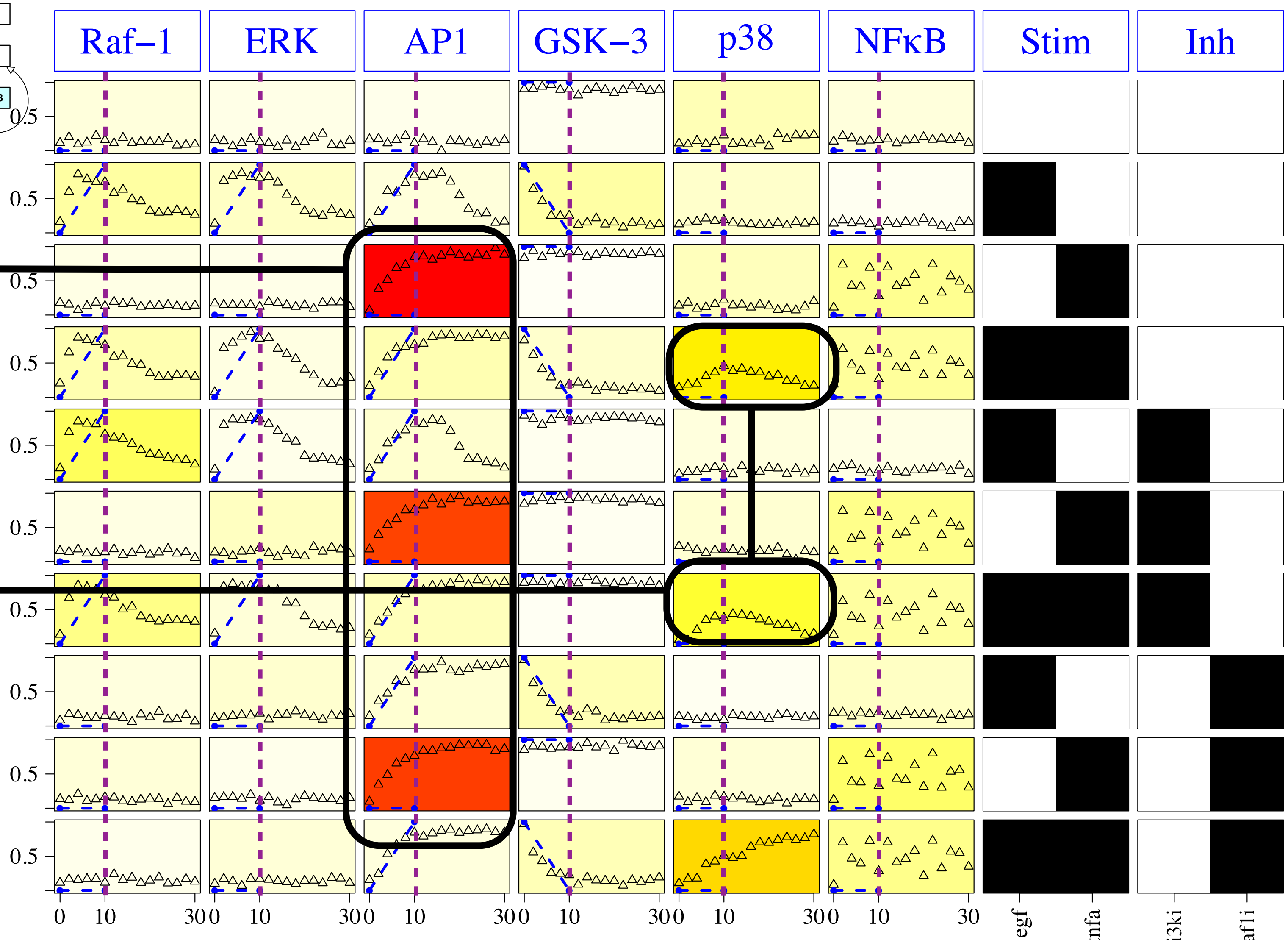
Training to data recovers structure



Can not explain data due to missing links

Identifies strong active links (except feedbacks)

Does not identify weak effects



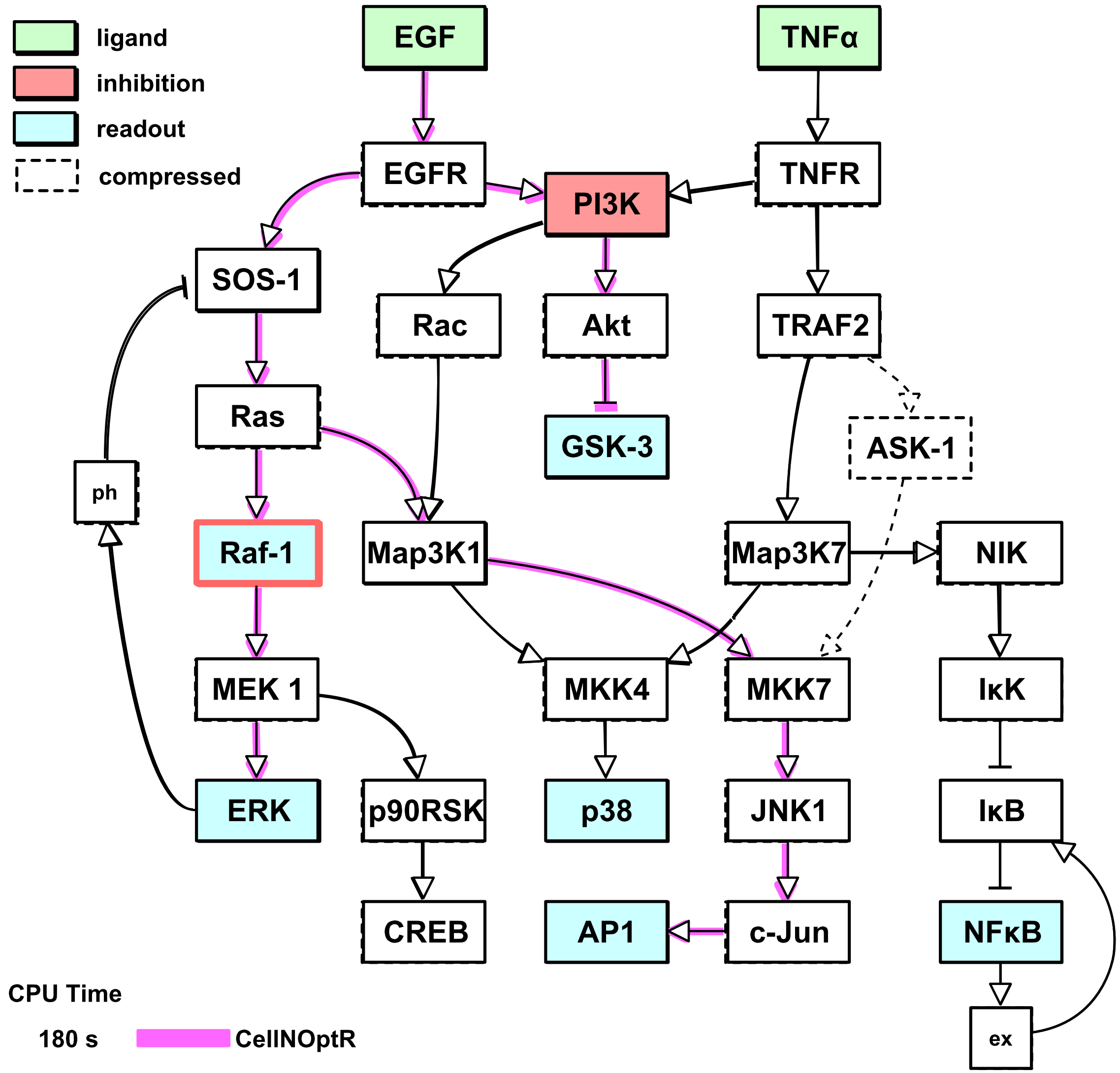
MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Boolean 1-steady state recovers strong active links except feedback

Can not explain data due to missing links

Identifies strong active links (except feedbacks)

Does not identify weak effects



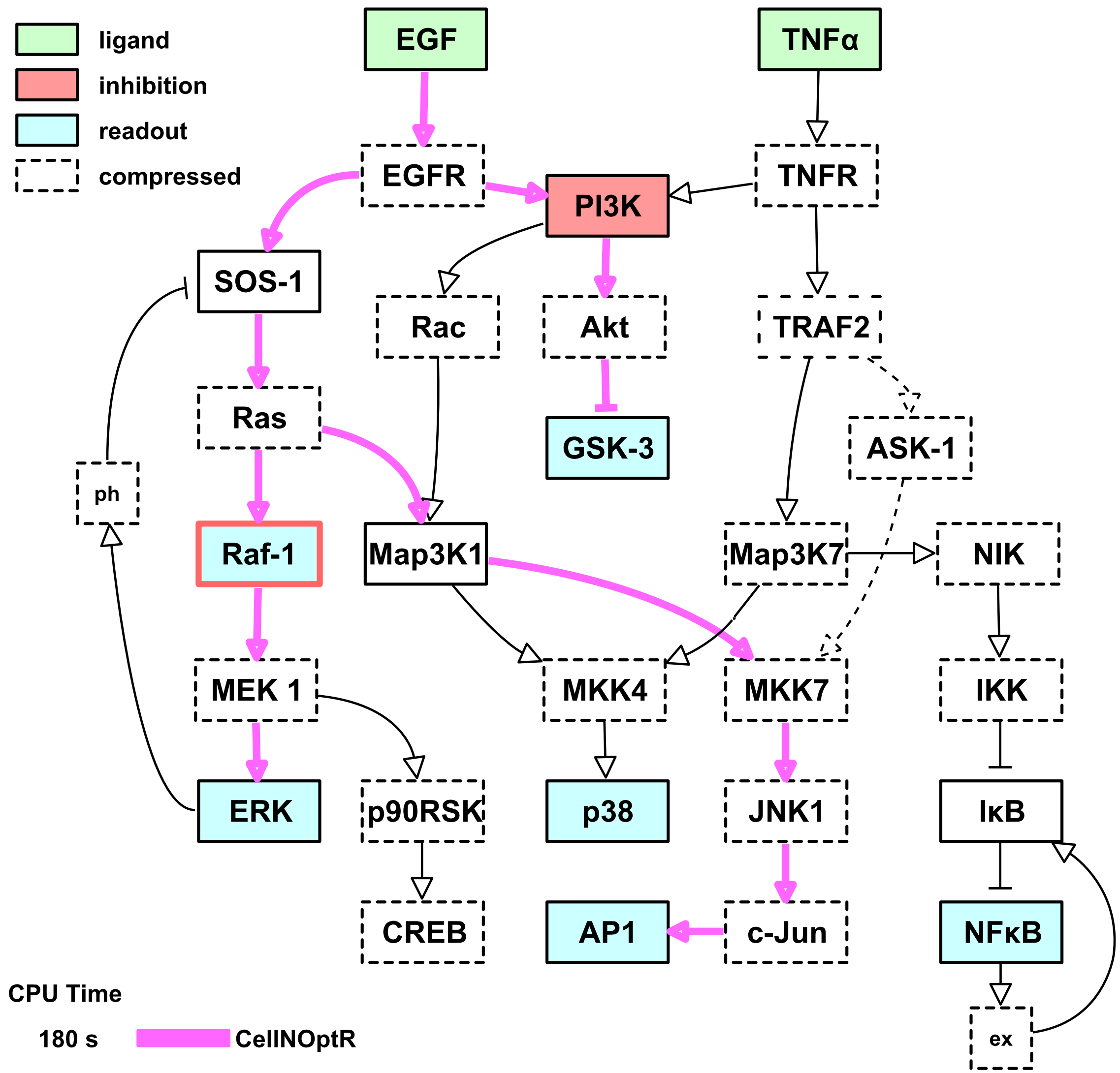
MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Boolean 1-steady state recovers strong active links except feedback

Can not explain data due to missing links

Identifies strong active links (except feedbacks)

Does not identify weak effects



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Constrained Fuzzy Logic can handle quantitative differences

- Boolean modeling can **not** describe **quantitative** aspects (e.g. intermediate activation)

Morris MK, Saez-Rodriguez J, Clarke D, Sorger PK, Lauffenburger DA, PloS Comp Bio 2011

Constrained Fuzzy Logic can handle quantitative differences

- Boolean modeling can **not** describe **quantitative** aspects (e.g. intermediate activation)

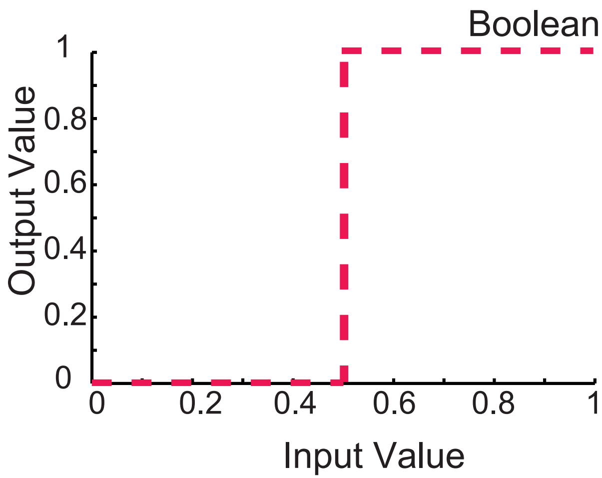
- Fuzzy logic can model quantitative signaling data (Aldridge et al. *Plos Comp. Bio.* 2009)

Morris MK, Saez-Rodriguez J, Clarke D, Sorger PK, Lauffenburger DA, PloS Comp Bio 2011

Constrained Fuzzy Logic can handle quantitative differences

- Boolean modeling can **not** describe **quantitative** aspects (e.g. intermediate activation)

- Fuzzy logic can model quantitative signaling data (Aldridge et al. *Plos Comp. Bio.* 2009)

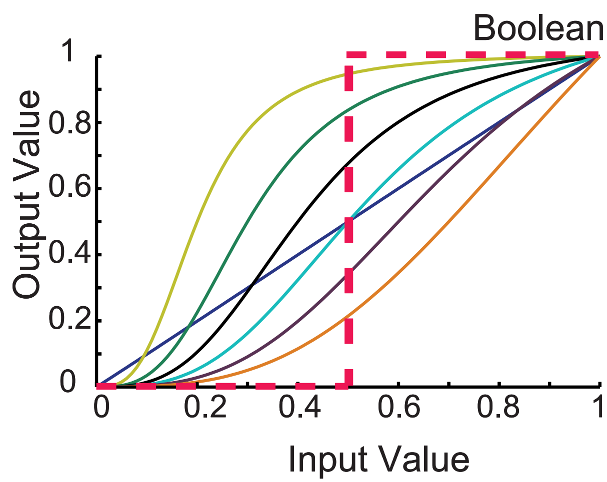


Morris MK, Saez-Rodriguez J, Clarke D, Sorger PK, Lauffenburger DA, PloS Comp Bio 2011

Constrained Fuzzy Logic can handle quantitative differences

- Boolean modeling can **not** describe **quantitative** aspects (e.g. intermediate activation)

- Fuzzy logic can model quantitative signaling data (Aldridge et al. *Plos Comp. Bio.* 2009)

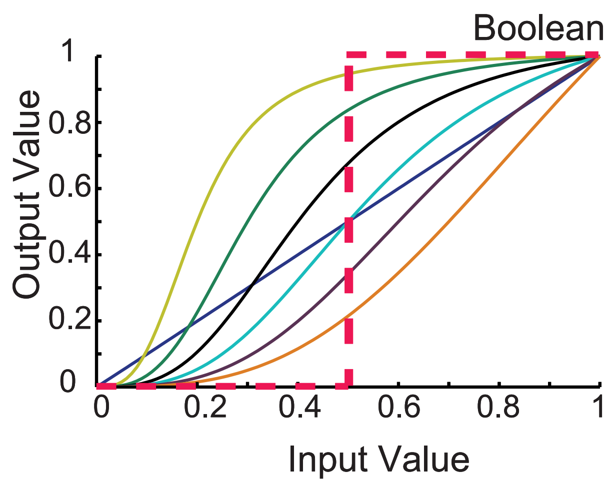


Morris MK, Saez-Rodriguez J, Clarke D, Sorger PK, Lauffenburger DA, PloS Comp Bio 2011

Constrained Fuzzy Logic can handle quantitative differences

- Boolean modeling can **not** describe **quantitative** aspects (e.g. intermediate activation)

- Fuzzy logic can model quantitative signaling data (Aldridge et al. *Plos Comp. Bio.* 2009)



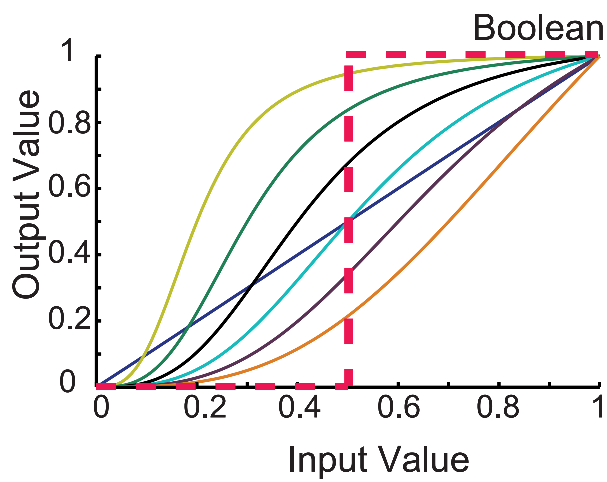
- **Constrained Fuzzy Logic**

Morris MK, Saez-Rodriguez J, Clarke D, Sorger PK, Lauffenburger DA, PloS Comp Bio 2011

Constrained Fuzzy Logic can handle quantitative differences

- Boolean modeling can **not** describe **quantitative** aspects (e.g. intermediate activation)

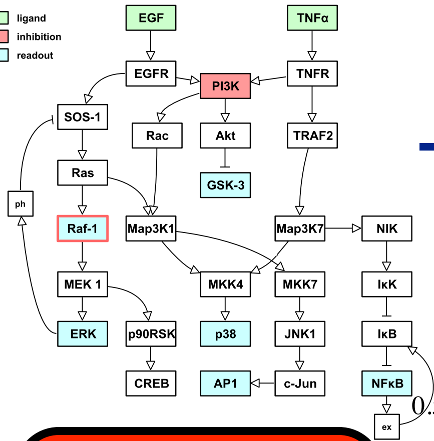
- Fuzzy logic can model quantitative signaling data (Aldridge et al. *Plos Comp. Bio.* 2009)



- **Constrained Fuzzy Logic**

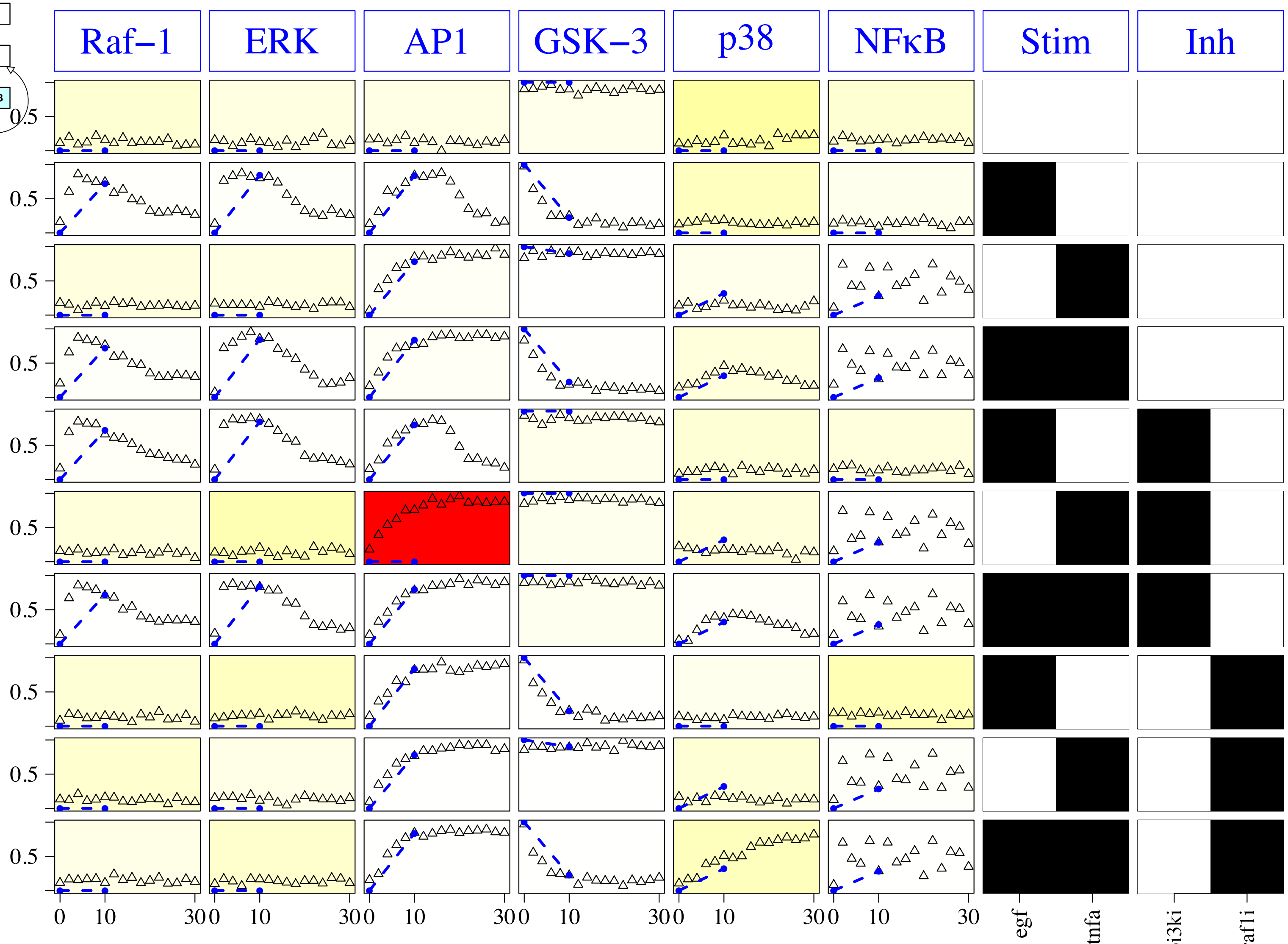
Morris MK, Saez-Rodriguez J, Clarke D, Sorger PK, Lauffenburger DA, PloS Comp Bio 2011

fit of fuzzy toy model



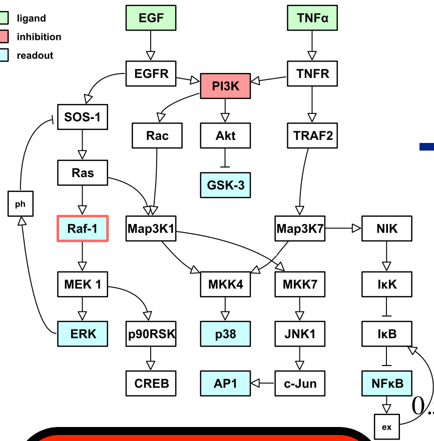
Can not explain data due to missing links

Identifies strong active links (except feedbacks)



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

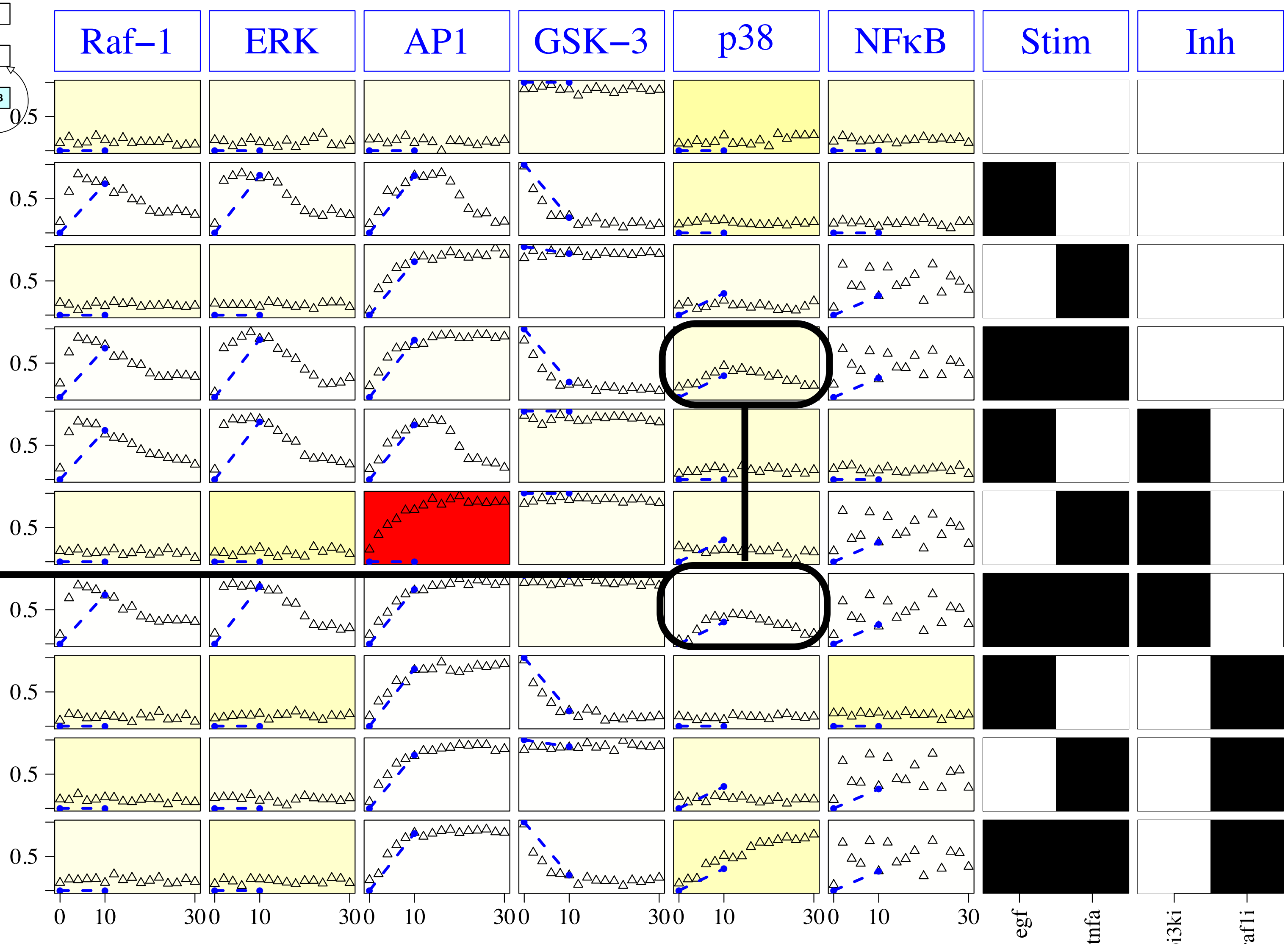
fit of fuzzy toy model



Can not explain data due to missing links

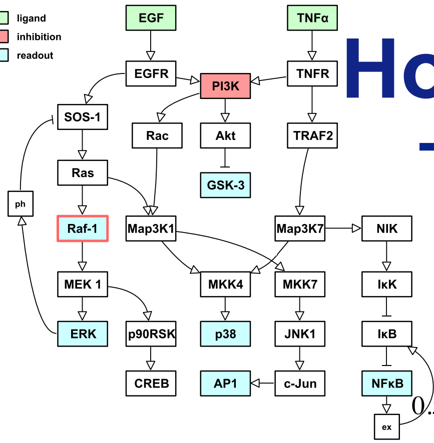
Identifies strong active links (except feedbacks)

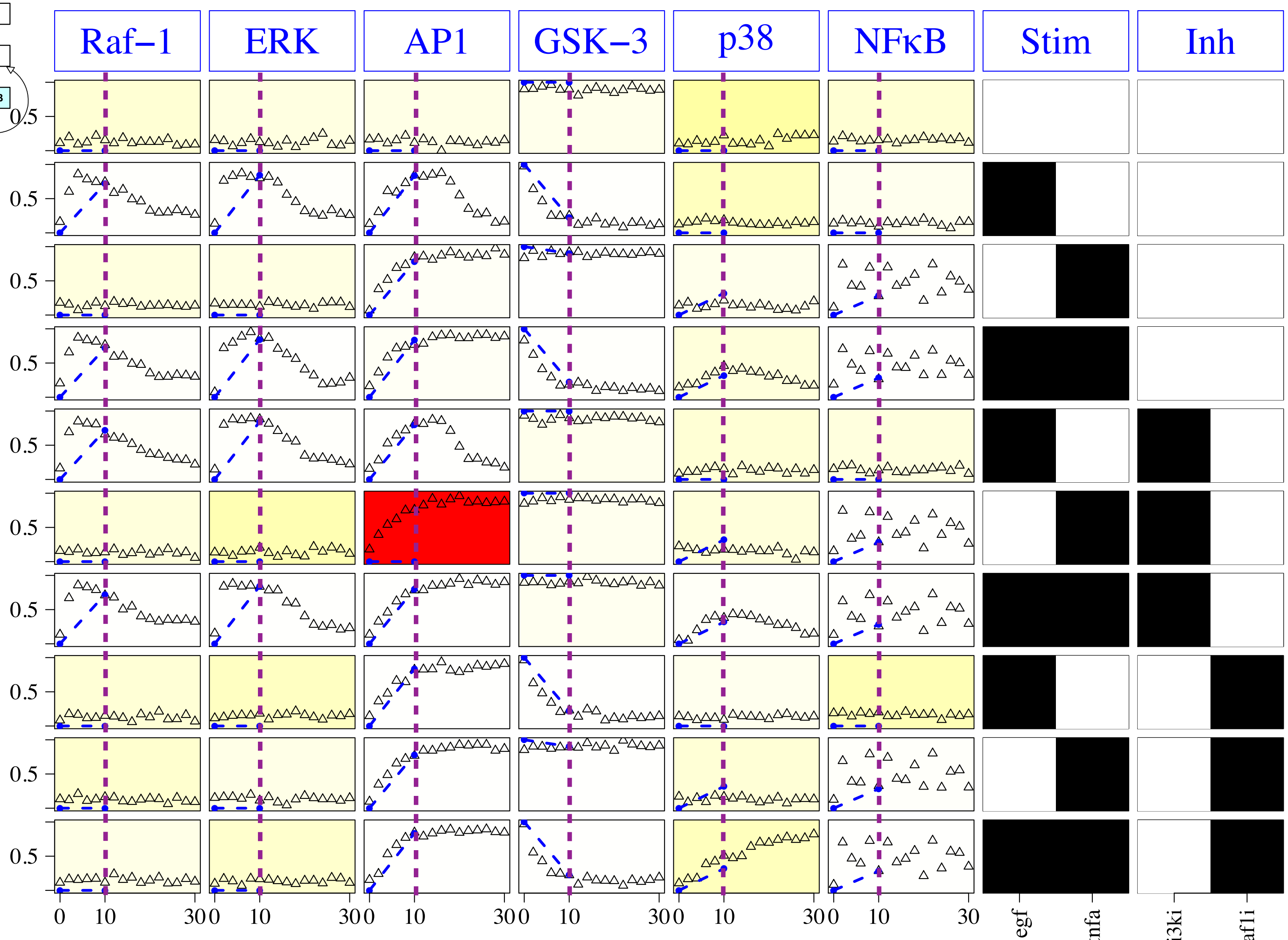
Identifies weak active links (except feedbacks)



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

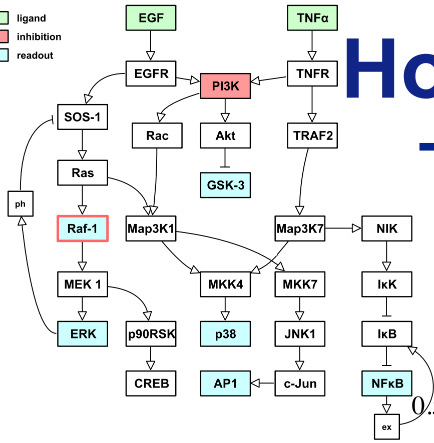
How to model feedback effects?

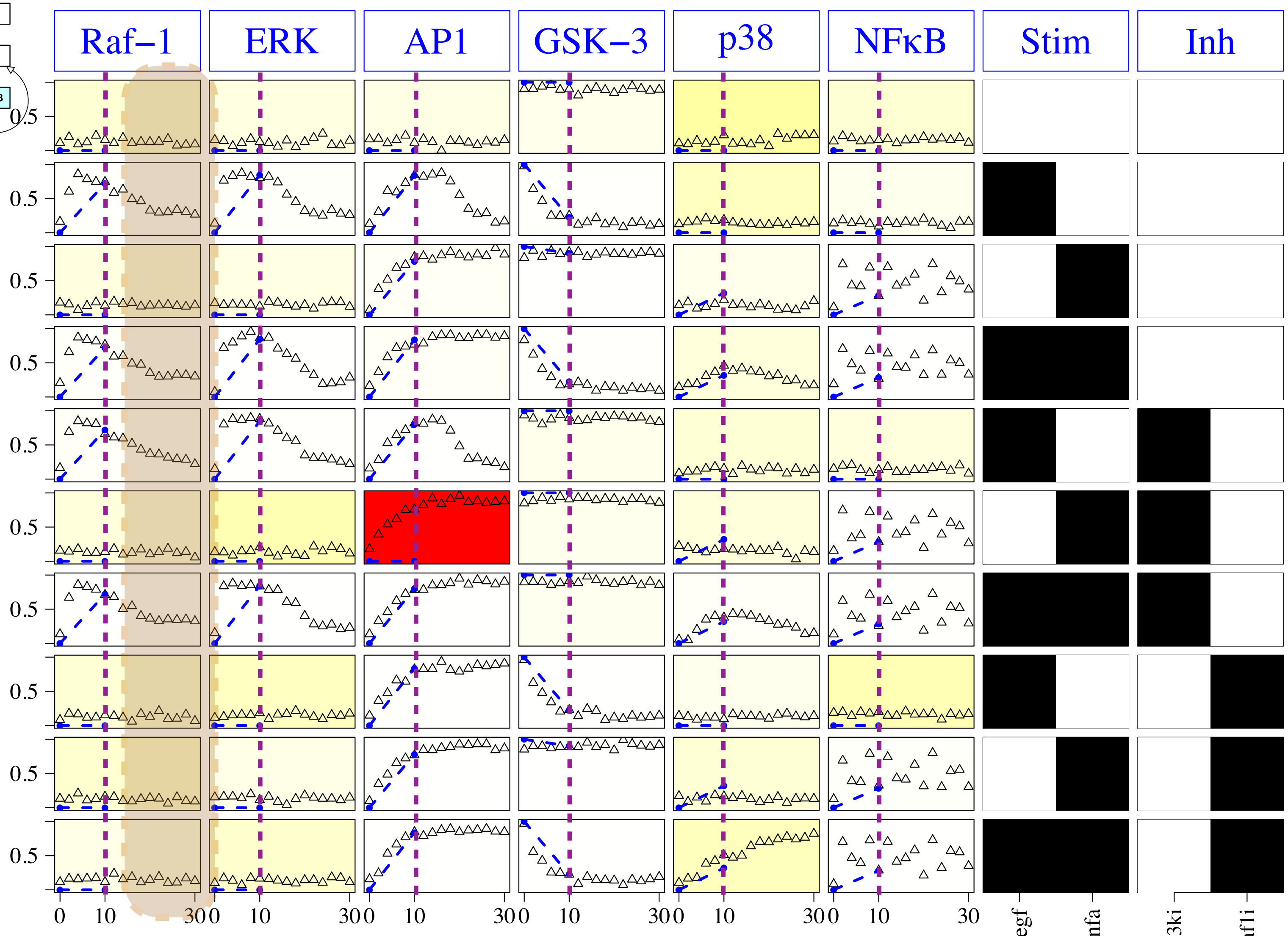




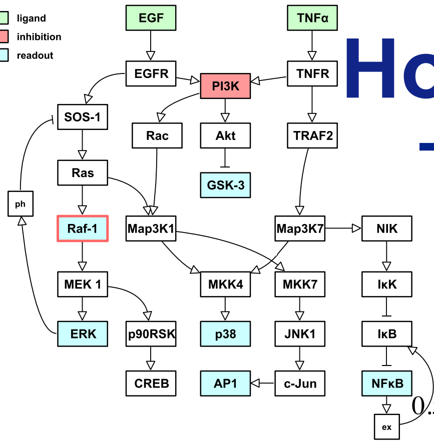
MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

How to model feedback effects?

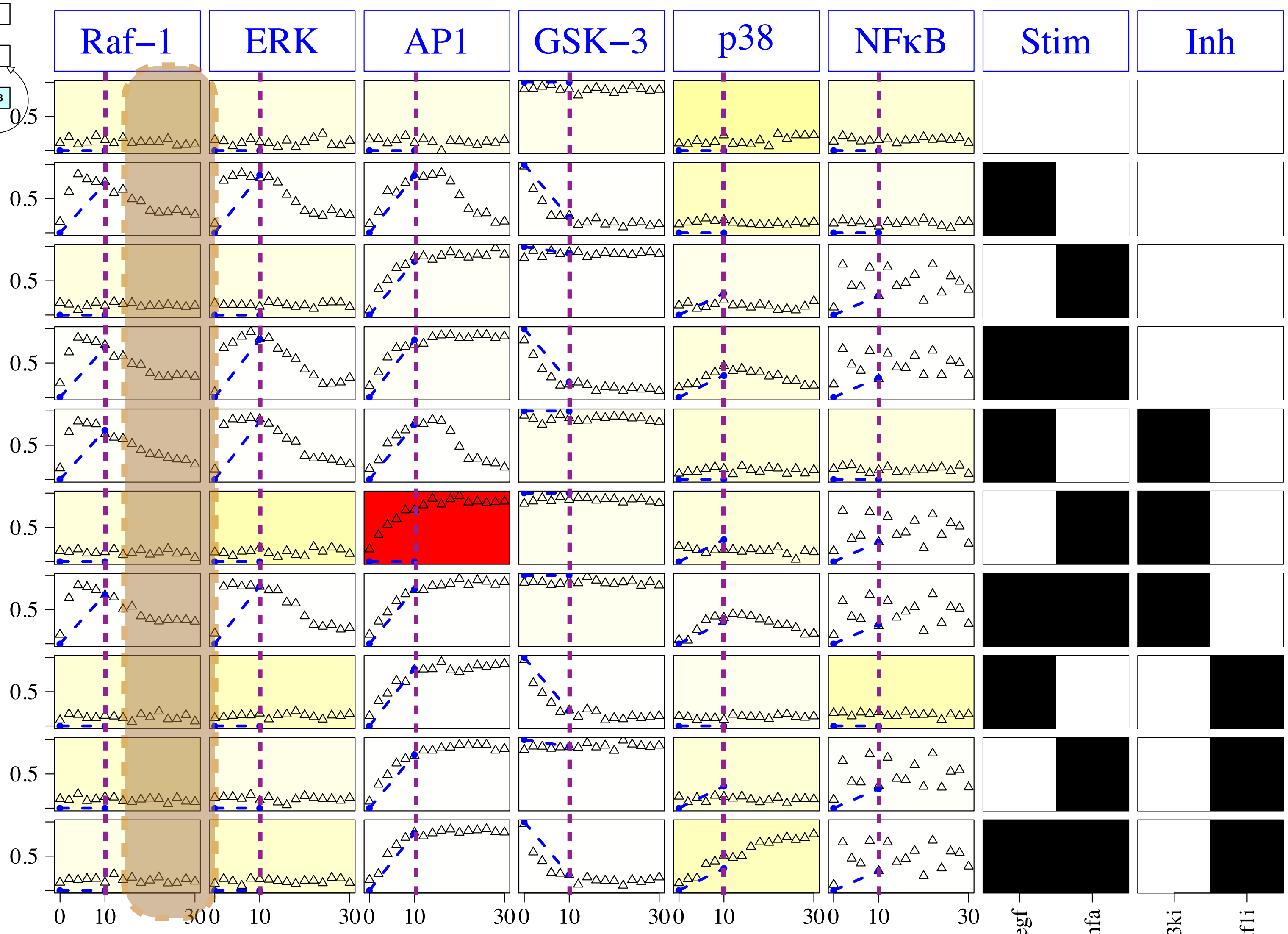




MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

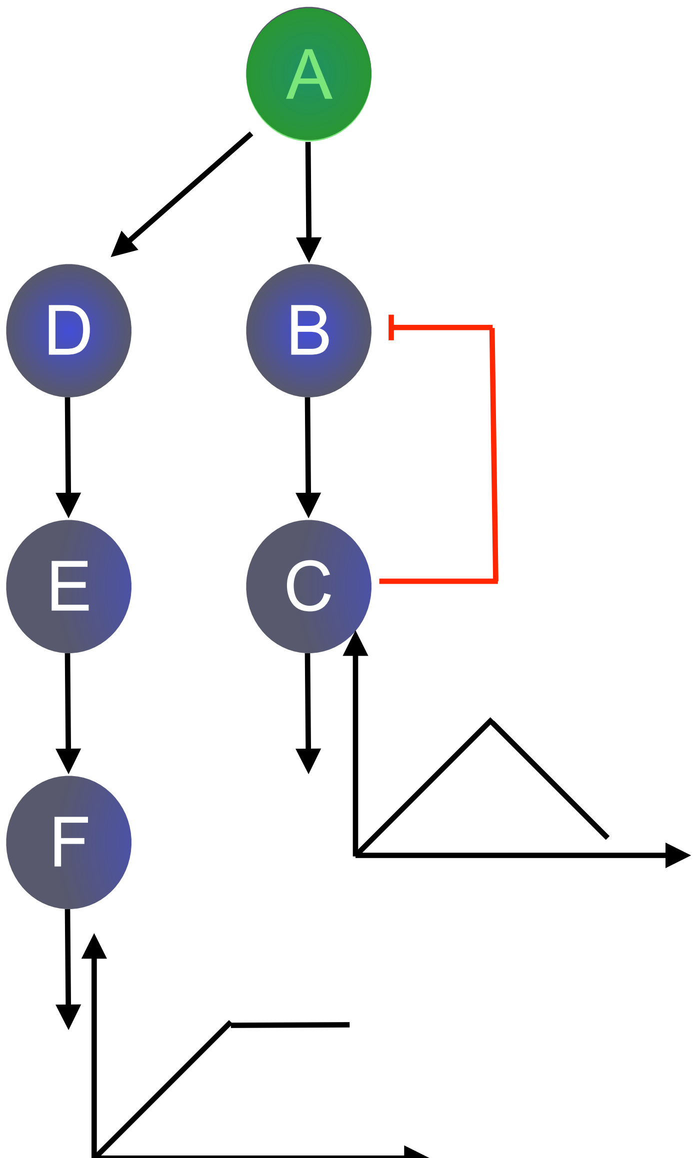


How to model feedback effects?



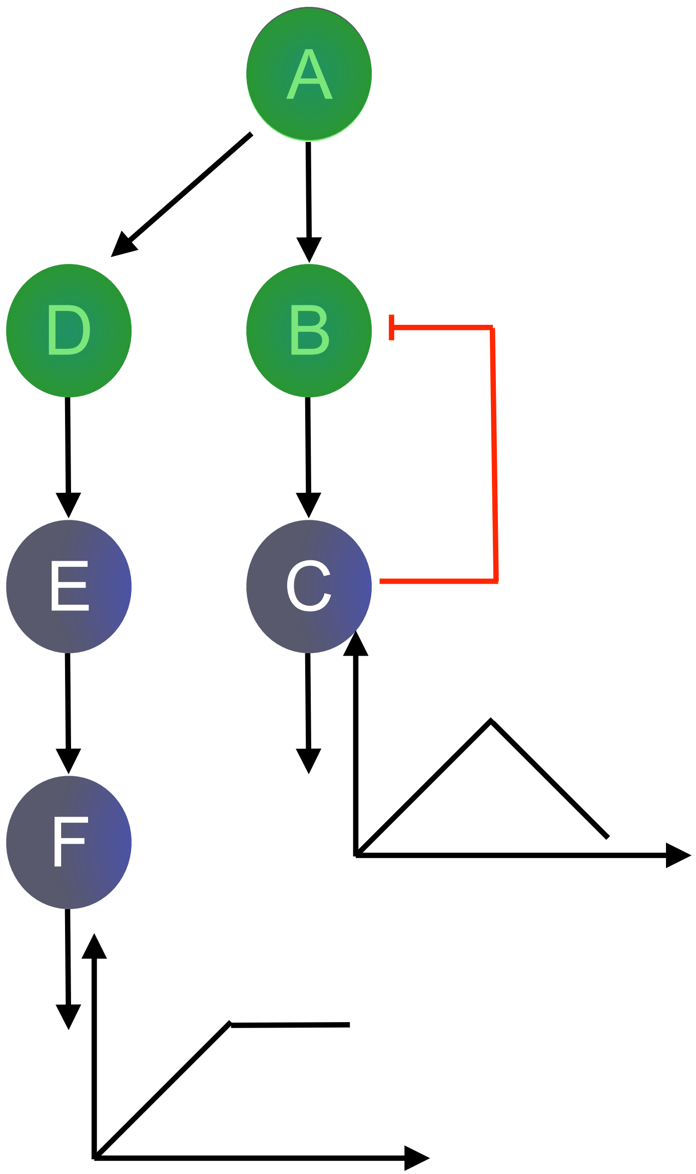
MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Approximation of transient behaviour using multiple time-scales



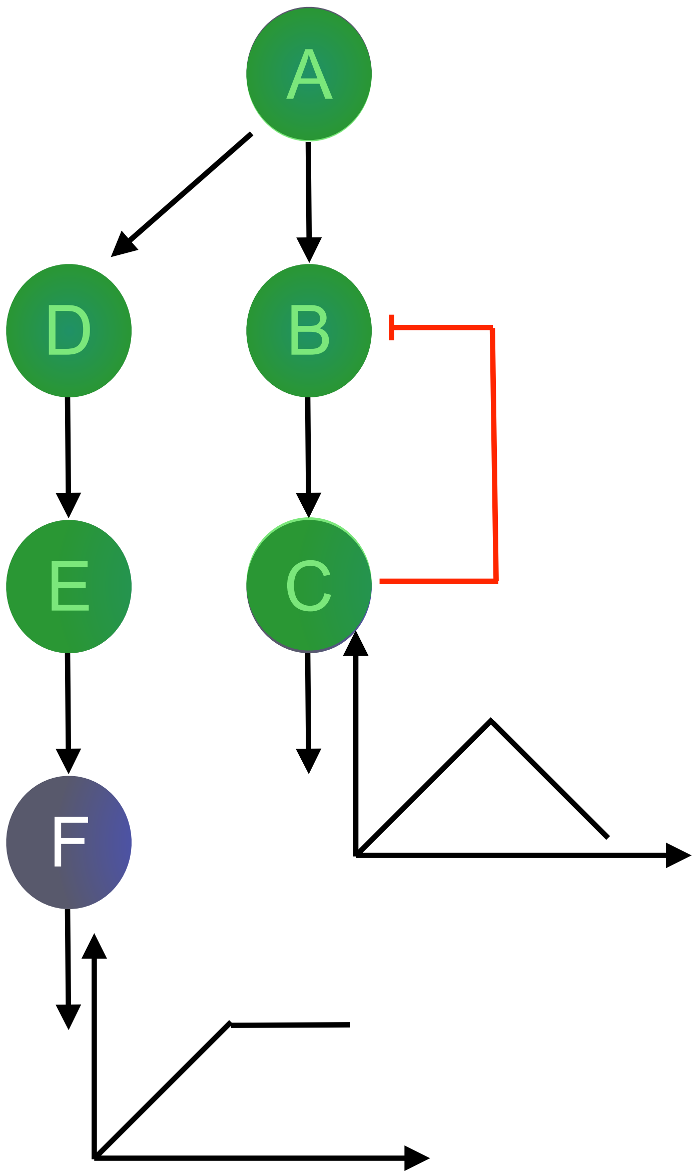
- (i) Train τ =1 → get early events
- (ii) Train τ =2 → find gates not active at τ =1 that explain evolution from τ =1 to τ =2

Approximation of transient behaviour using multiple time-scales



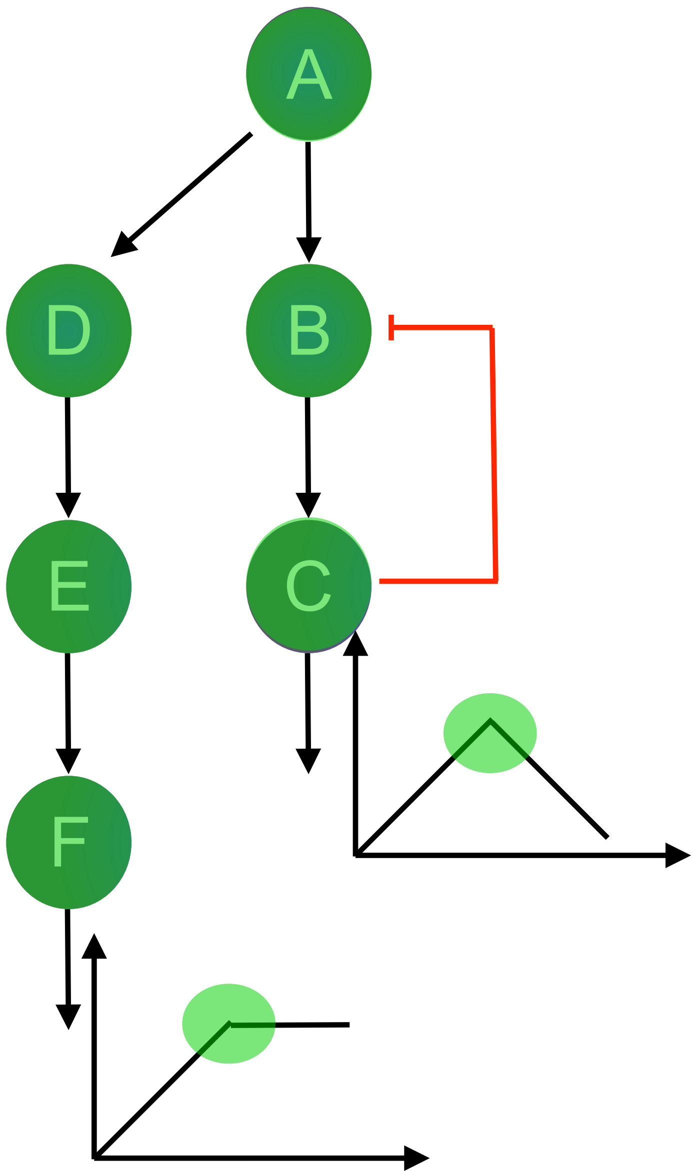
- (i) Train τ =1 → get early events
- (ii) Train τ =2 → find gates not active at τ =1 that explain evolution from τ =1 to τ =2

Approximation of transient behaviour using multiple time-scales



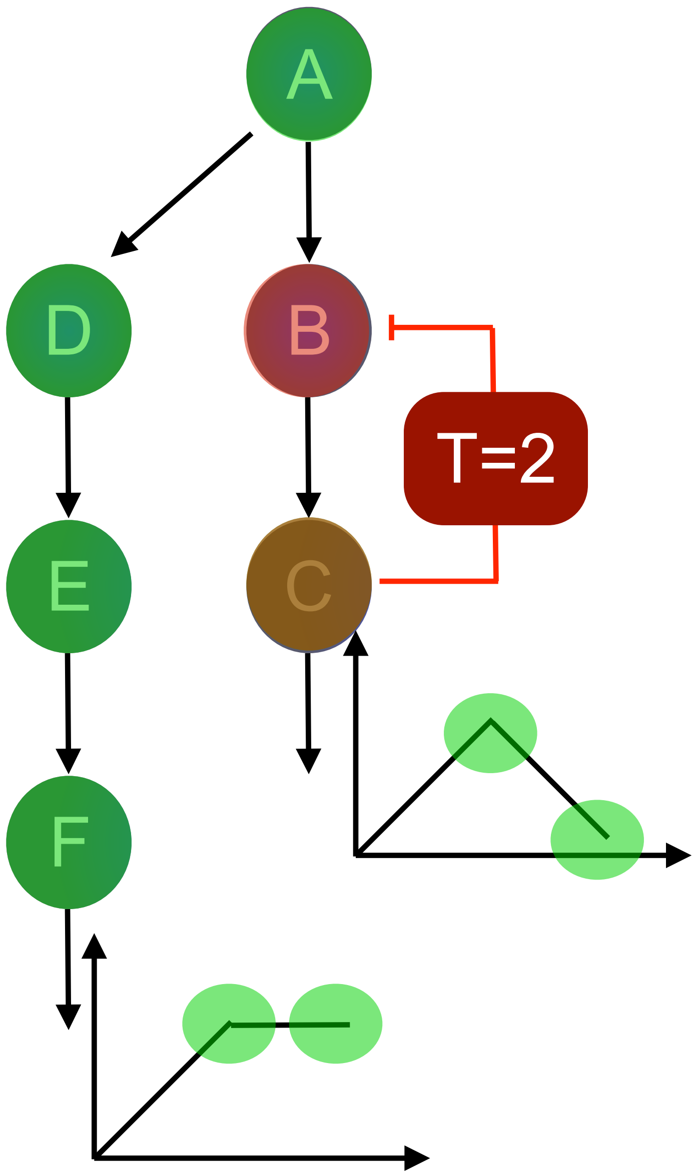
- (i) Train τ =1 → get early events
- (ii) Train τ =2 → find gates not active at τ =1 that explain evolution from τ =1 to τ =2

Approximation of transient behaviour using multiple time-scales



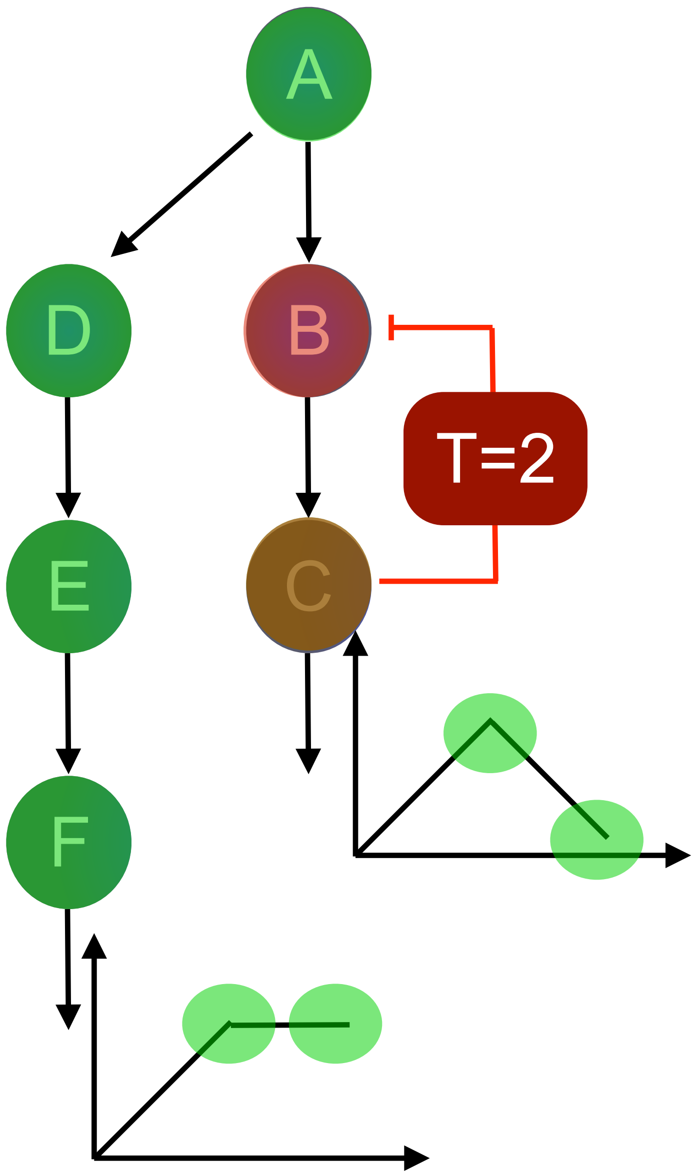
- (i) Train τ =1 → get early events
- (ii) Train τ =2 → find gates not active at τ =1 that explain evolution from τ =1 to τ =2

Approximation of transient behaviour using multiple time-scales



- (i) Train τ =1 → get early events
- (ii) Train τ =2 → find gates not active at τ =1 that explain evolution from τ =1 to τ =2

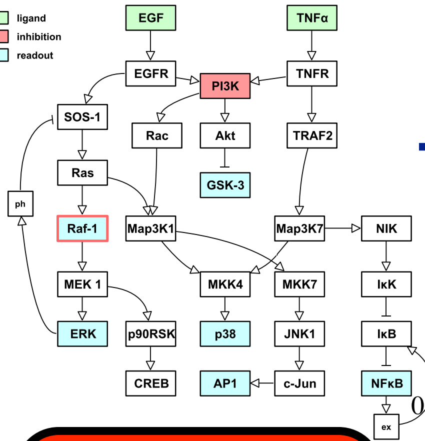
Approximation of transient behaviour using multiple time-scales



- (i) Train τ =1 → get early events
- (ii) Train τ =2 → find gates not active at τ =1 that explain evolution from τ =1 to τ =2

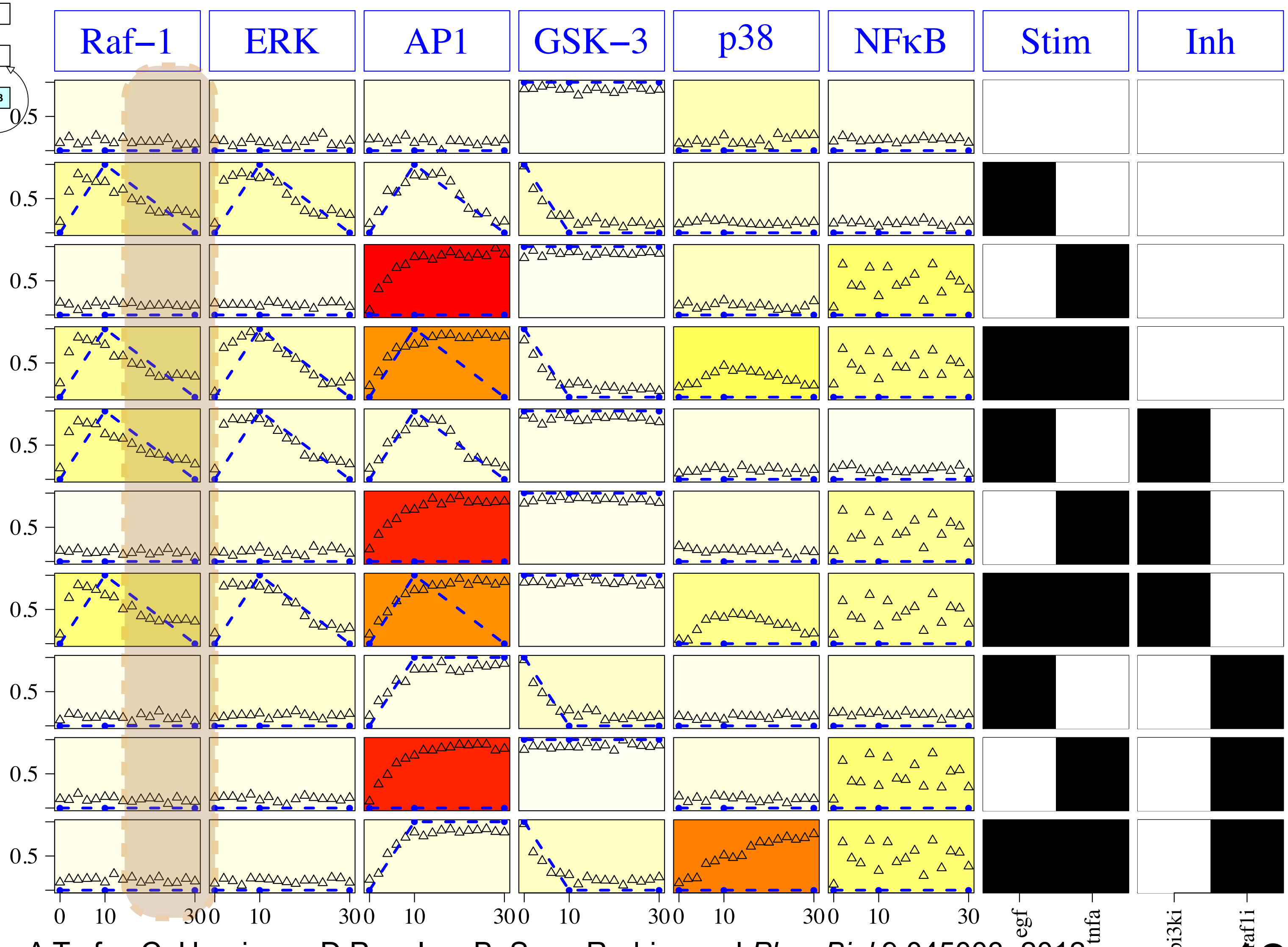
Rough approximation of dynamics, still computationally efficient

Multiple pseudo-steady-state captures feedbacks that lead to transient signals



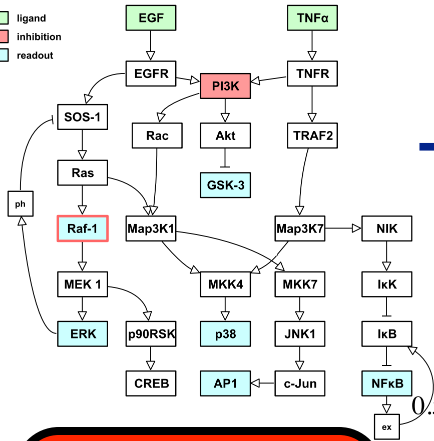
Can not explain data due to missing links

Identifies strong active links



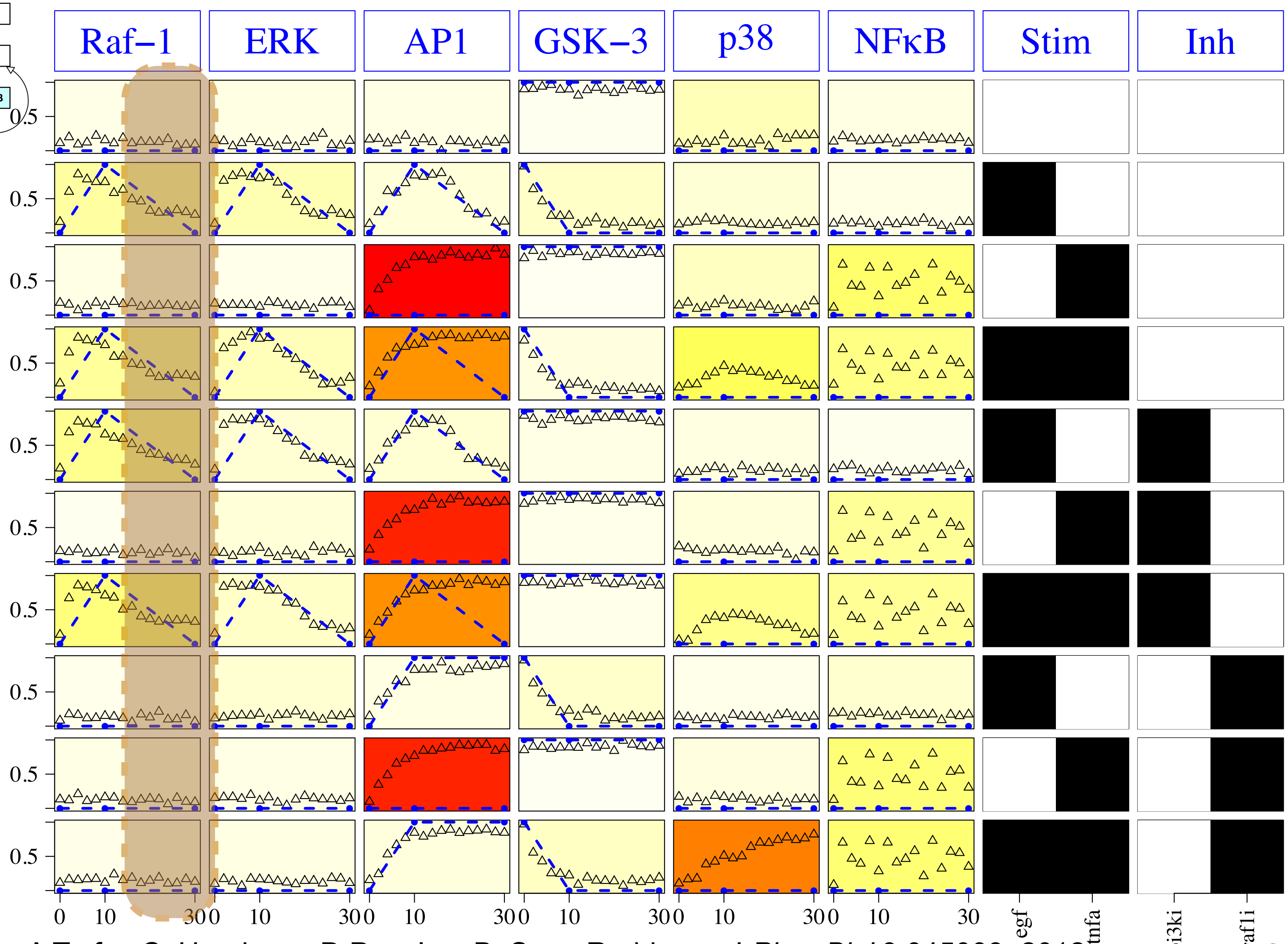
MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Multiple pseudo-steady-state captures feedbacks that lead to transient signals

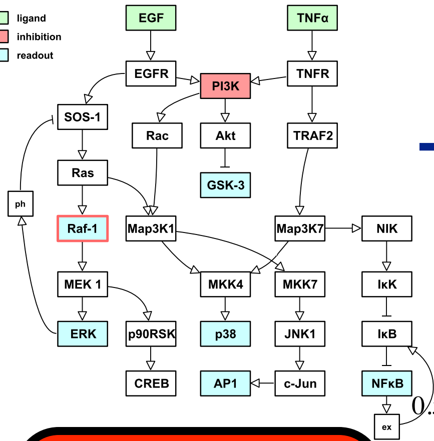


Can not explain data due to missing links

Identifies strong active links



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

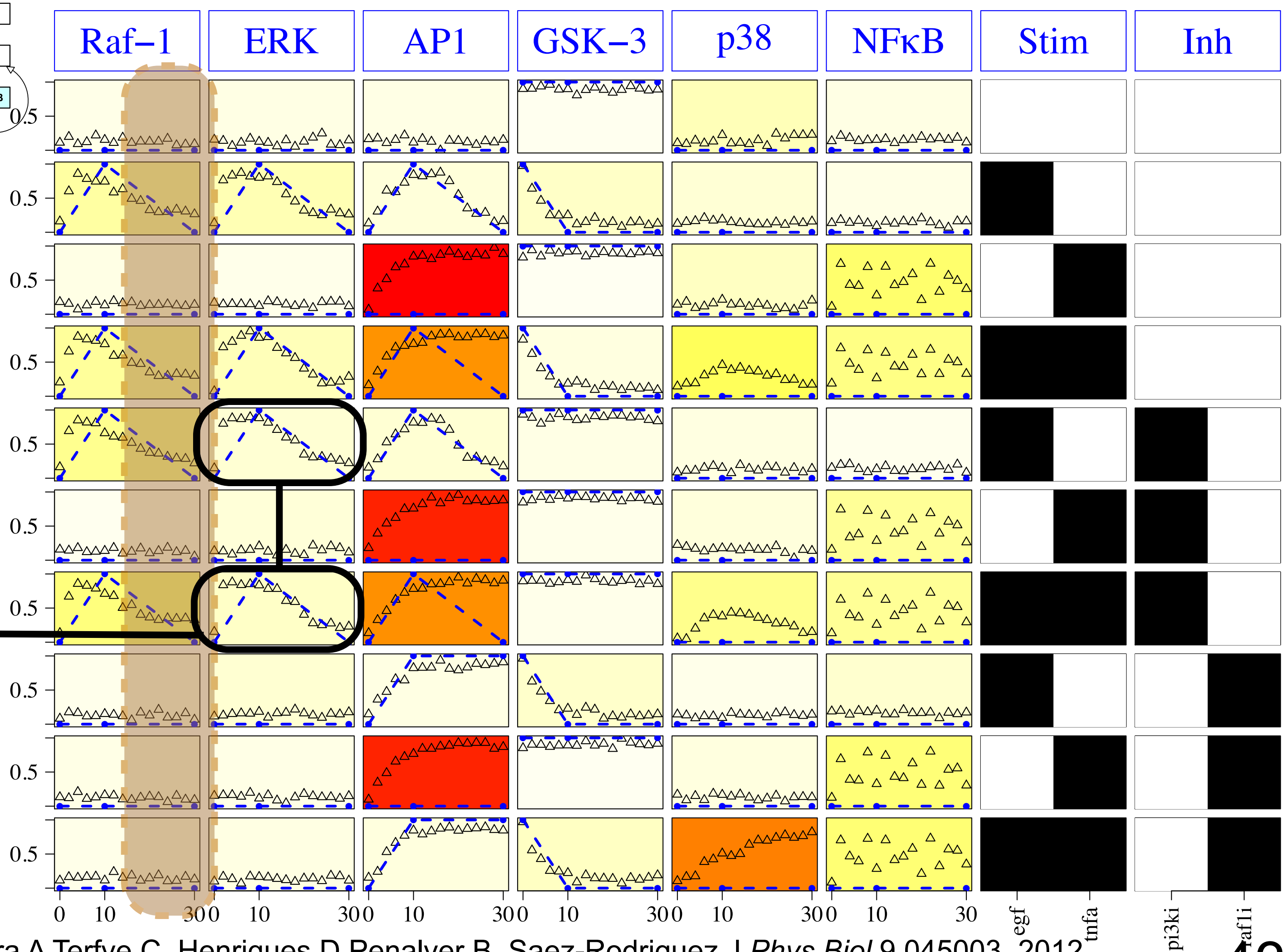


Multiple pseudo-steady-state captures feedbacks that lead to transient signals

Can not explain data due to missing links

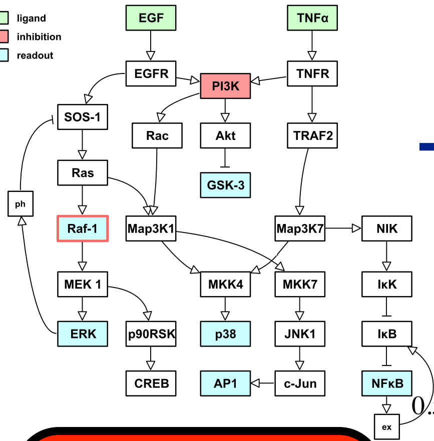
Identifies strong active links

Captures strong feedbacks



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Multiple pseudo-steady-state captures feedbacks that lead to transient signals

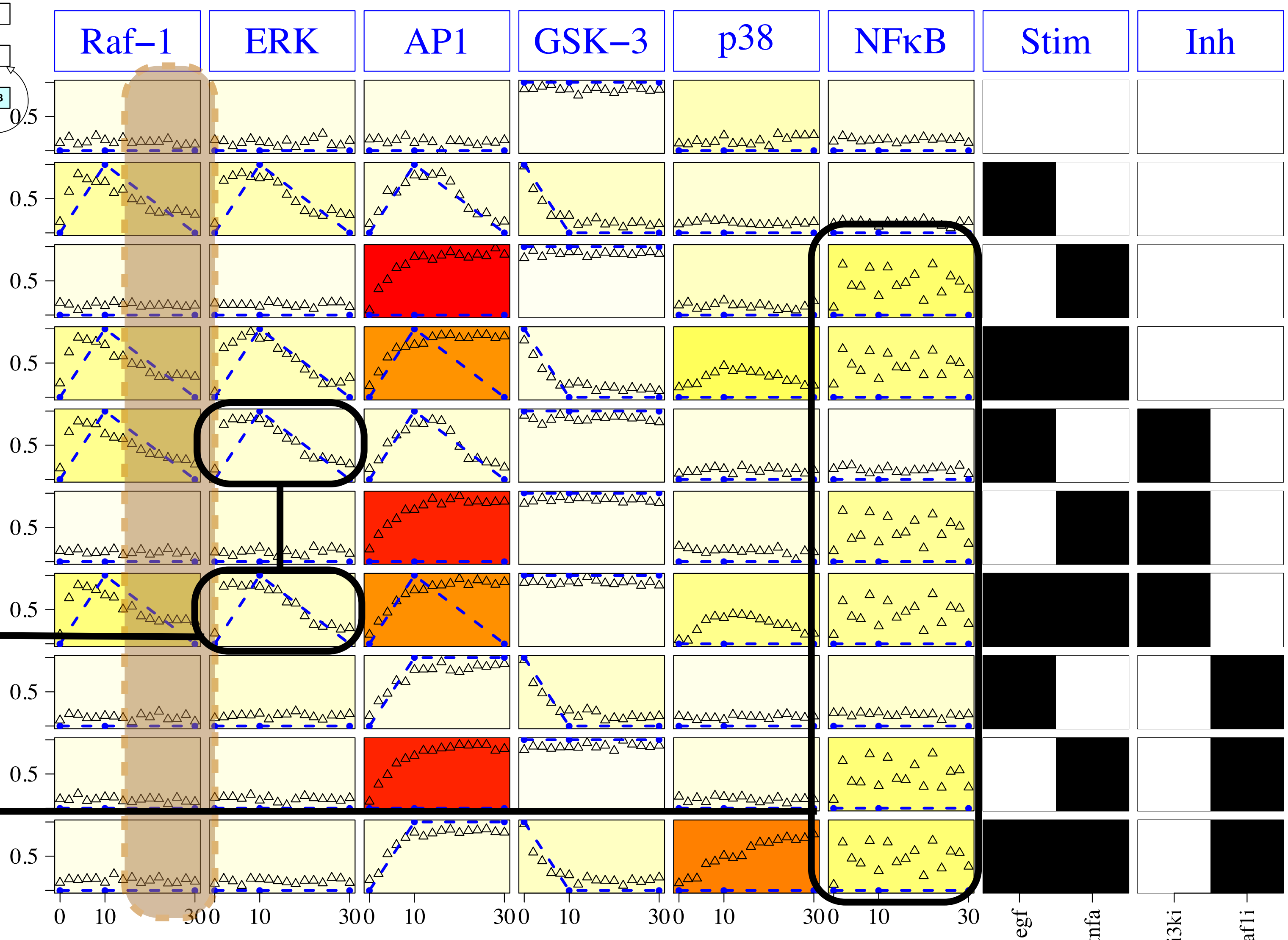


Can not explain data due to missing links

Identifies strong active links

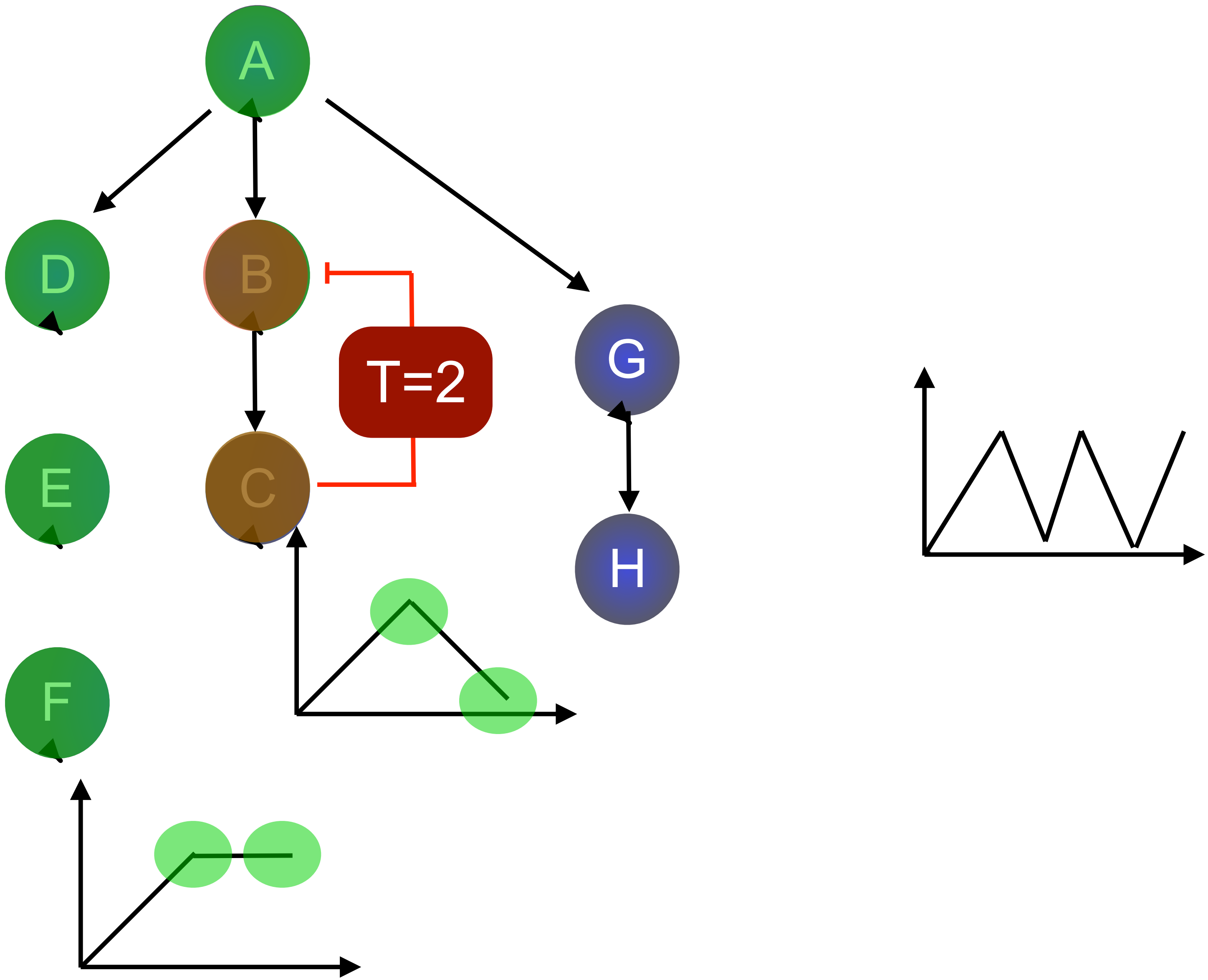
Captures strong feedbacks

Can not handle oscillations

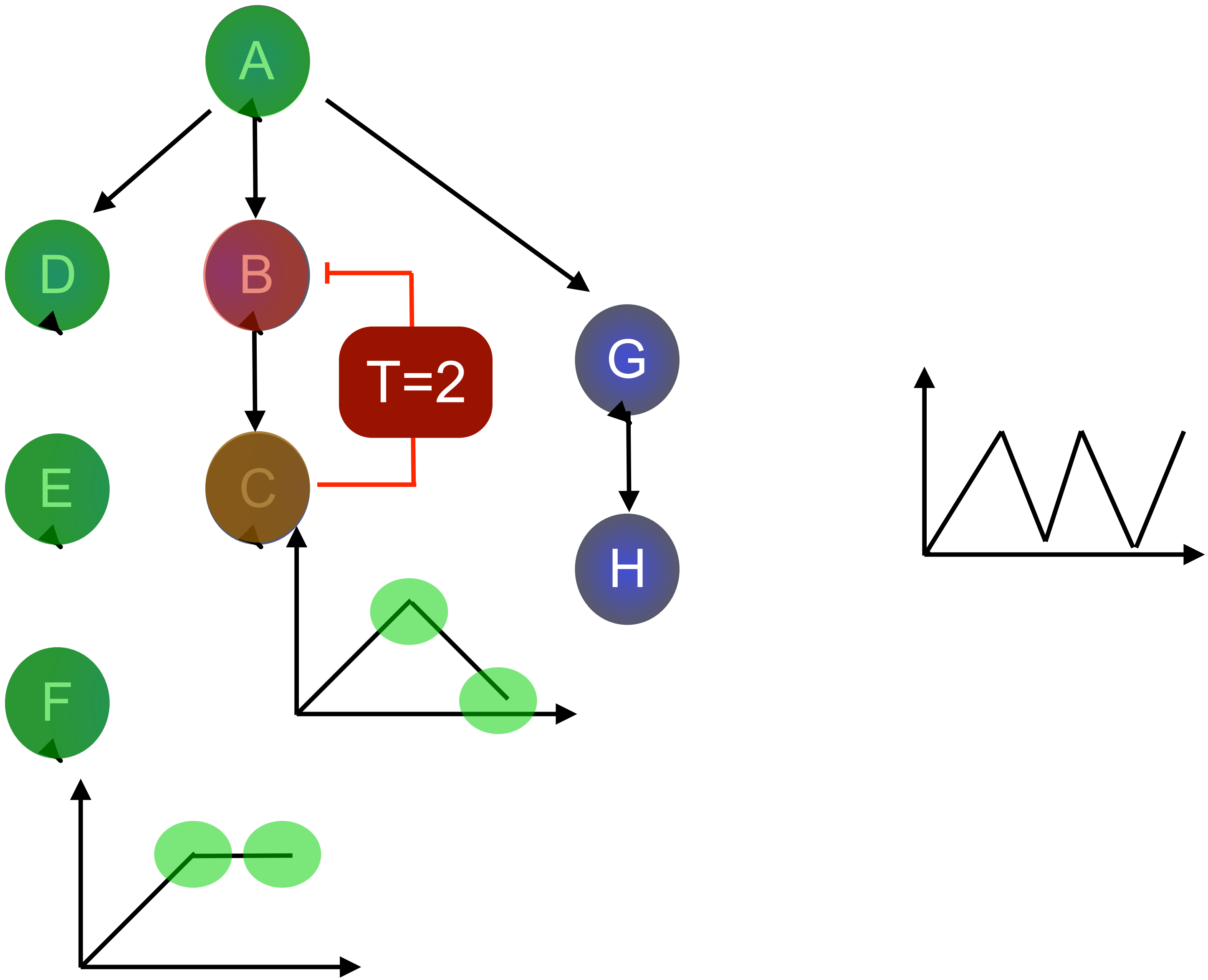


MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

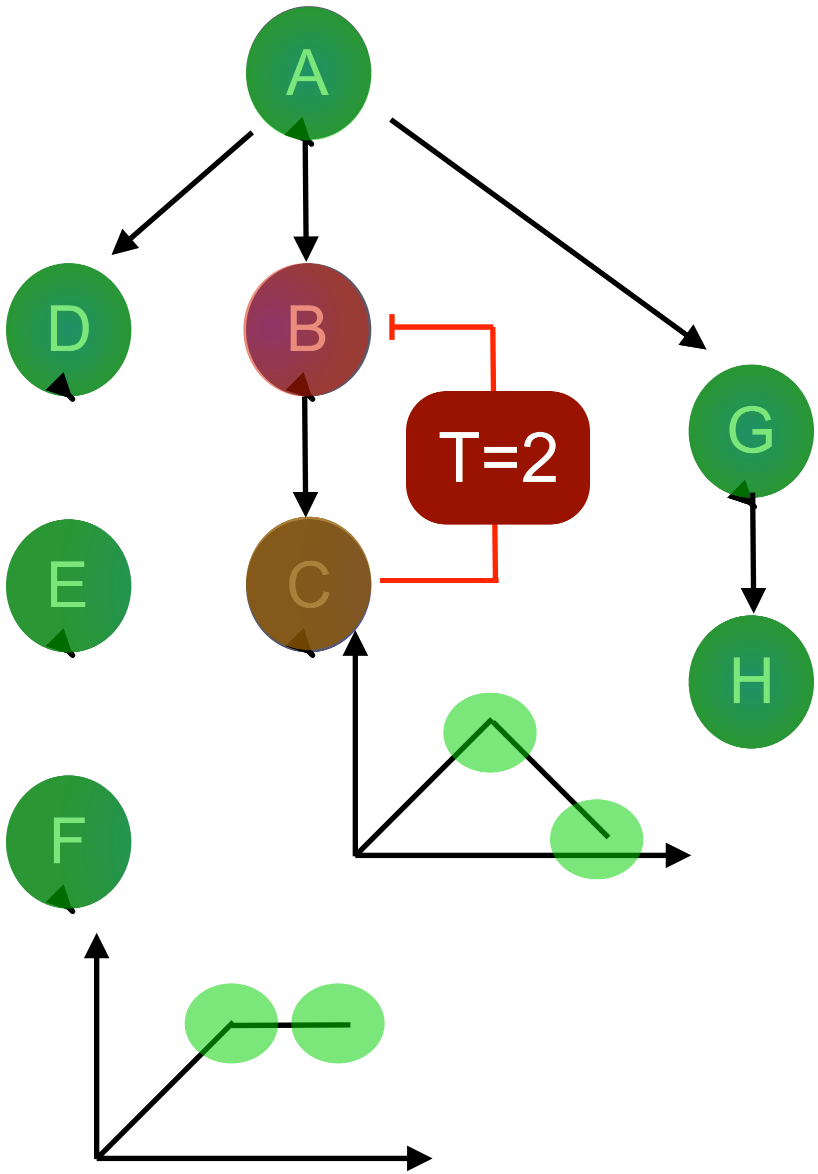
Approximation of dynamics using synchronous simulation & multiple time-scales

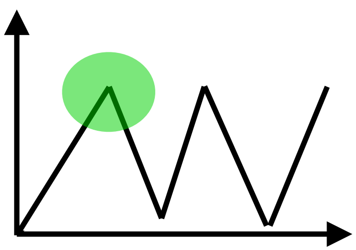


Approximation of dynamics using synchronous simulation & multiple time-scales

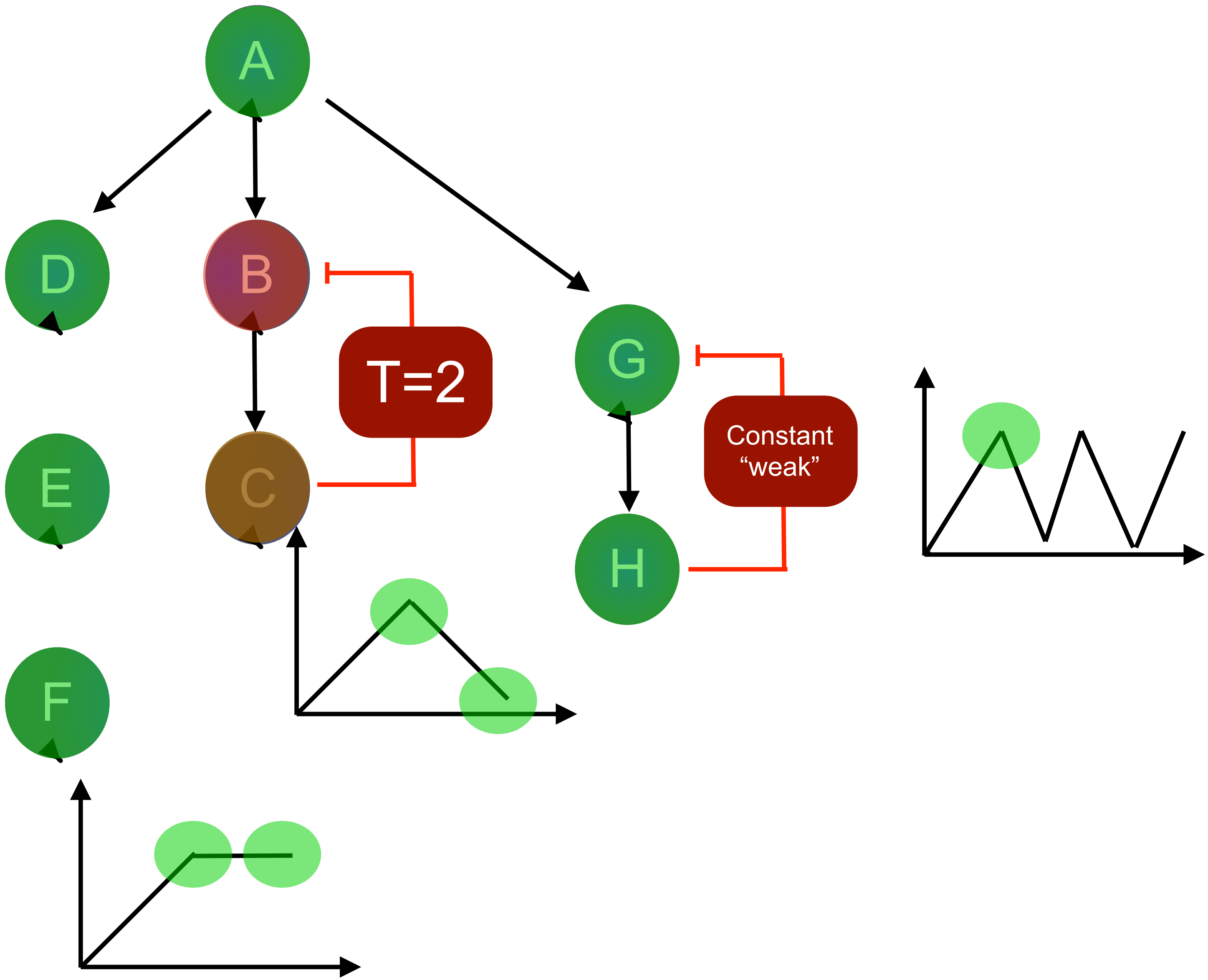


Approximation of dynamics using synchronous simulation & multiple time-scales

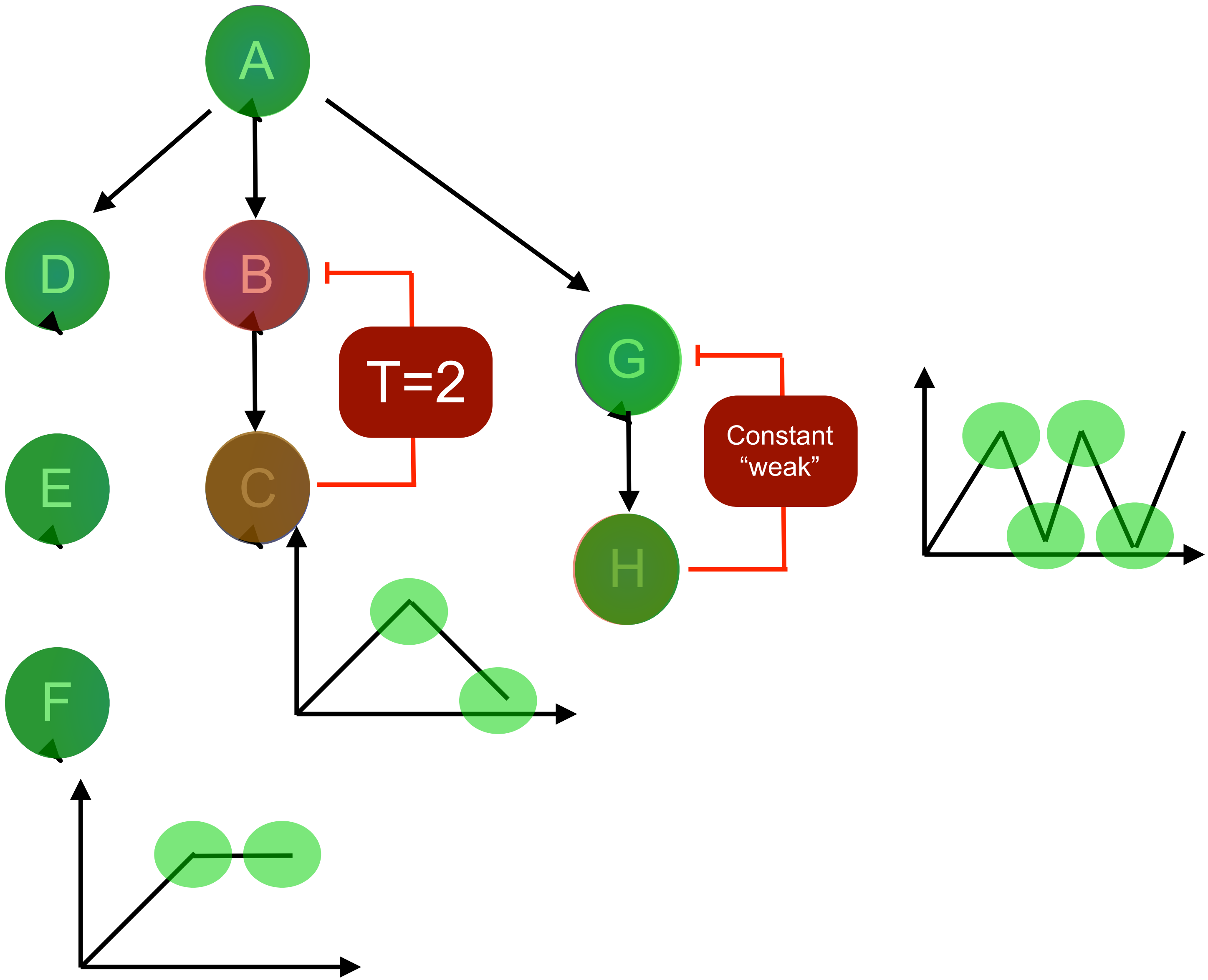




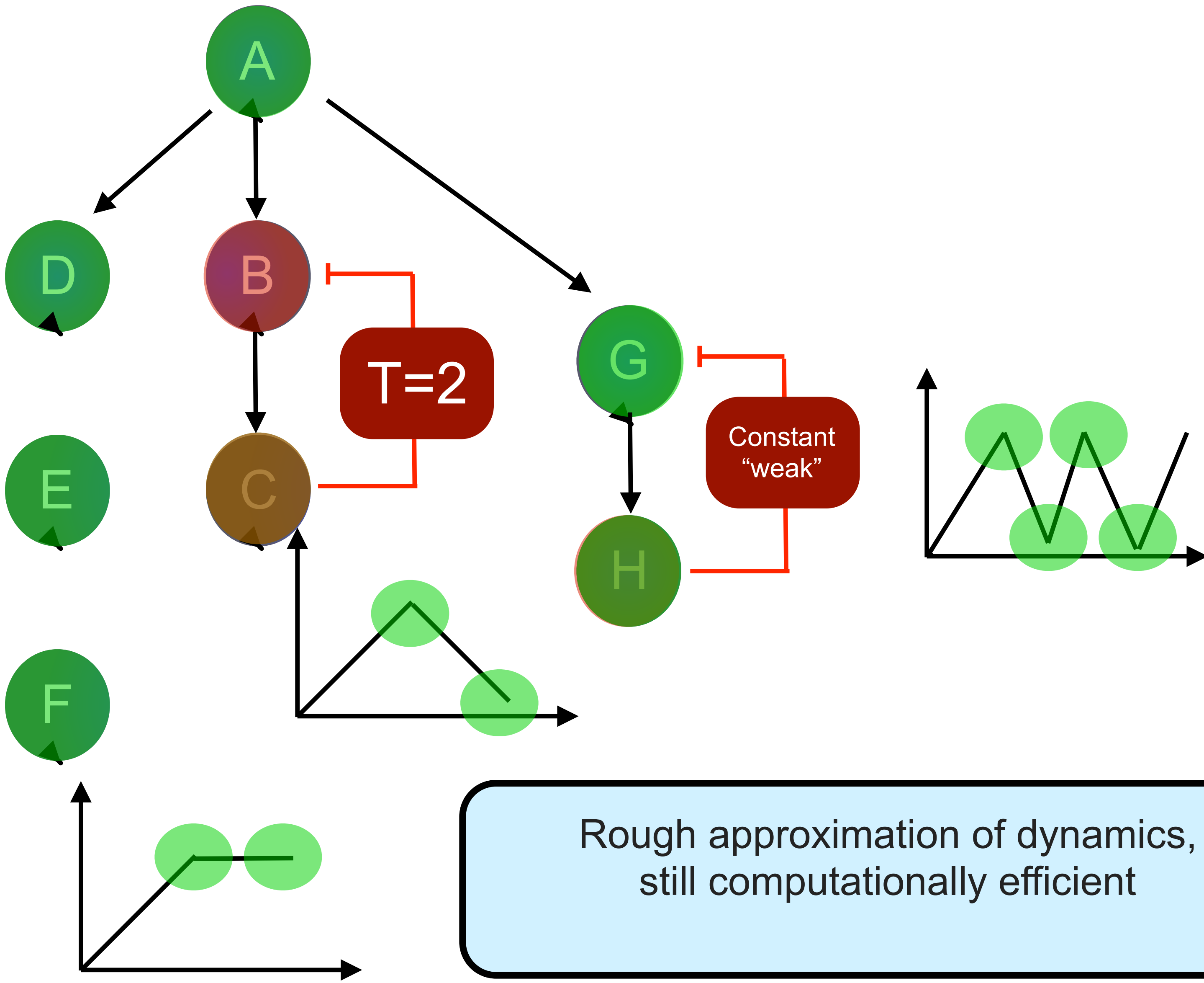
Approximation of dynamics using synchronous simulation & multiple time-scales



Approximation of dynamics using synchronous simulation & multiple time-scales

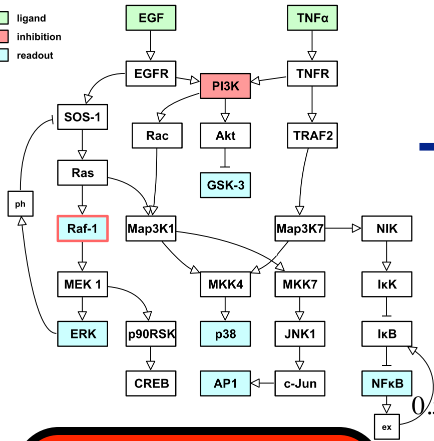


Approximation of dynamics using synchronous simulation & multiple time-scales



Rough approximation of dynamics,
still computationally efficient

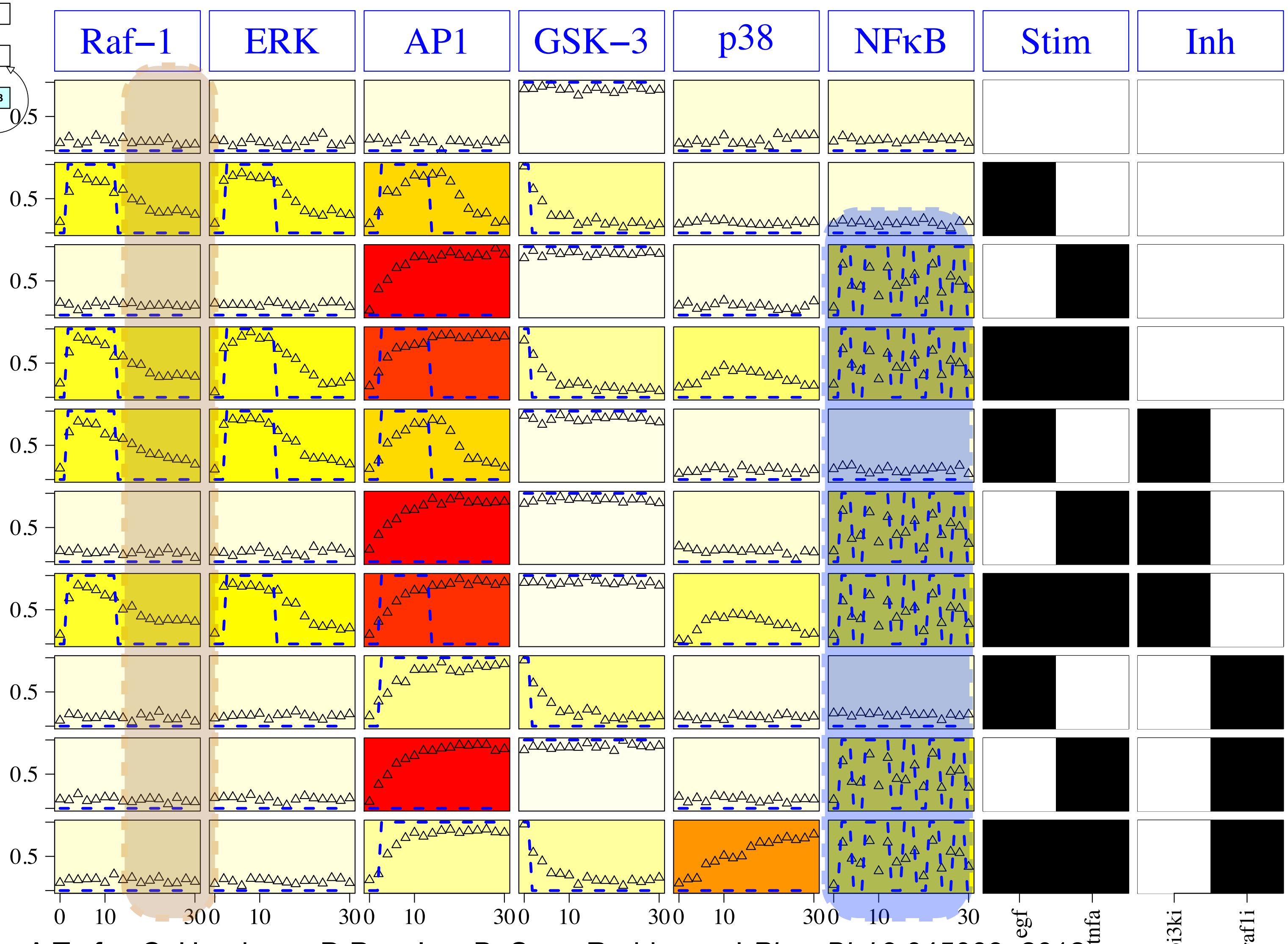
Synchronous simulation captures oscillatory behaviour



Can not explain data due to missing links

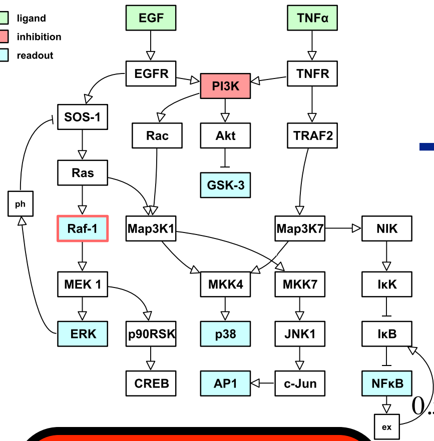
Identifies strong active links

Captures strong feedbacks



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

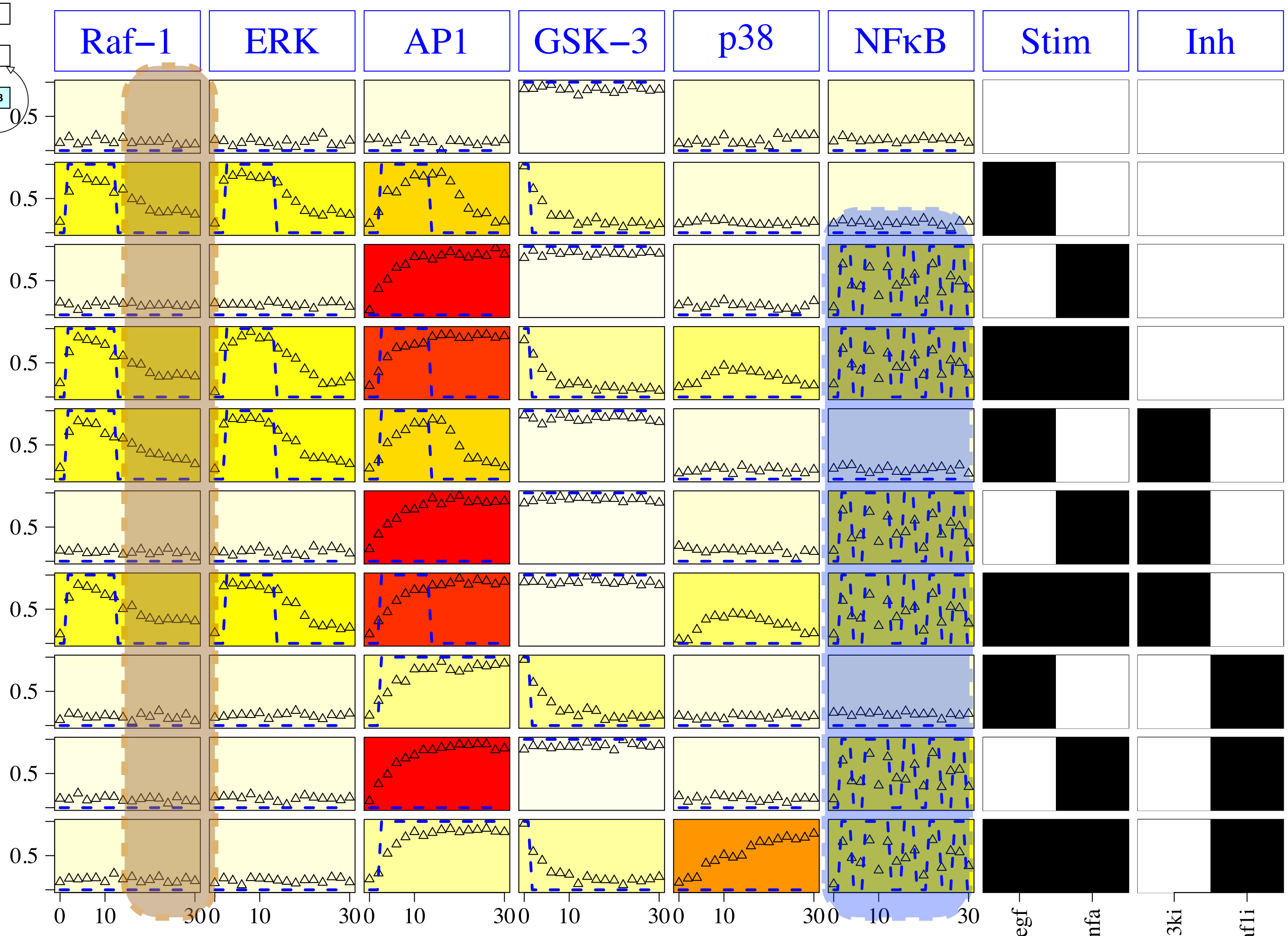
Synchronous simulation captures oscillatory behaviour



Can not explain data due to missing links

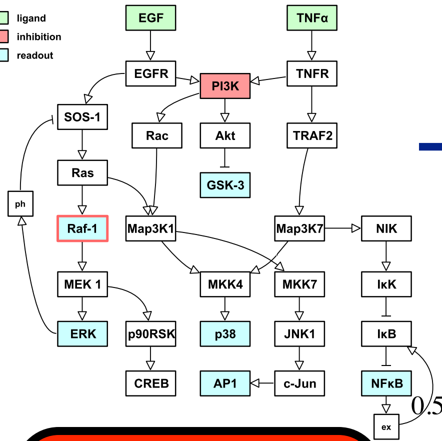
Identifies strong active links

Captures strong feedbacks



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

Synchronous simulation captures oscillatory behaviour

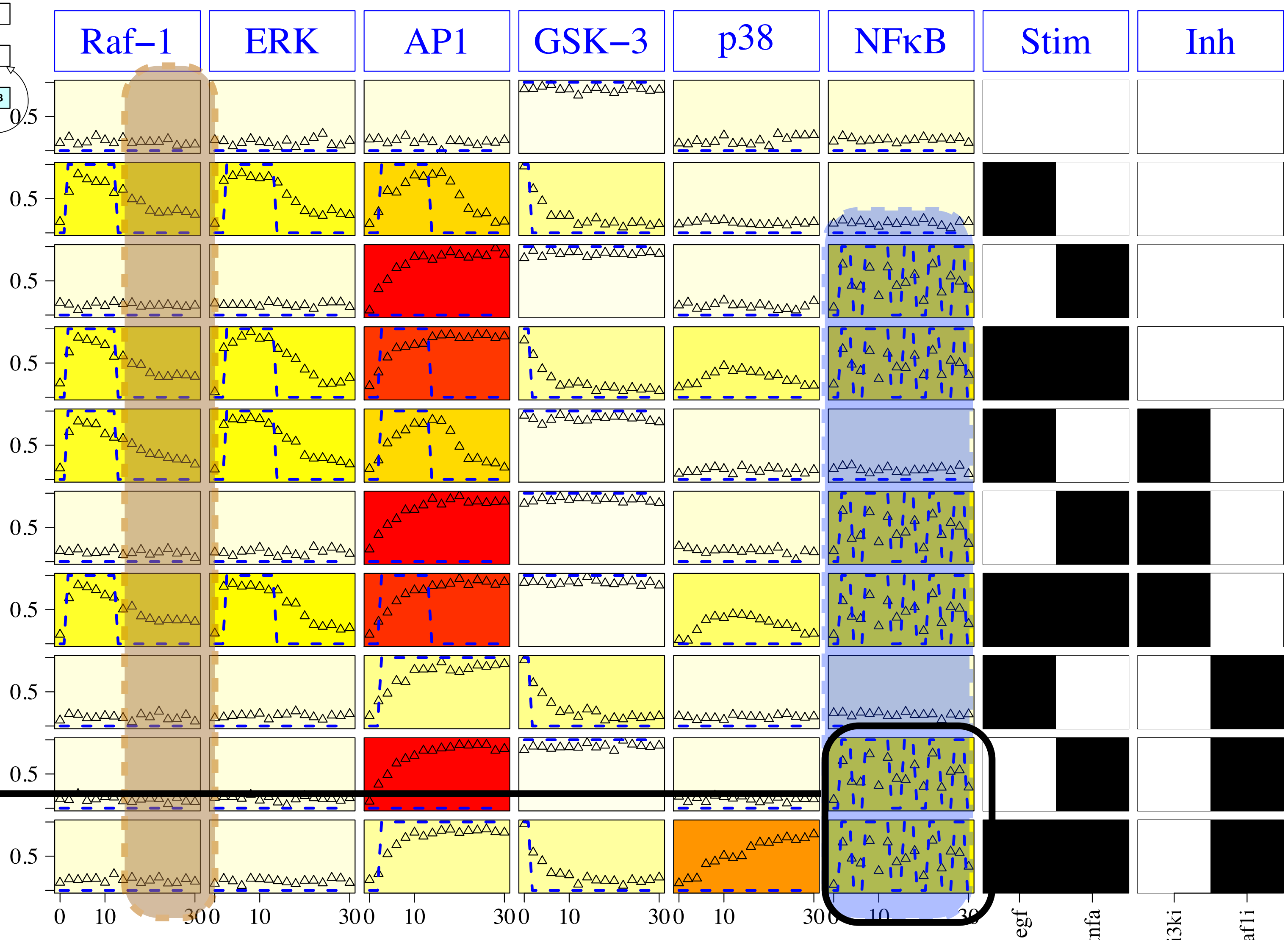


Can not explain data due to missing links

Identifies strong active links

Captures strong feedbacks

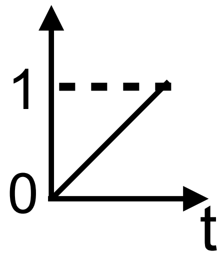
Captures 'weak' feedbacks



MacNamara A Terfve C Henriques D Penalver B Saez-Rodriguez J *Phys Biol* 9 045003, 2012

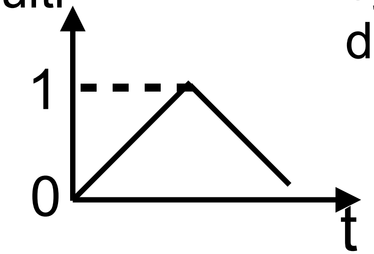
From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state

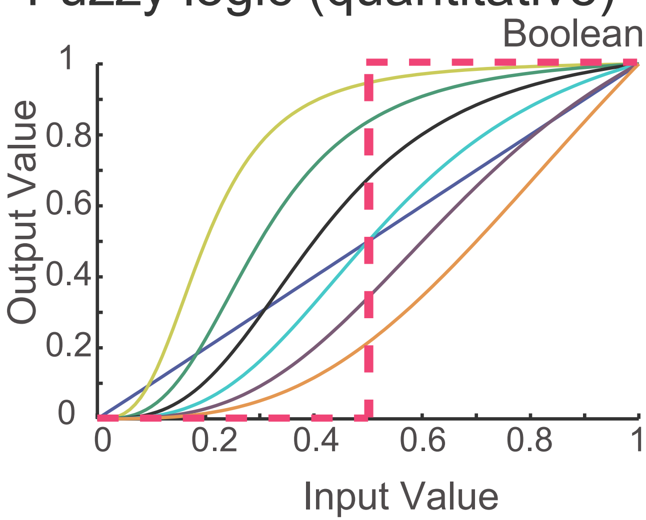


Boolean multi time-scale

sync. dynamics



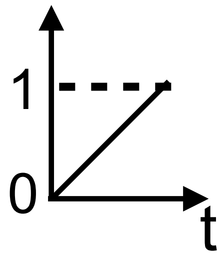
Fuzzy logic (quantitative)



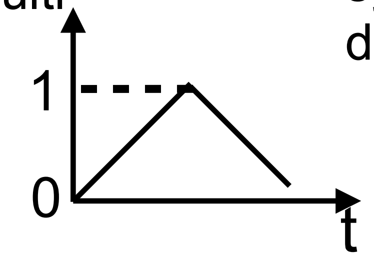
Morris et al., PloS Comp Bio 2011

From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state



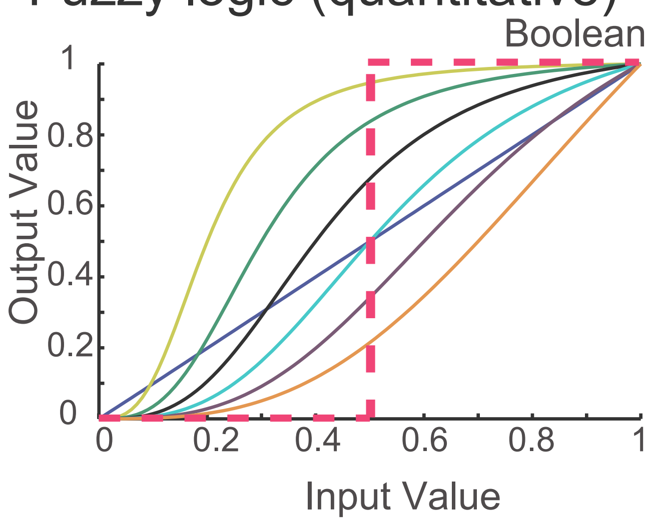
Boolean multi time-scale sync. dynamics



Camille Terfve

Aidan MacNamara

Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

Logic-based ODEs

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

$$\mathrm{f}(\overline{x_i}) = \frac{\overline{x_i}^n}{(\overline{x_i}^n + k^n)}$$

- Transform into differential equation

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

- E.g. a AND b inactivate C

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

Wittman D, Krumsiek J, Saez-Rodriguez J, Lauffenburger DA, Klamt S, Theis FJ, *BMC Sys Bio* 2009

Logic-based ODEs

- Convert Boolean update function *Bi* into a *continuous homologue* $\bar{B}_i$ using multivariate polynomial interpolation
  - **Accuracy** (same behavior as *Bi* for 0/1 → same monotony & steady state behavior)
  - Good **analytical** properties (smoothness)
  - **Minimal** and **unique**

IDEA:
ODE model
mathematically 'well-behaved'
that matches the Boolean model
when states are 0 or 1

- Make non linear replacing variable with Hill function

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

- Transform into differential equation

$$\bar{x}_i(t+1) = \bar{B}_i(\bar{x}_{i1}(t), \bar{x}_{i2}(t), \ldots, \bar{x}_{iN}(t))$$

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

- E.g. a AND b inactivate C

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

Wittman D, Krumsiek J, Saez-Rodriguez J, Lauffenburger DA, Klamt S, Theis FJ, *BMC Sys Bio* 2009

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+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) = ((eg _inh)

d/dt(pi3k) = ((e _tauinv)*
(1-pi3k_inh)

d/dt(ikb) = ((tn
(pi3k^ikb_n_pi
+ikb_k_tnfa^ik
(1+ikb_k_pi3k
*pi3k^ikb_n_p

d/dt(gsk3) = ((
gsk3_tauinv)*(

d/dt(erk12) = (
ikb^erk12_n_i
(raf^erk12_n_
+erk12_k_ikb^ af)*
(1+erk12_k_ra _ikb)-
erk12) * erk12

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

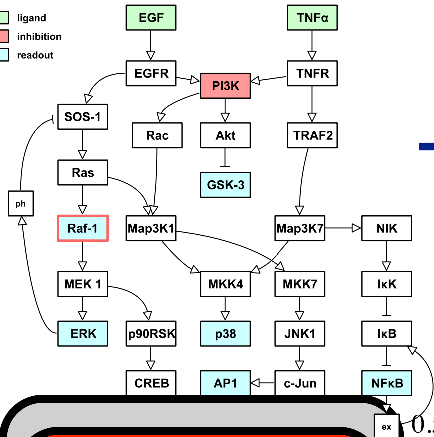
$$\frac{d}{dt}c = \frac{1}{\tau}\left(\frac{a^{n_a} * (1 + k_a{}^{n_a}) * (1 - b^{n_b}) * (1 + k_b{}^{n_b})}{(a^{n_a} + k_a{}^{n_a}) * (b^{n_b} + k_b{}^{n_b})} + \frac{(1 - a^{n_a}) * (1 + k_a{}^{n_a}) * b^{n_b} * (1 + k_b{}^{n_b})}{(a^{n_a} + k_a{}^{n_a}) * (b^{n_b} + k_b{}^{n_b})} + \frac{a^{n_a} *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

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)

Fit of ODE model



Can not explain data due to missing links

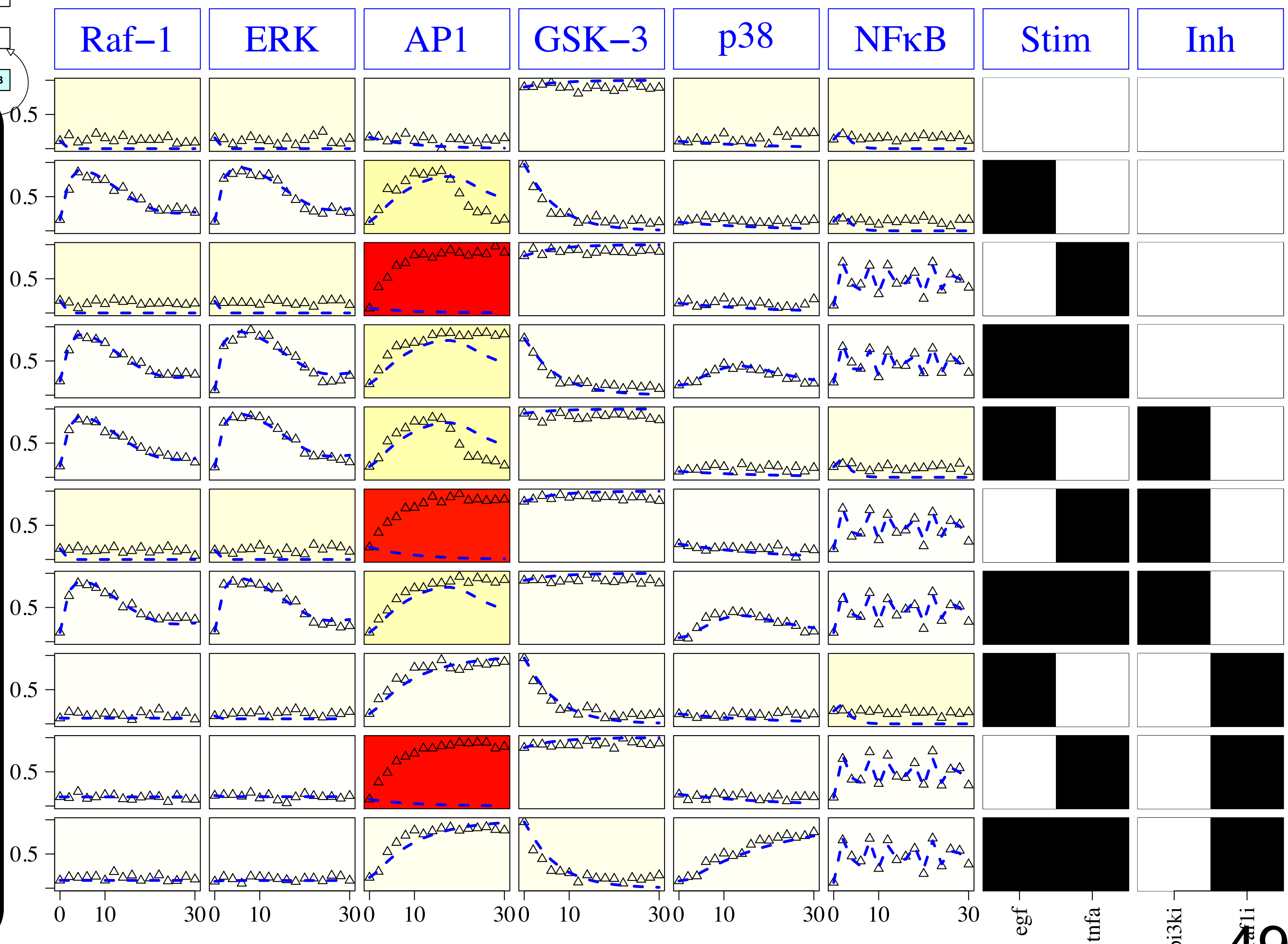
Identifies strong active links

Identifies weak active links

Captures strong feedbacks

Captures 'weak' feedbacks

ODE



Different methods capture different aspects

Can not explain data due to missing links

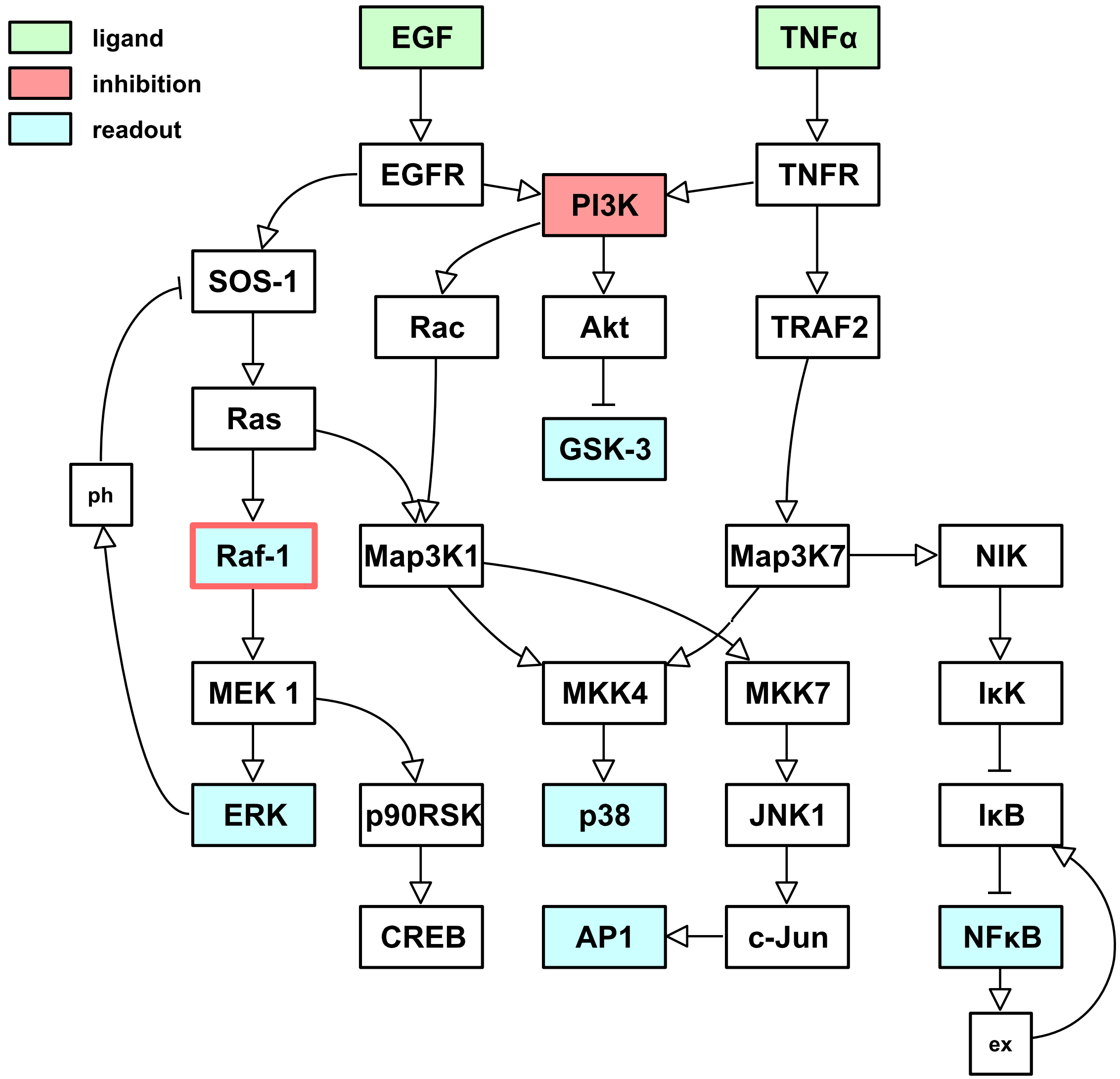
Identifies strong active links

Identifies weak active links

Captures strong feedbacks

Captures 'weak' feedbacks

ODE



Different methods capture different aspects

Can not explain data due to missing links

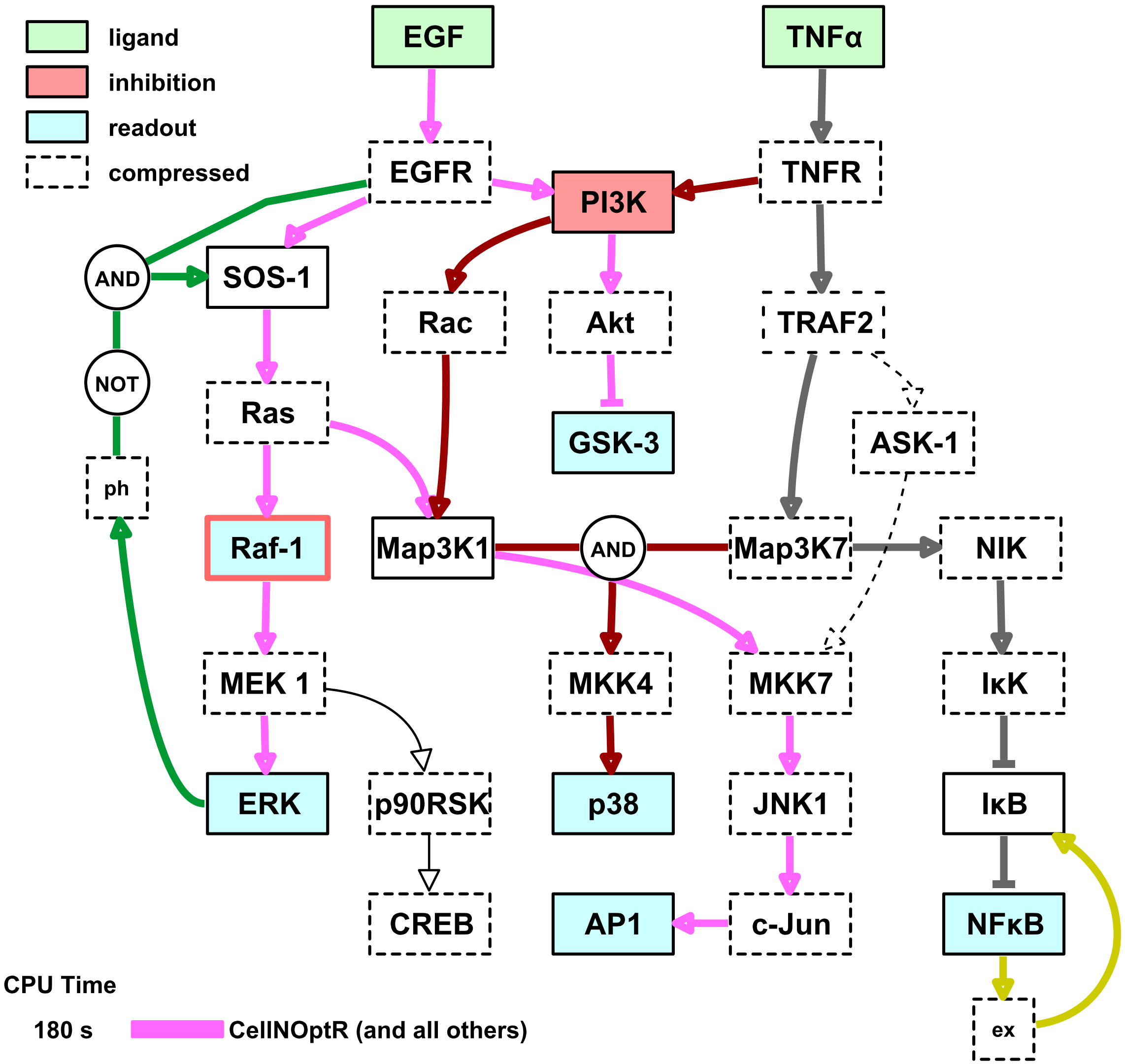
Identifies strong active links

Identifies weak active links

Captures strong feedbacks

Captures 'weak' feedbacks

ODE

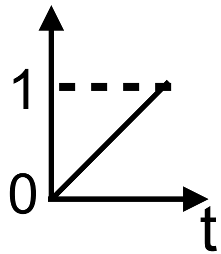


CPU Time

180 s	CellNOptR (and all others)
240 s	CellNOptR 2 steady state
1200 s	CNORcfl
2000 s	CNORdt
-	CNORdt & CNORcfl

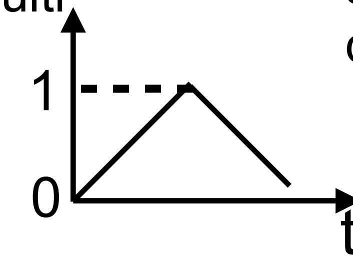
From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state



Boolean multi time-scale

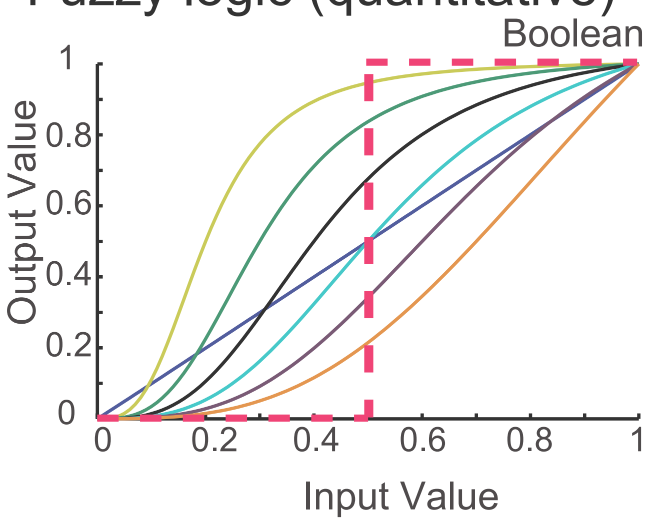
sync. dynamics



Camille Terfve

Aidan MacNamara

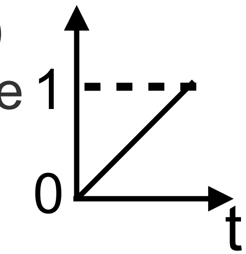
Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

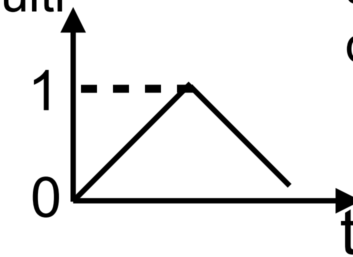
From Boolean to continuous and dynamic models within CellNOpt

Boolean (binary) logic steady state



Boolean multi time-scale

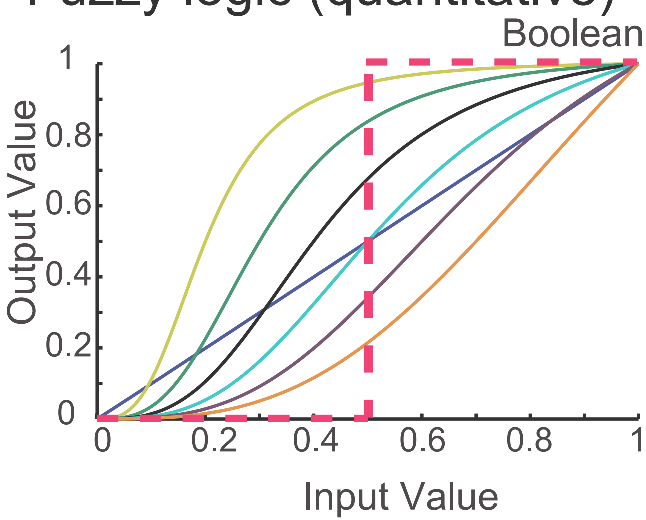
sync. dynamics



Camille Terfve

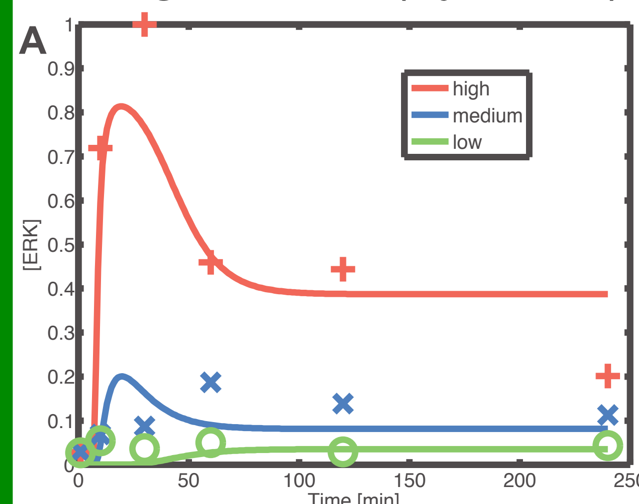
Aidan MacNamara

Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

Logic ODEs (dynamic)



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

Identify structure + parameters Henriques et al. Bioinformatics 2015

w. J Banga & J. Egea,

David Henriques

From Boolean to continuous and dynamic models within CellNOpt

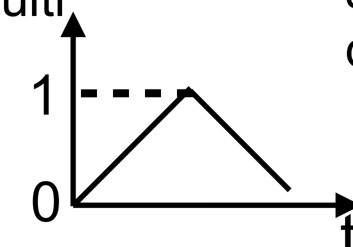
Boolean (bina
logic steady

+ detail

- scope

Boolean multi time-scale

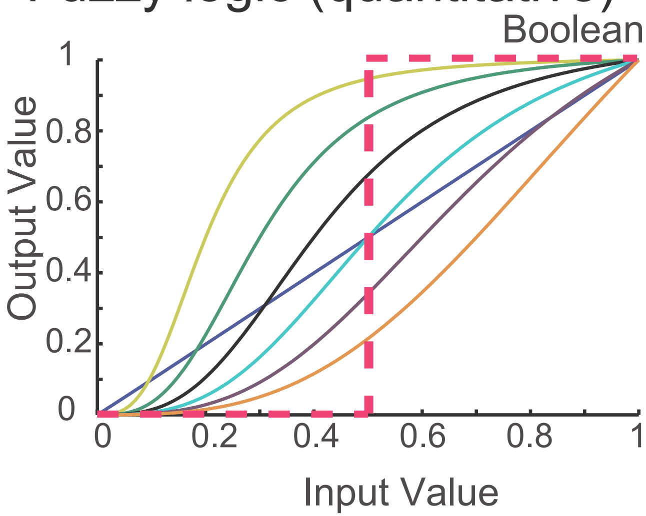
sync. dynamics



Camille Terfve

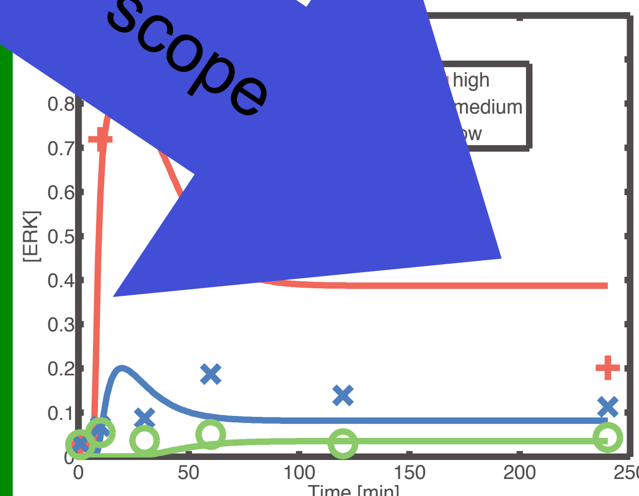
Aidan MacNamara

Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

ODE (dynamic)



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

Identify structure + parameters Henriques et al. Bioinformatics 2015

w. J Banga & J. Egea,

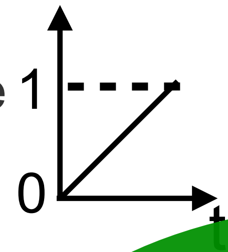
David Henriques

From Boolean to continuous and dynamic models within CellNOpt

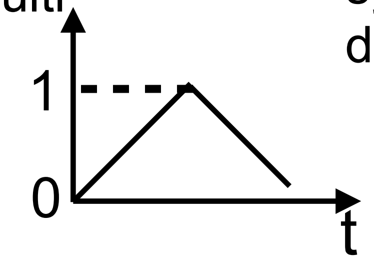
COMBINE

RWTH AACHEN

Boolean (binary) logic steady state



Boolean multi time-scale



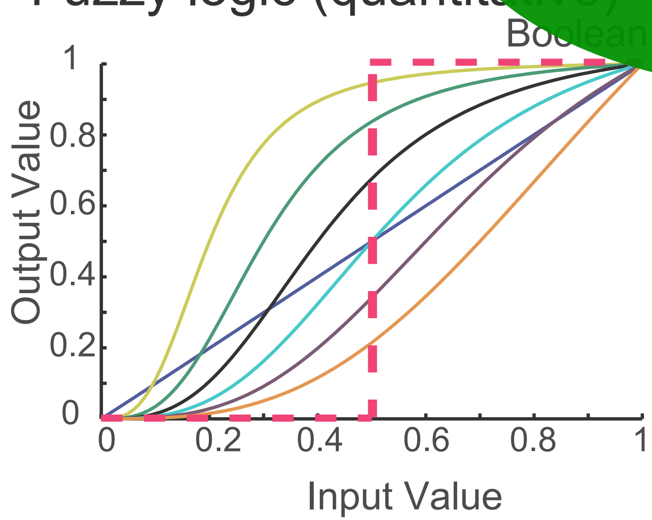
sync. dynamics

Camille Terfve

Aidan MacNamara

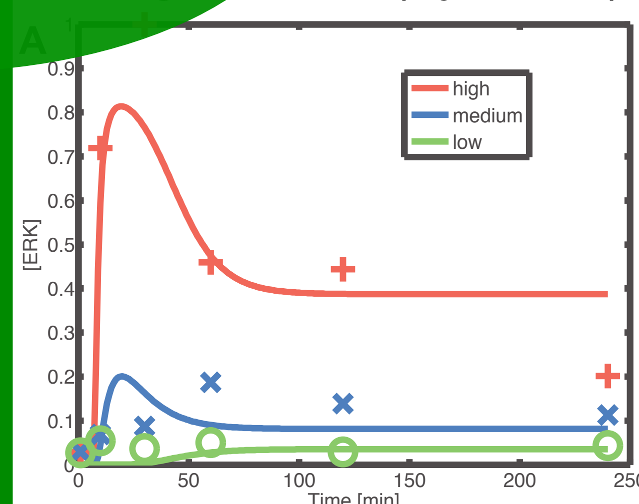
CellNOpt

Fuzzy logic (quantitative)



Morris et al., PloS Comp Bio 2011

Logic ODEs (dynamic)



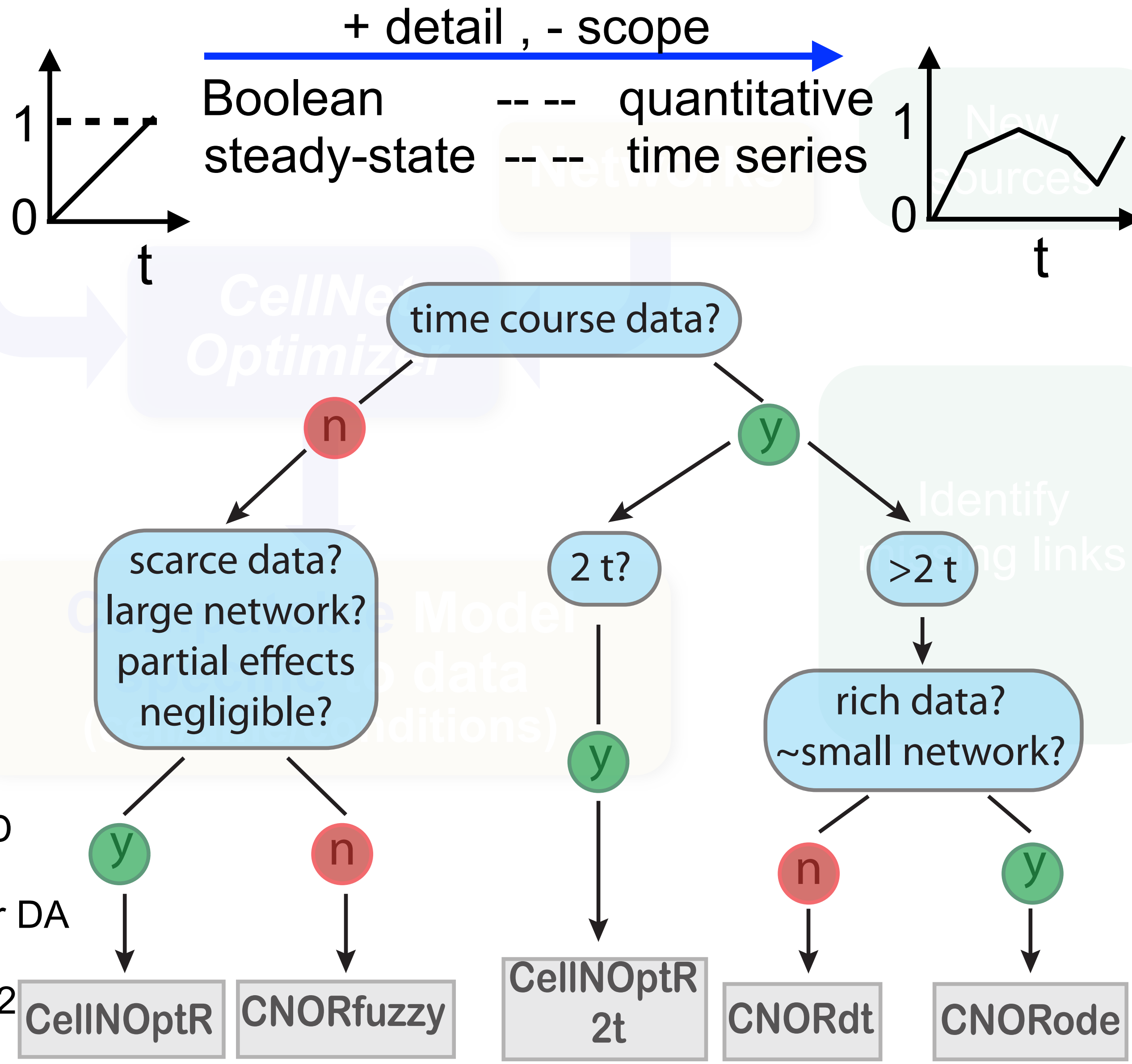
MEIGO: Global optimization in R/Matlab Egea et al. BMC Bioinf 2014

Identify structure + parameters Henriques et al. Bioinformatics 2015

w. J Banga & J. Egea,

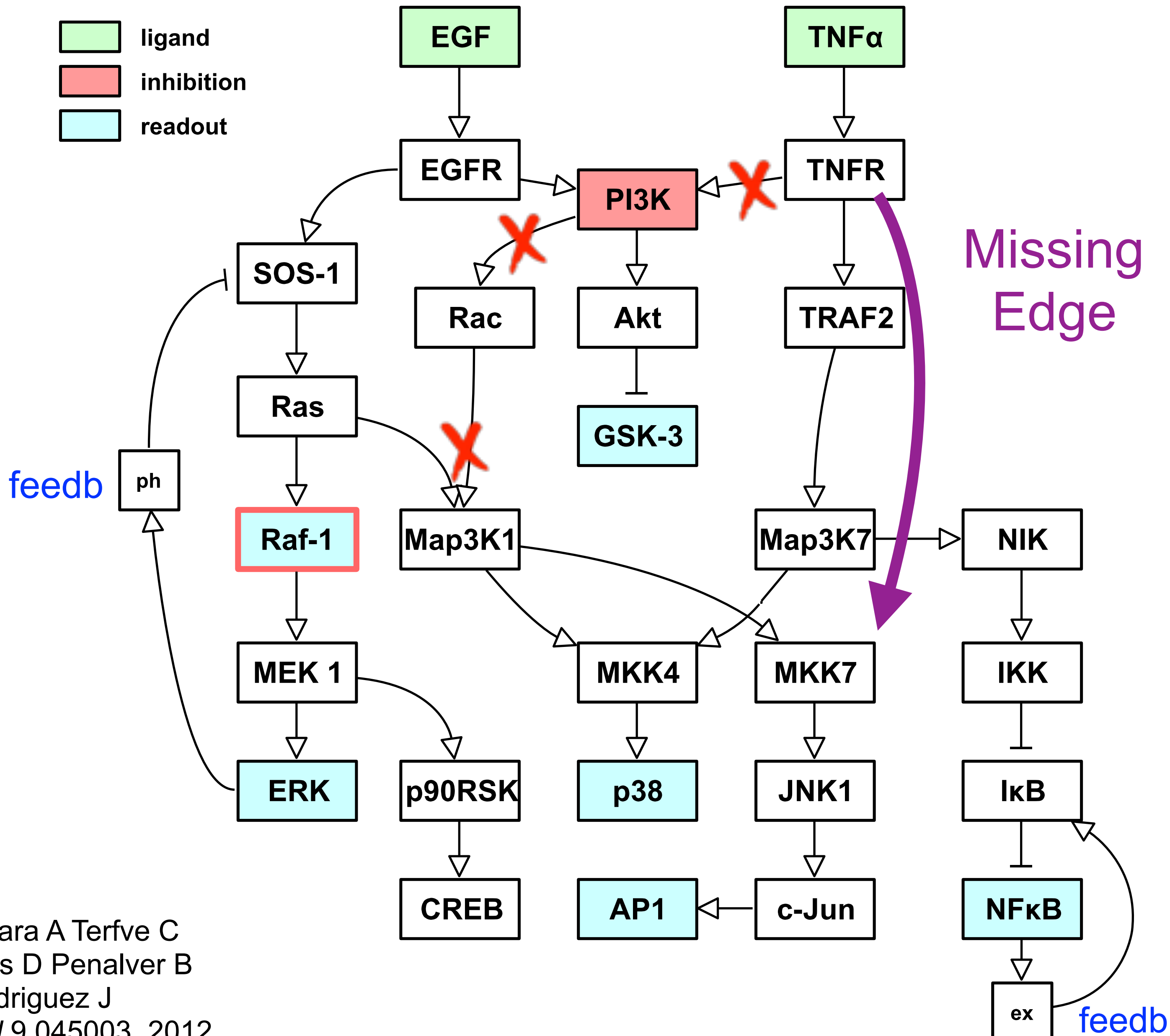
David Henriques

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?



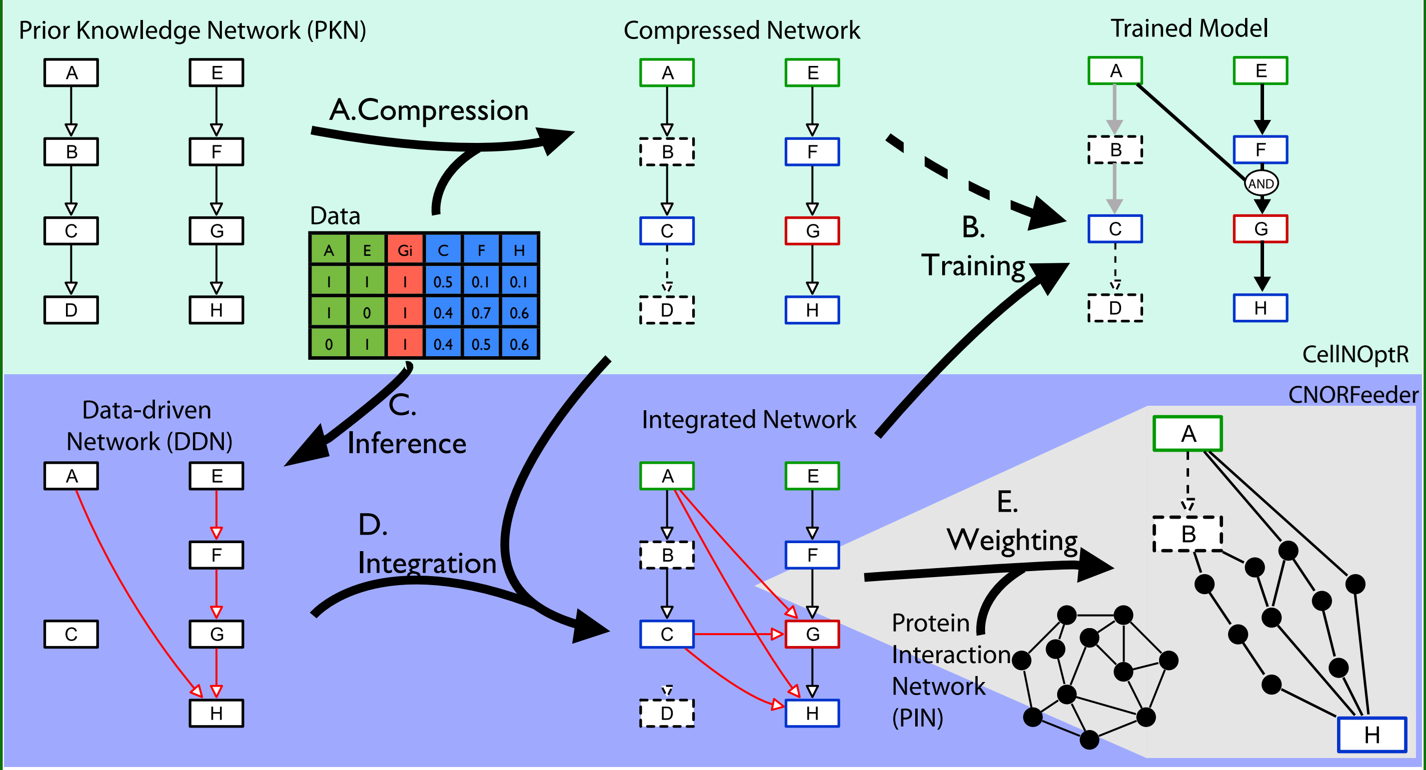
MacNamara A Terfve C
Henriques D Penalver B
Saez-Rodriguez J
Phys Biol 9 045003, 2012

How to deal with incomplete prior knowledge?

CNOFeed: Link CellNOpt to methods to infer new links

Federica Eduati

More types

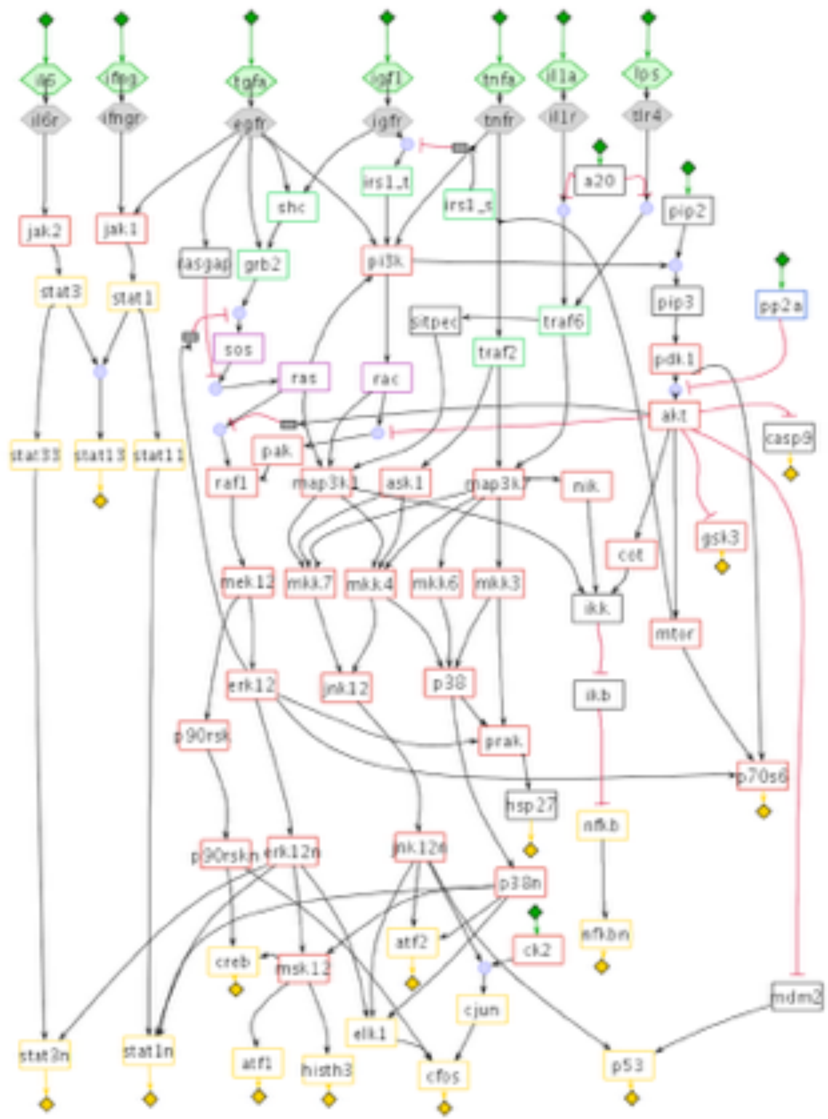


A	E	Gi	C	F	H
1	1	1	0.5	0.1	0.1
1	0	1	0.4	0.7	0.6
0	1	1	0.4	0.5	0.6

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

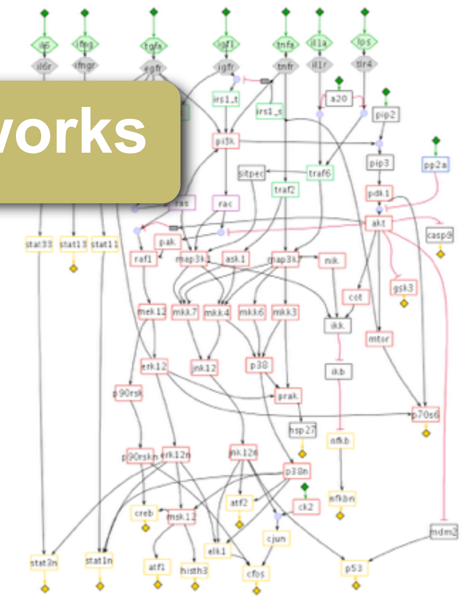
Logic modeling to link protein signaling networks with functional signaling data

Logic modeling to link protein signaling networks with functional signaling data

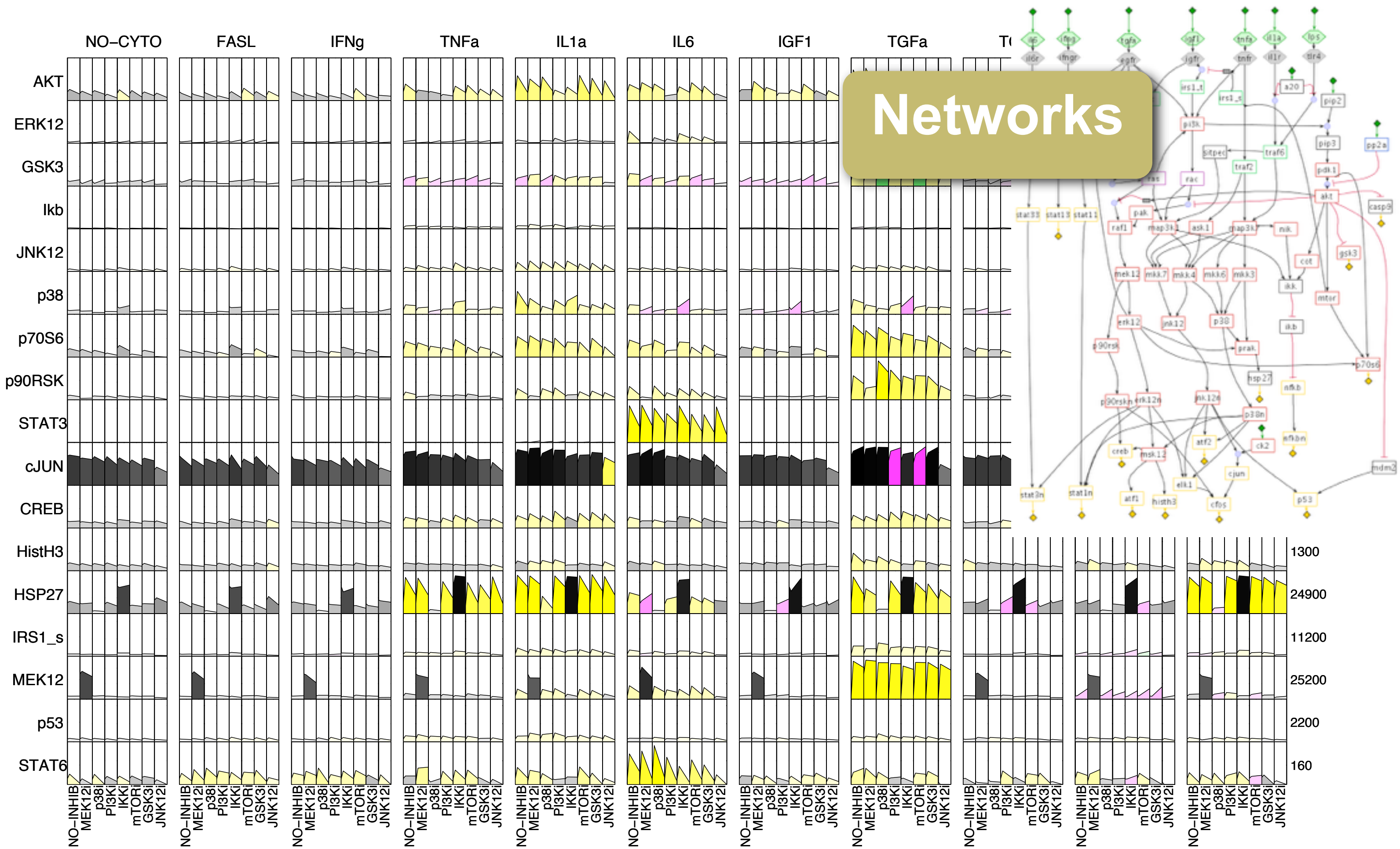


Logic modeling to link protein signaling networks with functional signaling data

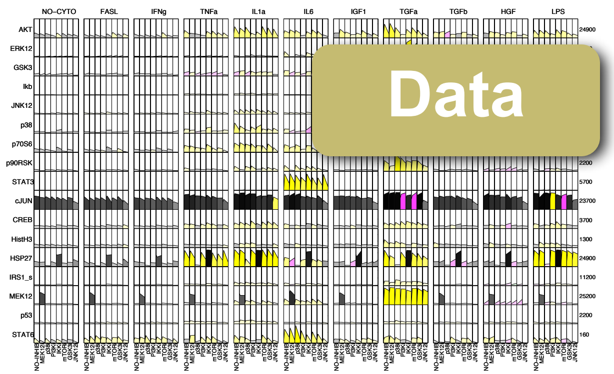
Networks

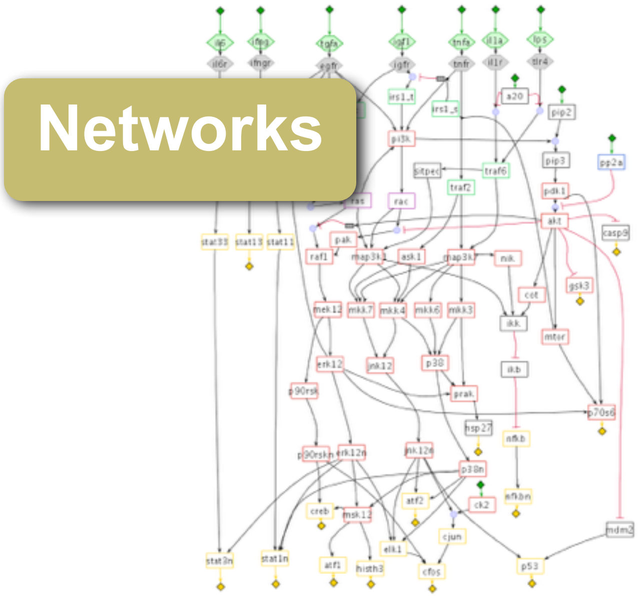


Logic modeling to link protein signaling networks with functional signaling data

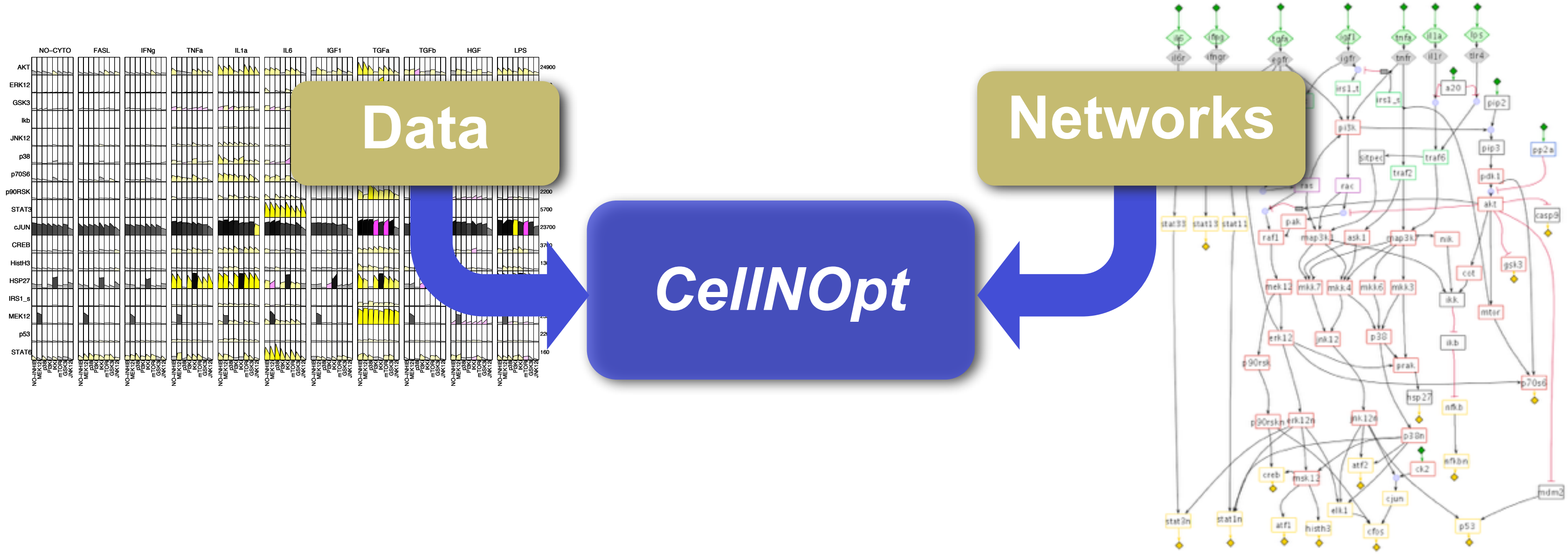


Logic modeling to link protein signaling networks with functional signaling data

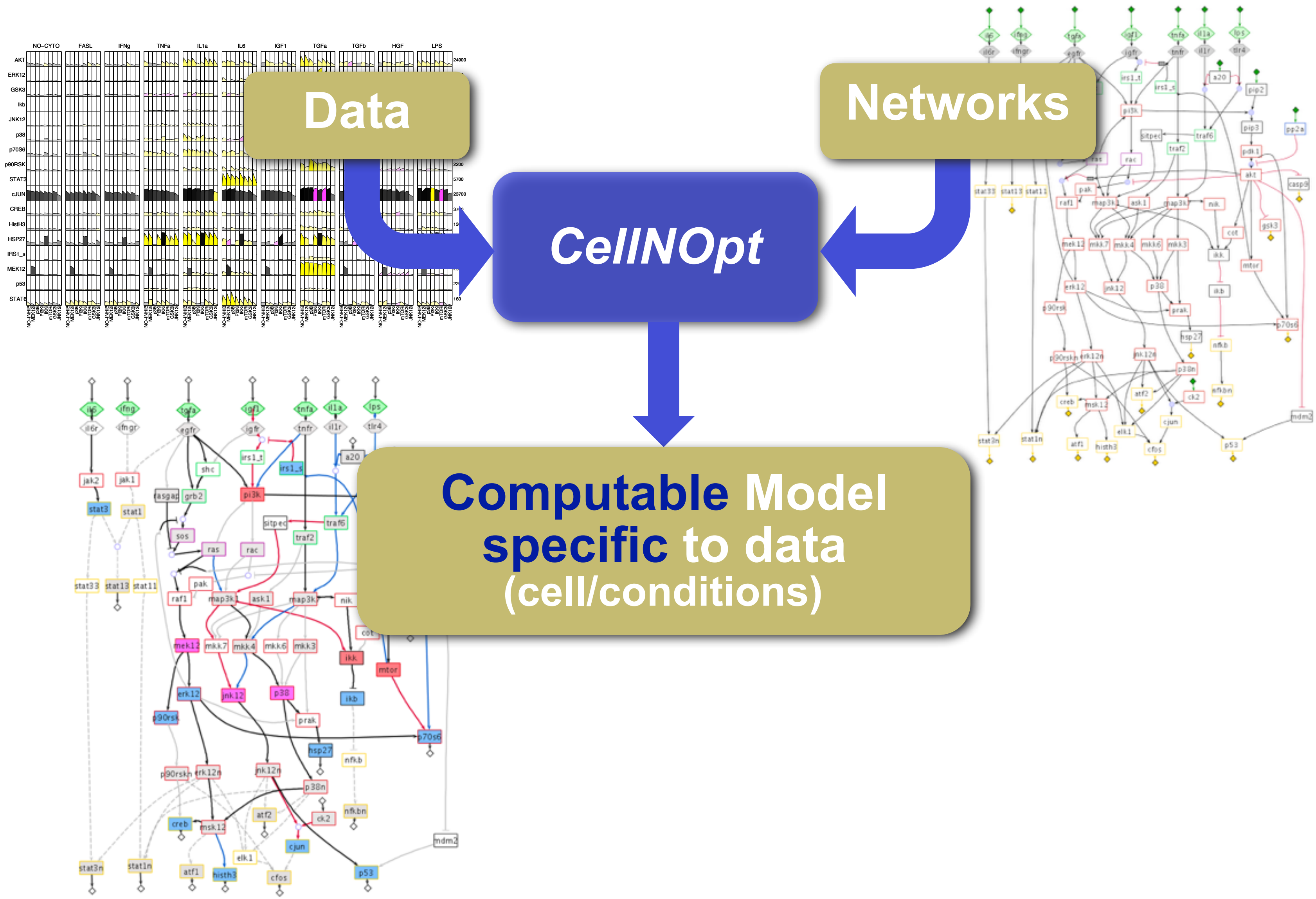


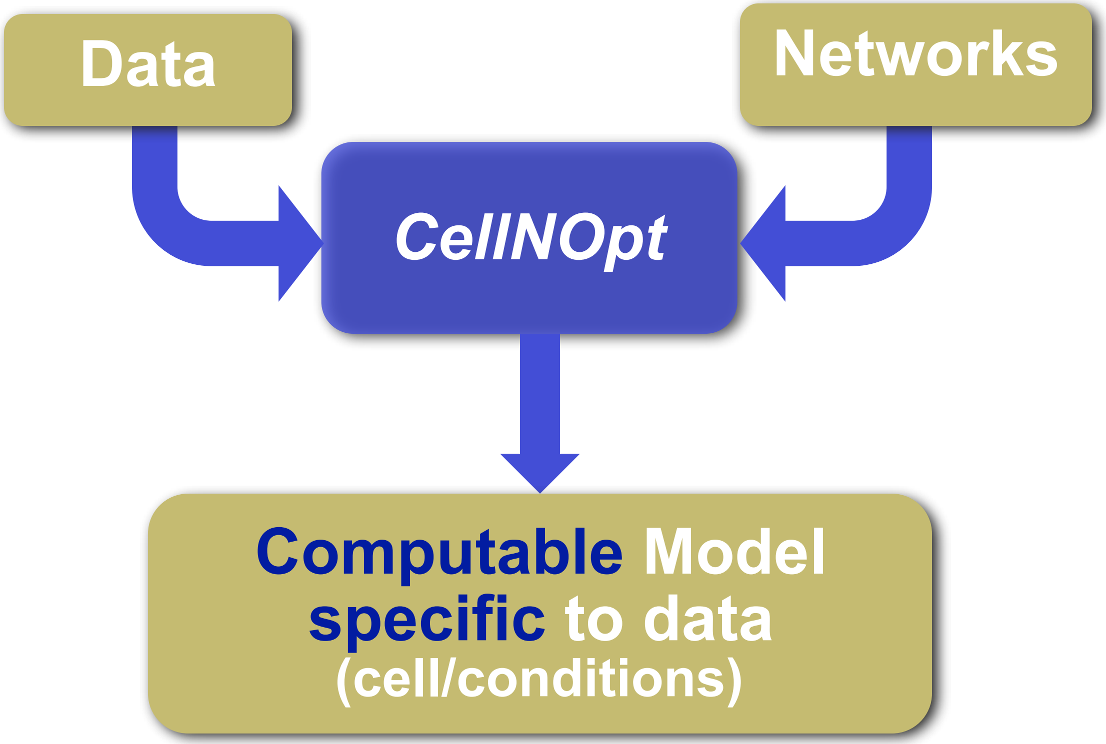


Logic modeling to link protein signaling networks with functional signaling data



Logic modeling to link protein signaling networks with functional signaling data




Data
Networks
CellNOpt
Computable Model
specific to data
(cell/conditions)

Where to get information from pathways?

Data

Networks

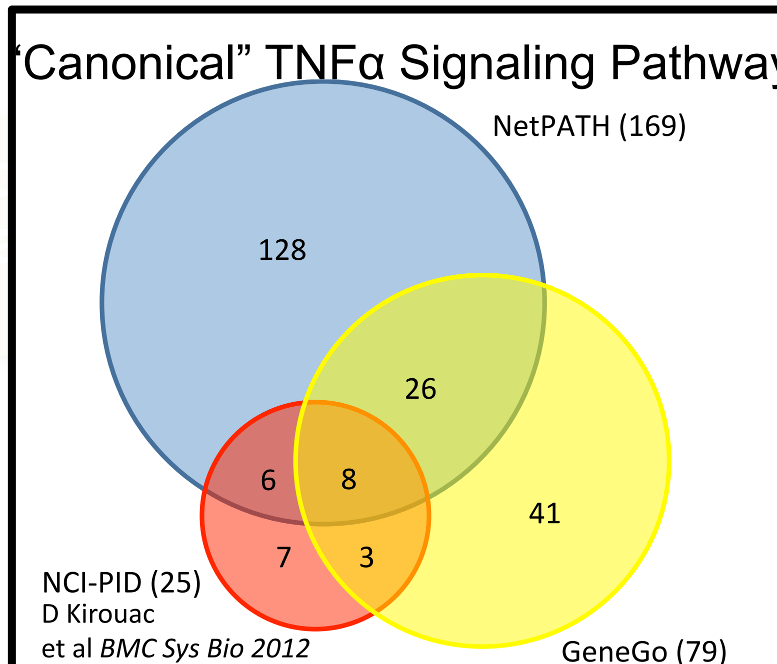
Multiple sources

Databases of curated pathways (Reactome, KEGG, Wiki-pathways, NCI-PID, Atlas Cancer, SignaLink, Path2Models, ...) incomplete, low overlap, different qualities

Important: standards (SBML, etc.)

Your own expertise/ literature review

Solution: OmniPath
in-house integrated repository of 24 pathway databases (~7k proteins,~21k PMIDs)
(free open-source; available



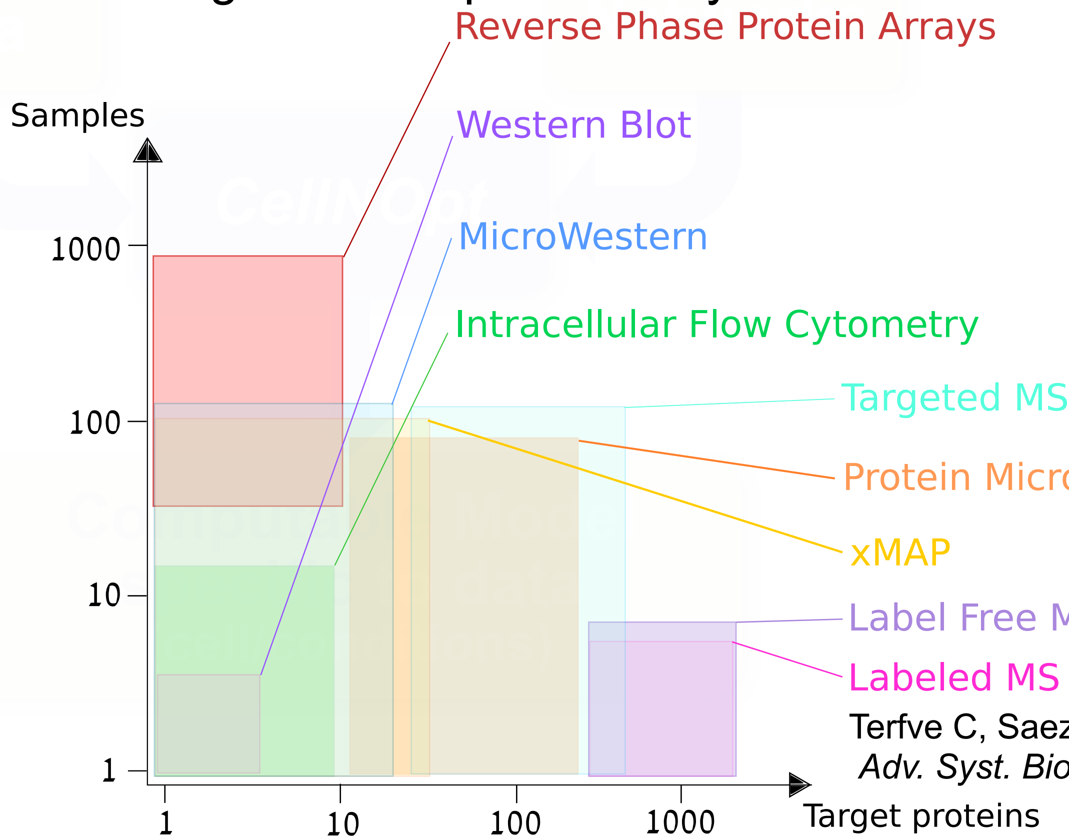
Use different phospho-proteomic & perturbation methods

Different data types

Different <u>measurement techniques</u>:

- Single cell: Imaging, flow cytometry, CytoF

  Broad coverage: Mass Spectrometry



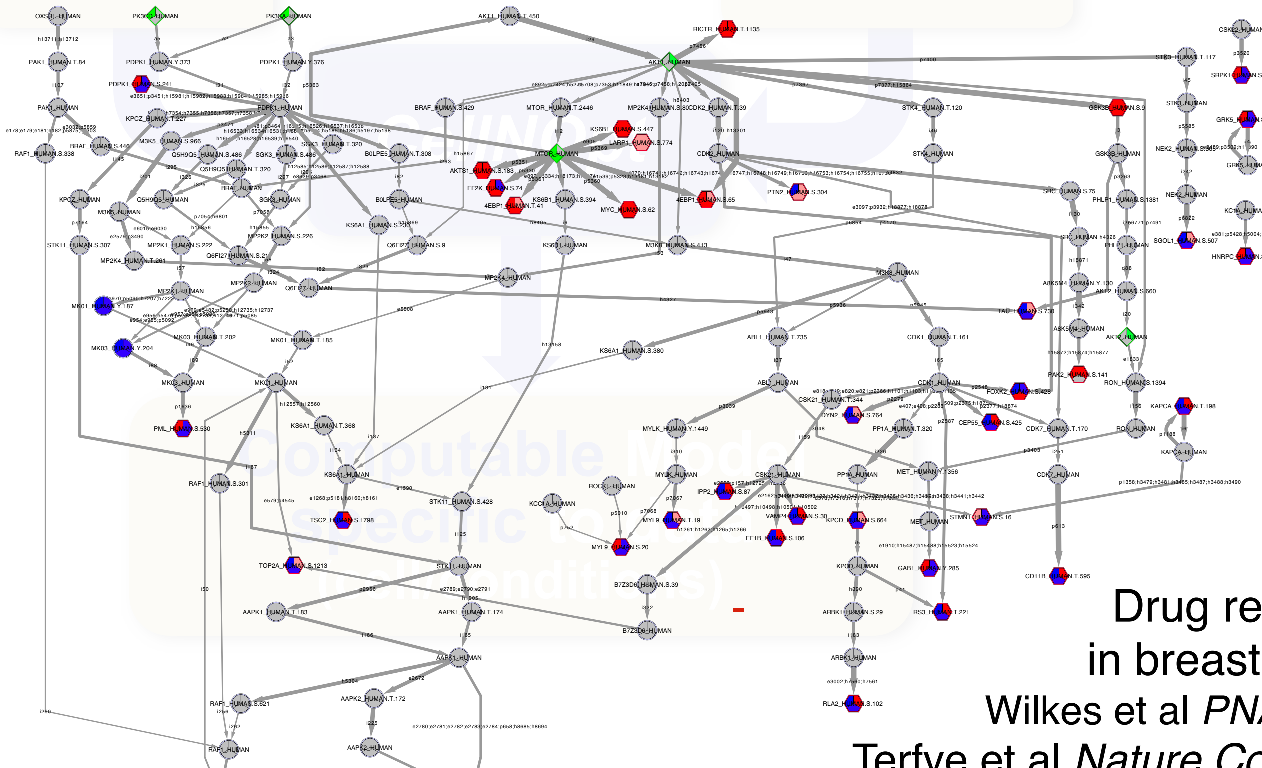
Terfve C, Saez-Rodriguez J, *Adv. Syst. Biol.*, 2012

<u>Orthogonal perturbations</u> - e.g. integration with siRNA screen to identify targets of phosphatases (Sacco et al. *Mol. Sys. Bio.*, 2012)

Use different phospho-proteomic & perturbation methods

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

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

Tool for modeling MS P-proteomics:
www.cellnopt.org/PHONEMeS

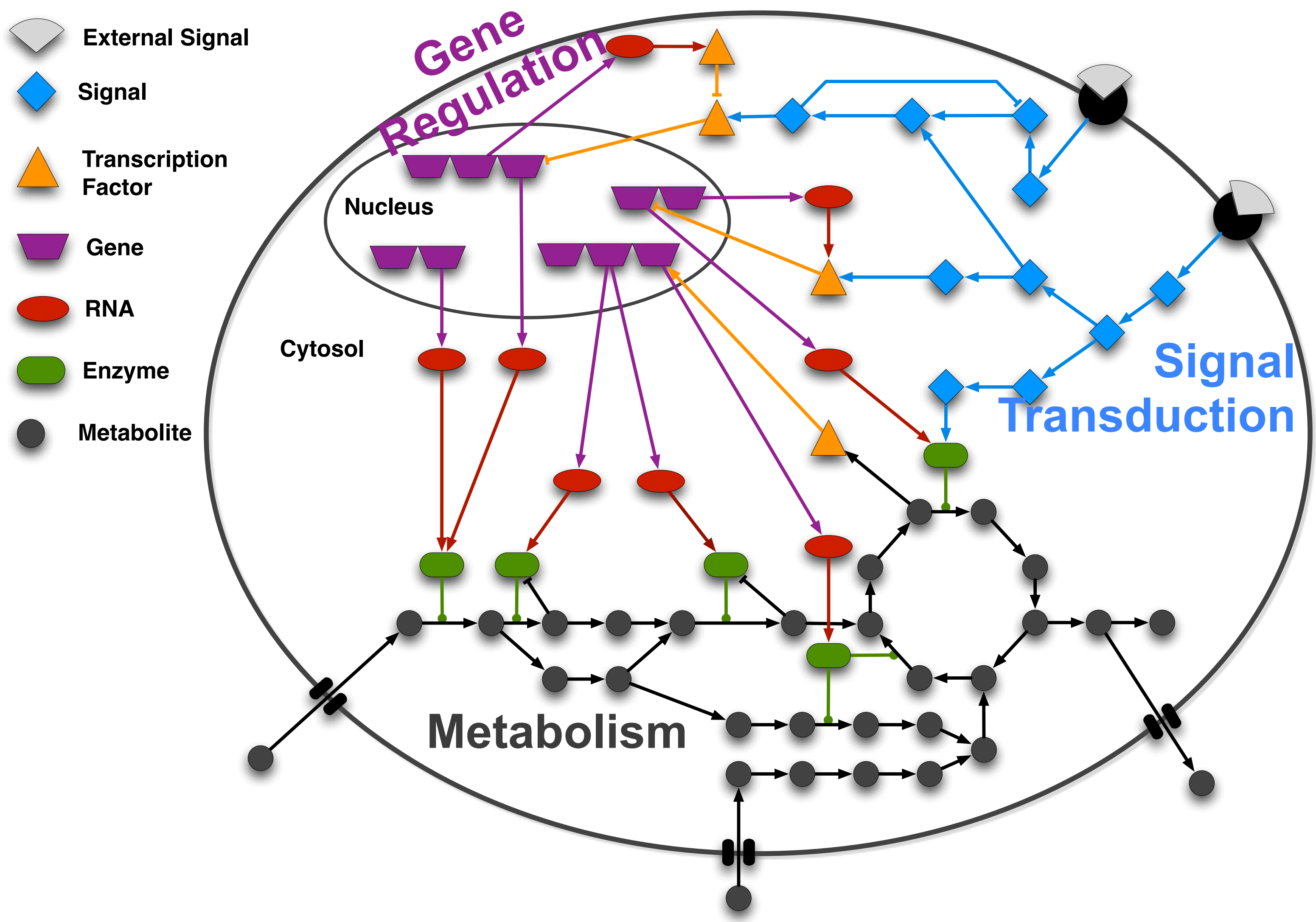
Use different phospho-proteomic & perturbation methods

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)

Models to link signalling to gene regulatory and metabolic networks



Emanuel Gonçalves
Michael Schubert
Luz Garcia

E Gonçalves *et al*. *Molecular Biosystems*, 9:1576-1583, 2013
Ryll et al *Biosystems* 2014:

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

A real example

Comparison of primary hepatocytes to 4 HCC cell lines

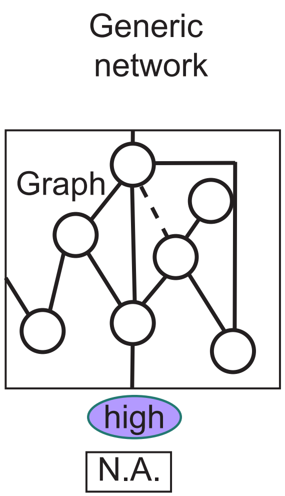
Primar

HepG

Hep3

Huh7

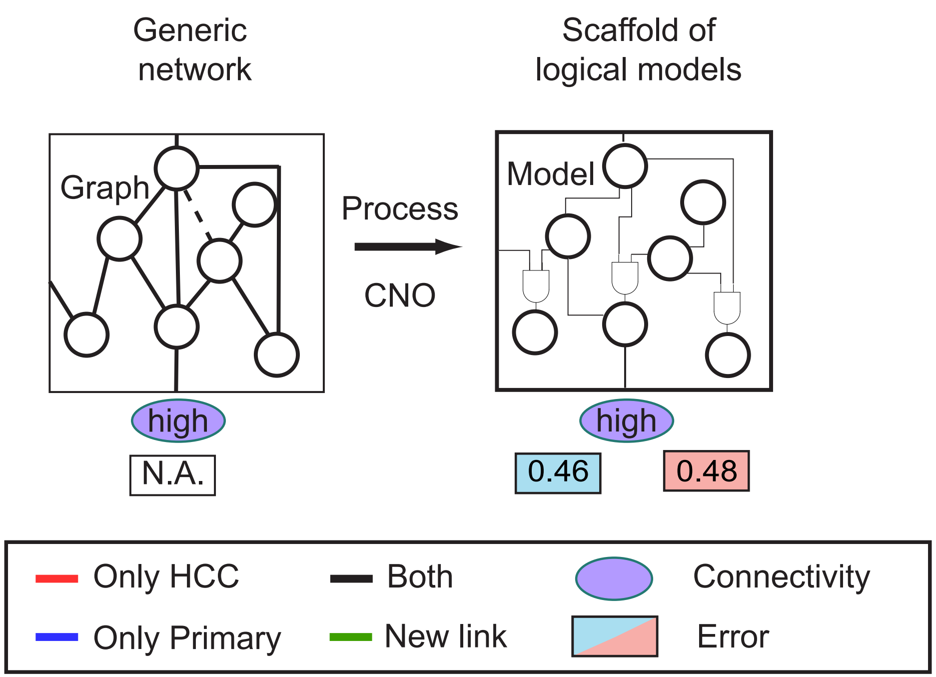
Focus



Saez-Rodriguez J, Alexopoulos LG, Zheng M, Morris MK, Lauffenburger DA, Sorger PK, *Cancer Research* 71(16), 2011

Comparison of primary hepatocytes to 4 HCC cell lines

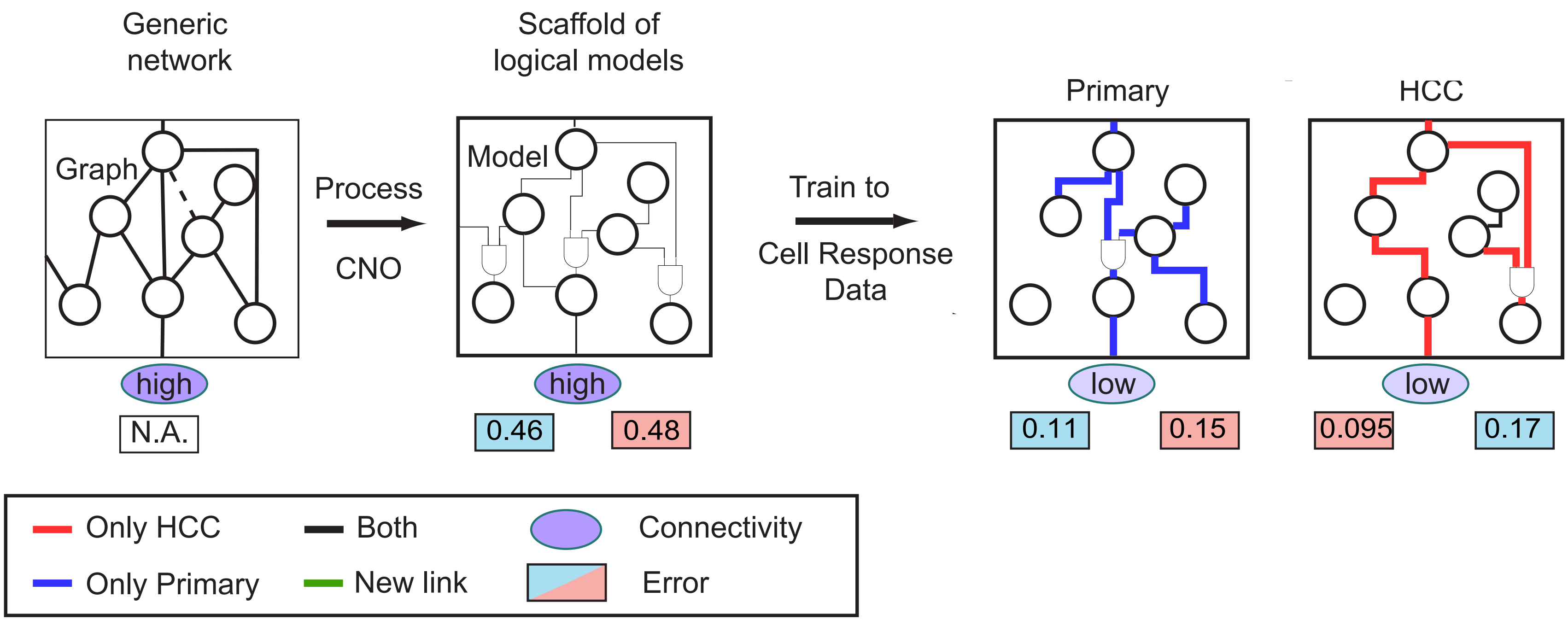
Primar | HepG | Hep3 | Huh7 | Focus



Saez-Rodriguez J, Alexopoulos LG, Zheng M, Morris MK, Lauffenburger DA, Sorger PK, *Cancer Research* 71(16), 2011

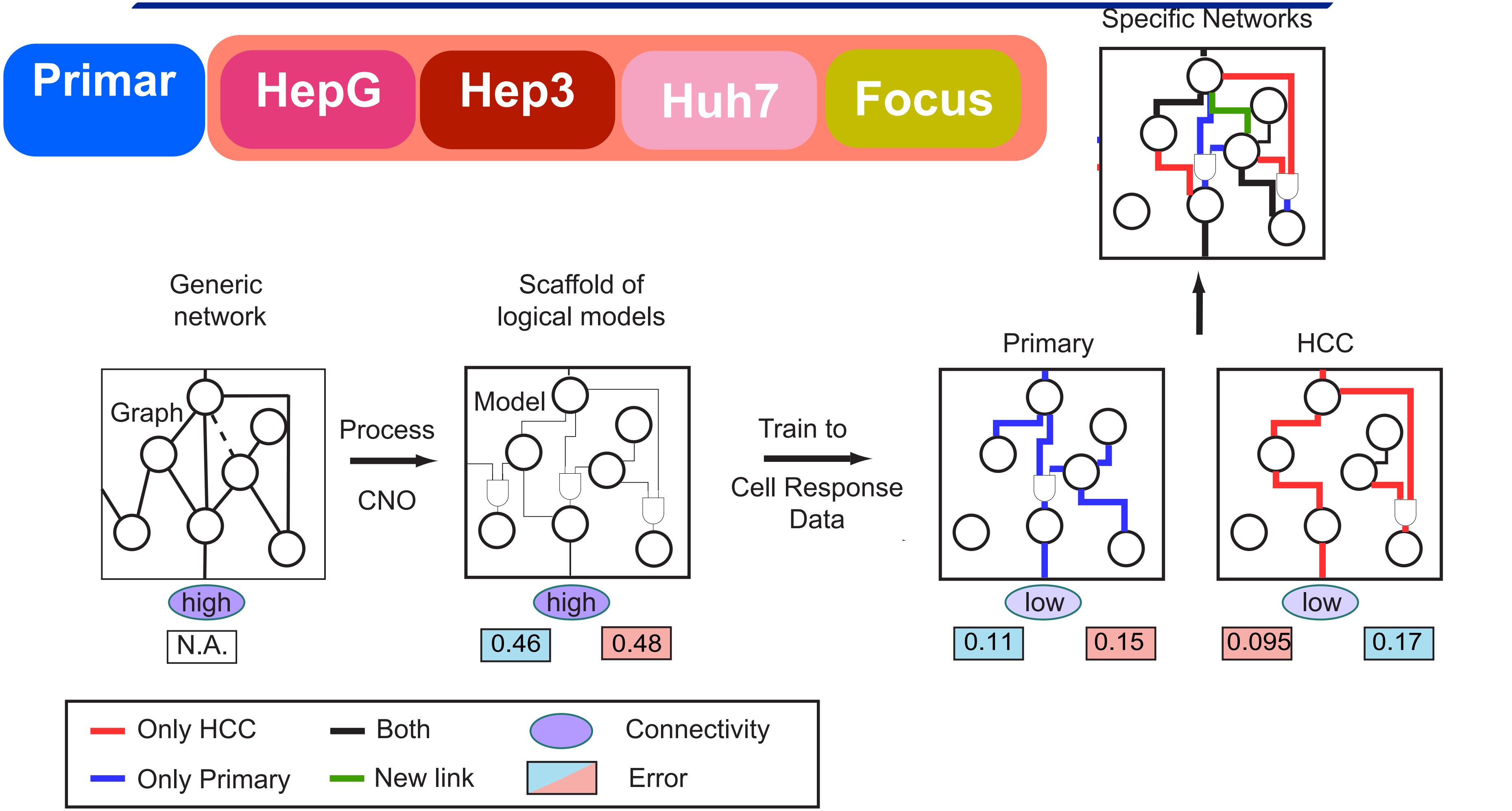
Comparison of primary hepatocytes to 4 HCC cell lines

Primar HepG Hep3 Huh7 Focus



Saez-Rodriguez J, Alexopoulos LG, Zheng M, Morris MK, Lauffenburger DA, Sorger PK, *Cancer Research* 71(16), 2011

Comparison of primary hepatocytes to 4 HCC cell lines

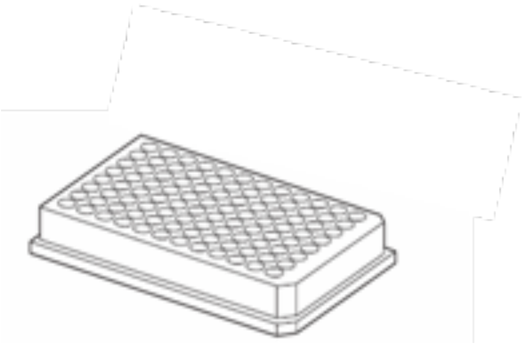


Saez-Rodriguez J, Alexopoulos LG, Zheng M, Morris MK, Lauffenburger DA, Sorger PK, *Cancer Research* 71(16), 2011

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

Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

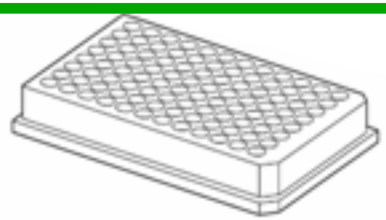
An example of a perturbation-based high-throughput data sets



Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.,* 9, 1849, 2010

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

Cue	➜ 7 **extracellular ligands** ➜ 7 specific **chemical inhibitors** (drugs)



Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

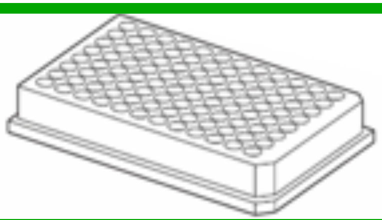
An example of a perturbation-based high-throughput data sets

Cue

➜ 7 **extracellular ligands**

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

at different times after stimulation



Signal

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

Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

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

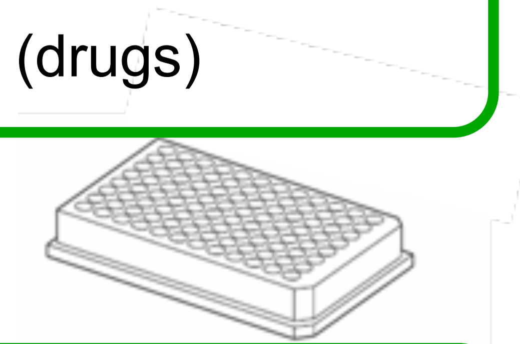
Primary human hepatocytes & HCC cell lines (HepG2, Hep3B, Huh7, Focus)

Cue

➜ 7 **extracellular ligands**

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

at different times after stimulation



Signal

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

using Luminex/xMAP (bead-based ELISA)

Response

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

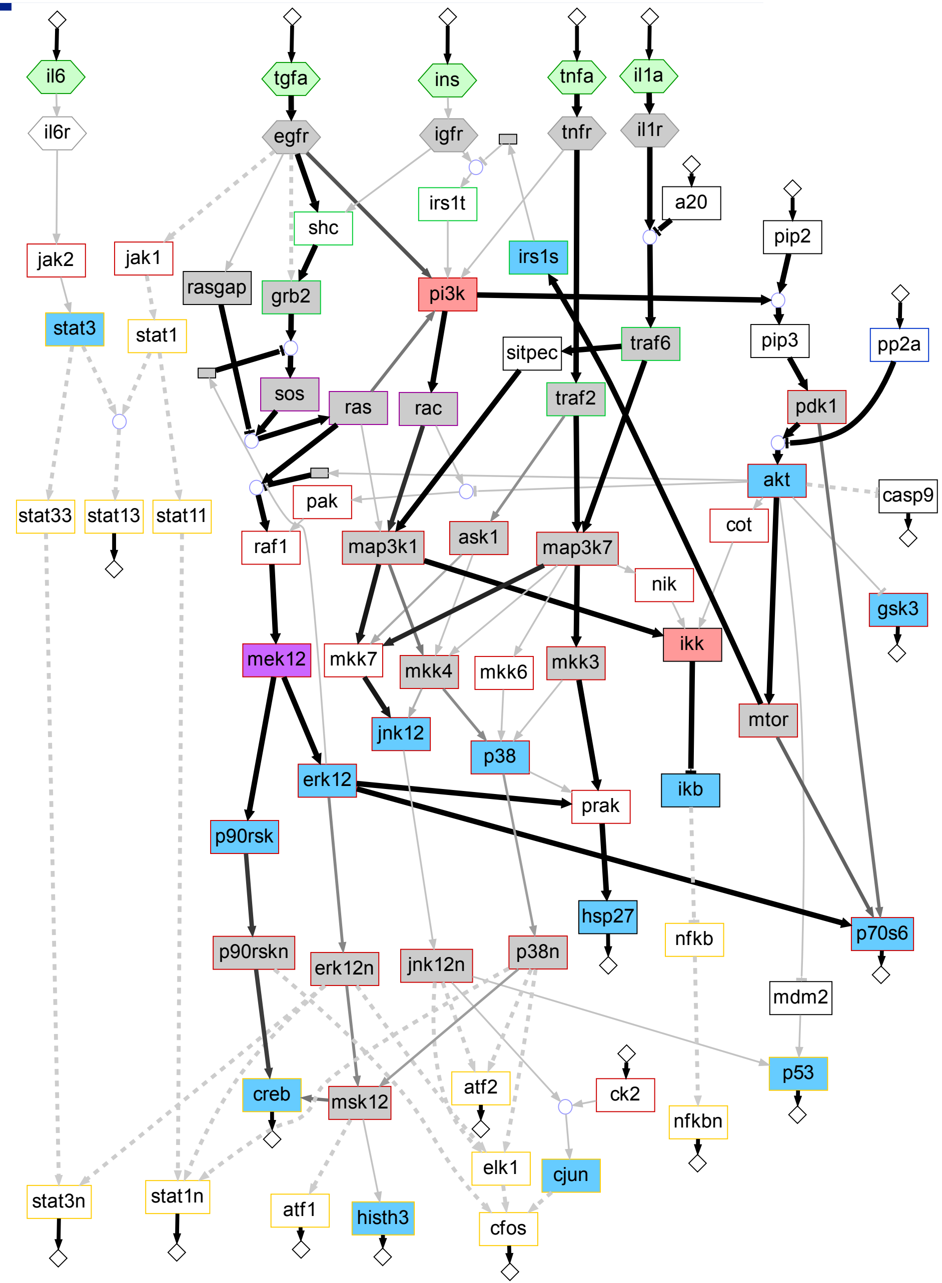
Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Prot.*, 9, 1849, 2010

COMBINE

Primar

Primary

RWTH AACHEN



Stimulus

Perturbation

Readout

Perturb&Read

Kept

Removed

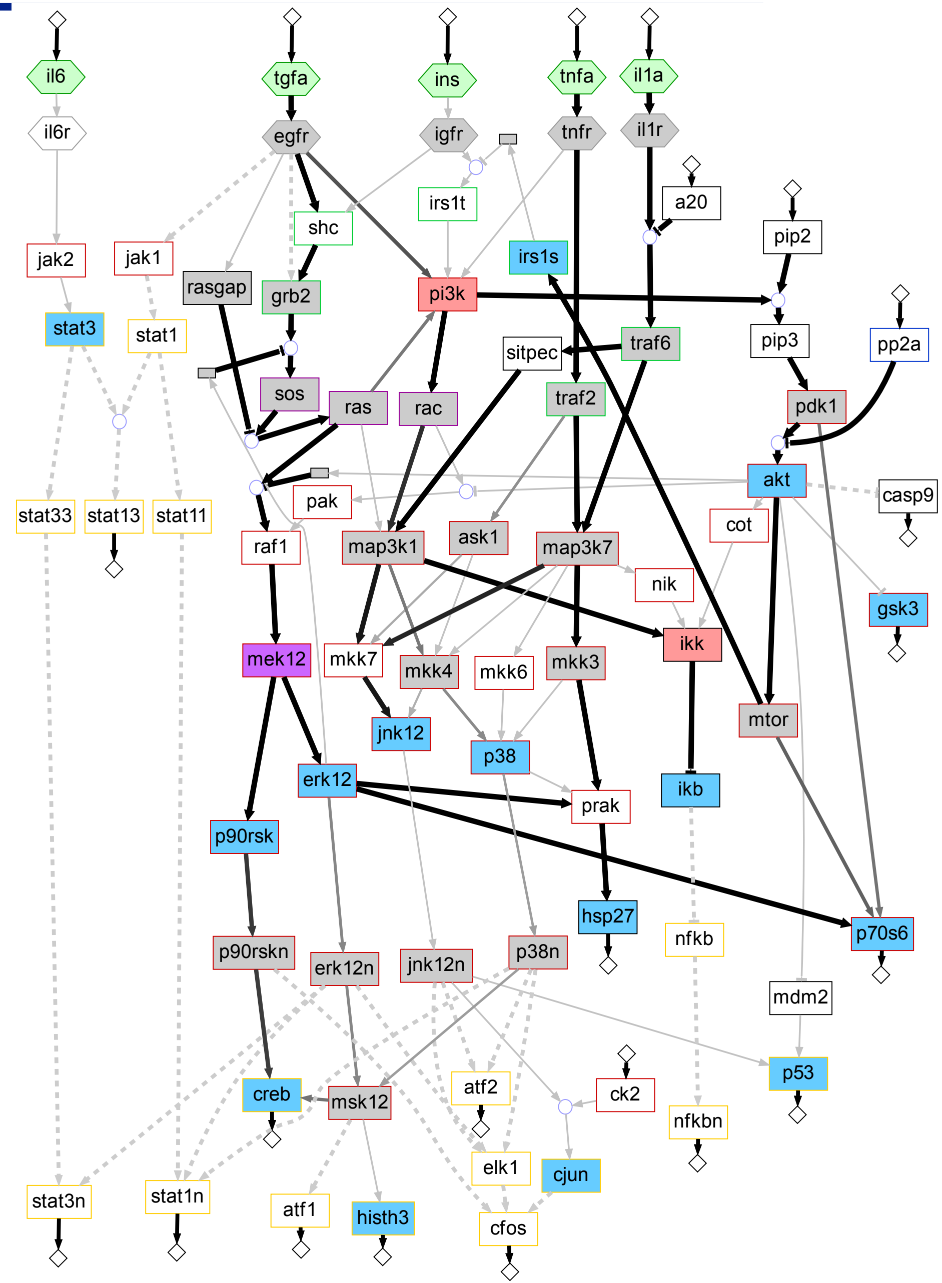
No effect

COMBINE

Primary

Primary

RWTH AACHEN



Stimulus

Perturbation

Readout

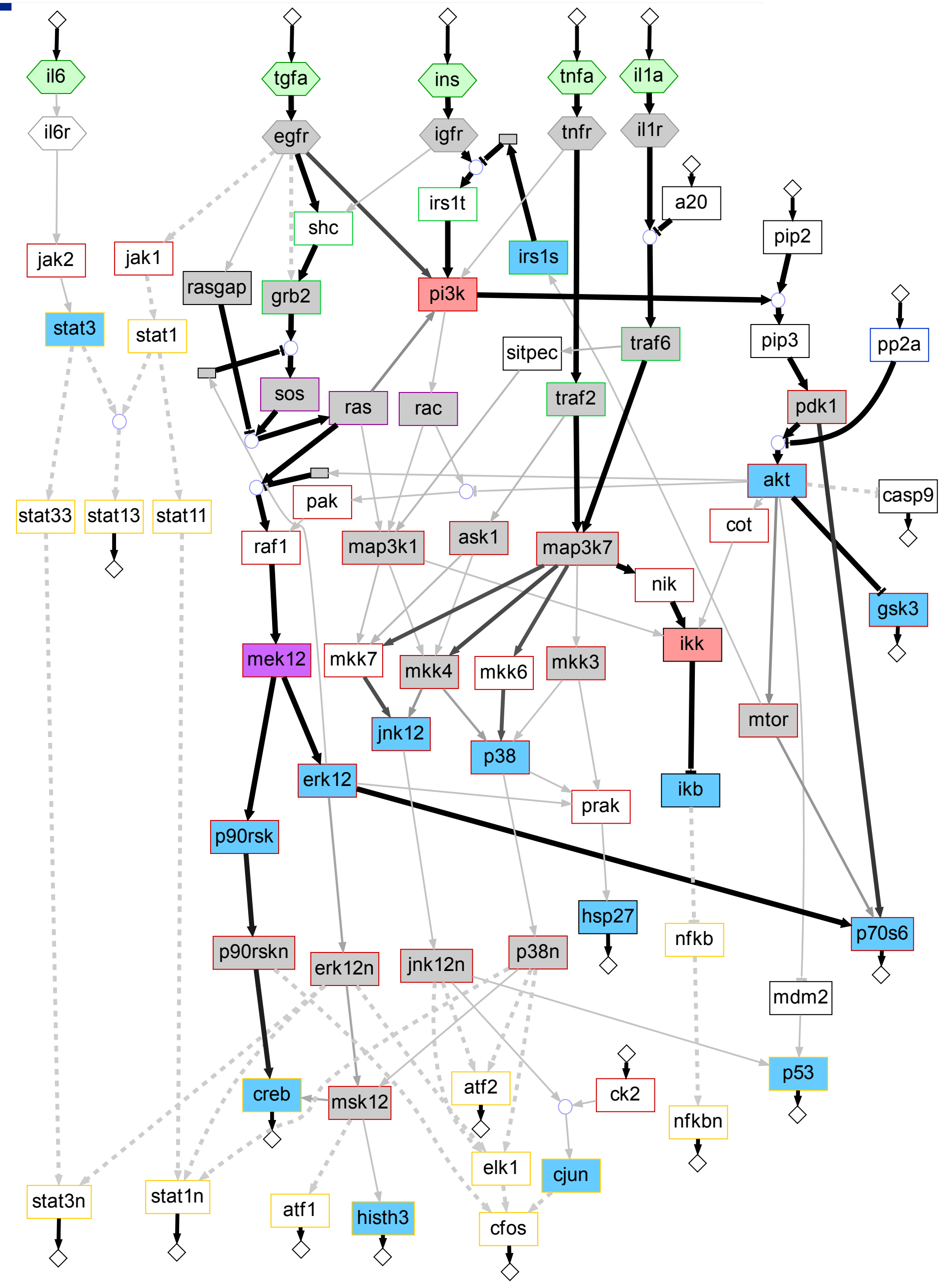
Perturb&Read

Kept

Removed

No effect

HepG2



Stimulus

Perturbation

Readout

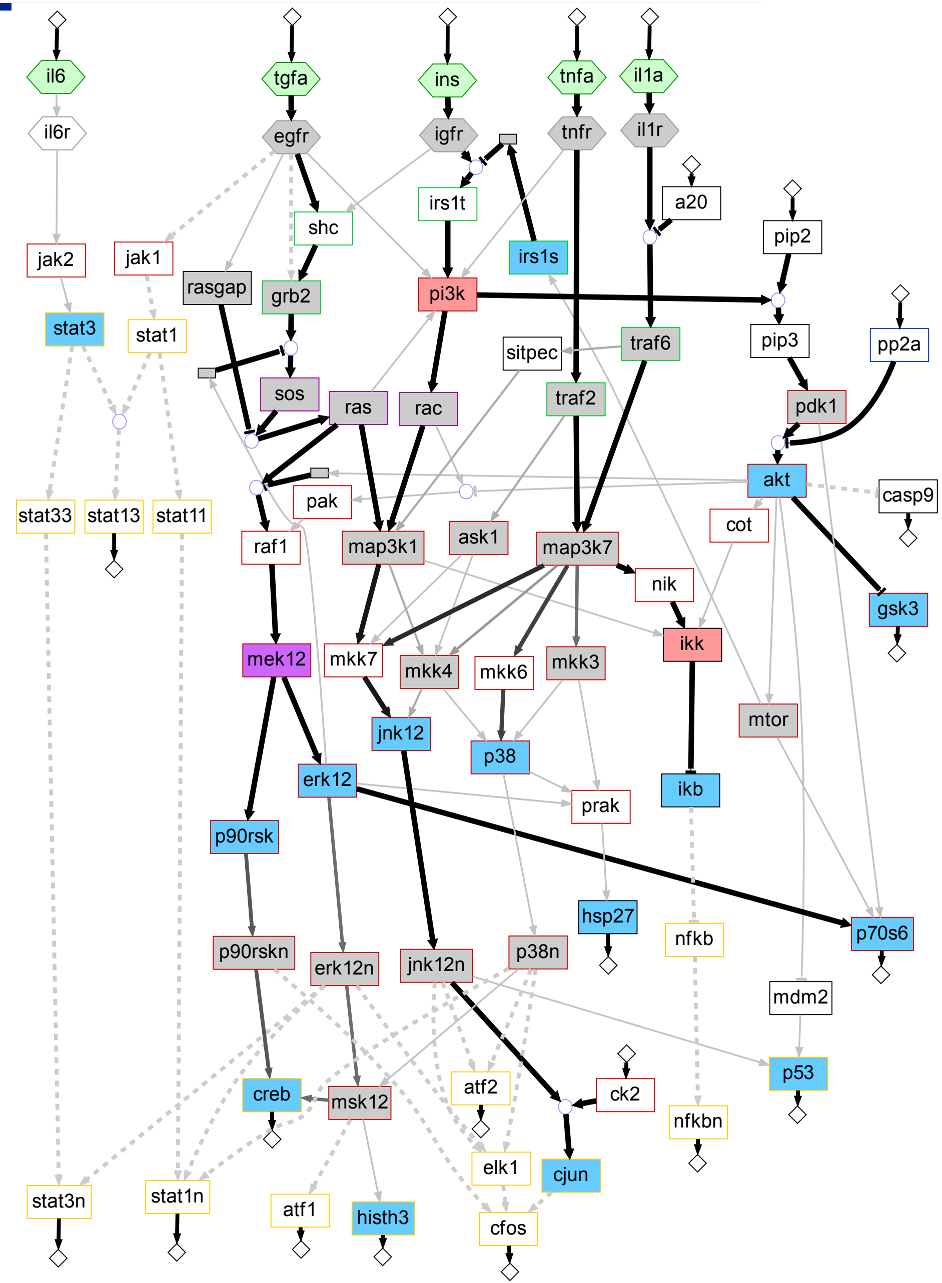
Perturb&Read

Kept

Removed

No effect

Hep3B



Stimulus

Perturbation

Readout

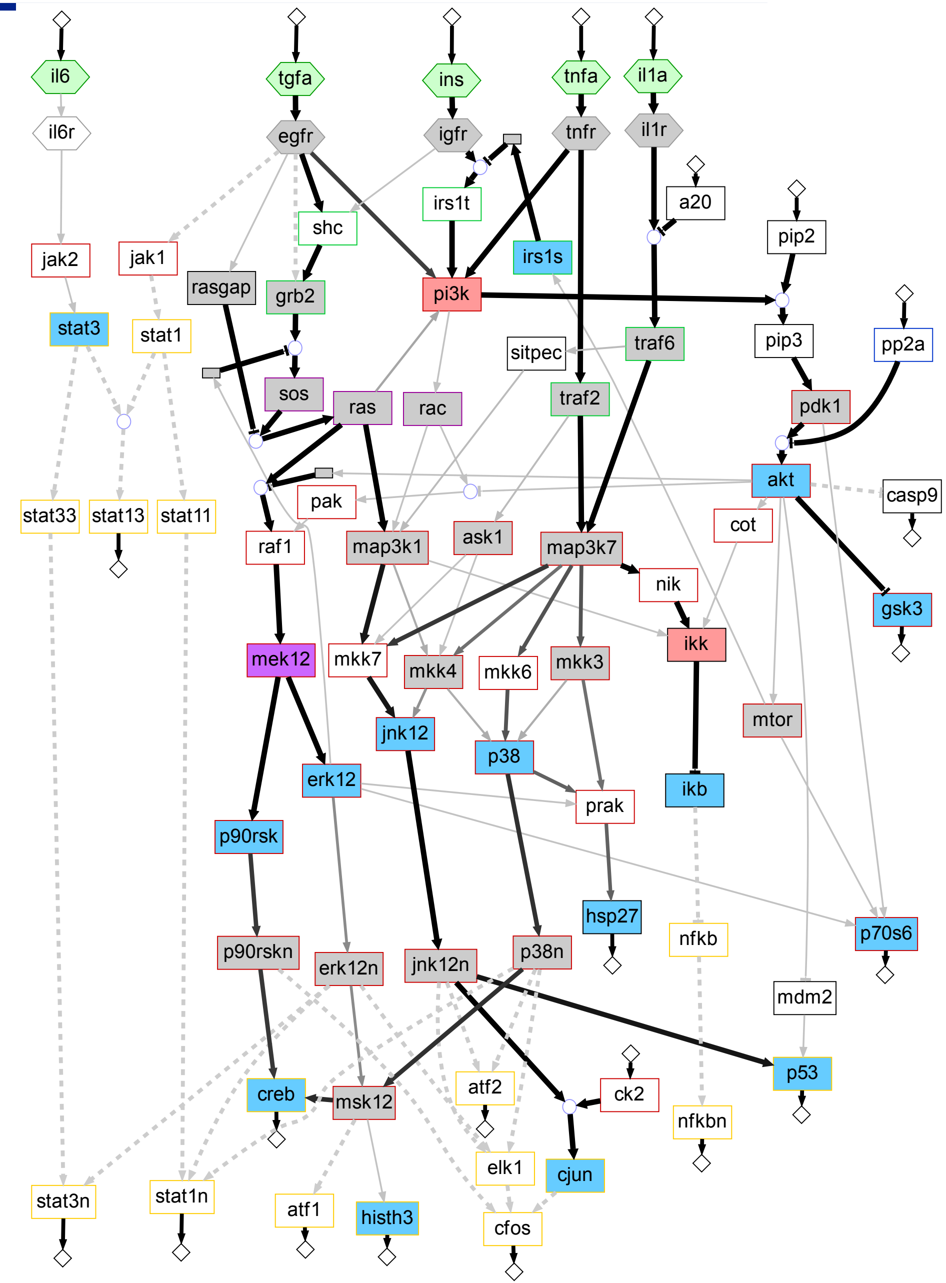
Perturb&Read

Kept

Removed

No effect

Huh7



Stimulus

Perturbation

Readout

Perturb&Read

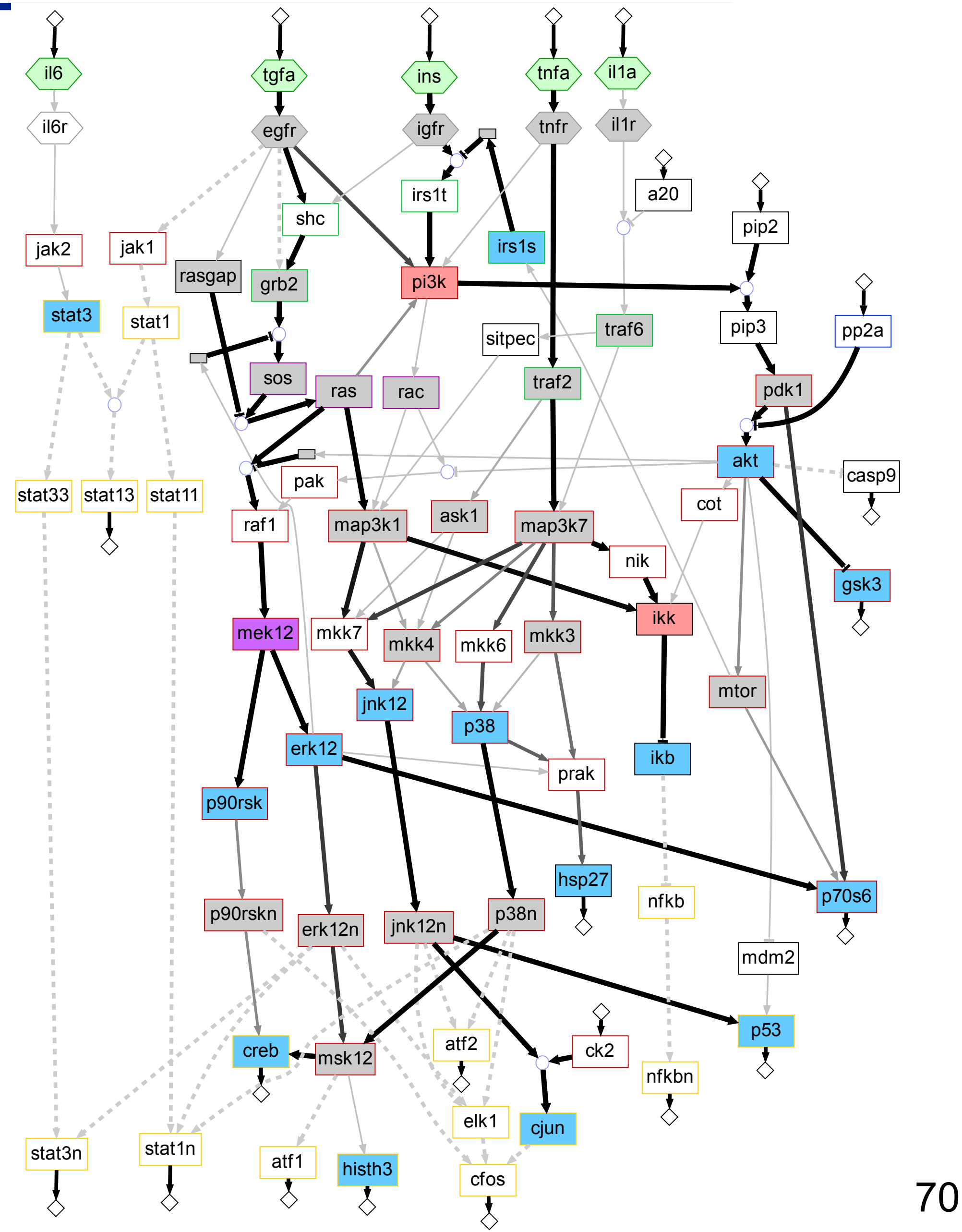
Kept

Removed

No effect

COMBINE

Focus



Stimulus

Perturbation

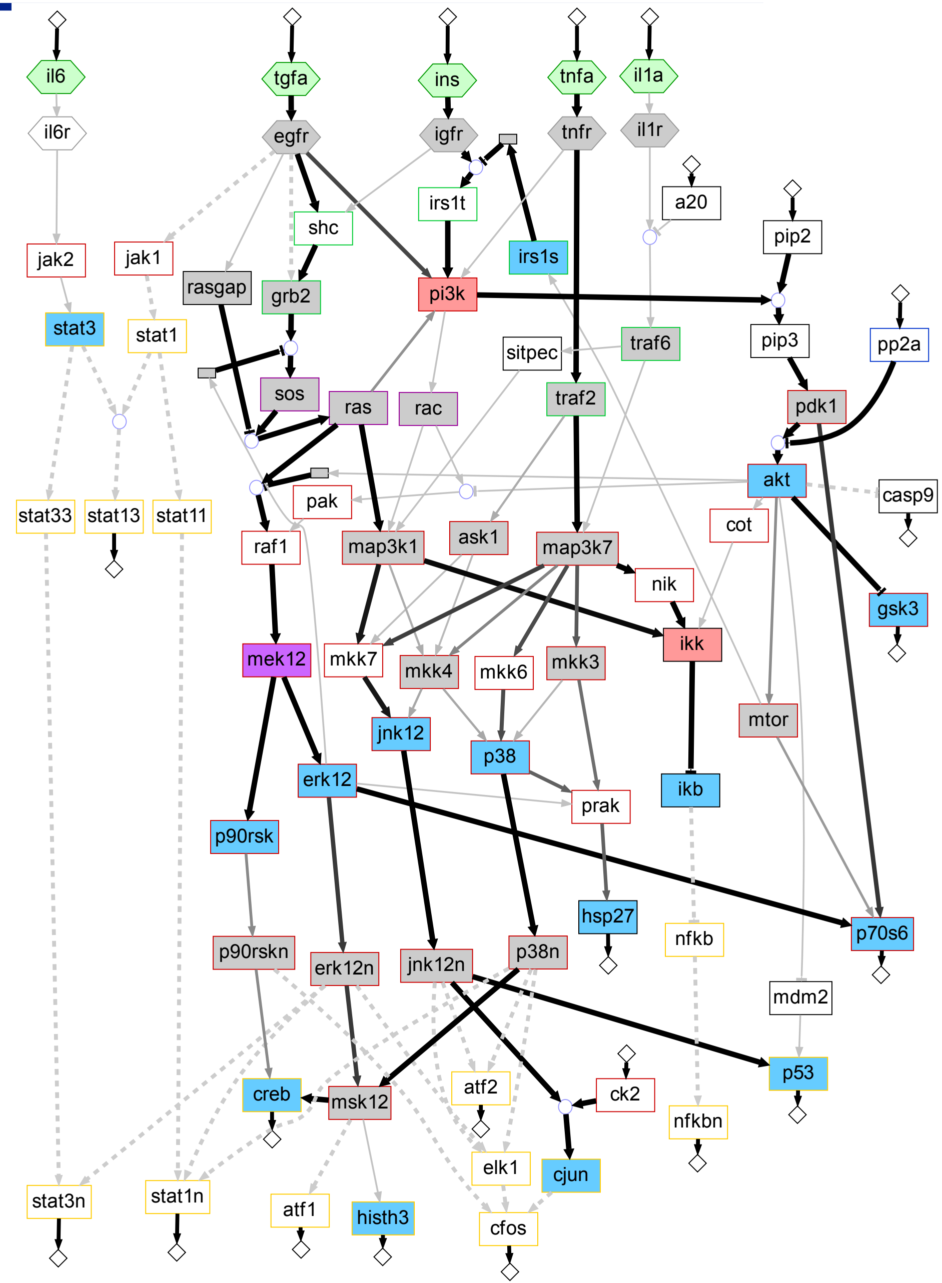
Readout

Perturb&Read

Kept

Removed

No effect



Stimulus

Perturbation

Readout

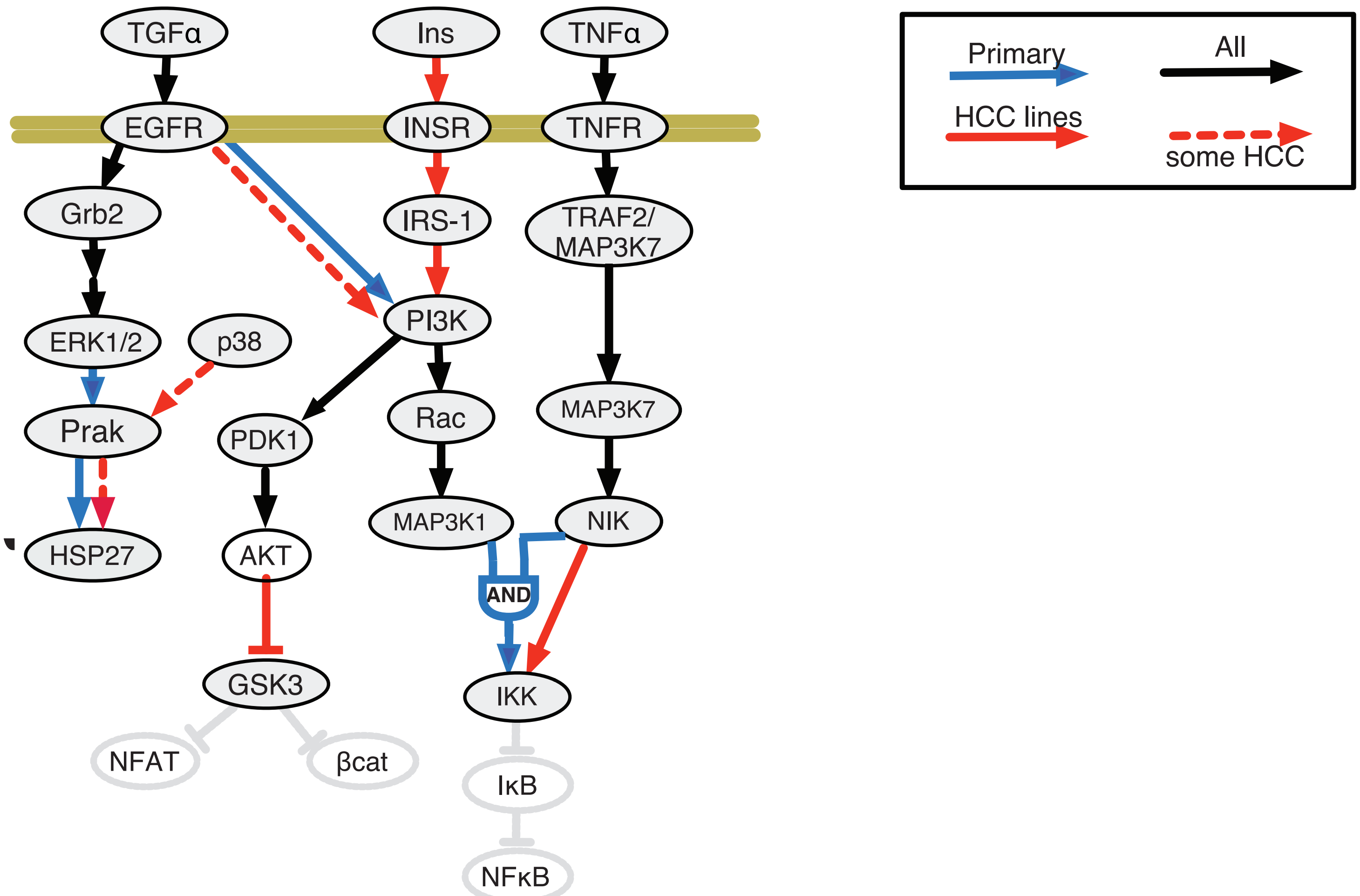
Perturb&Read

Kept

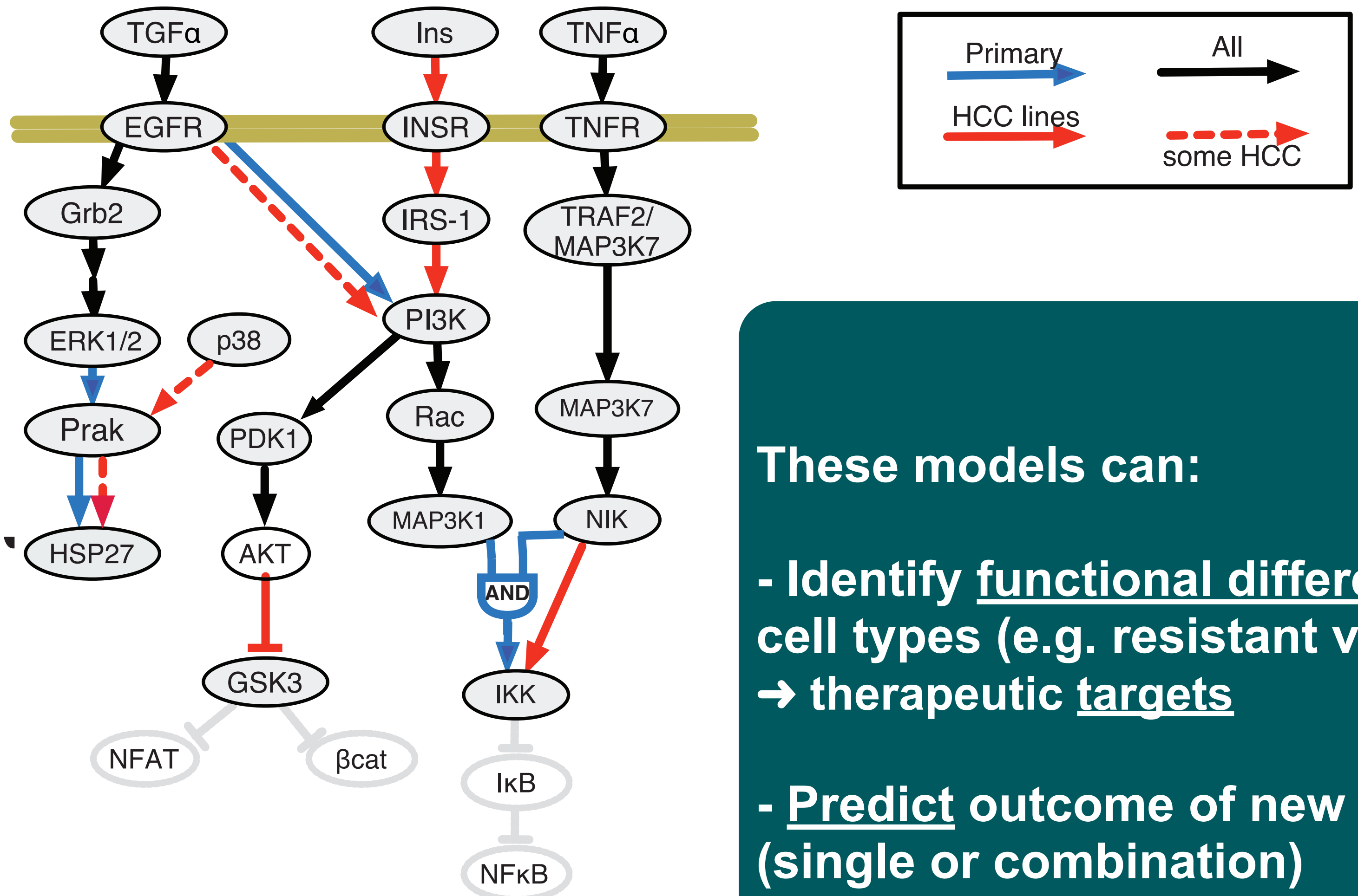
Removed

No effect

Logic models provide mechanistic insight in signal deregulation



Logic models provide mechanistic insight in signal deregulation



These models can:

- Identify <u>functional differences</u> between cell types (e.g. resistant vs. sensitive)
➜ therapeutic <u>targets</u>

- <u>Predict</u> outcome of new perturbations (single or combination)

- Characterize targets and <u>mode of action</u> of drugs (Mitsos et al. *PloS C.B.* 2009)

Acknowledgements

Francesco Iorio
Michael Menden
Michael Schubert
Luz Garcia Alonso
Damien Arnol
Denes Turei
Aidan MacNamara

Federica Eduati
Thomas Cokelaer
Marti Bernardo
Johannes Stephan
Camille Terfve
Angeliki Kalamara
Fatemeh Zamanzad
Emanuel Gonçalves

CoLoMoTo consortium sp. Claudine Chaouiya (IGC) Tomas Helikar(UN)

Ioannis Melas Leo Alexopoulos (NTUA) Alexander Mitsos (RWTH)

Regina Samaga   Steffen Klamt (MPI)

Jose Egea (U. Cartagena)
Alex Villlaverde David Henriques Julio Banga (CSIC)

Carito Guziolowski   Santiago Videla   Anne Siegel (CNRS)

Melody Morris Doug Lauffenburger (MIT)

Jeremy Muhlich Peter K Sorger (HMS)

EMBL-EIPOD   MRC   EU-7FP   GSK

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

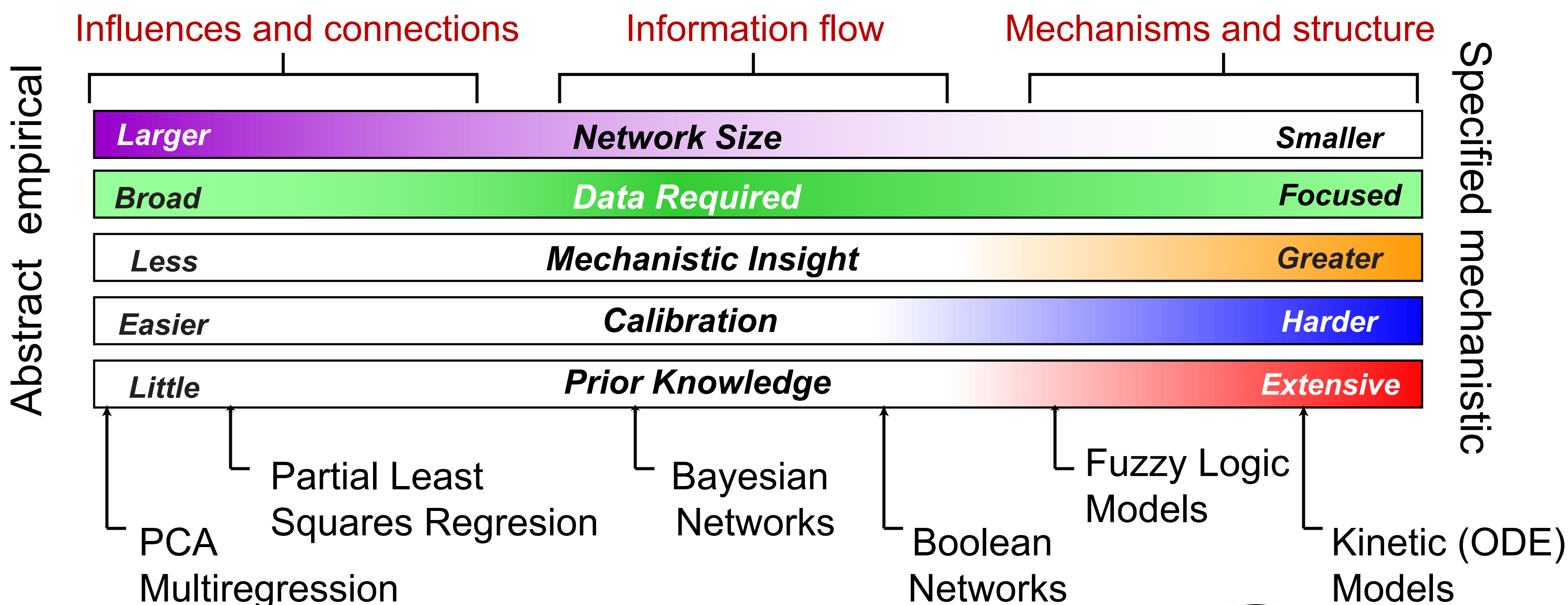
How to...

- 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

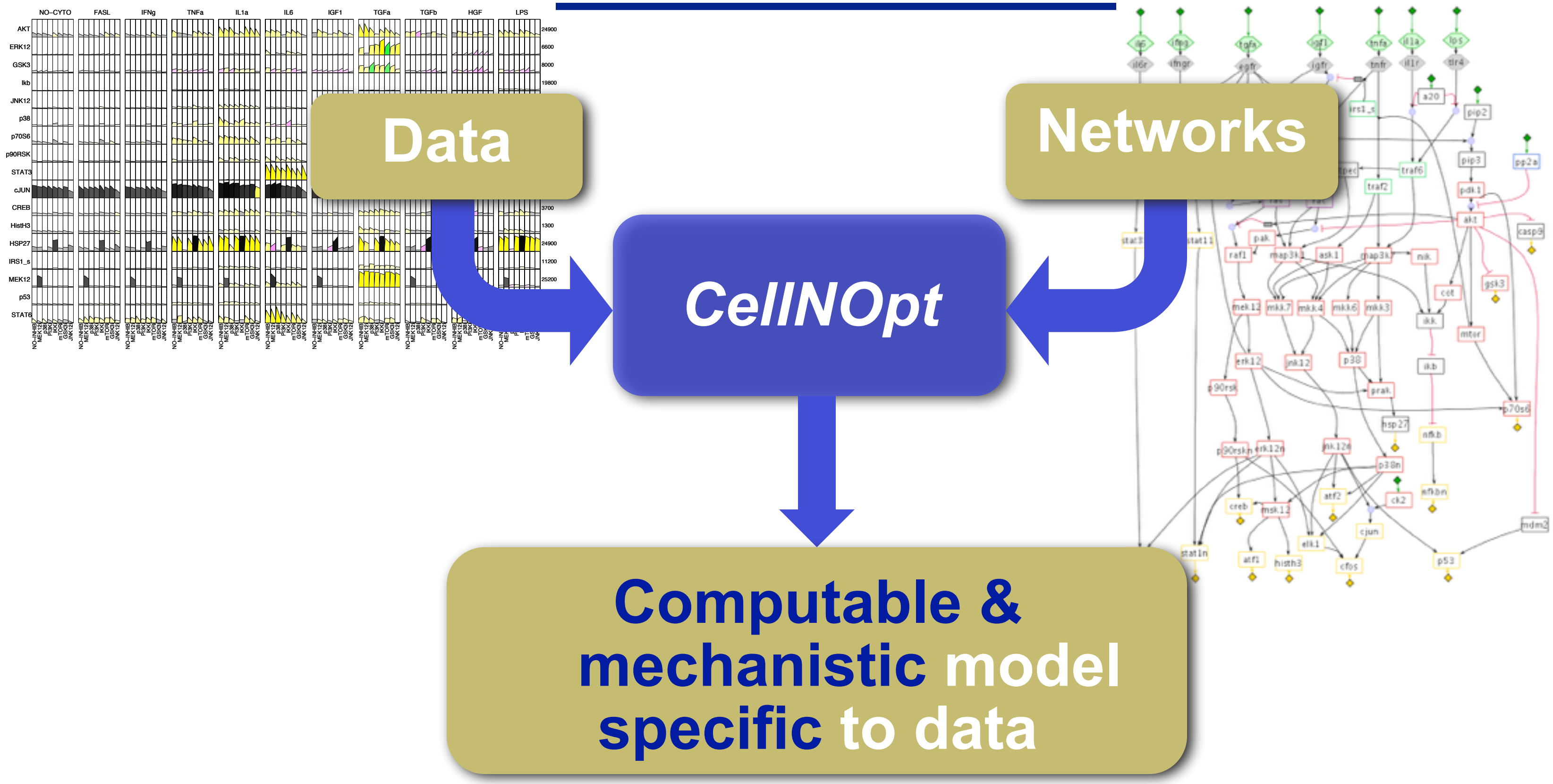
Used logic modelling & applications to signalling, but general principles hold for other modelling approaches & applications

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