

FlexFlux, a flexible Java framework to perform flux balance analysis with regulatory constraints

Lucas Marmiesse Remi Peyraud Ludovic Cottret

INRA CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM),
UMR441, F-31326 Castanet-Tolosan, France

January 21, 2014

Ralstonia solanacearum

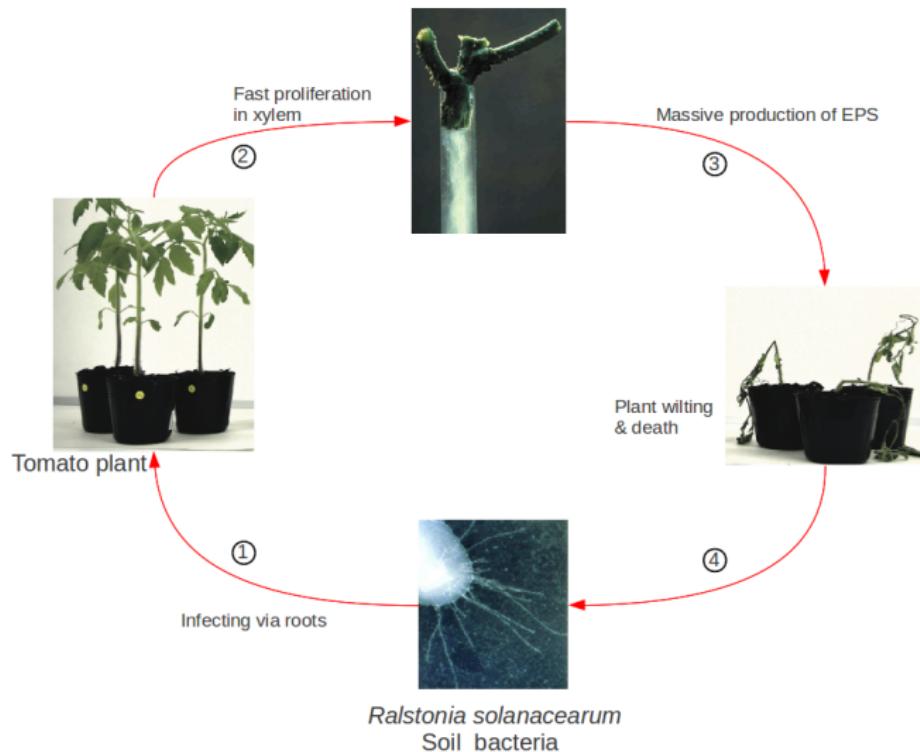


Figure: Infectious cycle of *Ralstonia solanacearum*

Ralstonia solanacearum's metabolic network

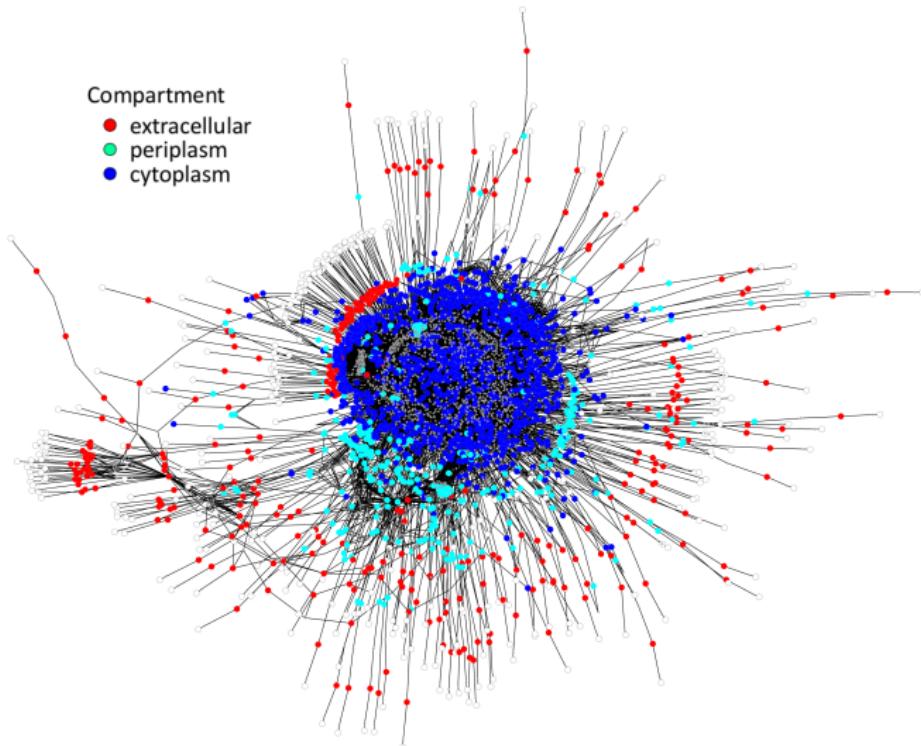


Figure: Reconstruction of *Ralstonia solanacearum*'s metabolic network. It contains 2304 reactions and 1289 unique metabolites

Biological questions

- ▶ How does the bacteria regulate its virulence functions ?
- ▶ Which part of the metabolism is used during the different stages of the infection ?
- ▶ How does the bacteria use plant's resources to its advantage ?

What analysis to make ?

- ▶ Analyze reactions fluxes.
- ▶ Flux variability.
- ▶ Time dependent analyses.
- ▶ Multi-objective analyses.

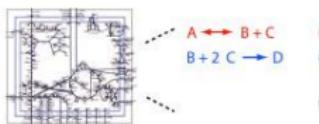
Flux balance analysis (FBA)

What is FBA ?

- ▶ Mathematical approach to analyze genome scale metabolic networks.
- ▶ Based on constraints.
- ▶ No kinetic parameters needed.
- ▶ Well validated by different methods.

Flux balance analysis

a Curate metabolic reactions



b Formulate S matrix

Reactions						
Metabolites	A	B	C	D	...	n
A	1	-1	0	0	...	0
B	0	1	0	0	...	0
C	0	0	1	-2	...	0
D	0	0	0	1	...	0
...	m

c Apply mass balance constraints

$$\begin{matrix} S \quad (m \times n) \\ \begin{bmatrix} 1 & -1 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & -2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \end{matrix} * \begin{matrix} v \quad (n \times 1) \\ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \end{bmatrix} \end{matrix} = 0 \rightarrow \begin{matrix} m \text{ mass balance} \\ \text{equations} \\ -v_1 + v_2 = 0 \\ v_1 - v_3 + \dots = 0 \\ v_3 - 2v_4 + \dots = 0 \\ v_5 + \dots = 0 \\ \dots \end{matrix}$$

d Define objective function Z

$$Z = c^T \begin{matrix} v \quad (n \times 1) \\ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \end{bmatrix} \end{matrix} * \begin{matrix} \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \\ \text{sets reaction 1 as} \\ \text{the objective} \end{matrix}$$

e Optimize Z using linear programming

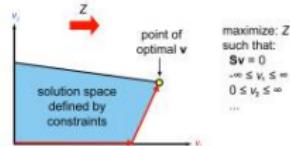


Figure: Creation of a FBA problem.¹

¹Jeffrey D. Orth et al. "What is flux balance analysis?" In: *Nature Biotechnology* (2010).

Example of FBA on a simple network

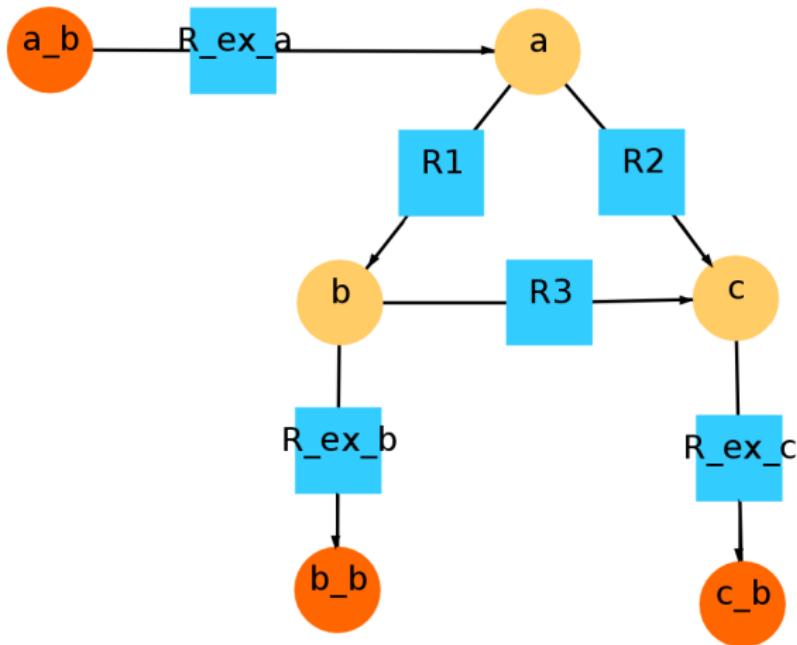
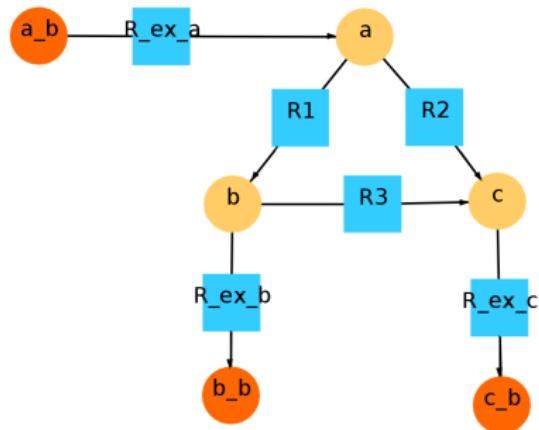


Figure: Simple network example for FBA.

Example of FBA on a simple network

	R1	R2	R3	R_ex_a	R_ex_b	R_ex_b
a	-1	-1	0	1	0	0
b	1	0	-1	0	-1	0
c	0	1	1	0	0	-1

Table: Stoichiometry matrix



$$-R_1 + -R_2 + R_{\text{ex_a}} = 0.0$$

$$R_1 + -R_3 + -R_{\text{ex_b}} = 0.0$$

$$R_2 + R_3 - R_{\text{ex_c}} = 0.0$$

Example of FBA on a simple network

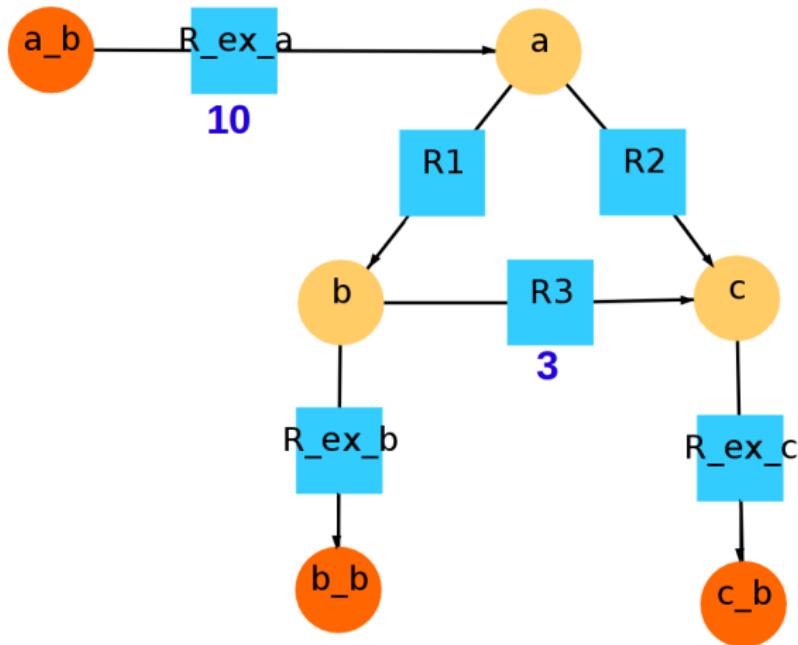


Figure: Constraints on fluxes.

Example of FBA on a simple network

Objective function : maximize R_{ex_c}

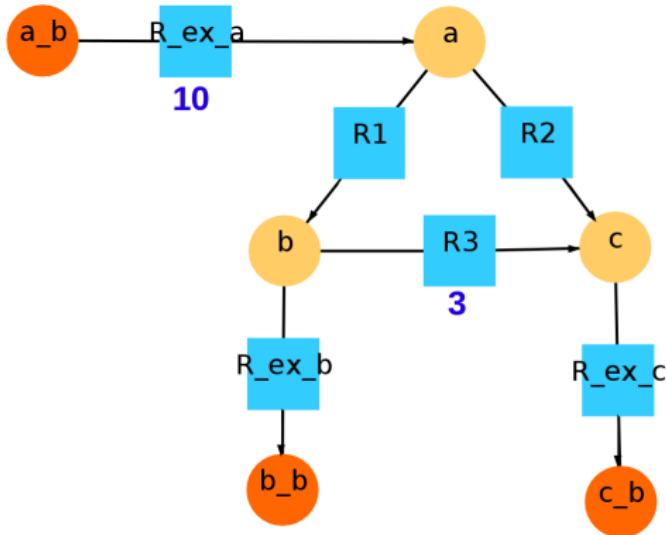


Figure: Constraints on fluxes.

Example of FBA on a simple network

Objective function : maximize R_{ex_c}

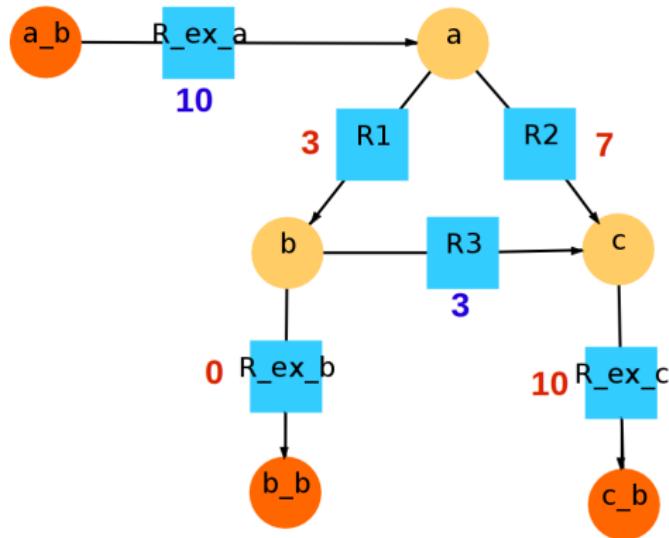


Figure: Optimization of R_{ex_c} .

Example of FBA on a simple network

Objective function : maximize R_{ex_b}

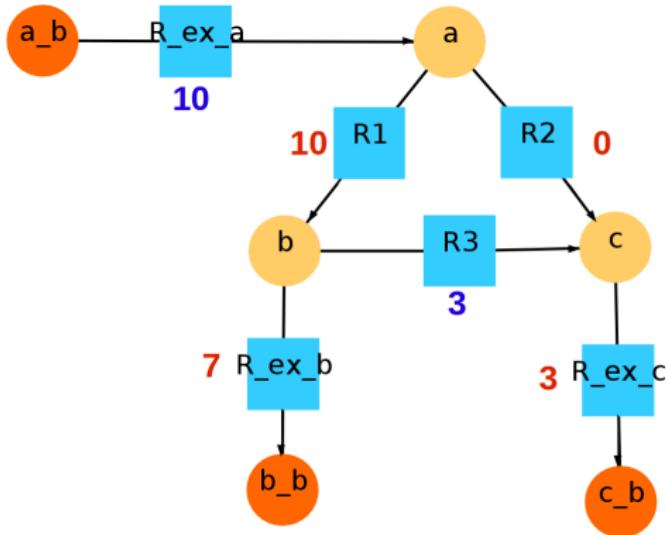


Figure: Optimization of R_{ex_b} .

State of the art

Our needs

- ▶ Easy to use.
- ▶ Regulation.
- ▶ Modularity.
- ▶ Speed.
- ▶ Java.

FlexFlux

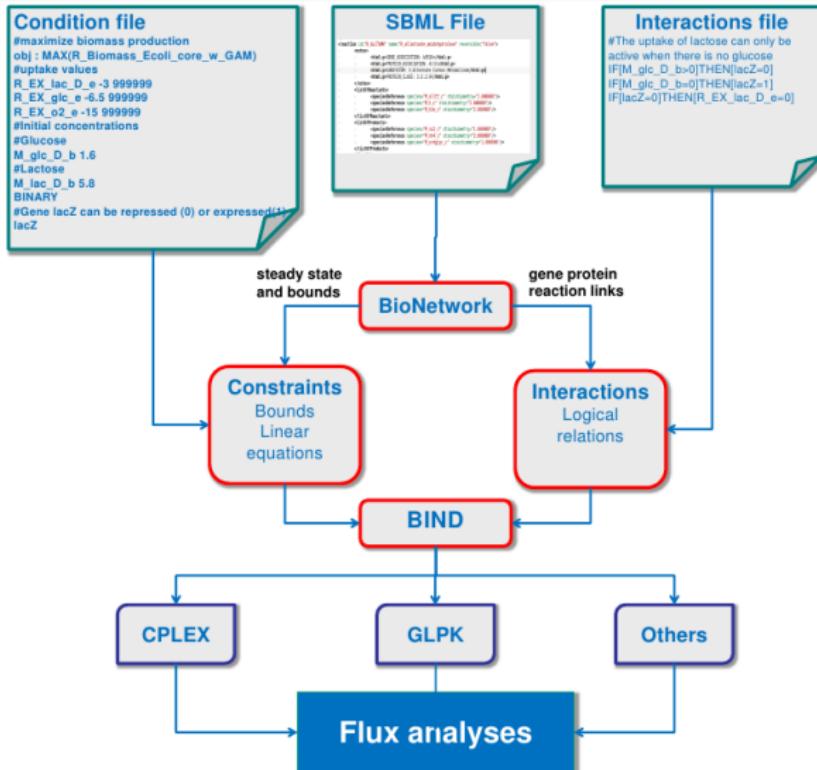
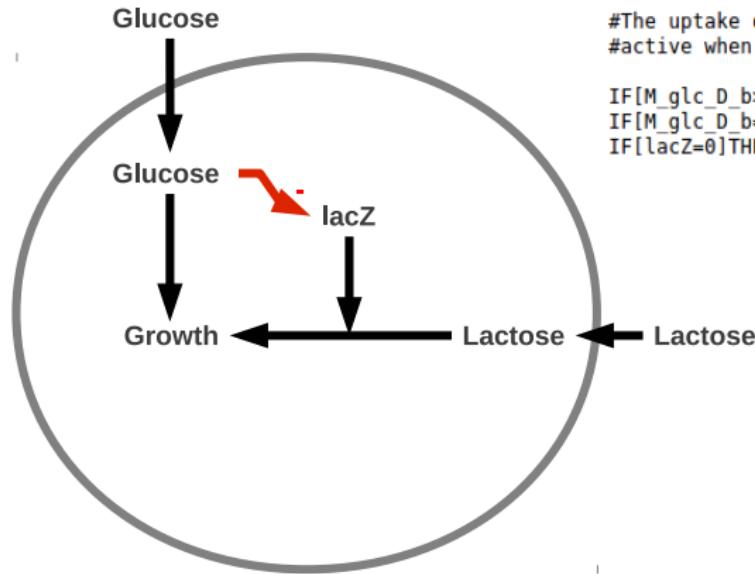


Figure: Architecture of FlexFlux

Interactions in FlexFlux

- ▶ They are written like this : IF[..]THEN[..], example :
IF [G1>0] THEN [R1=4]
- ▶ They can be more complicated :
IF [(G1>0 AND G2>0) OR G3=0] THEN [R1=4]
- ▶ If the first part is true, the second part is translated into a constraint.
- ▶ Some solvers can handle the interactions themselves.

FlexFlux permits to model regulations easily



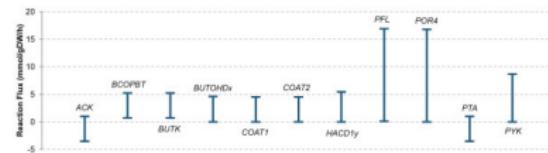
#The uptake of lactose can only be
#active when there is no glucose

```
IF[M_glc_D_b>0]THEN[lacZ=0]  
IF[M_glc_D_b=0]THEN[lacZ=1]  
IF[lacZ=0]THEN[R_EX_lac_D_e=0]
```

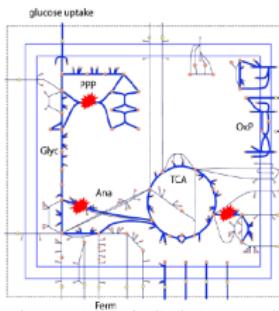
Figure: Example of a time dependent analysis easy to set up

9 functions already developed in FlexFlux

- ▶ Optimization of an objective.
- ▶ Flux variability analysis (FVA).

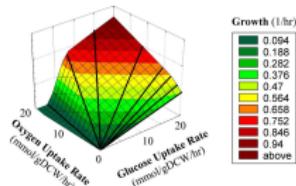


- ▶ Comparison of two FVA's in two conditions.
- ▶ KO analysis.

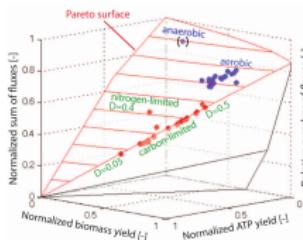


9 functions already developed in FlexFlux

- ▶ Dead reactions.
- ▶ Phenotypic phases analysis.²



- ▶ Pareto analysis.³



- ▶ Time dependent analysis.

²Natalie C Duarte et al. "Integrated analysis of metabolic phenotypes in *Saccharomyces cerevisiae*". In: *BMC Genomics* (2004).

³Robert Schuetz et al. "Multidimensional Optimality of Microbial Metabolism." In: *Science* (2012).

Flux minimization

- ▶ Condition file :

```
obj : MAX(R_Biomass)  
obj : MIN(FluxSum)
```

- ▶ Command line :

```
FlexFluxFBA -s coli.xml -cond conditions.txt -plot
```

Flux minimization

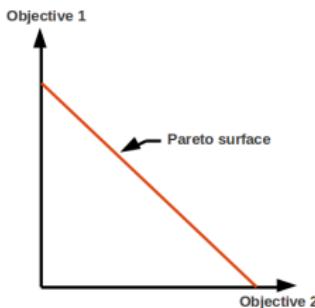
FBA results	
obj : 497.814344	
Search for an entity : <input type="text" value="R_ex"/>	
Entity name	Value ▼
R_EX_h2o_e	30,588
R_EX_co2_e	14,782
R_EX_h_e	5,855
R_EX_ac_e	0,349
R_EX_tma_e	0
R_EX_g3ps_e	0
R_EX_12ppd_R_e	0
R_EX_for_e	0
R_EX_glyalid_e	0
R_EX_gam_e	0
R_EX_but_e	0
R_EX_cmp_e	0
R_EX_tcynt_e	0
R_EX_g3pe_e	0
R_EX_isetac_e	0
R_EX_ser_L_e	0
R_EX_asn_L_e	0
R_EX_sucr_e	0
R_EX_26dap_M_e	0
R_EX_etoh_e	0
R_EX_fuc_L_e	0
R_EX_34dhpac_e	0
R_EX_thymd_e	0
R_EX_frulys_e	0
R_EX_lysls_e	0

Figure: Flux minimization, graphical result.

Pareto analysis

Principle

- ▶ The goal is to test cell objectives.
- ▶ Entry : Objectives and experimental values.
- ▶ We calculate a theoretical surface which corresponds to the optimization of these objectives in every condition.

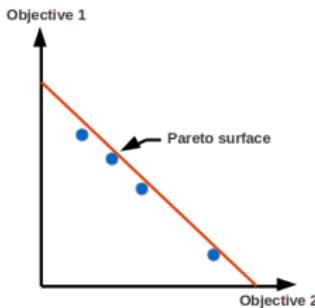


- ▶ The closer our experimental values are to the surface, the more the objectives are optimized.
- ▶ Result : The most optimized objectives.

Pareto analysis

Principle

- ▶ The goal is to test cell objectives.
- ▶ Entry : Objectives and experimental values.
- ▶ We calculate a theoretical surface which corresponds to the optimization of these objectives in every condition.

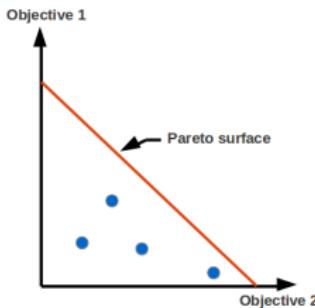


- ▶ The closer our experimental values are to the surface, the more the objectives are optimized.
- ▶ Result : The most optimized objectives.

Pareto analysis

Principle

- ▶ The goal is to test cell objectives.
- ▶ Entry : Objectives and experimental values.
- ▶ We calculate a theoretical surface which corresponds to the optimization of these objectives in every condition.



- ▶ The closer our experimental values are to the surface, the more the objectives are optimized.
- ▶ Result : The most optimized objectives.

Pareto analysis

- ▶ First implementation.
- ▶ We need an additional file containing objectives and experimental values.

```
MIN(FlexSum) MAX(atpYield) MAX(R_Biomass_Ecoli_core_w_GAM)
```

```
1667.4 213.0 6.4  
1635.9 191.0 7.9  
1640.4 171.0 7.4  
1696.2 176.0 7.2  
1657.4 173.0 7.4  
1642.9 207.0 6.9  
1616.7 185.0 7.7  
. .
```

Pareto analysis

- ▶ Condition file :

```
#Pour normaliser toutes les valeurs expérimentales  
R_EX_glc_e -100  
  
atpYield 0 999999  
  
EQUATIONS  
  
#definitions of variables for objectives  
  
atpYield = R_PYK + -1*R_PPCK + R_PTAr
```

- ▶ Command line :

```
FlexFluxPareto -s coli.xml -cond conditions.txt -exp expFile.txt -plot
```

Pareto analysis

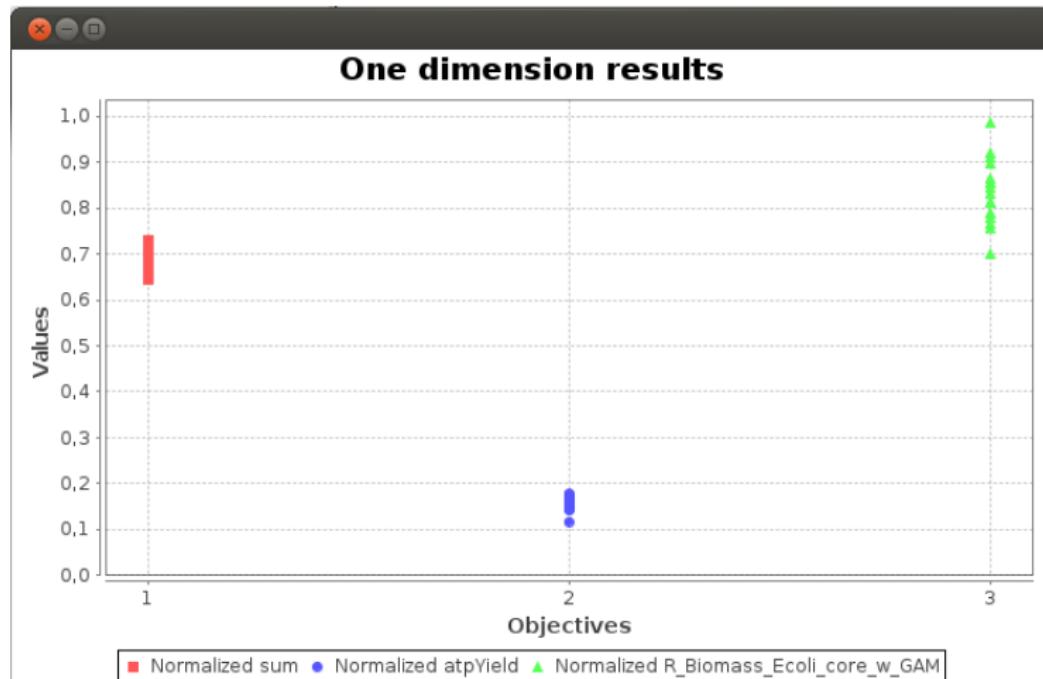


Figure: Pareto analysis, graphical result.

Pareto analysis

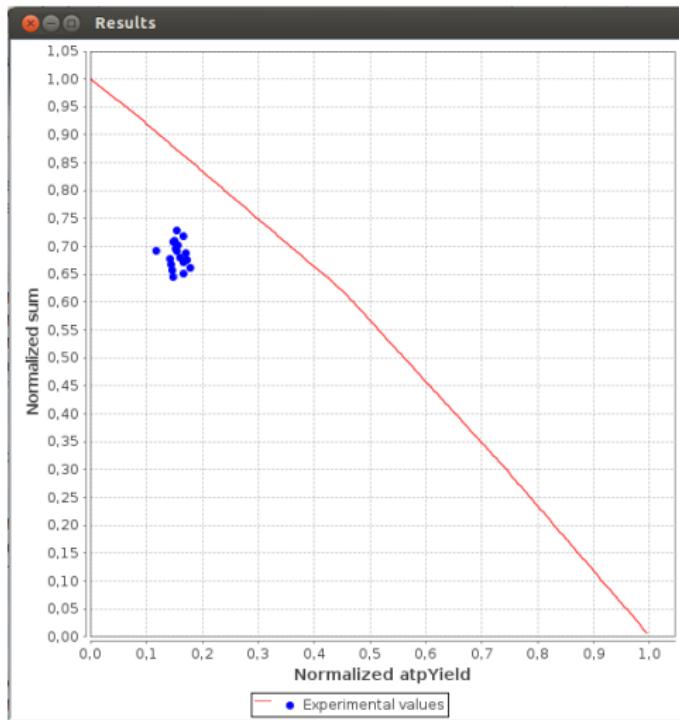


Figure: Pareto analysis, graphical result.

Pareto analysis

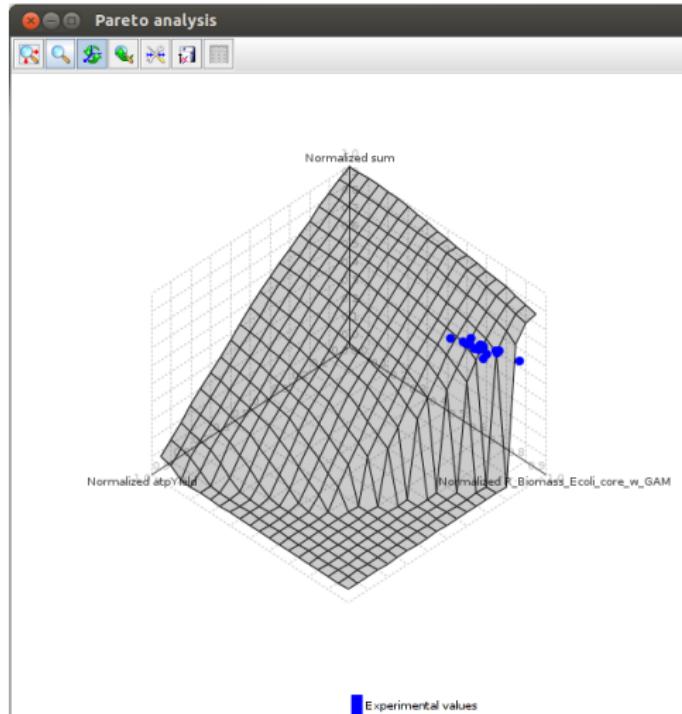
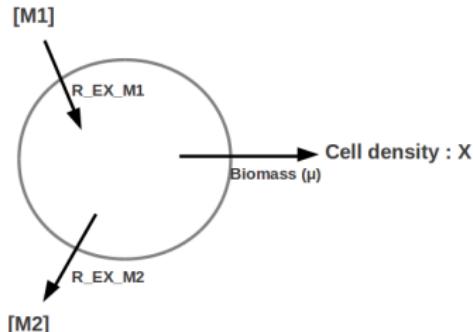


Figure: Pareto analysis, graphical result.

Time dependent analysis : calculations at every step⁴



- ▶ Maximal uptake of each external metabolite :
$$R_{EX_M_{max}} = \frac{[M]}{X \cdot \Delta_t}$$
- ▶ Optimization of the model.
- ▶ New cell density : $X = X_{-1} \cdot e^{\mu \Delta_t}$
- ▶ New external metabolites concentrations :
$$[M] = [M_{-1}] - \frac{R_{EX_M}}{\mu} X (1 - e^{\mu \Delta_t})$$

⁴Amit Warma et al. "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110". In: *Appl Environ Microbiol* (1994).

Time dependent analysis

► Command line :

```
FlexFluxRFBA -x 0.011 -s coli.xml -cond condition.txt -int interactions.txt -bio R_Biomass -plot
```

```
#maximize biomass production
obj : MAX[R_Biomass_Ecoli_core_w_GAM]

#uptake values
R_EX_lac_D_e -3 999999
R_EX_glc_e -6.5 999999
R_EX_o2_e -15 999999

#Initial concentrations
#Glucose
M_glc_D_b 1.6
#Lactose
M_lac_D_b 5.8

BINARY

#Gene lacZ can be repressed (0)
#or expressed(1)
lacZ
```

The uptake of lactose can only be active when there is no glucose

```
IF[R_glc_D_b=0]THEN[lacZ=0]
IF[R_glc_D_b=0]THEN[lacZ=1]
IF[lacZ=0]THEN[R_EX_lac_D_e=0]
```

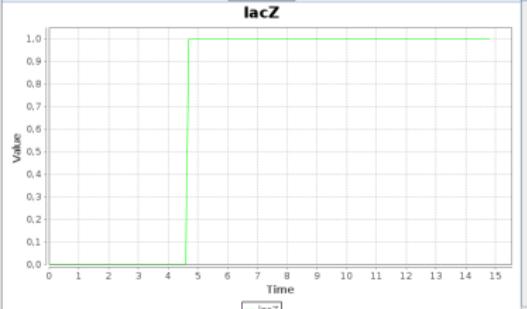
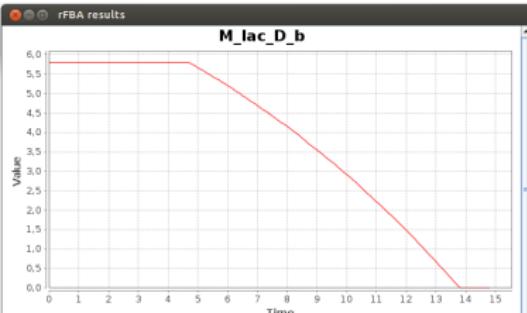
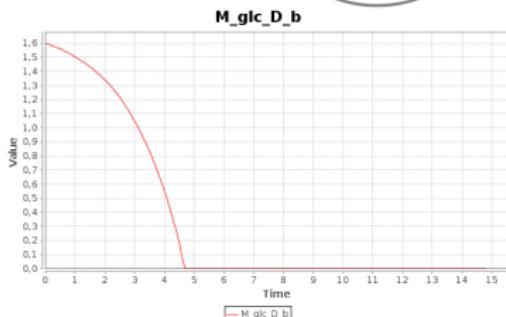
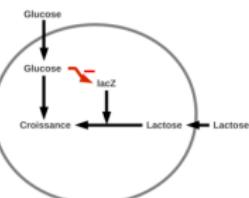


Figure: Time dependent analysis graphical result.

Time dependent analysis : regulation pathways of *Ralstonia*

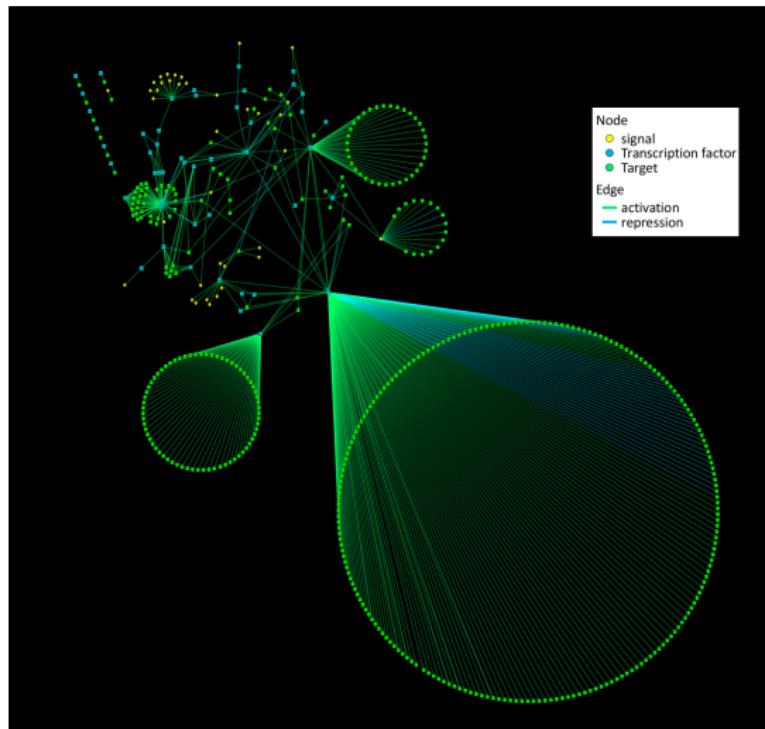


Figure: Regulation network of *Ralstonia*.

Time dependent analysis : regulation pathways of *Ralstonia*

- Hypothesis : 3-OH PAME concentration allows *Ralstonia* to count its sisters.

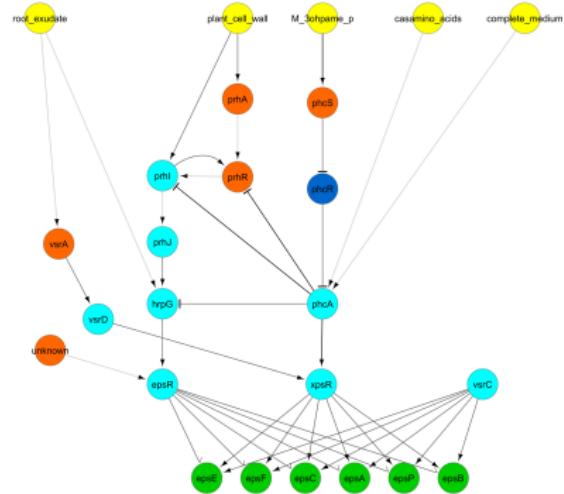
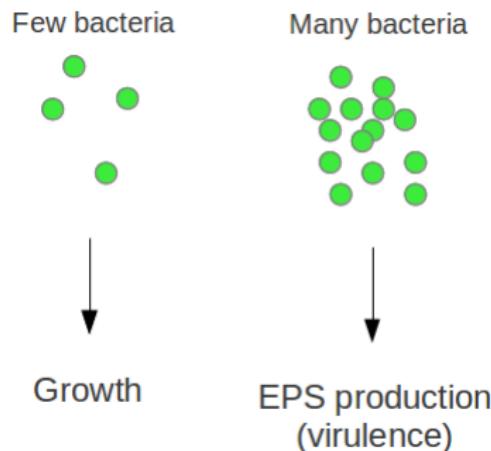


Figure: Interactions controlling EPS production in *Ralstonia*.

Time dependent analysis : regulation pathways of *Ralstonia*

- ▶ Condition file :

```
#On maximise la biomasse
obj : MAX(R_BIOMASS)

#Valeur des réactions d'échange
R_EX_glu_L_e_ -3.8 999999
R_EX_3OHPAMES_e_ 0.000148

#Concentrations initiales
M_glu_L_b 20
M_3ohpame_b 0
M_EPS_b 0
```

- ▶ Interaction file :

```
.IF [RSc2942 AND RSc2941 AND RSc2940] THEN [RSc1889=1]
IF [RSp0873=0] THEN [RSc1892=1]
IF [RSp0873] THEN [RSc1900=1]
IF [RSp0873] THEN [RSc1901=1]
IF [RSp0873=0] THEN [RSc1918=1]
.
```

- ▶ Command line :

```
FlexFluxRFBA -x 0.011 -s ralstonia.xml -cond conditions.txt -int interactions.txt
-bio R_BIOMASS -plot
```

Activation of EPS production by an external signal

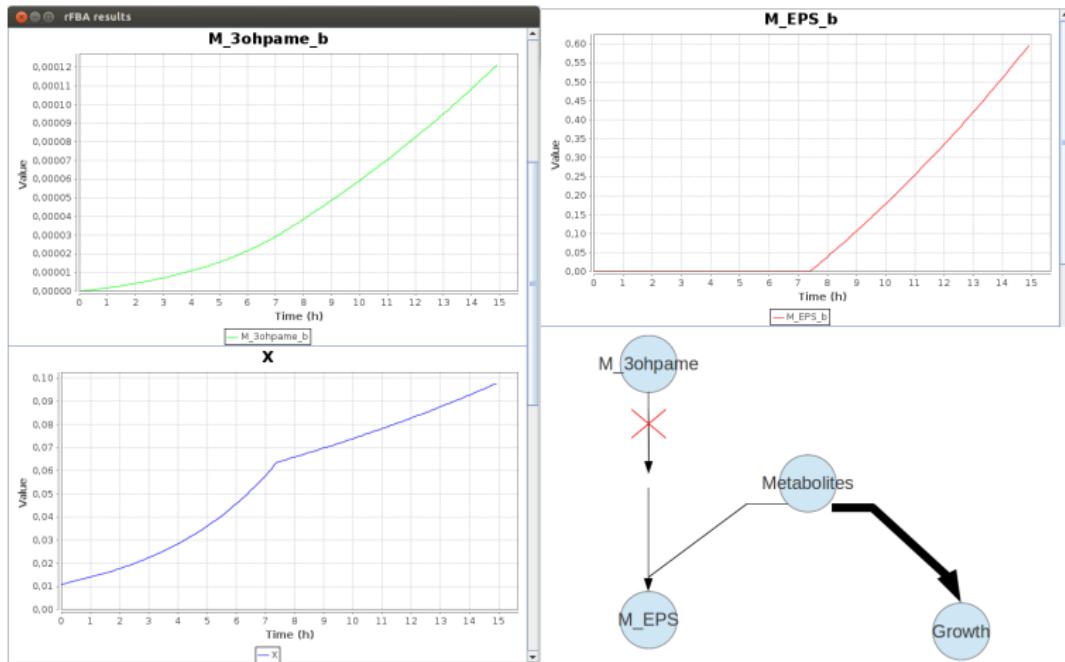


Figure: Activation of EPS production by an external signal

Activation of EPS production by an external signal

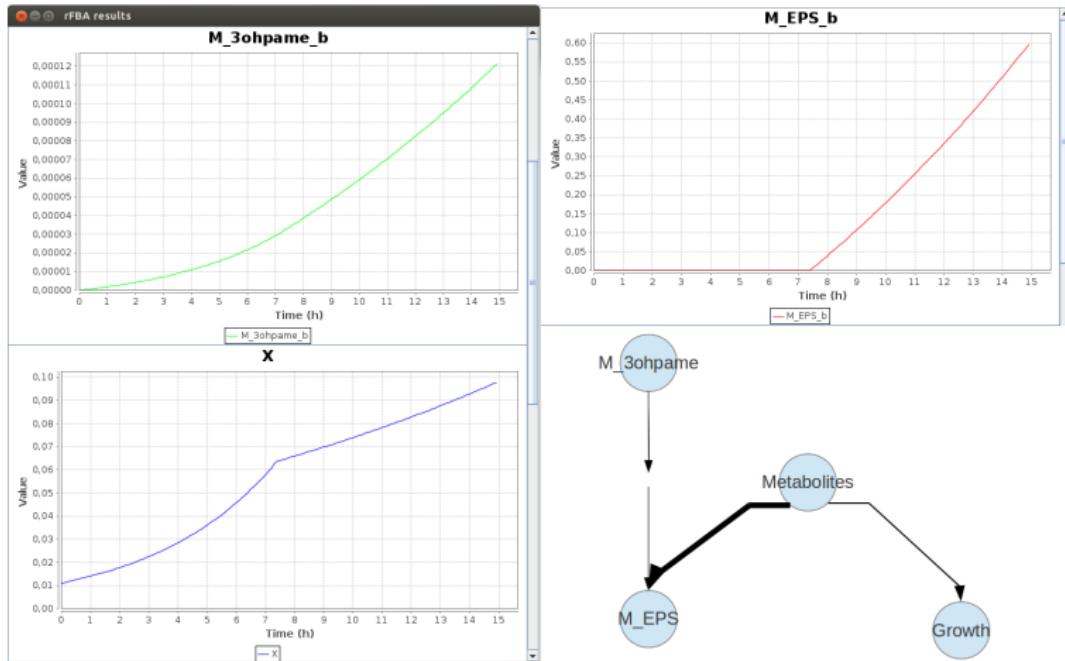
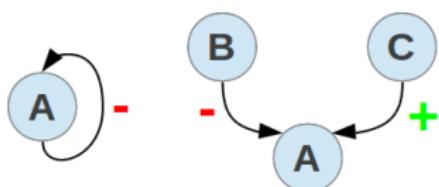


Figure: Activation of EPS production by an external signal

Regulatory network modelling

Some interaction motifs are not easy to model with FlexFlux

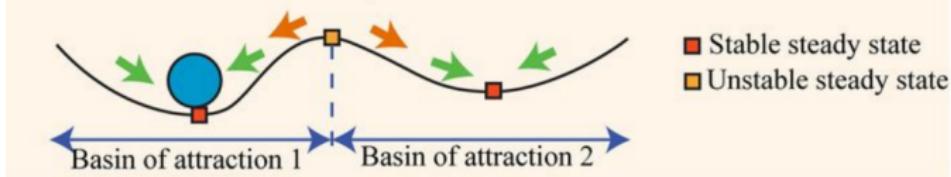


- ▶ Negative autoregulation.
- ▶ Gene with several inputs.

Regulatory network modelling⁵

- ▶ Model : mathematical representation of a system.
- ▶ Dynamic models : understand how a network behaves over time under various conditions.
- ▶ Importance of steady states and stability.

Bistable case: two stable steady states and an unstable one



⁵Alistair M Middleton et al. "Modeling regulatory networks to understand plant development: small is beautiful." In: *The Plant cell* (2012).

Regulatory network modelling

- ▶ State S of the model at time t is the value of each variable at time t .

$$S(t) = \{X_1(t), X_2(t), \dots, X_n(t)\}$$

- ▶ Our assumption : we can predict $S(t+1)$ from $S(t)$, the future is predictable.

Regulatory network modelling

Different ways of modelling

- ▶ **Boolean network.**
- ▶ **Ordinary differential equations (ODEs).**
- ▶ Stochastic models.
- ▶ Other.

Boolean networks

- ▶ Variables represent the state of a gene.
- ▶ Values are either 0 or 1.
- ▶ Each variable has an input list that will determine its future state.
- ▶ As the number of states is finite, the network will eventually cycle : this cycle is an attractor.
- ▶ The steady states are the attractors of length 1.

Boolean networks

How to update the network

- ▶ Synchronous update
 - ▶ All genes are updated at each time step.
 - ▶ Criticized because genes do not change their states all at the same moment.
 - ▶ Usually used to find steady states.
- ▶ Asynchronous update
 - ▶ One component is updated at a time.
 - ▶ Different update approaches (order, random, time delay ...).
 - ▶ Used to study the dynamics of the system.

ODEs

- ▶ The model gives the rate change of each variable.
- ▶ Rate of change of variable X is function of other variable concentrations.
- ▶ $\frac{dX}{dt} = \text{X production rate} - \text{X decay rate}$
- ▶ Requires parameters.

ODEs

Example of equations for gene X.

Kp = production constant.

Kd = degradation constant.

- ▶ If the gene is not regulated :

$$\frac{dX}{dt} = Kp - Kd * X$$

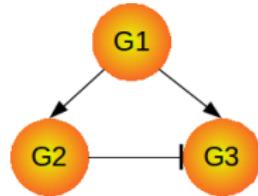
- ▶ If the transcription of X is activated by Y :

$$\frac{dX}{dt} = Kp * \frac{K_{xy} * Y}{1 + K_{xy} * Y} - Kd * X$$

- ▶ If the transcription of X is repressed by Y :

$$\frac{dX}{dt} = Kp * \frac{1}{1 + K_{xy} * Y} - Kd * X$$

Example



Boolean model

- ▶ $G1(t+1) = 1$
- ▶ $G2(t+1) = G1(t)$
- ▶ $G3(t+1) = G1(t) \text{ AND NOT}[G2(t)]$

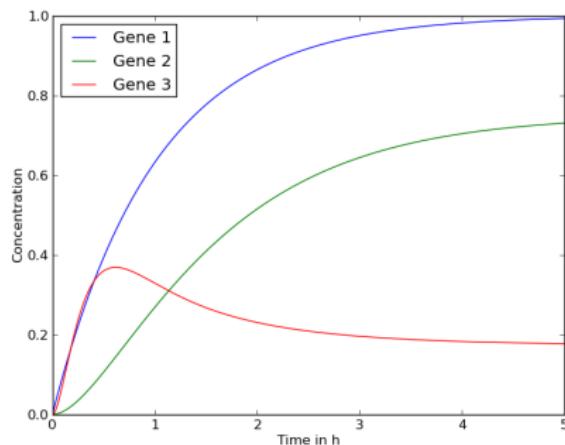
	t0	t1	t2	t3	t4
G1	0	1	1	1	1
G2	0	0	1	1	1
G3	0	0	1	0	0

A steady state is reached

Example

ODE model

- ▶ $\frac{dG_1}{dt} = Kp_{g1} - Kd_{g1} * G_1$
- ▶ $\frac{dG_2}{dt} = Kp_{g2} * \frac{K_{21} * G_1}{1 + K_{21} * G_1} - Kd_{g2} * G_2$
- ▶ $\frac{dG_3}{dt} = Kp_{g3} * \frac{K_{31} * G_1}{1 + K_{31} * G_1} * \frac{1}{1 + K_{32} * G_2} - Kd_{g3} * G_3$



Example

Results comparison

- ▶ Both models lead to a similar steady state.
- ▶ The boolean model is simpler but only gives qualitative results.
- ▶ The ODE based model gives quantitative results but they highly depend on the chosen parameters.
- ▶ Increasing complexity doesn't necessarily give better predictions.

Conclusion and prospects

- ▶ FlexFlux is a Java tool for metabolic flux analysis.
- ▶ Packaging :
 - ▶ Complete documentation (web site)⁶.
 - ▶ Easy installation on Windows and Linux.
- ▶ Provides answers to complex biological problems.

Prospects

- ▶ Integration in the web interface Metexplore.
- ▶ Improvement of regulation network modelling.
- ▶ Support of models containing plant and pathogen.

⁶<http://lipm-bioinfo.toulouse.inra.fr/flexflux/>