BaSyLiCA*

A web interface for automatic process of Live Cell Array data using R.

AÏCHAOUI-DENÈVE LESLIE - MIG - INRA

*BaSysBio Live Cell Array

Outline

I. BaSysBio European project
    1. Research themes
    2. Partners
II. Live Cell Array
    1. Data recovery
    2. Curves processing
III. Interface
    1. Tool structure
    2. Demo
IV. Perspectives

BaSysBio
European Project

BaSysBio*

*Bacillus Systems Biology

Towards an understanding of dynamic transcriptional regulation at global scale in bacteria: a systems biology approach

- **Objective: Understand genetic regulations in bacteria across the cell:**
  - **linkages between genetic and metabolic networks.**
  - **responses of bacteria in terms of pathogenicity.**

- **15 European partners including modelers and biologists.**
  - **Data production at all levels of the cell: transcriptome, proteome, Fluxome, metabolome, live cell arrays, Chip-Chip, prot-prot interactions, etc.**
  - **Models and tools developments to understand and integrate biological complexity.**

- **MIG (Mathématique, Informatique et Génome):**
  1. **Genome structural annotation (A. Leduc, P. Nicolas, P. Bessières).**
  2. **Formalization and analysis of biological systems (A. Goelzer, V. Fromion).**
  3. **Data processing/analysis (L. Aïchaoui, V. Fromion).**

INRA BaSysBio MIG Mathématique Informatique & Génome

Partners (involved)

Data treatments

Jouy-en-Josas (MIG)- France

Philippe Bessières
Pierre Nicolas
Aurélie Leduc

Vincent Fromion
Anne Goelzer
Leslie Aichaoui

Grignon - France

Stephane Aymerich
Matthieu Jules
Ludovic Le Chat
Dominique Le Coq

Gröningen - Netherlands

Jan Maarten van Dijl
Emma Denham
Sjouke Piersma

LCA Technology

York - England

Tony Wilkinson
Mark Fogg
Rachel Adamson

Dublin - Ireland

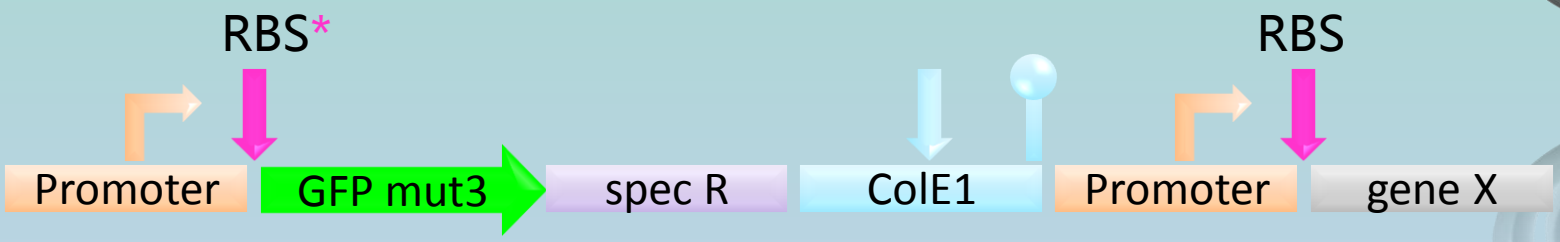
Kevin Devine
Eric Botella
Annette Hansen
Sebastian Hübner

INRA BaSysBio MIG Mathématique Informatique & Génome

Live Cell Array Data

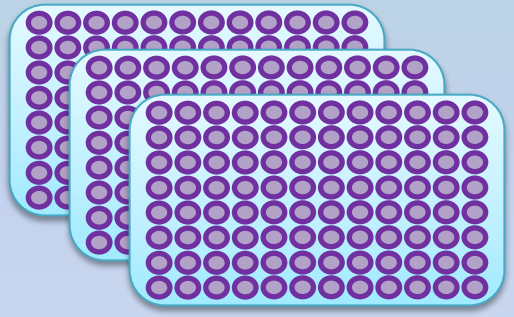
Live Cell Array (LCA)

Development of a new technology for *B. subtilis*, the LCA



96-well plate

Library of Px-gfp strains



High-thoughput reading

Data storage and analysis

OD → GFP

High-throughput :
Real-time monitoring of promoter activities in living cells

INRA BaSysBio MIG Mathématique Informatique & Génome

Needs

- A database to store data from LCA
- A useR!-friendly interface to entering data into the database
- A useR!-friendly interface to process data and retrieve results and graphs

Dynamics with LCA

- LCA Objectives:
  - High-throughput
  - Real-time monitoring of promoter activities in living cells
  - Organizing regulations in a temporal hierarchy
- In order to test perturbations on a coherent subset of fusions:
  - Mutants
  - Growth conditions

$$\frac{d}{dt}[GFP_{Cell}(t)] = \rho P_{activity}(t) - \mu(t)\,[GFP_{Cell}(t)]$$

$$\left(\frac{GFP_{Total}(t)}{OD(t)}\right) = [GFP_{Cell}(t)]$$

$$\frac{\frac{d}{dt}[GFP_{Total}(t)]}{OD(t)} = \rho P_{activity}(t)$$

Model

$$\begin{cases} \widehat{OD}_{(k+1)} = \widehat{OD}_{(k)} + \hat{\mu}_{(k)} OD_{(k)} \delta_t \\ \hat{\mu}_{(k+1)} = \hat{\mu}_{(k)} + \omega_{(k)} \delta_t \\ \widehat{GFP}_{(k+1)} = \widehat{GFP}_{(k)} + \hat{a}_{(k)} OD_{(k)} \delta_t \\ \hat{a}_{(k+1)} = \hat{a}_{(k)} + \omega_{2(k)} \delta_t \end{cases}$$

$$\begin{pmatrix} \widehat{OD}_{(k+1)} \\ \hat{\mu}_{(k+1)} \\ \widehat{GFP}_{(k+1)} \\ \hat{a}_{(k+1)} \end{pmatrix} = \begin{pmatrix} 1 & OD_{(k)}\delta_t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & OD_{(k)}\delta_t \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \widehat{OD}_{(k)} \\ \hat{\mu}_{(k)} \\ \widehat{GFP}_{(k)} \\ \hat{a}_{(k)} \end{pmatrix}$$

A_k

Kalman Filter

Time Update ("Predict")

(1) Project the state ahead

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

(2) Project the error covariance ahead

$$P_k^- = AP_{k-1}A^T + Q$$

Measurement Update ("Correct")

(1) Compute the Kalman gain

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

(2) Update estimate with measurement z_k

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

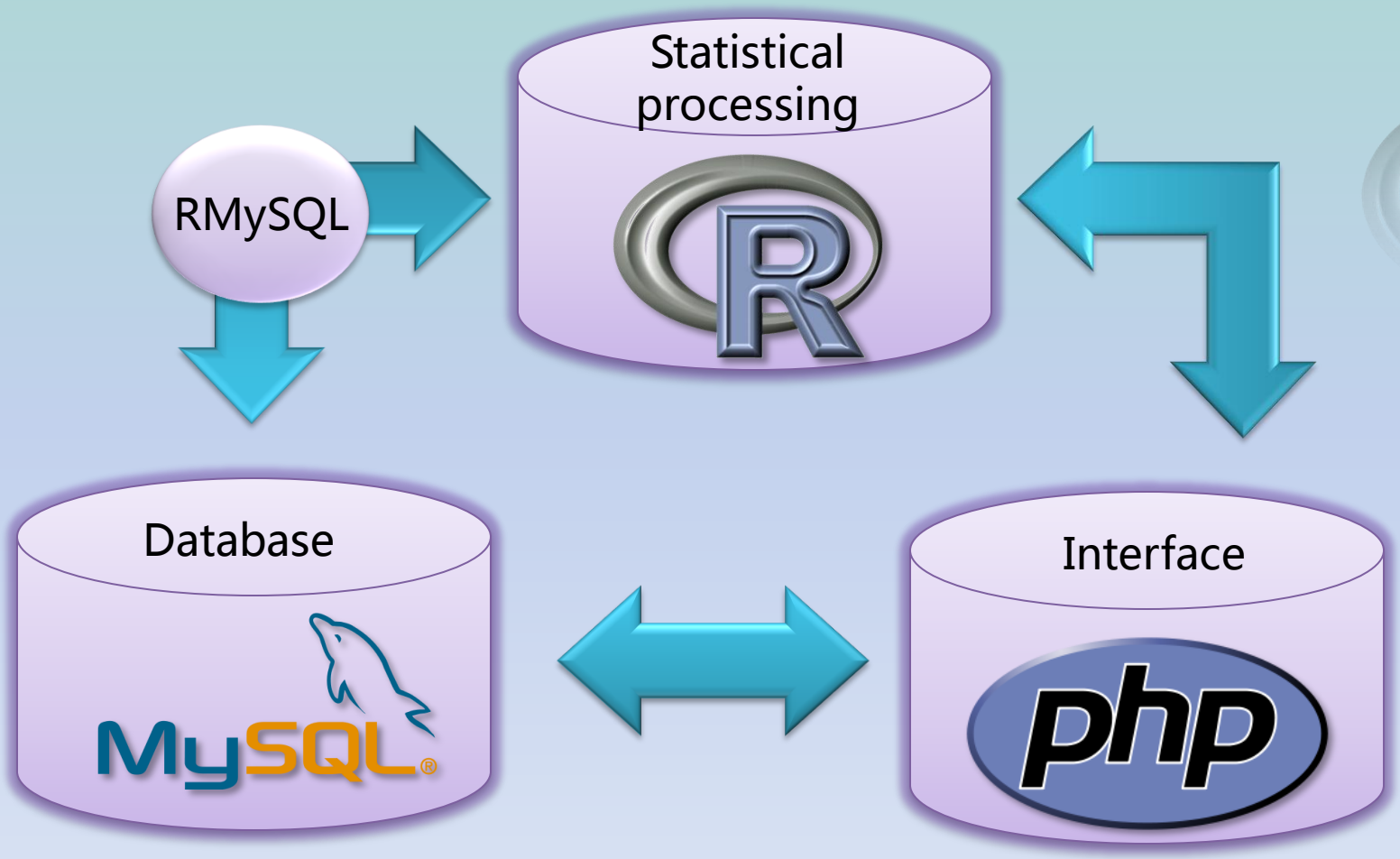
(3) Update the error covariance

$$P_k = (I - K_k H)P_k^-$$

Initial estimates for $\hat{x}_{k-1}$ *and* P_{k-1}

Interface

Interactions



Interaction examples

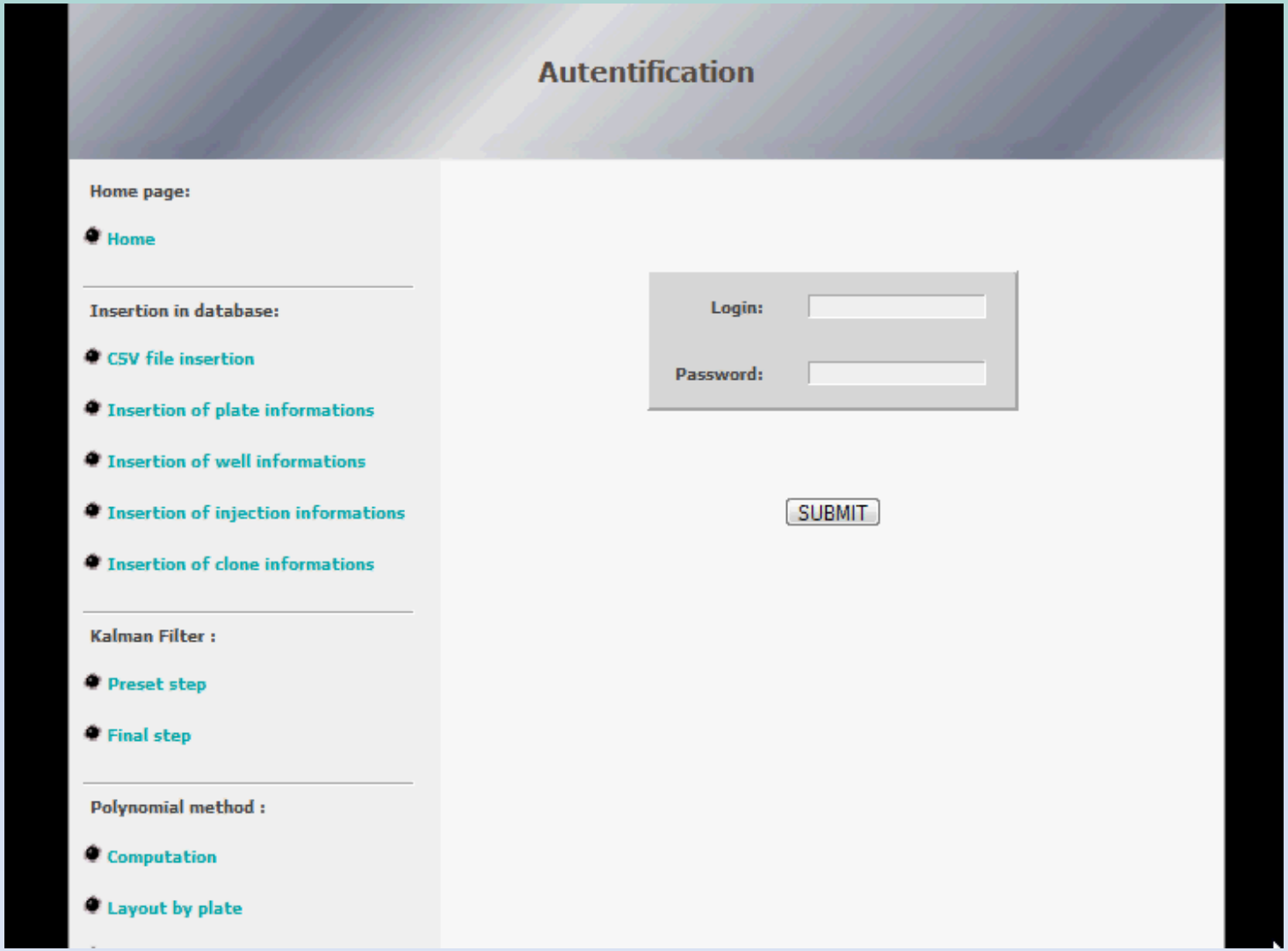
R call in PHP

```
system("C:/.../R/R-2.10.0/bin/Rcmd" BATCH C:/.../.../script.R)
```

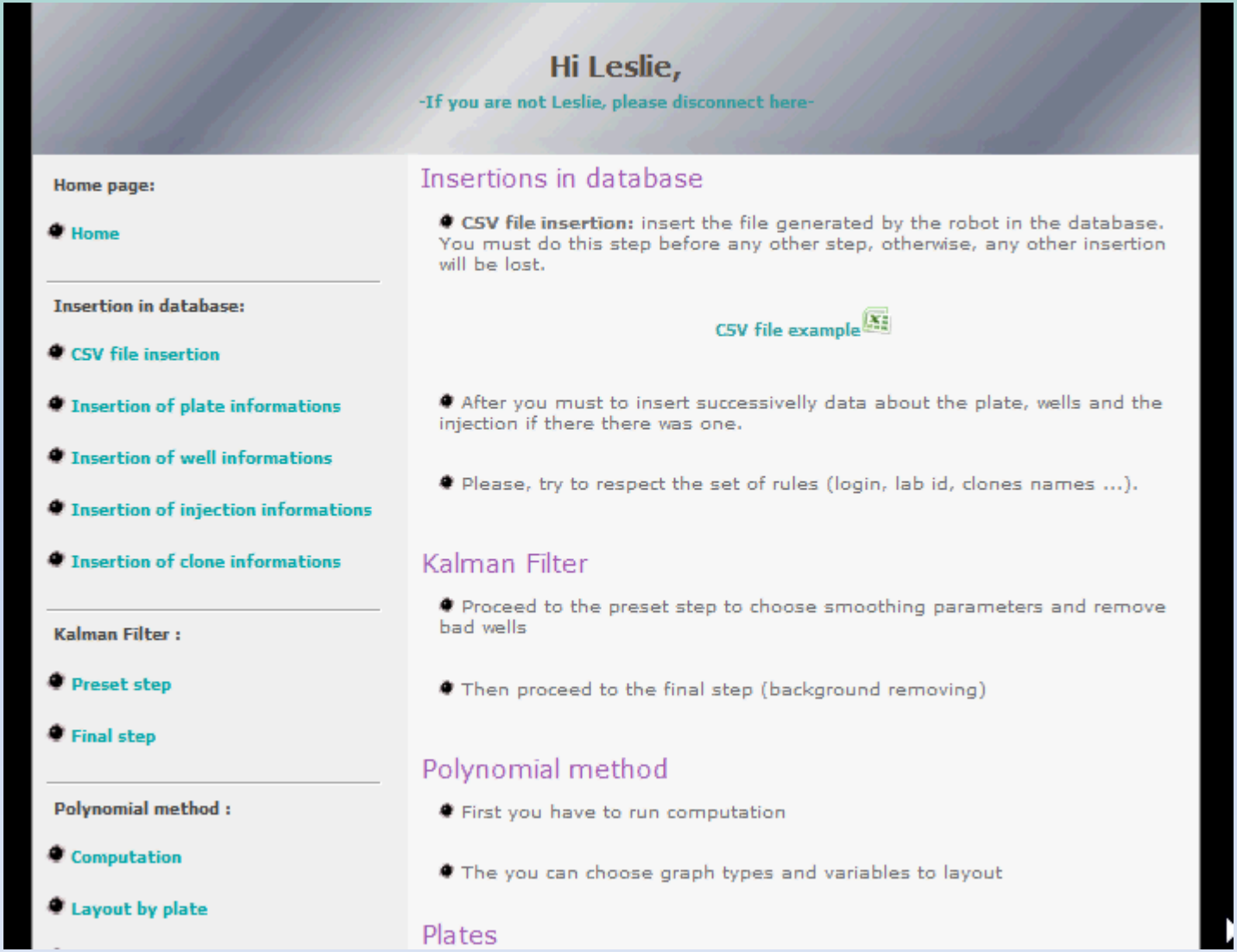
Connexion to MySQL database

```
library(RMySQL)
m=dbDriver("MySQL")
chaCurv=dbConnect(m,dbname=dsnc,user=uidc,password=pwdc)
reqbg=paste("SELECT well_pk FROM well WHERE backgd='yes'
      AND platew_pk='",nplate,"'",sep="")
wbg=dbGetQuery(chaCurv,reqbg)
dbDisconnect(chaCurv)
```

Data insertion



Kalman filter



INRA BaSysBio MIG Mathématique Informatique & Génome

R Graph example

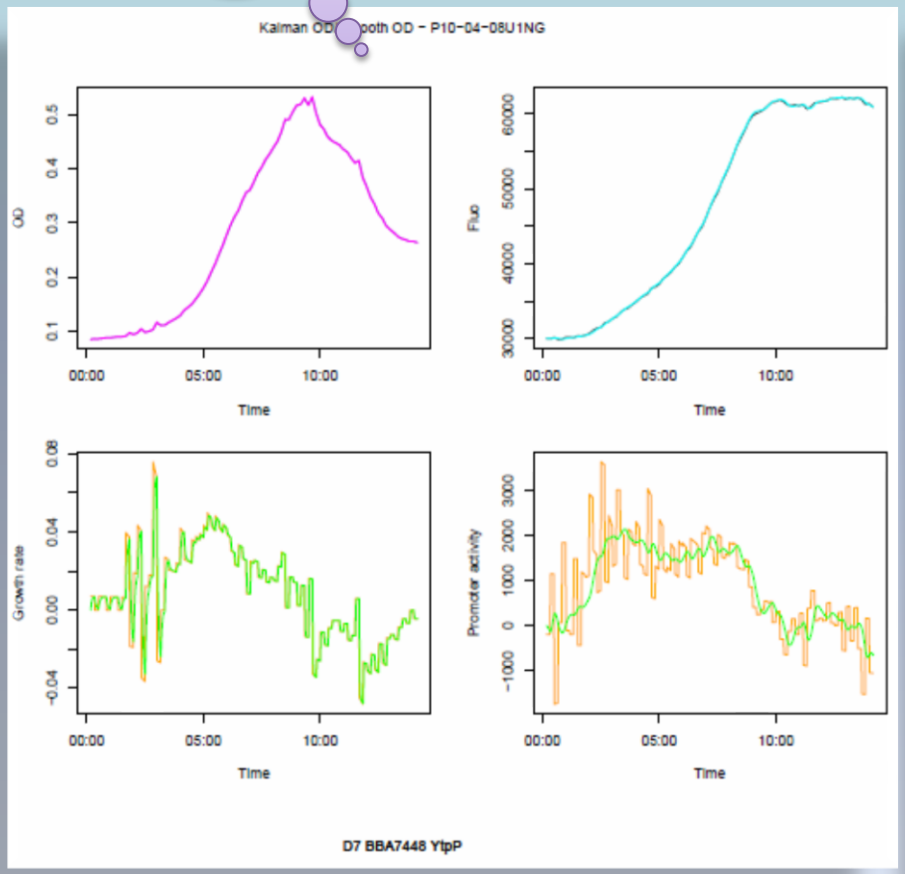
A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE	PEROXYDE
B1 BSB168 Gly5 M9 GLUCOSE PEROXYDE	B2 BSB168 Gly5 M9 GLUCOSE PEROXYDE	B3 BSB168 Gly5 M9 GLUCOSE PEROXYDE	B4 BBA7439 ResA M9 GLUCOSE PEROXYDE	B5 BBA7439 ResA M9 GLUCOSE PEROXYDE	B6 BBA7439 ResA M9 GLUCOSE PEROXYDE	B7 BBA7454 ArsC M9 GLUCOSE PEROXYDE	B8 BBA7454 ArsC M9 GLUCOSE PEROXYDE	B9 BBA7454 ArsC M9 GLUCOSE PEROXYDE	B10 BBA7442 YneN M9 GLUCOSE PEROXYDE	B11 BBA7442 YneN M9 GLUCOSE PEROXYDE	B12 BBA7442 YneN M9 GLUCOSE PEROXYDE
C1 BBA7499 yoeB M9 GLUCOSE PEROXYDE	C2 BBA7499 yoeB M9 GLUCOSE PEROXYDE	C3 BBA7499 yoeB M9 GLUCOSE PEROXYDE	C4 BBA7443 YkvV M9 GLUCOSE PEROXYDE	C5 BBA7443 YkvV M9 GLUCOSE PEROXYDE	C6 BBA7443 YkvV M9 GLUCOSE PEROXYDE	C7 BBA7460 dnaK M9 GLUCOSE PEROXYDE	C8 BBA7460 dnaK M9 GLUCOSE PEROXYDE	C9 BBA7460 dnaK M9 GLUCOSE PEROXYDE	C10 BBA7484 yuiA M9 GLUCOSE PEROXYDE	C11 BBA7484 yuiA M9 GLUCOSE PEROXYDE	C12 BBA7484 yuiA M9 GLUCOSE PEROXYDE
D1 BBA7446 YdbP M9 GLUCOSE PEROXYDE	D2 BBA7446 YdbP M9 GLUCOSE PEROXYDE	D3 BBA7446 YdbP M9 GLUCOSE PEROXYDE	D4 BBA7488 dhbF M9 GLUCOSE PEROXYDE	D5 BBA7488 dhbF M9 GLUCOSE PEROXYDE	D6 BBA7488 dhbF M9 GLUCOSE PEROXYDE	D7 BBA7448 YtpP M9 GLUCOSE PEROXYDE	D8 BBA7448 YtpP M9 GLUCOSE PEROXYDE	D9 BBA7448 YtpP M9 GLUCOSE PEROXYDE	D10 BBA7475 dps M9 GLUCOSE PEROXYDE	D11 BBA7475 dps M9 GLUCOSE PEROXYDE	D12 BBA7475 dps M9 GLUCOSE PEROXYDE
E1 BBA7490 ykuN M9 GLUCOSE PEROXYDE	E2 BBA7490 ykuN M9 GLUCOSE PEROXYDE	E3 BBA7490 ykuN M9 GLUCOSE PEROXYDE	E4 BBA7449 YvgZ M9 GLUCOSE PEROXYDE	E5 BBA7449 YvgZ M9 GLUCOSE PEROXYDE	E6 BBA7449 YvgZ M9 GLUCOSE PEROXYDE	E7 BBA7472 clpX M9 GLUCOSE PEROXYDE	E8 BBA7472 clpX M9 GLUCOSE PEROXYDE	E9 BBA7472 clpX M9 GLUCOSE PEROXYDE	E10 BBA7491 ypbR M9 GLUCOSE PEROXYDE	E11 BBA7491 ypbR M9 GLUCOSE PEROXYDE	E12 BBA7491 ypbR M9 GLUCOSE PEROXYDE
F1 BBA7510 yvgN M9 GLUCOSE PEROXYDE	F2 BBA7510 yvgN M9 GLUCOSE PEROXYDE	F3 BBA7510 yvgN M9 GLUCOSE PEROXYDE	F4 BBA7739 Yaat M9 GLUCOSE PEROXYDE	F5 BBA7739 Yaat M9 GLUCOSE PEROXYDE	F6 BBA7739 Yaat M9 GLUCOSE PEROXYDE	F7 BBA7489 fhuD M9 GLUCOSE PEROXYDE	F8 BBA7489 fhuD M9 GLUCOSE PEROXYDE	F9 BBA7489 fhuD M9 GLUCOSE PEROXYDE	F10 BBA7738 yjbG M9 GLUCOSE PEROXYDE	F11 BBA7738 yjbG M9 GLUCOSE PEROXYDE	F12 BBA7738 yjbG M9 GLUCOSE PEROXYDE
G1 BBA7470 gsiB M9 GLUCOSE PEROXYDE	G2 BBA7470 gsiB M9 GLUCOSE PEROXYDE	G3 BBA7470 gsiB M9 GLUCOSE PEROXYDE	G4 BBA7651 relA M9 GLUCOSE PEROXYDE	G5 BBA7651 relA M9 GLUCOSE PEROXYDE	G6 BBA7651 relA M9 GLUCOSE PEROXYDE	G7 BBA7465 mrgA M9 GLUCOSE PEROXYDE	G8 BBA7465 mrgA M9 GLUCOSE PEROXYDE	G9 BBA7465 mrgA M9 GLUCOSE PEROXYDE	G10 BBA7571 yurL M9 GLUCOSE PEROXYDE	G11 BBA7571 yurL M9 GLUCOSE PEROXYDE	G12 BBA7571 yurL M9 GLUCOSE PEROXYDE
H1 BBA7735 yrzF M9 GLUCOSE PEROXYDE	H2 BBA7735 yrzF M9 GLUCOSE PEROXYDE	H3 BBA7735 yrzF M9 GLUCOSE PEROXYDE	H4 BBA7684 yfhF M9 GLUCOSE PEROXYDE	H5 BBA7684 yfhF M9 GLUCOSE PEROXYDE	H6 BBA7684 yfhF M9 GLUCOSE PEROXYDE	H7 BBA7450 yvbW M9 GLUCOSE PEROXYDE	H8 BBA7450 yvbW M9 GLUCOSE PEROXYDE	H9 BBA7450 yvbW M9 GLUCOSE PEROXYDE	H10 BSB168 Gly5 M9 GLUCOSE PEROXYDE	H11 BSB168 Gly5 M9 GLUCOSE PEROXYDE	H12 BSB168 Gly5 M9 GLUCOSE PEROXYDE

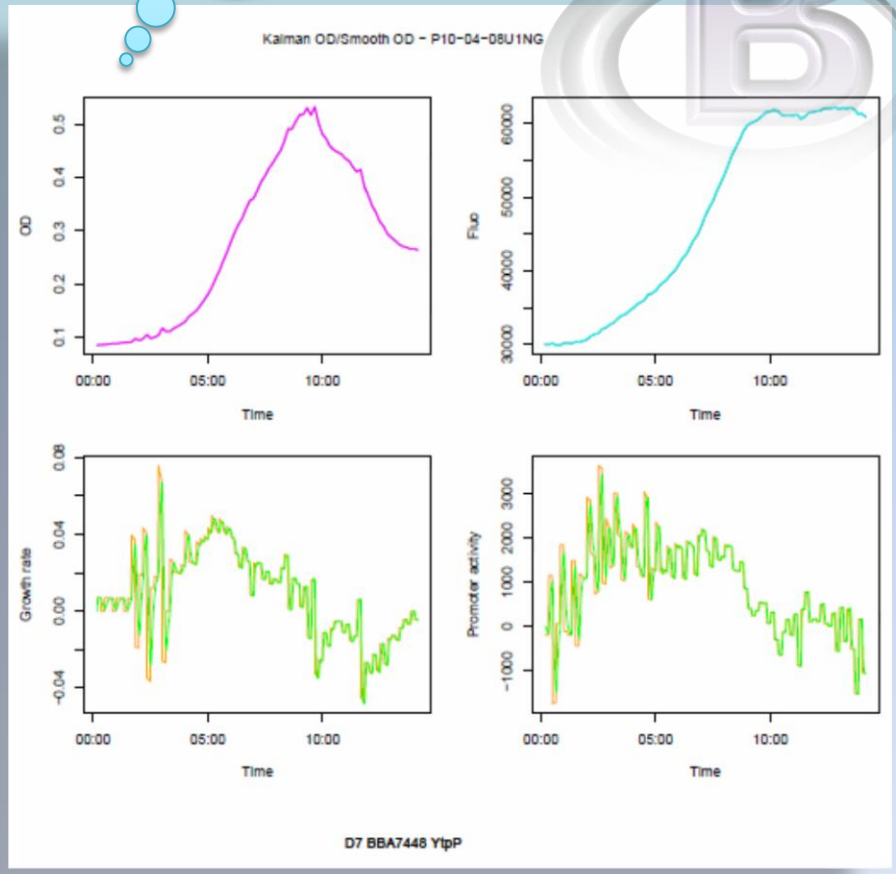
INRA BaSysBio MIG Mathématique Informatique & Génome

R Graph example

Very smoothed

Less smoothed





INRA BaSysBio MIG Mathématique Informatique & Génome

Perspectives

Perspectives

- Automatic classification of promoter activity curves
- Development of an installer
- Tool distribution to the LCA community

Aknowledgments

- The BaSysBio Consortium
- Vincent Fromion and Anne Goelzer for their assistance
- Pascal Neveu and Anne Tireau for their help on the first version of the interface
- Matthieu Jules and Ludovic Le Chat for their help in biology
- Hugo Devillers for his advices in R
- Thibault De Maillard for his help in javascript

Bibliography

- David A. James and Saikat DebRoy (2009). **RMySQL: R interface to the MySQL database**, http://biostat.mc.vanderbilt.edu/wiki/Main/RMySQL.

- Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, Surette MG, Alon U. (2004). **Just-in-time transcription program in metabolic pathways**. Nat Genet. 36, 486-91.

- Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, Alon U. (2006), **A comprehensive library of fluorescent transcriptional reporters for Escherichia coli**. Nat Methods. 3, 623-8.

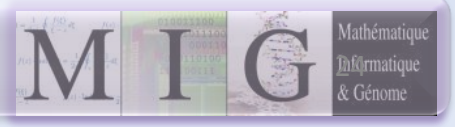
- G. Welch and G. Bishop **An Introduction to the Kalman Filter**, University of North Carolina at Chapel Hill, Department of Computer Science

INRA
Working group on R

http://ciam.inra.fr/r4ciam/

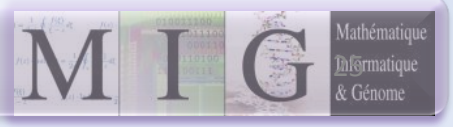

INRA


BaSysBio


M I G
Mathématique
Informatique
& Génome


INRA


BaSysBio


M I G
Mathématique
Informatique
& Génome