DFG research center Berlin
mathematics for key technologies

veritas
iustitia
libertas
Freie Universität Berlin

A theory of optimal differential gene expression

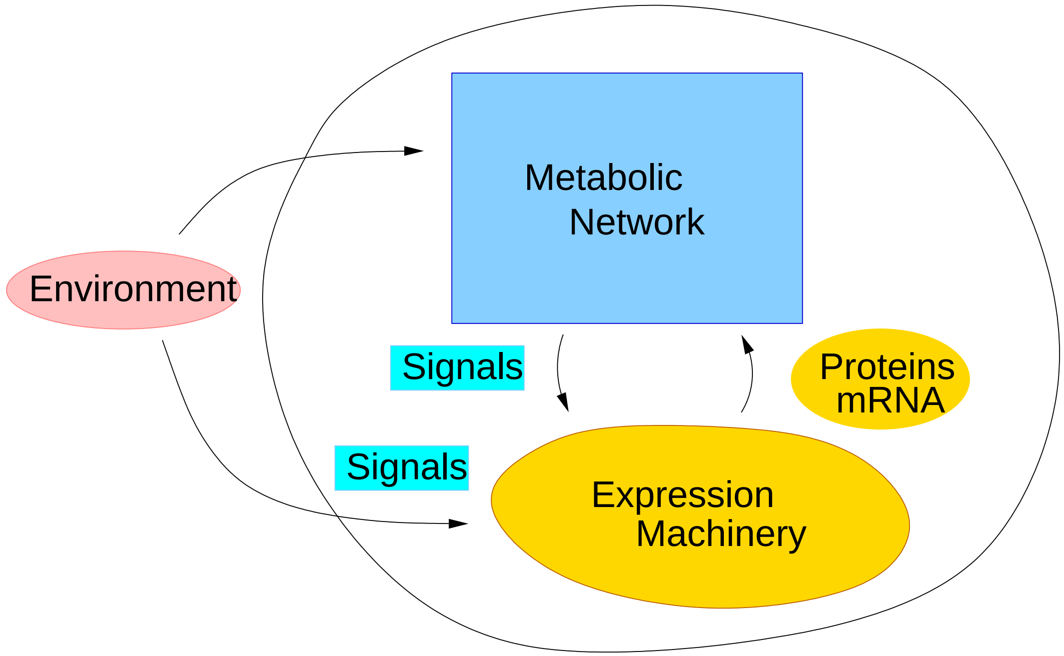
Wolfram Liebermeister, Edda Klipp,
Stefan Schuster, Reinhart Heinrich

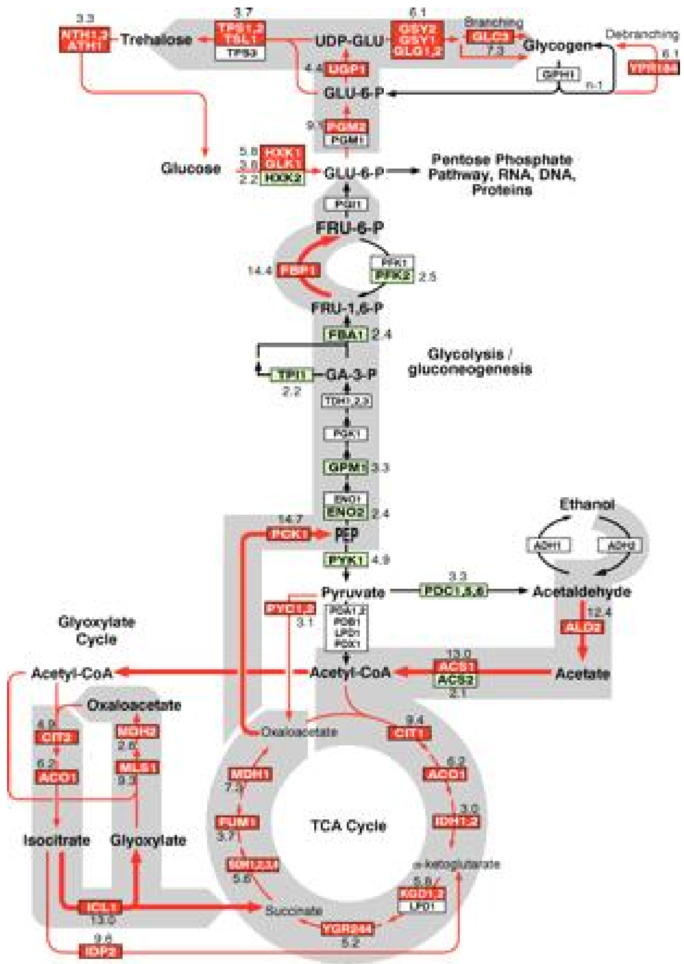
IPCAT2003

Fifth international workshop on information processing in cells and tissues

September 10, 2003

Schematic view of the cell





DeRisi et al., 1997

A model for optimal gene expression patterns

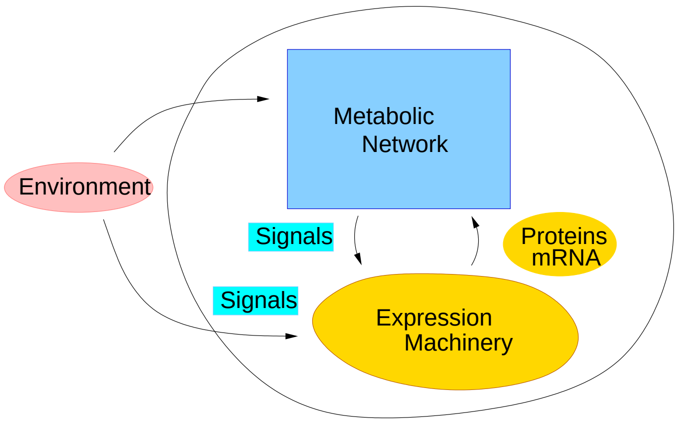
Hypotheses for the Model

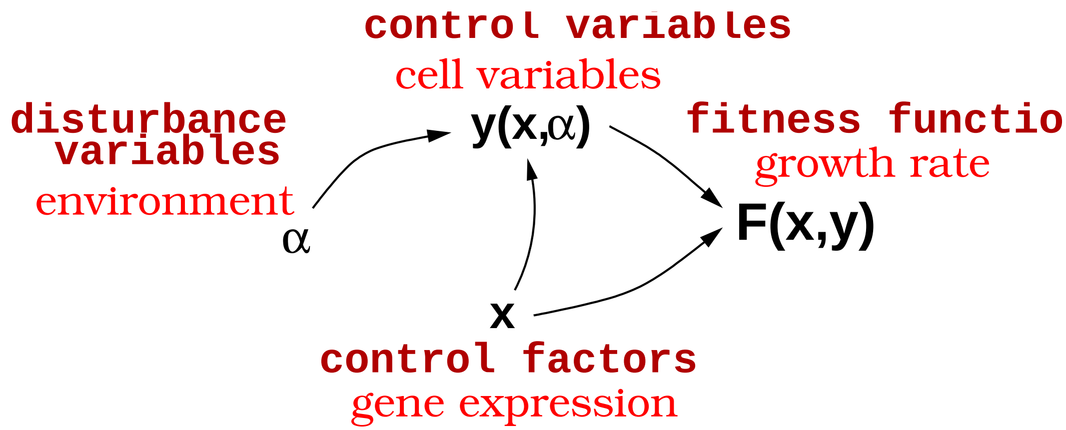
- Genes act as regulators.
  $\longrightarrow$ Differential gene expression ensures optimal control of cell processes.

- Genes can control several fitness-relevant processes at a time
  Gene expression itself is costly
  $\longrightarrow$ Find optimal compromise

- Evolution has physically realised an optimal expression behaviour in the signalling network and in the regulatory sequences.

The Model





The optimality principle

Expression x behaves always such as to maximise the fitness $F(x, y(x, \alpha))$.

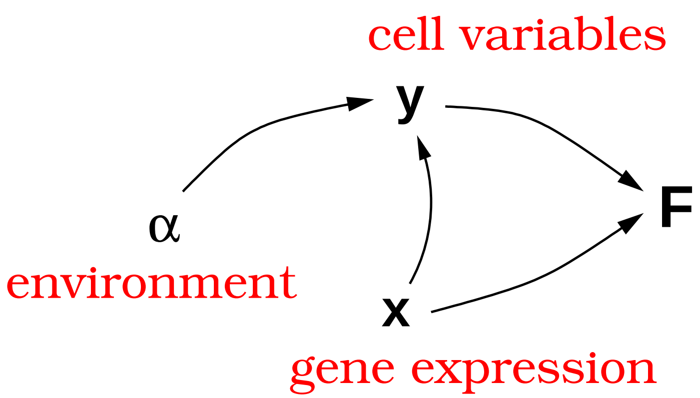
Differential expression

What is the best reaction of gene expression to perturbations?

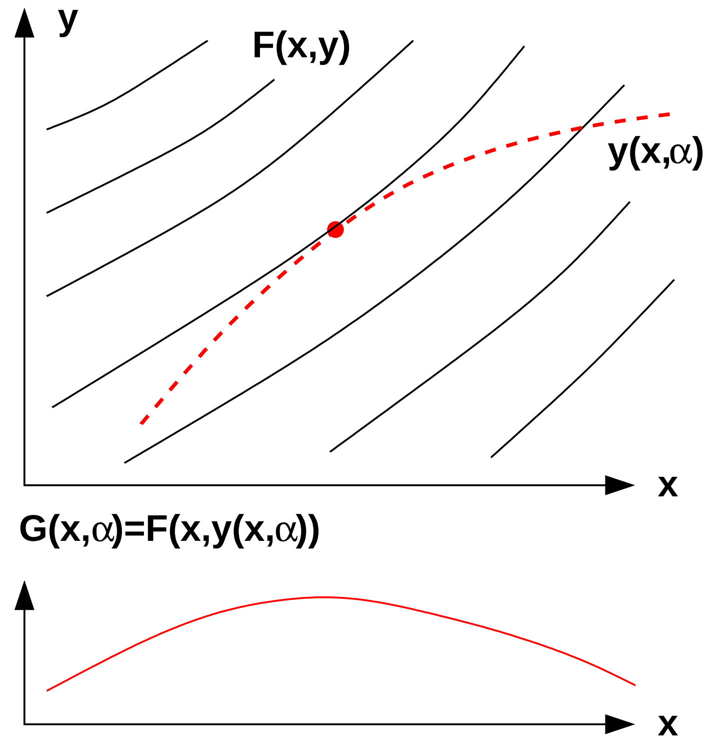
Here

Consider stationary states, small perturbations

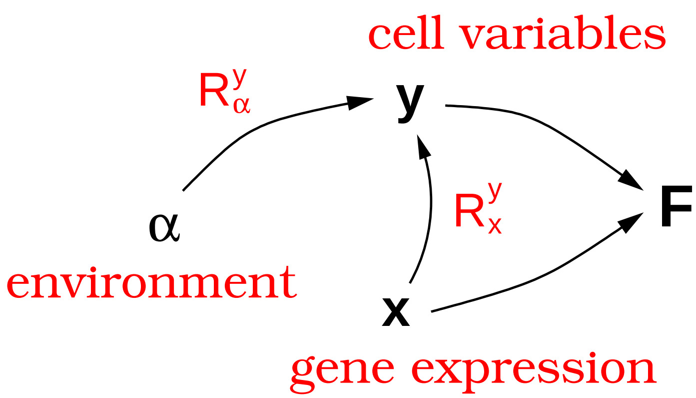
Optimal response to perturbations

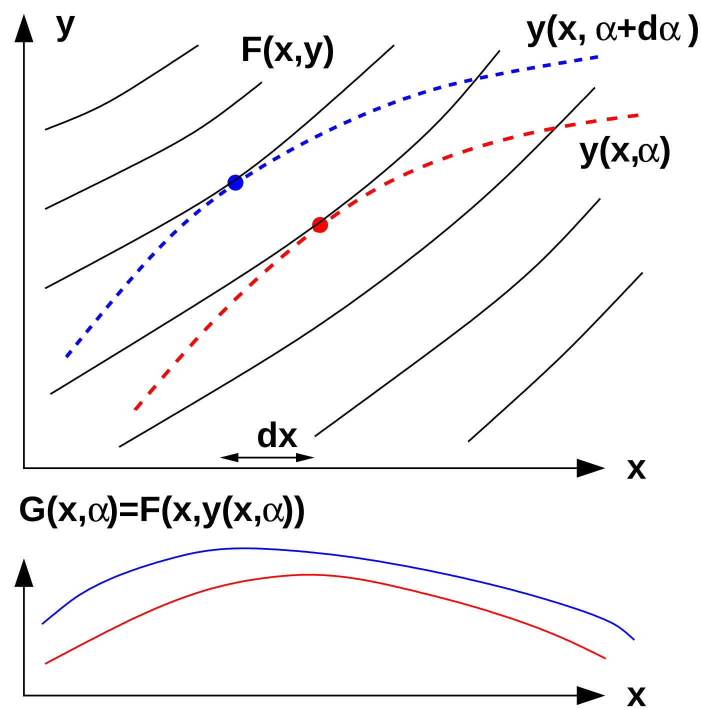


Effective fitness $G(x, \alpha) = F(x, y(x, \alpha))$



Optimal response to perturbations





Effective fitness $G(x, \alpha) = F(x, y(x, \alpha))$

Optimal response to perturbations

- Locally optimal state
- Consider a small perturbation of y or x
- Find the response d$\bar{x}$ for reaching a new optimal state.
- Optimality condition: d$G_x = 0$

Local Approximation

Response coefficients

$R^y_x = \frac{\partial y}{\partial x}, \quad R^y_\alpha = \frac{\partial y}{\partial \alpha}$

Fitness curvatures $F_{xx}, \; F_{yy}$

Optimal response

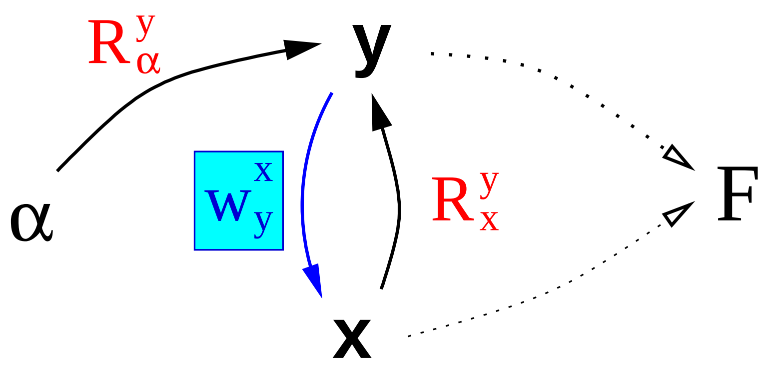
$$\mathrm{d}\bar{x} = -G_{xx}^{-1} \left({R^y_x}^T \; F_{yy} \, \mathrm{d}\hat{y} + F_y^T \mathrm{d}\hat{R}^y_x \right)$$

Results: Consequences from the optimality principle

1. Optimal linear gene programs
2. Symmetric response to gene deletions
3. Expression patterns and response coefficients
4. Sum rules for the control of metabolism

(1) Optimal linear gene programs

The optimal response to perturbations dα can be realised by a linear feedback mechanism (gene program).



$$
\begin{aligned}
\mathrm{d}y &= R_\alpha^y\, \mathrm{d}\alpha + R_x^y\, \mathrm{d}x \\
\mathrm{d}x &= w_y^x\, \mathrm{d}y
\end{aligned}
$$

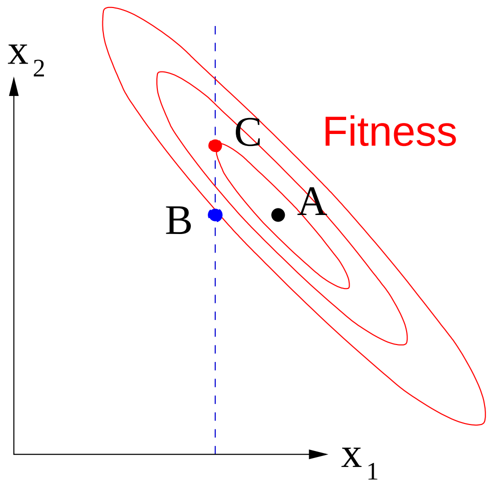
Optimal feedback

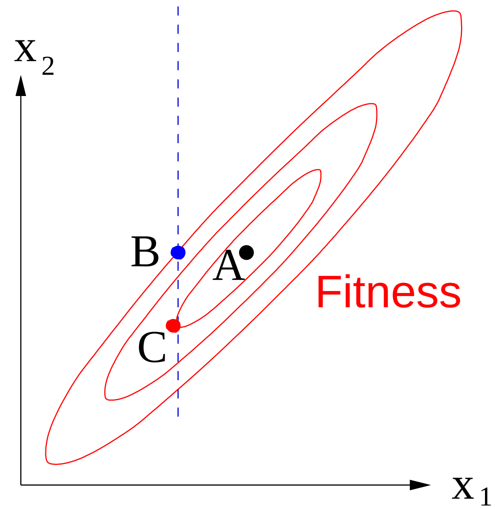
$$w_y^x = -F_{xx}^{-1} {R_x^y}^T F_{yy}$$

- Large influence of a gene $\longrightarrow$ strong feedback
- Similar influence of genes $\longrightarrow$ common feedback

(2) Symmetric response in deletion experiments

The expression of a gene is decreased by a deletion: $x_i \rightarrow x_i + \mathrm{d}\hat{x}_i$.





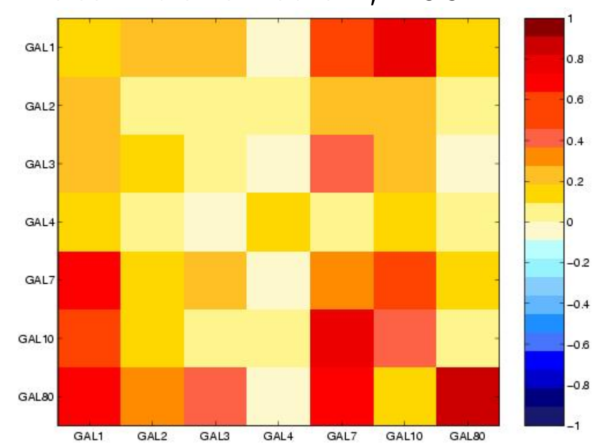
Model prediction

for quadratic expression data matrix X

$$X = G_{xx}^{-1}\ D$$

where G_{xx}^{-1} is symmetric and D is diagonal.

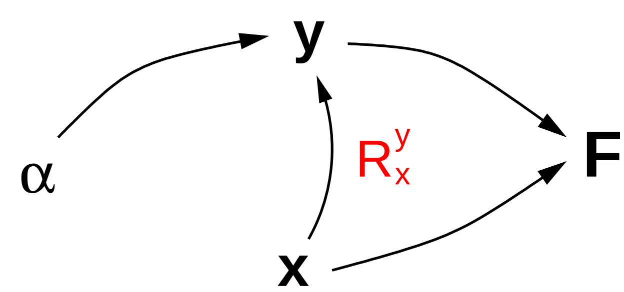
Data: Ideker et al., 2001



(3) Linear superposition of response coefficients

Optimal regulation profile
to yield change dy:

$$\mathrm{d}\bar{x} = F_{xx}^{-1}\ {R_x^y}^T\ (R_x^y\ F_{xx}^{-1}\ {R_x^y}^T)^{-1}\ \mathrm{d}y$$



Assumption:
F_{xx} is isotropic (scalar) $\longrightarrow$

$$\begin{aligned} \mathrm{d}\bar{x} &= {R_x^y}^T\ \mathrm{d}\mathbf{m} \\ &= {R_x^{y_1}}^T\ \mathrm{d}m_1 + {R_x^{y_2}}^T\ \mathrm{d}m_2\ + \ldots \end{aligned}$$

Optimal expression profile is a linear combination
of response coefficient profiles.

Prediction:

Response coefficient profiles should appear
as linear components in expression data.

(4) Optimal control of metabolic systems

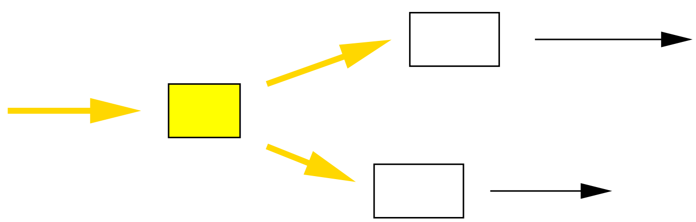
From the **theorems of metabolic control theory** follow relations between the optimal expression profiles and the structure of the metabolic network.

Supposed, the fitness depends only on ...

Fluxes

Connectivity theorem: $C^J \epsilon L = 0$

$\rightarrow \; \mathrm{d}\bar{x}^T \epsilon L = 0$

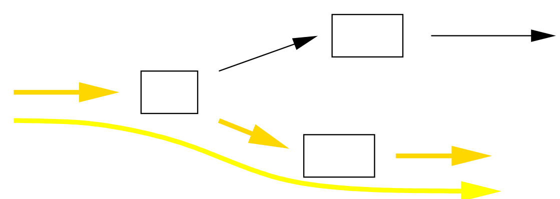


$\epsilon_1 \, \mathrm{d}\bar{x}_1 + \epsilon_2 \, \mathrm{d}\bar{x}_2 + \epsilon_3 \, \mathrm{d}\bar{x}_3 = 0$

Concentrations

Summation theorem: $C^S K = 0$

$\rightarrow \; \mathrm{d}\bar{x}^T K = 0$



$\sum_{i \in \text{El. Mode}} \mathrm{d}\bar{x}_i = 0$

Conclusions

- Assumption:
  Expression patterns serve as regulators of cell functions.

- Optimal gene expression patterns are predicted from
  (1) a cell model (2) a fitness function (3) external perturbations.

- Small perturbations: gene functions were described by
  response coefficients from metabolic control analysis

- General properties of expression patterns were derived,
  in particular relations to the metabolic network.

- Linear components within optimal expression profiles describe
  control of distinct fitness-relevant cell variables.
  $\rightarrow$ consistent with linear statistical models (e.g., PCA or ICA)