Structure of small genetic networks and evolution \textit{in silico}

Paul François Hervé Rouault, E. Siggia

Laboratoire de Physique Statistique, CNRS & ENS, Paris.
Center for Physics and Biology, The Rockefeller U., NYC.

Les Houches, april 07.
Genetic networks.

- Dynamics in a cell: bistability, oscillations (circadian, ...)
- Spatial patterns (C. elegans, somites, ...)
- Coordinated evolution of several genes/proteins.
- Design of synthetic modules.
A synthetic genetic switch

Switching can be induced by an IPTG or a temperature pulse.

Bistability requires dimerizations (or other interactions).
A synthetic genetic ring oscillator

The oscillation is based on three genes that repress each other in a circle ("rock-scissor-paper").
'Toggle' switch : mathematical analysis

\[
\frac{dA}{dt} = \frac{\alpha}{B_0 + B^\nu} - \delta_A A
\]

\[
\frac{dB}{dt} = \frac{\beta}{A_0 + A^\mu} - \delta_B B
\]

\(\nu, \mu\) must be strictly higher than 1 to have bistability, which requires at least four (and not two) elementary reactions. [Cherry and Adler, J. Theor. Biol. (2000)]
An overrepresented motif in transcriptional networks

The “feedforward loop” is overrepresented in the transcriptional networks of *E. Coli* and *S. Cerevisiae* (Milo et al., *Science* 298: 824-827(2002)).

**Function:** a persistence detector?
What are the designs that achieve a given function?

Can one sample them and add desired constraints (robustness,...) ?

Easyness of creation, evolvability,...?

Blueprints of useful networks.
Design by evolution/selection \textit{in silico}.

The inverse of the statistical approach: from the desired task to the network.

To design modules performing given tasks (e.g. switches and oscillators), without imposing \textit{a priori} any structure to the network, one evolves a collection of virtual “cells”.

One computer ’cell’ consists in
One computer ‘cell’ consists in
- a collection of genes $a$
- and associated proteins $A$

First implementation: transcription and translation condensed in one single step.
mRNA are included in the present version.
One computer 'cell' consists in
- a collection of genes
- and associated proteins

First implementation: transcription and translation condensed in one single step.
mRNA are included in the present version.

- transcriptional regulations
- post-transcriptional regulations.
<table>
<thead>
<tr>
<th>Representation</th>
<th>Corresponding equations</th>
</tr>
</thead>
</table>
| ![Diagram of a → A](image1) | \[
\frac{d}{dt}[A] = \tau_A[a] - \delta_A[A]
\] |
| ![Diagram of a + B](image2) | \[
\begin{align*}
\frac{d}{dt}[a] &= \theta[a:B] - \gamma[a][B] \\
\frac{d}{dt}[a:B] &= \gamma[a][B] - \theta[a:B] \\
\frac{d}{dt}[A] &= \tau_A[a] + \tau'_A[a:B]
\end{align*}
\] |
The evolution *in silico*. 

Cells
The evolution *in silico*.
The evolution \textit{in silico}.
The evolution *in silico*.

Elimination
The evolution *in silico*.

Duplication
The evolution *in silico*.

Mutations
Possible mutations

- The modification of a kinetic constant in an existing reaction

or the addition of

- A new transcriptional regulation

- A new post-transcriptional regulation

- A new gene

The process is iterated over several generations.
Fitness function for oscillators

Two concentrations are fixed $A_1$ and $A_2$.

ODEs are integrated

For $t = T/2, 3T/2, 5T/2, ...$ fitness is given by the integral $(A - A_1)^2$.

For $t = T, 2T, 3T, ...$ fitness is given by the integral $(A - A_2)^2$. 
Fitness evolution

- Generation 49
- Generation 99
- Generation 199
- Generation 249
- Generation 260
- Final State

Graph showing the evolution of concentration over generations and time, with a downward trend in score over generations.
The oscillating network
A purely biochemical oscillator

\[
\begin{align*}
A + A &\rightarrow A + A \\
A + A &\rightarrow A + B \\
B &\rightarrow A \\
A &\rightarrow A + B \\
B &\rightarrow A \\
a &\rightarrow b
\end{align*}
\]
A created bistable switch

Very different from two genes with reciprocal inhibition
A created bistable switch
A created bistable switch
A second type of switch
A second type of switch
A second type of switch
Transcriptional switches
Comparison with real networks

First switch: lactose operon, with allolactose binding to lac repressor.

Proposed in 1961 by Monod and Jacob (based on Lac operon) as an alternative to reciprocal inhibition (Delbrück, 1949)!
Comparison with real networks

First switch: lysis/lysogeny switch in a “simple” phage

CI transcriptionally represses Mor, Mor lifts this repression by binding to CI and preventing CI DNA binding.
An ingenious refinement: CI represses itself so as to avoid continuous production of repressor. This works because CI is a very stable protein and also because self-repression of CI is much weaker than CI repression of Mor.
Comparison with real networks

Second switch: development of competence in *B. subtilis*, Comk activates itself and is repressed by MecA.

Smits et al (2005), Maamar and Dubnau (2005): experimental demonstration that this stripped network is indeed a switch!
Endogeneous oscillator: the circadian clock
Circadian activities of whole animals and single cells
Liu et al, Cell (1997)
The core structure of circadian clocks


Organism
Neurospora Crassa
Drosophila
Mammals

Activators A
WC-1, WC-2
dCLK
CLOCK, BMAL

Repressors B
FRQ
PER, TIM
PER, CRY

The created networks are working examples without delays or high Hill coefficients ⇒ motivation for new models of the circadian rhythms [for Neurospora, P. François Biophys. J. 88, 2369 (2005)].
A core genetic circuit: the Mixed Feedback Loop

A loop combining transcriptional and post-transcriptional interaction (i.e. protein-protein interaction) is at the core of several of these networks.

This Mixed Feedback Loop has now been found to be over-represented in S. Cerevisiae and E.Coli (Yeger-Lotem et al, PNAS 2004).
Mathematical analysis of the MFL

Reduced parameters: \( \rho_0 = \beta \rho_f / (\rho_A \delta_r), \rho_1 = \beta \rho_b / (\rho_A \delta_r) \)

A small parameter: \( \delta_r / \sqrt{\rho_A \gamma} \)

( P. François and V. Hakim, PRE (2005) )
The algorithm finds known (with complete description) and original designs.

**An important lesson:** The post-transcriptional interactions play a crucial role: the function of the networks cannot be understood at all by focusing only on the transcriptional regulations (*protein sequestration* in a complex appears to be a particularly important mechanism).
From single cells to multicellular pattern formation

• Can one evolve more complicated structures?

• Spatial patterns, morphogenesis?

• A test case: segmentation
Segmentation in a static gradient

• Drosophila early embryogenesis is very well studied (Thieffry and Perkins’s lectures!): a nice bench mark.

• Evolution of a collection of “organisms” (hundred cell each).

• Fitness: maximise the number of “segments”: the number of jumps in the concentration of a test protein.

• Only transcriptional interactions for numerical tractability.
Evolution of segmentation in a static gradient
Evolution of segmentation in a static gradient

• Cascades of repressors reminiscent of Drosophila segmentation network.

• Feedforward loop that gives a general way to express a gene at an intermediate gradient concentration
  Real example: Dorsal activates rhomboid and less efficiently snail. Snail represses rhomboid. Rhomboid expressed at intermediate Dorsal concentration (neurogenic ectoderm)

• Reasonable results. What does one find for a dynamic gradient (short germ insects, somites,...)?
Segmentation in vertebrates: somite formation

Segmentation and oscillations


Evolution of segmentation in a temporal gradient

• Dynamics as the gradient sweeps across the array.

• Score: numbers of jumps in a concentration of a protein after the gradient has disappeared.

• Again, only transcriptional interactions for numerical tractability (delays to account for intermediate steps).
Evolution of segmentation in a temporal gradient
Segmentation as an oscillating/bistable transition?

Oscillations when external signal (FGF8?) is high, bistability when it is low. Bistability encodes the oscillation phase at the time of transition in a binary (and cell autonomous) way.
Some conclusions and perspectives

• A simple model of sequential segmentation. Hopefully, useful in different contexts (somites, short-germ insects, other segmented structures e.g. limbs ?,...)

• Very constrained evolutionary path *in silico*: first bistability to have persistent activation after disappearance of the gradient, then repressors added (for creation of high/low boundaries), finally negative feedback and oscillations. Real evolution?

• Early appearance of sequential segmentation? Multiple interconversion between the two modes of segmentation? (new phylogeny: hymenoptera -long germ wasp Nasonia- at the base of holometabolous insects -include short and long germ).
The End (for today).

Thank you!
Temperature compensation

Selection of activation energies for temperature compensation:
• a $10^\circ K$ increase: $T: 300^\circ K \rightarrow 310^\circ K$
• the kinetic constants increase $> 30\%$,
• period change $< 3\%$. 

![Graph showing temperature compensation](image)
Fitness function for the switches

The desired two stable states are chosen \((A_1, B_1)\) and \((A_2, B_2)\). ODEs are integrated, the “fitness” is given by the integral \((A - A_1)^2 + (B - B_1)^2\). Pulse of B protein ODEs are integrated, the fitness is given by the integral \((A - A_2)^2 + (B - B_2)^2\).
Transcriptional regulations

\[ A \rightarrow {} \]

\[ B \rightarrow {} \]

\[ B \rightarrow {} \]

\[ a \rightarrow {} \]

\[ a \rightarrow {} \]

\[ B \leftrightarrow a \]

\[ B \leftrightarrow a \]
Post-transcriptional regulations
Transcriptional switches

![Diagram of transcriptional switches with nodes A and B, showing regulatory interactions between 'a' and 'b' genes.](image)
A second type of switch
A second type of switch
A second type of switch
'Toggle' switch: mathematical analysis

\[
\begin{align*}
\frac{dA}{dt} &= \frac{\alpha}{B_0 + B^\nu} - \delta_A A \\
\frac{dB}{dt} &= \frac{\beta}{A_0 + A^\mu} - \delta_B B
\end{align*}
\]

\(\nu, \mu\) must be strictly higher than 1 to have bistability, which requires at least four (and not two) elementary reactions.

[Cherry and Adler, J. Theor. Biol. (2000)]