

# Inferring regulatory networks from spatial and temporal gene expression patterns

Y. Fomekong Nanfack<sup>1</sup>, Boaz Leskes<sup>1</sup>, Jaap Kaandorp<sup>1</sup> and Joke Blom<sup>2</sup>

<sup>1</sup>) Section Computational Science, University of Amsterdam, Kruislaan 403, NL-1098 SJ Amsterdam  
<sup>2</sup>) Center for Mathematics and Computer Science (CWI), Kruislaan 413, NL- 1098 SJ Amsterdam



## 1. Introduction

Regulatory networks play a fundamental role in many biological processes. In this project we aim to develop a model for simulating regulatory networks that are capable of quantitatively reproducing spatial and temporal expression patterns in developmental processes (case studies: *Drosophila* and sponges). An important option for the analysis of regulatory control systems are simulation models in combination with an optimization algorithm, due to the uncertainty value of parameters. Here we present a model based approach for inferring networks from spatial and temporal expression patterns.

## 2. Early development of *Drosophila* embryo

Here, we briefly discuss the body plan formation of *Drosophila* embryo. Fig 2.1 shows the basic outline of the anteroposterior (A/P) axis regulatory hierarchical network of genetic interactions that together regulate the process of development. In our model we simulate the formation of the second stripe eve 2 (Fig 2.2)

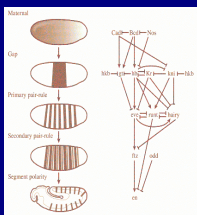


Fig 2.1. left: different stages of development of five classes of A-P axis genes. Right: selected members of the regulatory network. © 2001 by S.B Carroll

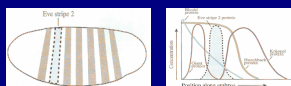


Fig 2.2. Illustration of the Eve stripe 2 formation. Left: formation of the eve stripe 2 controlled by the gene eve-skipped. Right: location of Bicoid, Hunchback (who activate eve) and Giant and Krüppel (repress eve). © 2001 by S.B Carroll

## 3. Acquisition of quantitative Data

From an image of a *Drosophila* embryo stained for three different genes, we construct a binary nuclear mask and find the protein concentration of each gene within all nuclei. The binary nuclear mask is an image in which all the different nuclei have been isolated by an image segmentation and edge detection. The nuclei contour is found mainly by a watershed mathematical morphology operation. The resulting image is the nuclei without overlap.

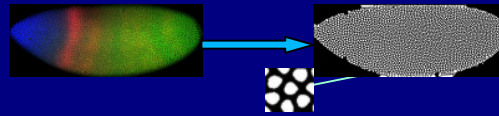


Fig 5.1 Construction of a binary nuclei mask. Above: left, an original image of a drosophila embryo stained for three genes. Right, the final nuclear mask after segmentation. Below a crop of the binary nuclear mask where we can see the isolated nuclei.

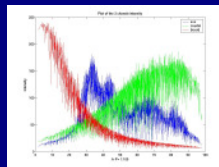


Fig 5.2 This figure gives the gene concentration along the A-P axis of the drosophila image in figure 5.1. This is obtained by quantifying the pixel value in each nuclei for all the 3 channels. The Bicoid concentrate on the left and Caudal on the right (see Fig 2.2, b)

## 4. Mathematical Model

The model is a generalization of the standard connectionist model used for modeling genetic interaction [2, 3]. It assumes that three basic processes govern gene product concentration:

$$\left( \begin{array}{l} \text{time rate of change of} \\ \text{gene product concentration} \end{array} \right) = \text{regulation} + \text{Diffusion} + \text{Decay}$$

**Model Abstraction:** here we decouple in time the regulation and diffusion. The genetic regulation is modelled by a set of differential equation (eq. 4.1) in which protein can activate/inhibit each other, and diffuse between the cells. Cells are discrete and allow us to support the general case in two or three dimensions and integrate migrating cells.

$$\frac{\partial g_{ij}}{\partial t} = \underbrace{R_j \Phi(h_{ij})}_{\text{regulation}} + \underbrace{D \nabla^2 g_{ij}}_{\text{diffusion}} - \underbrace{\lambda_j g_{ij}}_{\text{decay rate}}$$

$$\text{where } h_{ij} = \sum_{k=1}^{N_i} w_{jk} g_{ik} + h_i \text{ and } i=1, \dots, N_c \text{ and } j=1, \dots, N_g$$

$g_{ij}$  is the intracellular concentration of gene  $j$  within cell  $i$ .

and  $W_{ab}$  is the matrix describing the effect of gene  $a$  on gene  $b$

$$4.1$$

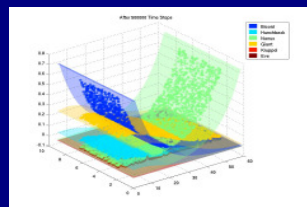


Fig 4.1 concentration gradients for a 6-genes network. This 3D graph shows that our model is capable of reproducing the spatial temporal expression patterns in 2D

## 5. Parameter Optimization

Real expression data procured the quantitative value of the concentration of a number of proteins in a large number of nuclei cells at a number of different developmental time. It is a main issue to have a set of parameters for Equation 4.1 that gives the closest possible fit to the real expression data. The optimization process allows to find such parameters. We are using a simulated annealing approach to find the value of these parameters.

The simulated annealing used is based on the Lam schedule cooling temperature [5] which is accurate and rapid as it decreases the temperature after every move (Equation 5.1) The Cost function, which evaluates the quality of the current solution used is the least squares means of the deviation of the solution of the expression data. (Equation 5.2).

$$S_n = S_{n-1} + \left( \frac{\lambda}{\sigma(S_n)} \right) \left( \frac{1}{S_n^2 \sigma^2(S_n)} \right) * \left( \frac{4 \rho_0 (1 - \rho_0)^2}{(2 - \rho_0)^2} \right)$$

where  $S_n = 1/T_n$  is the inverse temperature at step  $n$   $\rho_0 = 0.44$ ,  $\lambda$  is a quality parameter and  $\sigma(S_n)$  is the standard deviation.

$$E_{\text{pattern}} = \sum (g_{ij}(t)_{\text{model}} - g_{ij}(t)_{\text{data}})^2$$

$g_{ij}$  is the intracellular concentration of gene  $j$  within cell  $i$  at time  $t$

$$5.2$$

In Fig 6.1, the two plots illustrate the simulation of the model described by equation 4.1. The left figure gives the solution from an initial set of parameters considered as the target. The right figure illustrates the simulation from a given set of parameters optimized. A close look at the parameter  $W$  from the model and the data (see yellow box) shows that in some cases, there is a significant difference between the values. From this observation, it is interesting to see if the non-uniqueness of the parameters solution comes from the robustness of the model.

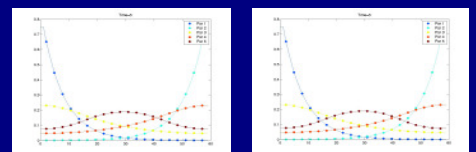


Fig 6.1. 5-genes network simulation. Left: Simulation of Eq 4.1 from a set of target initial parameters. Right: simulation of Eq 4.1 with parameters obtain from optimization

$$\begin{array}{l} w_{11}^t = 0 \quad w_{11}^m = -1.294 \\ w_{55}^t = 0.600 \quad w_{55}^m = 0.605 \end{array}$$

where  $t$ =target and  $m$ =model

## 6. Discussion

The model presented here is capable of reproducing the formation of eve stripe 2 in *Drosophila* in 2D (Fig. 4.1) and support 3D. We need to improve our model and find an answer to the non-uniqueness problem. So far only artificial data have been used and it is important to move to realistic biological data. Thus, in the data acquisition process, we'll need to integrate different images stained for 3 genes to obtain quantitative data for networks with more than 3 genes and used a more appropriate cost function in the optimization process.

## Acknowledgement

This project (635.100.0.10), is part of the Computational Life Science's project and is financed by the NWO as 3D-RegNet: simulation of developmental regulatory networks Research team is composed as follow:  
 PhD Students: Yves Fomekong Nanfack (UvA) & Maksat Ashrylalyev (CWI).  
 Staff: Jaap Kaandorp (UvA), Joke Blom (CWI), Peter Sloot (UvA) & Werner Müller (Johannes Gutenberg Universität-Mainz)  
 website: <http://www.science.uva.nl/research/scs/3D-RegNet/>

## References

- 1) S.B Carroll & al: From DNA to diversity: molecular genetics and the evolution of animal design (2001)
- 2) E. Mjølness & al: Connectionist model of development. J. Theo Biol 152 (1991) 429-453
- 3) J. Reinitz & al: Mechanism of eve stripe formation. Mechanism of Development 49 (1995) 133-158
- 4) Krul & al: Modelling Developmental Regulatory Networks. In Proceedings of Computational Science - ICCS 2003,
- 5) J. Lam & J.M. Deslome: An efficient simulated annealing: Derivation. Technical report 8816. Yale Electrical Engineering Department (1988a)