

Modeling, Simulating and Calibrating Genetic Regulatory Networks: An Application to *Drosophila* Development with Multi-Objective Optimization Techniques

Rui Dilão
NonLinear Dynamics Group, IST
<https://sd.ist.utl.pt>

- A **software tool** to model **genetic regulatory networks**.
(More details in the tutorial on Computational Biology)
- The **calibration** of a pattern formation model in *Drosophila* early development --- **parameter determination by a swarm technique**.
Non-determinacy of parameters.
- Evolutionary computation techniques (**multi-objective**) to calibrate the parameters of a morphogenesis model of *Drosophila* early development.
Pareto fronts. Non-unicity of parameter solutions.
- Making **predictions** about protein regulation.

A software tool to model genetic regulatory networks. Applications to the modeling of threshold phenomena and of spatial patterning in *Drosophila*.

R. Dilão and D. Muraro, *PLoS ONE*, 5 (5) (2010) 1-10 (e10743).

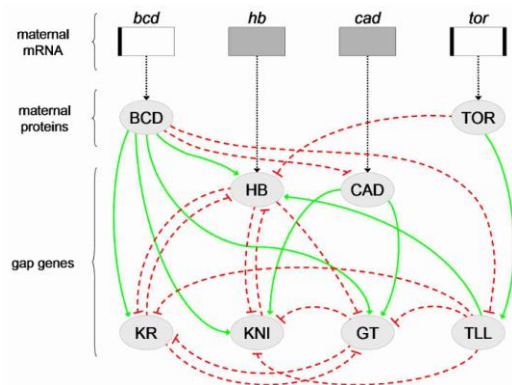
We present a **general methodology** in order to **build mathematical models of genetic regulatory networks**. This approach is based on the mass action law and on the Jacob and Monod operon model.

--- The mathematical models are built symbolically by **Mathematica software package GeneticNetworks**. This package accepts as input the interaction graphs of the transcriptional activators and repressors and, as output, gives the mathematical model in the form of a system of ordinary differential equations. All the relevant biological parameters are chosen automatically by the software.

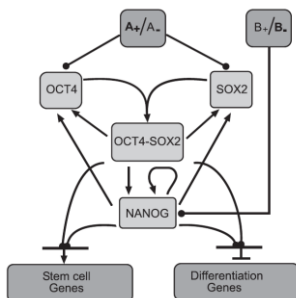
--- We show that **threshold effects** in biology **emerge from the catalytic properties of genes** and its **associated conservation laws**.

---We show that spatial patterning in embryology can be obtained as a **dynamic threshold effect**.

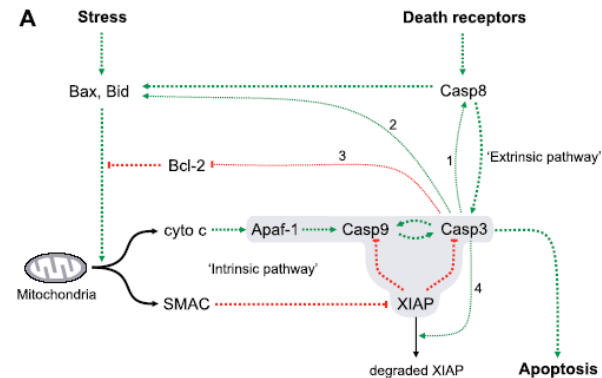
To develop a computational and an analytical tool to analyse regulatory networks, enabling the calibration with biological parameters and protein concentrations.



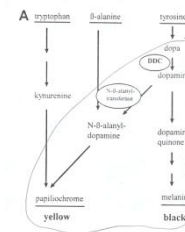
F. Alves and R. Dilão, *J. Theoretical Biology*, 241 (2006) 342-359.



V. Chickarmane, C. Troein, U. A. Nuber, H. M. Sauro, C. Peterson, *Transcriptional Dynamics of the Embryonic Stem Cell Switch*, *PLOS Comp. Biology*, 2(9) (2006) 1080-1092.



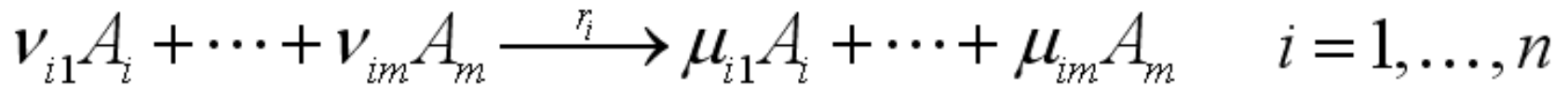
S. Legewie, N. Blüthgen, H. Herzel, *Mathematical Modeling Identifies Inhibitors of Apoptosis as Mediators of Positive Feedback and Bistability*, *PLOS Comp. Biology*, 2(9) (2006) 1061-1073.



Koch et al. 98, *Development*

The mass action law

An ensemble of m chemicals A_i that react in a media according to the mechanisms:



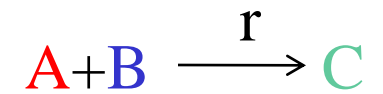
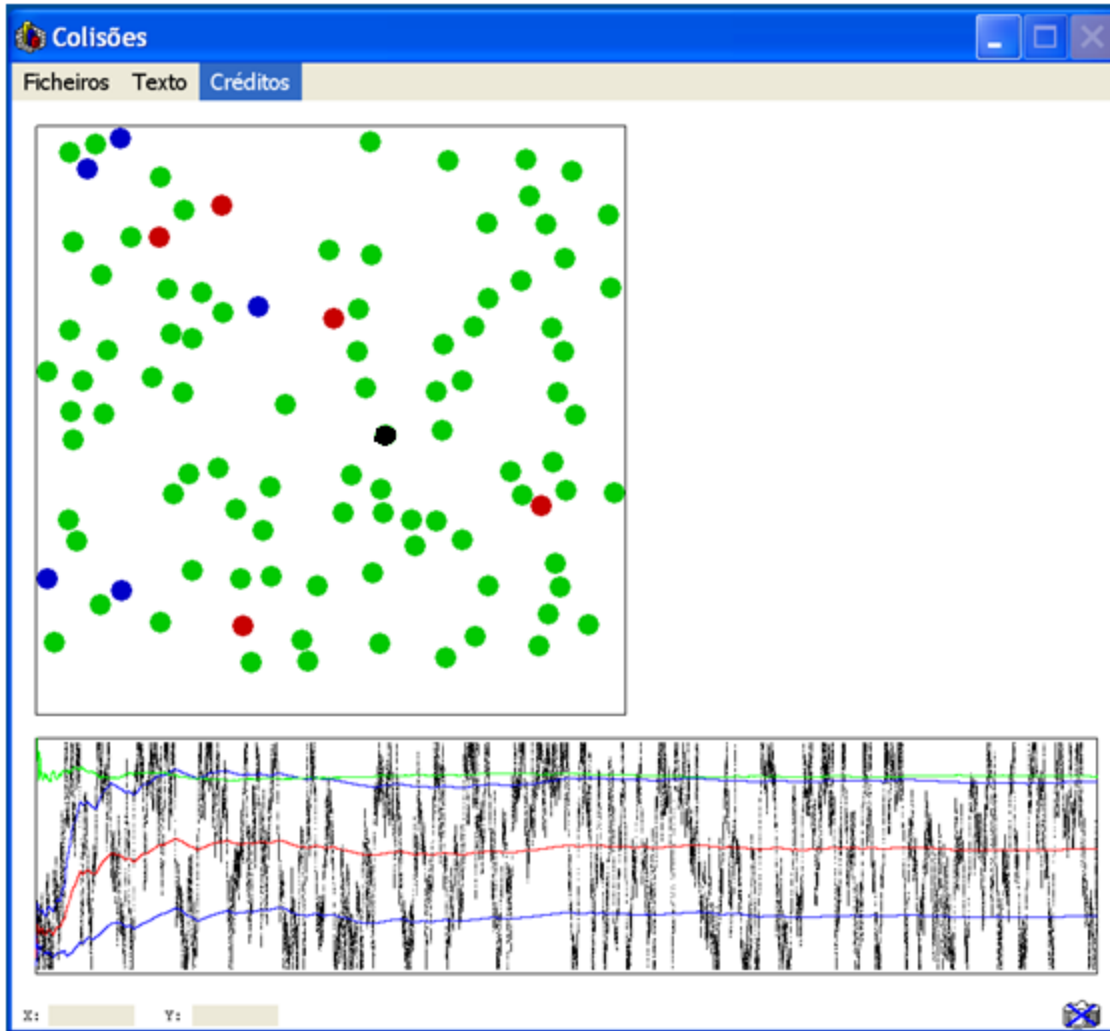
There are m chemicals and n reactions. Assuming a well stirred media, and that the chemicals have a **Brownian type** motion, we have the reaction evolution laws:

$$a = [A] \quad \frac{da_j}{dt} = \sum_{i=1}^n r_i (\mu_{ij} - \nu_{ij}) a_1^{\nu_{i1}} \cdots a_m^{\nu_{im}} \quad j = 1, \dots, m$$

$$\frac{dA}{dt} = \Gamma \omega \quad \omega = (a_1^{\nu_{11}} \cdots a_m^{\nu_{1m}}, \dots, a_1^{\nu_{n1}} \cdots a_m^{\nu_{nm}}) \quad \Gamma \text{ is a } n \times m \text{ matrix, with rank } r$$

Conservation laws:
$$\sum_{k=1}^m \sigma_{kl} A_k = a_l \quad l = 1, \dots, m - r$$

The microscopic mechanism associated with the mass action law



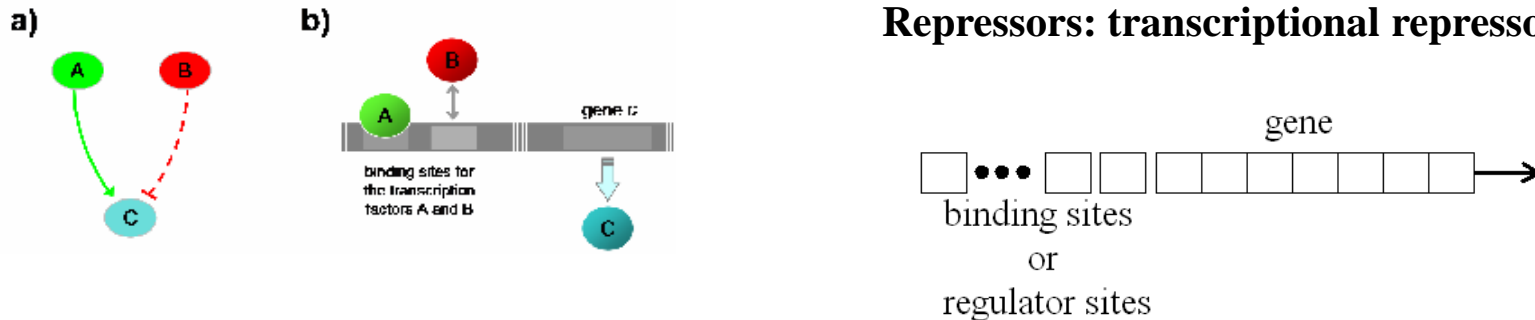
$$\frac{dC}{dt} = rA.B$$

Transcriptomics -- cis-regulation of gene expression

F. Alves and R. Dilão, A simple framework to describe the regulation of gene expression in prokaryotes, *Comptes Rendus - Biologies*, 328 (2005) 429-444.

This approach is based on the operon model (Jacob and Monod, 1961) as a paradigm of genetic regulation in bacteria.

Activators: transcriptional activators
Repressors: transcriptional repressors



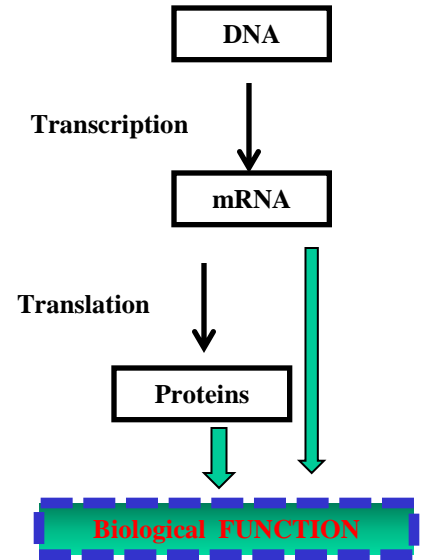
Model assumptions:

- Genes are considered templates for protein production.
- Transcriptional and translational mechanisms are described by one overall rate constant.
 - We only consider bi-molecular mass action law, (no *ad-hoc* regulatory functions in intermediate states like Michaelis-Menton type functional forms).
- Regulation (activation, repression, competition) occurs only through the binding sites.

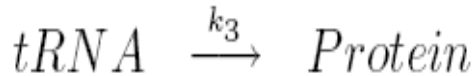
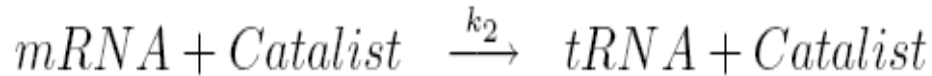
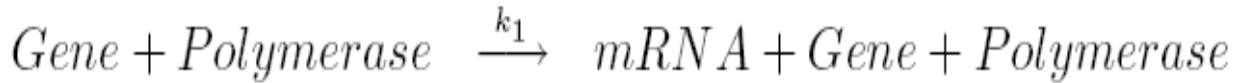
Choices:

- To make, *ab initio*, the mathematical framework as simple as possible in such a way that any regulatory genetic network can be described in this framework.
- We don't want to introduce had hoc threshold effects (at the end we will have dynamic threshold effects).

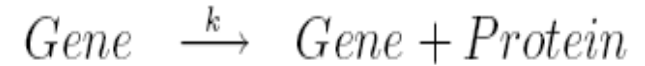
Regulation of gene expression



Free protein production: basic model



Simplified model that will be used



(conservation of information)

A gene is a catalytic substance

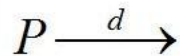
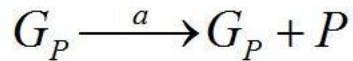
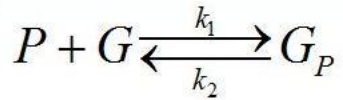
(R. Dilão and D. Muraro, *PLoS ONE*, 5 (5) (2010) 1-10 (e10743))

$$\frac{d(Protein)}{dt} = k \cdot Gene - d \cdot Protein$$

$$\frac{d(Gene)}{dt} = 0$$

Dynamic threshold effects and conservation laws

Self-activation



$$\dot{G} = -k_1 PG + k_2 G_P$$

$$\dot{G}_P = k_1 PG - k_2 G_P$$

$$\dot{P} = aG_P - dP - k_1 PG + k_2 G_P$$

Conservation law $G(t) + G_P(t) = G(0) + G_P(0) = c_2$

$$\dot{G} = -k_1 PG + k_2(c_2 - G)$$

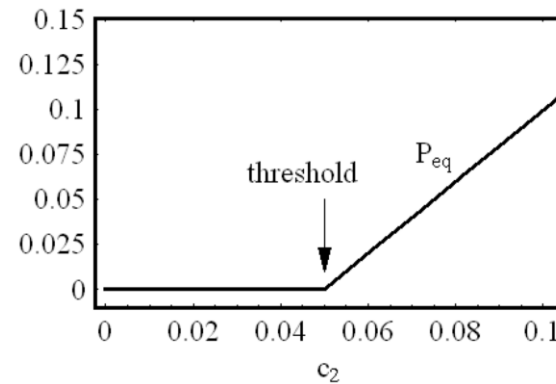
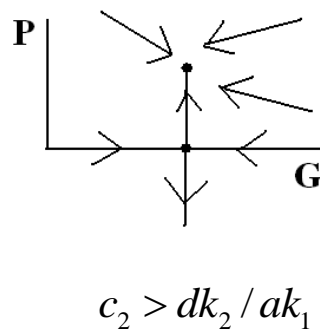
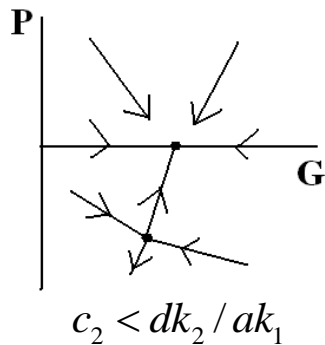
$$\dot{P} = a(c_2 - G) - dP - k_1 PG + k_2(c_2 - G)$$

Fixed points

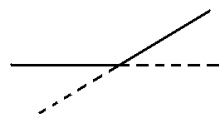
$$(c_2, 0) \quad \left(\frac{dk_2}{ak_1}, \frac{ak_1 c_2 - dk_2}{dk_1} \right)$$

Threshold condition:

$$c_2 > dk_2 / ak_1$$



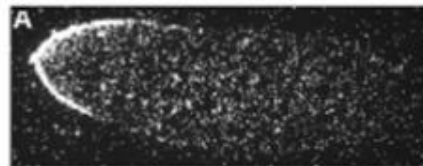
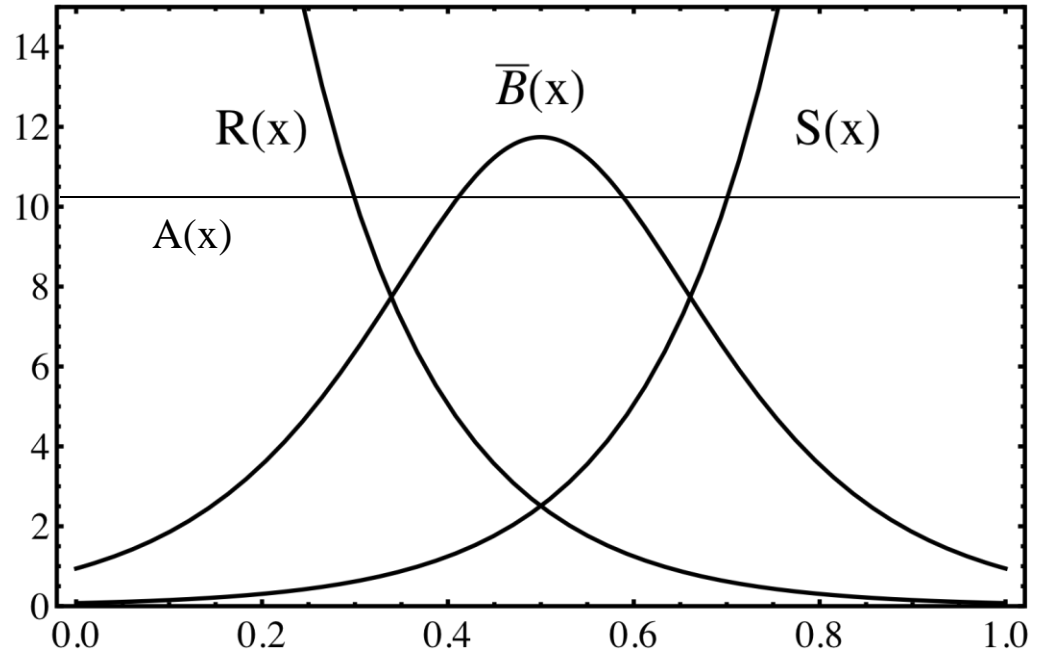
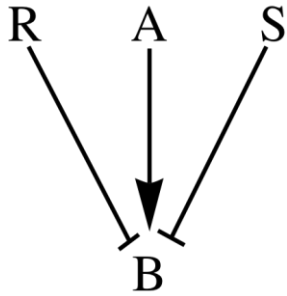
Transcritical Bifurcation



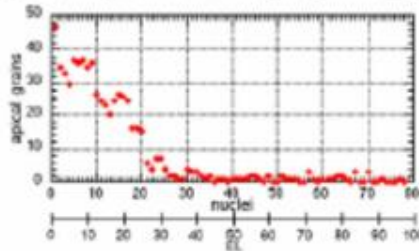
The threshold value depends of an initial condition

Spatial patterning without diffusion

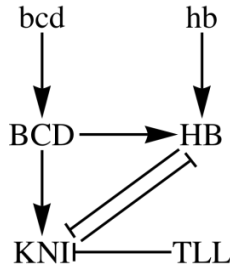
Spatial patterning obtained without diffusion (dynamic thresholds)



x



(biological) input



Software



model

```

HB[t] = -HB[t] khb + (khbcdon BCD[t] + (khbcdoff / HB[t]) KNI[t] KNI1,0[t]) -
  HB[t] Khbcd KNI1,0[t] - HB[t] Khbcd KNI1,1[t] - HB[t] Khbcd KNI1,2[t] +
  Khbcd KNI1,0[t] + Khbcd KNI1,1[t] + Khbcd KNI1,2[t] + Khbcd KNI1,3[t]
KNI[t] = -KNI[t] kkni + KNI[t] khbcd HB[t] + KNI[t] khbcd HB1,0[t] + khbcd HB1,0[t] + khbcd HB1,1[t] +
  (khbcdoff / HB[t]) KNI[t] KNI1,0[t] + (khbcdoff / HB[t]) KNI[t] KNI1,1[t] + (khbcdoff / HB[t]) KNI[t] KNI1,2[t] +
  (khbcdoff / HB[t]) KNI[t] KNI1,3[t] + khbcd HB1,0[t] + khbcd HB1,1[t] + khbcd HB1,2[t] + khbcd HB1,3[t]
HB1,0[t] = -khbcd HB1,0[t] + KNI[t] khbcd HB[t] + khbcd HB1,0[t] + khbcd HB1,1[t] + khbcd HB1,2[t] + khbcd HB1,3[t]
HB1,1[t] = -khbcd HB1,1[t] + KNI[t] khbcd HB[t] + khbcd HB1,0[t] + khbcd HB1,1[t] + khbcd HB1,2[t] + khbcd HB1,3[t]
HB1,2[t] = -khbcd HB1,2[t] + KNI[t] khbcd HB[t] + khbcd HB1,0[t] + khbcd HB1,1[t] + khbcd HB1,2[t] + khbcd HB1,3[t]
HB1,3[t] = -khbcd HB1,3[t] + KNI[t] khbcd HB[t] + khbcd HB1,0[t] + khbcd HB1,1[t] + khbcd HB1,2[t] + khbcd HB1,3[t]
KNI1,0[t] = -kkni KNI1,0[t] + HB[t] Khbcd KNI1,0[t] - HB[t] Khbcd KNI1,1[t] +
  TLL Kkni KNI1,0[t] + Kkni KNI1,0[t] + Kkni KNI1,1[t] + Kkni KNI1,2[t] + Kkni KNI1,3[t]
KNI1,1[t] = -kkni KNI1,1[t] + HB[t] Khbcd KNI1,0[t] - HB[t] Khbcd KNI1,1[t] +
  TLL Kkni KNI1,1[t] + Kkni KNI1,0[t] + Kkni KNI1,1[t] + Kkni KNI1,2[t] + Kkni KNI1,3[t]
KNI1,2[t] = -kkni KNI1,2[t] + HB[t] Khbcd KNI1,0[t] - HB[t] Khbcd KNI1,2[t] +
  TLL Kkni KNI1,2[t] + Kkni KNI1,0[t] + Kkni KNI1,1[t] + Kkni KNI1,2[t] + Kkni KNI1,3[t]
KNI1,3[t] = -kkni KNI1,3[t] + HB[t] Khbcd KNI1,0[t] - HB[t] Khbcd KNI1,3[t] +
  TLL Kkni KNI1,3[t] + Kkni KNI1,0[t] + Kkni KNI1,1[t] + Kkni KNI1,2[t] + Kkni KNI1,3[t]
  
```

Software

R. Alves-Pires, R. Dilão and D. Muraro, Kinetics: A Mathematica© Package to calculate and to analyze the equations of chemical kinetics.

R. Dilão and D. Muraro, GeneticNetworks: A Mathematica© Package for the modeling and simulation of Genetic Regulatory Networks.

Download:

<https://sd.ist.utl.pt/Download/download.html/GeneticNetworks.zip>

mRNA diffusion explains protein gradients in *Drosophila* early development

R. Dilão and D. Muraro, *Journal of Theoretical Biology*, 264 (2010) 847-853, doi:10.1016/j.jtbi.2010.03.012.

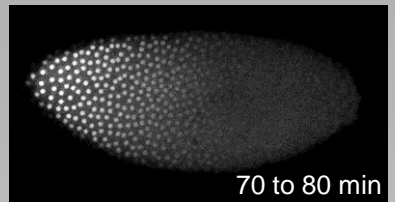
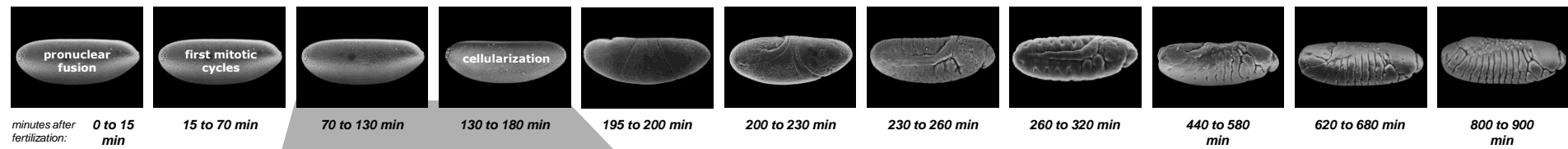
We propose a new model describing the production and the establishment of the stable gradient of the Bicoid protein along the antero-posterior axis of the embryo of *Drosophila*. In this model, we consider that ***bicoid* mRNA diffuses** along the antero-posterior axis of the embryo and the protein is produced in the ribosomes localized near the syncytial nuclei. Bicoid protein stays localized near the syncytial nuclei as observed in experiments.

We **calibrate** the parameters of the mathematical model with **experimental** data taken during the cleavage stages 11 to 14 of the developing embryo of *Drosophila*.

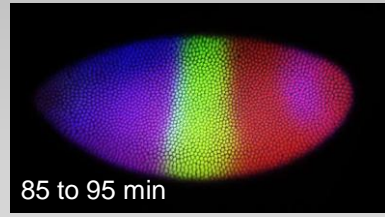
--- We obtain good agreement between the experimental and the model gradients, with **relative errors in the range 5-8%**.

--- The inferred diffusion coefficient of *bicoid* mRNA is in the range $4.6 \cdot 10^{-12}$ - $1.5 \cdot 10^{-11} \text{ m}^2\text{s}^{-1}$, in agreement with the theoretical predictions and experimental measurements for the diffusion of macromolecules in the cytoplasm.

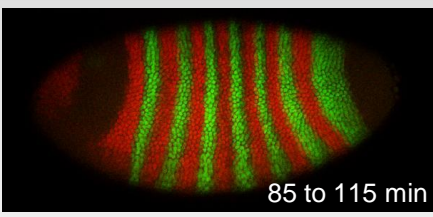
!!! The model based on the mRNA diffusion hypothesis is consistent with the known observational data, supporting the recent **experimental findings of the gradient of *bicoid* mRNA** in *Drosophila* [Spirov *et al.* (2009) *Development* 136:605-614].



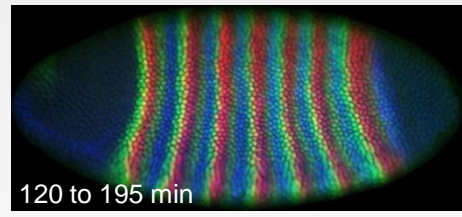
maternal-effect genes



gap genes

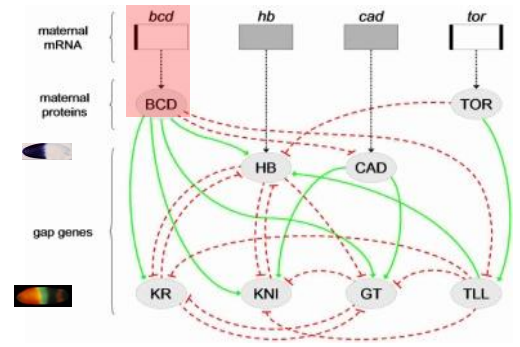


pair-rule genes



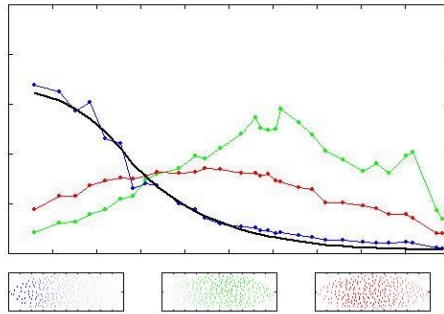
segment-polarity genes

Genetic (transcriptional) regulatory network

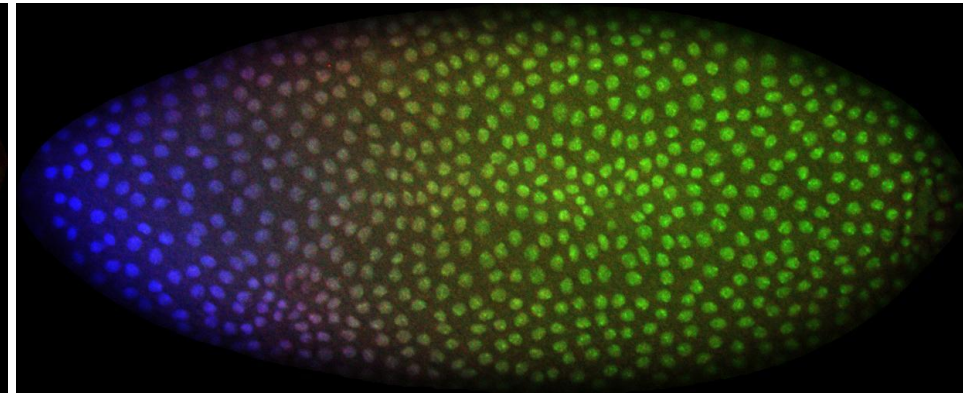
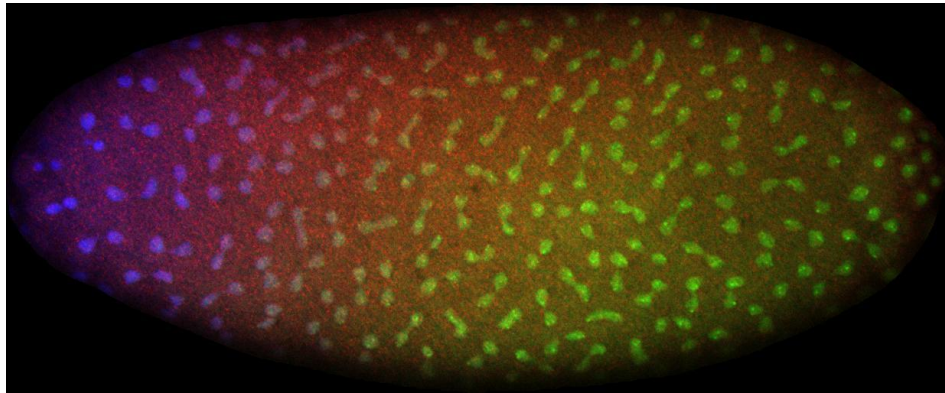
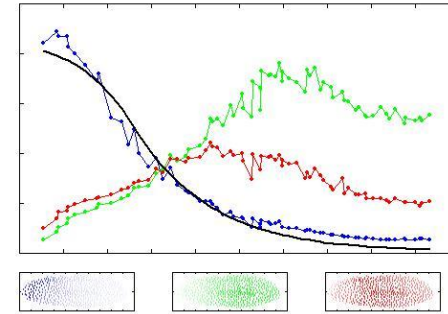


Alves & Dilão, JTB 2006

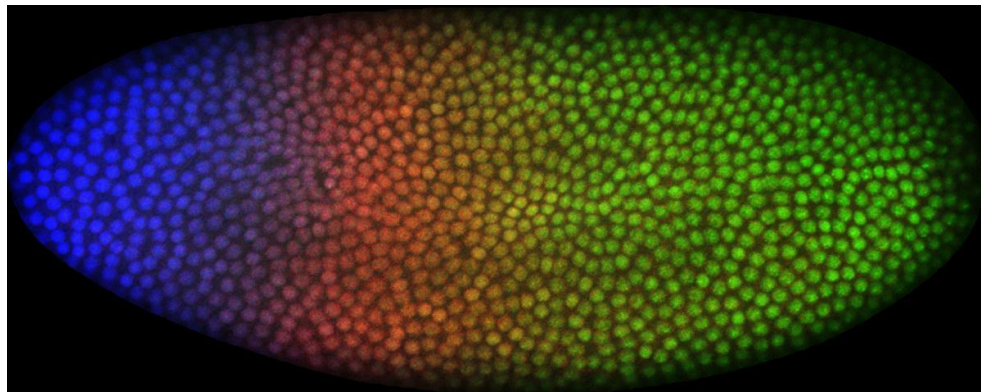
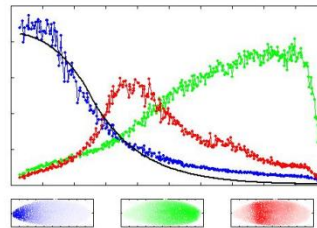
Cleavage cycle 11 (ab18)



Cleavage cycle 12 (ab17)



Cleavage cycle 13 (ab12)



In cleavage cycles 11, 12, 13 and 14, images show that the Bicoid protein is localised near the nucleus.

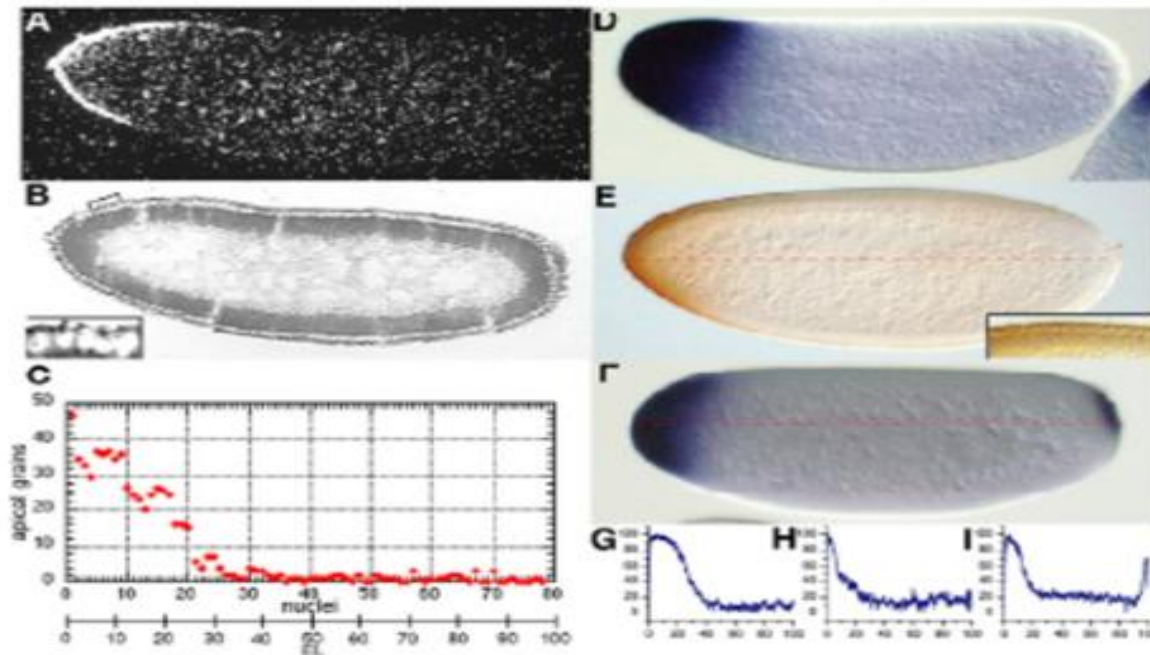
Bicoid protein do not diffuse!

Experimental evidence of mRNA diffusion:

Development 136, 605–614 (2009) doi:10.1242/dev.031195

(2009) Formation of the *bicoid* morphogen gradient: an mRNA gradient dictates the protein gradient

Alexander Spirov¹, Khalid Fahmy^{2,*}, Martina Schneider^{2,†}, Erich Frei³, Markus Noll^{3,‡} and Stefan Baumgartner^{2,‡}



Cha, et al., Cell (2001), "saltatory movements in injected mRNA bicoid with dispersion but without localization".

Forrest and Gavis, Curr. Biol. (2003), "mRNA nanos has diffusive like behaviour".

mRNA diffusion model:



$$\frac{\partial R}{\partial t} = -dR + D_r \frac{\partial^2 R}{\partial x^2} + \text{zero flux}$$

$$\frac{\partial B}{\partial t} = aR$$

a/d is the number of protein molecules produced by mRNA molecule

bicoid mRNA initial condition:

$$R(x, t = 0) = \begin{cases} A > 0 & \text{if } 0 \leq \ell_1 \leq x \leq \ell_2 \leq L \\ 0 & \text{otherwise} \end{cases}$$

Bicoid protein steady state:

$$B_{eq}(x) = a_1 \frac{(L_2 - L_1)}{L} + 2a_1 \sum_{n=1}^{\infty} \frac{1}{n\pi + \frac{a_2^2}{a_1}} \cos\left(\frac{n\pi x}{L}\right) \left(\sin\left(\frac{n\pi L_2}{L}\right) - \sin\left(\frac{n\pi L_1}{L}\right) \right)$$

$$B_{eq}(x) = 2 \frac{a_1}{e^{2a_2/L} - 1} \cosh\left(a_2 \frac{x}{L}\right) \left(\sinh\left(a_2 \frac{\ell_2}{L}\right) - \sinh\left(a_2 \frac{\ell_1}{L}\right) \right) + \frac{a_1}{2} \left(e^{-a_2(x+\ell_1)/L} - e^{-a_2(x+\ell_2)/L} \right) + I(x)$$

$$I(x) = \begin{cases} a_1 \left(e^{-a_2(\ell_1-x)/L} - e^{-a_2(\ell_2-x)/L} \right) / 2, & \text{if } x < \ell_1 \\ a_1 - \frac{a_1}{2} \left(e^{-a_2(x-\ell_1)/L} + e^{-a_2(\ell_2-x)/L} \right), & \text{if } \ell_1 \leq x \leq \ell_2 \\ a_1 \left(e^{-a_2(x-\ell_2)/L} - e^{-a_2(x-\ell_1)/L} \right) / 2, & \text{if } x > \ell_2 \end{cases}$$

Parameters to be determined from the experimental data.

$$a_1 = A \frac{a}{D}, \quad a_2^2 = d \frac{L^2}{D}, \quad \ell_1, \quad \ell_2$$

Free parameters: L and D

The model depends on 7 parameters
The solution depends on 4 parameters

Fitting the mRNA diffusion model with the experimental data for protein gradients --- calibration and validation of the model

Ad-hoc parameter values:

$$D = 10^{-11} \text{ m}^2\text{s}^{-1} \quad \text{bicoid mRNA}$$

$$L = 0.5 \times 10^{-3} \text{ m.}$$

Non-determinancy of parameters.

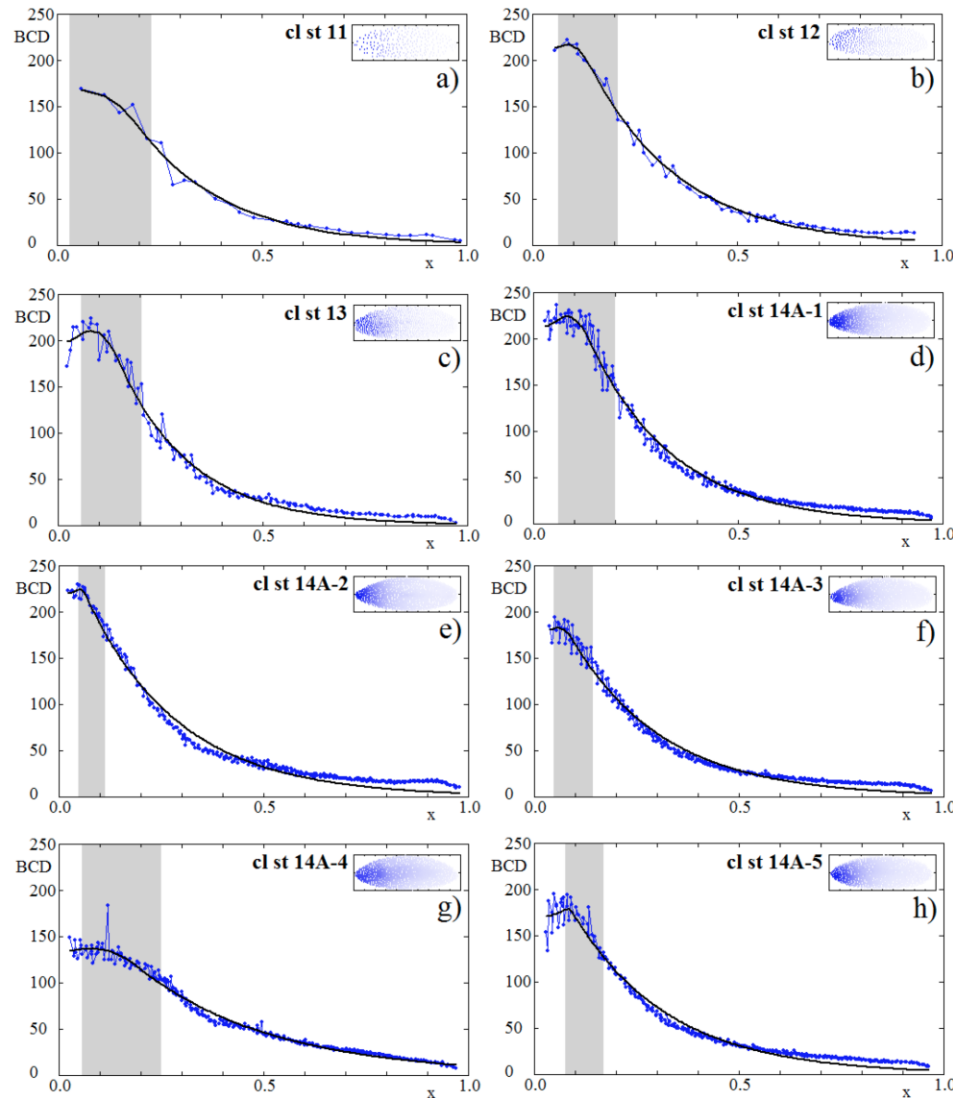


Table 1: Fitted model parameters for the protein Bicoid antero-posterior distributions

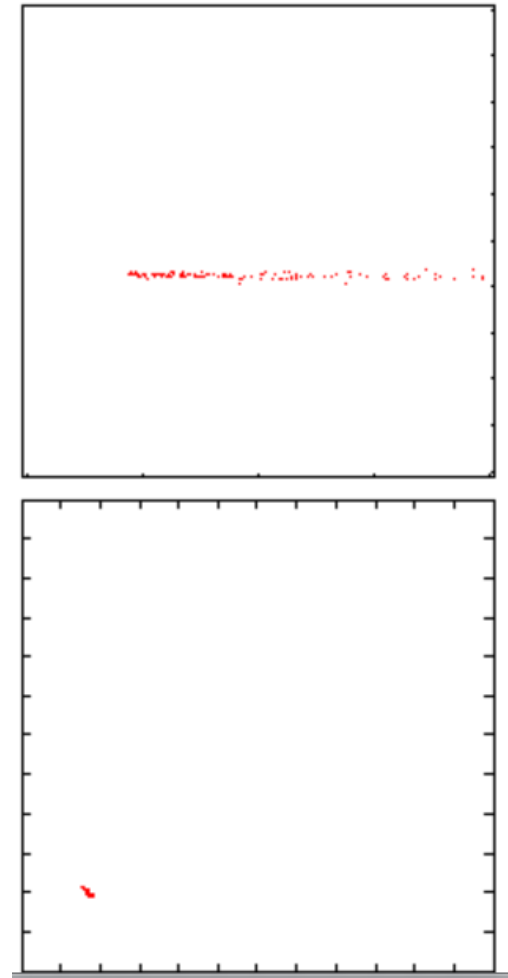
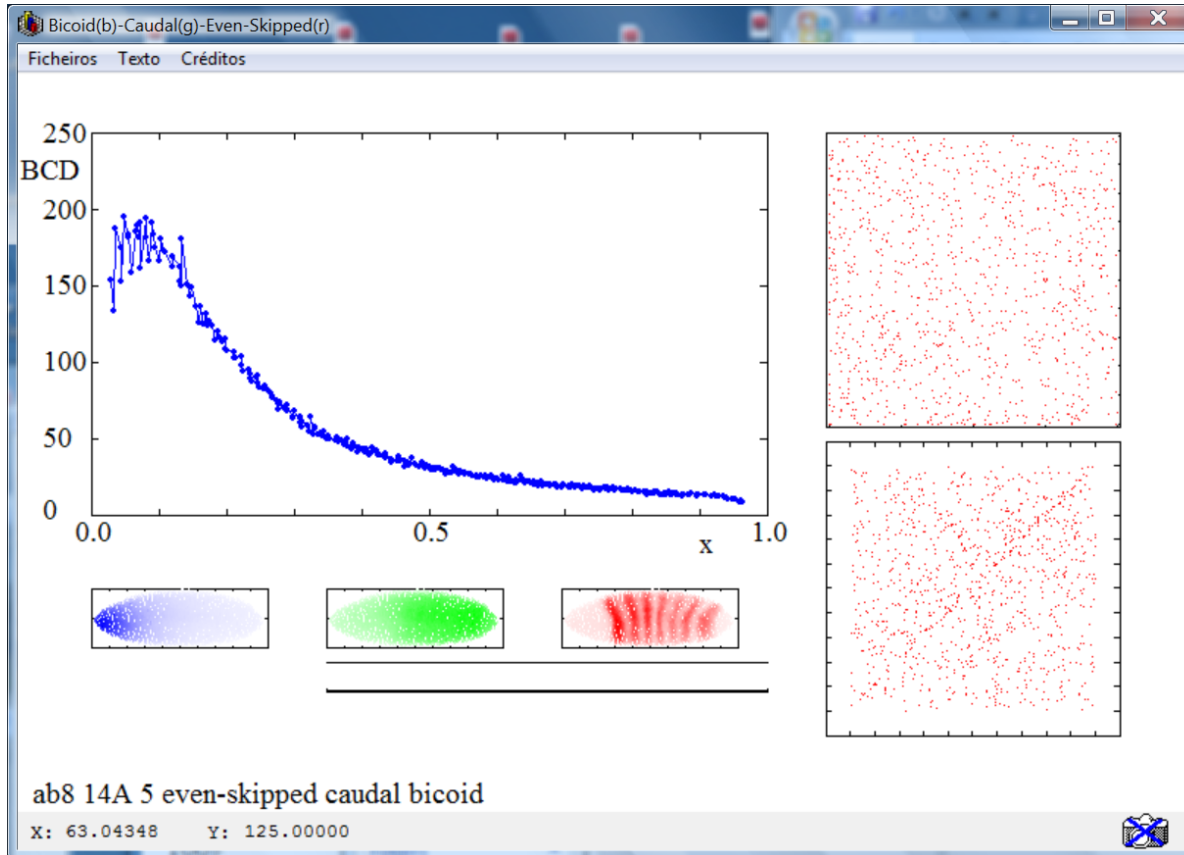
		a_1	a_2	ℓ_1/L	ℓ_2/L	$\sqrt{\chi_m^2/B_{max}^2}$	n	$d \text{ (s}^{-1}\text{)}$	$Aa(\ell_2 - \ell_1)/L$
a)	ab18 (11)	345.2	4.69	0.03	0.20	0.06	30	8.8×10^{-4}	5.2×10^{-2}
b)	ab17 (12)	894.4	4.50	0.06	0.14	0.06	70	8.1×10^{-4}	5.8×10^{-2}
c)	ab16 (13)	684.2	5.51	0.06	0.15	0.08	152	1.2×10^{-3}	7.4×10^{-2}
d)	ab12 (14-1)	927.6	4.82	0.06	0.14	0.08	309	9.2×10^{-4}	6.8×10^{-2}
e)	ab14 (14-2)	3414.7	4.38	0.05	0.07	0.08	314	7.7×10^{-4}	5.3×10^{-2}
f)	ab9 (14-3)	1191.6	4.39	0.05	0.10	0.07	343	7.7×10^{-4}	4.6×10^{-2}
g)	ad13 (14-4)	470.4	3.02	0.06	0.19	0.05	324	3.6×10^{-4}	2.2×10^{-2}
h)	ab8 (14-5)	3271.7	4.25	0.08	0.09	0.07	332	7.2×10^{-4}	2.4×10^{-2}

Fits with a swarm algorithm.

Mean relative error between experimental and theoretical predictions:

5% - 8%

Parameter swarming (slow convergence)



Swarm rule: the randomly chosen parameter is updated only if the fitness function decreases.

(There are other techniques for parameter estimation, specifically suited for **ill defined problems**: Covariance Matrix Adaptation Evolutionary Strategy)

Validation of a morphogenesis model of *Drosophila* early development by a multi-objective evolutionary optimization algorithm

R. Dilão, D. Muraro, M. Nicolau and M. Schoenauer,
In C. Pizzuti, M.D. Ritchie, and M. Giacobini (Eds.): *EvoBIO 2009*, Lecture Notes in Computer Science 5483, pp. 176–190, 2009.

Best Paper Nomination EvoBIO2009.

We apply evolutionary computation to calibrate the parameters of a morphogenesis model of *Drosophila* early development.

The model aims to describe the establishment of the steady gradients of **Bicoid and Caudal** proteins along the antero-posterior axis of the embryo of *Drosophila*.

The model equations consist of a system of non-linear parabolic partial differential equations (**PDE**) with initial and zero flux boundary conditions.

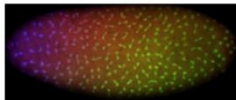
--- We compare the results of single- and multi-objective variants of the CMA-ES algorithm for **the model calibration** with the experimental data. Whereas the **multi-objective algorithm** computes a full approximation of the **Pareto front**, repeated runs of the single-objective algorithm give solutions that dominate (in the Pareto sense) the results of the multi-objective approach. We retain as best solutions those found by the latter technique.

--- From the biological point of view, all such solutions are all equally acceptable, and for our test cases, the **relative error between the experimental data and validated model solutions on the Pareto front** are in the range **3%-6%**.

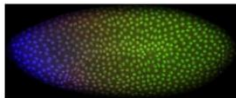
--- **This technique is general and can be used as a generic tool for parameter calibration problems.**

Further developments of the mRNA diffusion model with a multiobjective approach --- Pareto optimality.

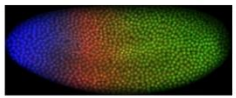
Embryo Name: ab18 Cleavage cycle 11
Bicoid: blue Even-Skipped: red Caudal: green



Embryo Name: ab17 Cleavage cycle 12
Bicoid: blue Even-Skipped: red Caudal: green



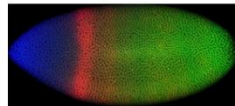
Embryo Name: ab16 Cleavage cycle 13
Bicoid: blue Even-Skipped: red Caudal: green



Embryo Name: ab12 Cleavage cycle 14A-1
Bicoid: blue Even-Skipped: red Caudal: green



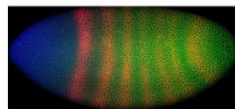
Embryo Name: ab14 Cleavage cycle 14A-2
Bicoid: blue Even-Skipped: red Caudal: green



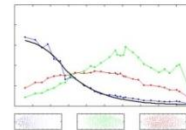
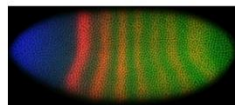
Embryo Name: ab9 Cleavage cycle 14A-3
Bicoid: blue Even-Skipped: red Caudal: green



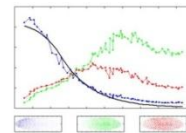
Embryo Name: ad13 Cleavage cycle 14A-4
Bicoid: blue Even-Skipped: red Caudal: green



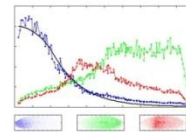
Embryo Name: ab8 Cleavage cycle 14A-5
Bicoid: blue Even-Skipped: red Caudal: green



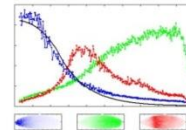
L1=0.0000, L2=0.2585, a1=5.3795, a2=219.0651
ab18 11 - even-skipped caudal bicoid



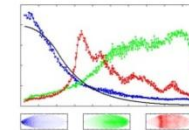
L1=0.0000, L2=0.2608, a1=5.7209, a2=265.2672
ab17 12 - even-skipped caudal bicoid



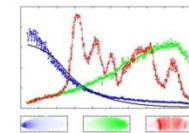
L1=0.0000, L2=0.2263, a1=6.2962, a2=263.9287
ab16 13 - even-skipped caudal bicoid



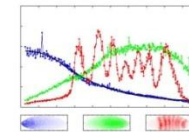
L1=0.0000, L2=0.2488, a1=6.4766, a2=264.7846
ab12 14A 1 - even-skipped caudal bicoid



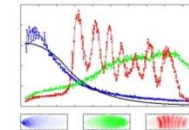
L1=0.0000, L2=0.2120, a1=5.7609, a2=279.6630
ab14 14A 2 - even-skipped caudal bicoid



L1=0.0000, L2=0.2377, a1=5.6038, a2=213.2764
ab9 14A 3 - even-skipped caudal bicoid



L1=0.0181, L2=0.2380, a1=3.3568, a2=279.1849
ad13 14A 4 - even-skipped caudal bicoid



L1=0.0000, L2=0.2484, a1=5.3397, a2=213.3544
ab8 14A 5 - even-skipped caudal bicoid

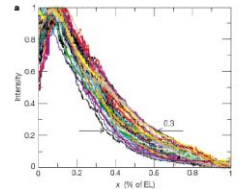
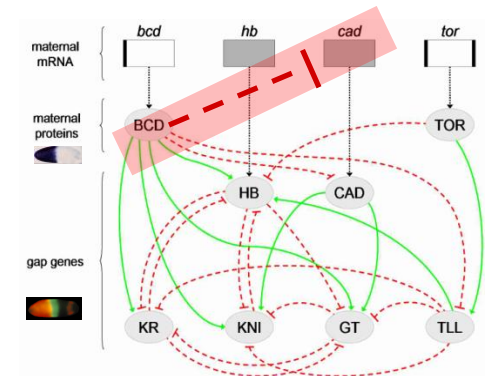
Initial distribution of mRNA

$$bcd(x, t = 0) = \begin{cases} B > 0, & \text{if } 0 < L_1 < x < L_2 < L \\ 0, & \text{otherwise} \end{cases}$$

$$cad(x, t = 0) = \begin{cases} C > 0, & \text{if } 0 < L_3 < x < L_4 < L \\ 0, & \text{otherwise} \end{cases}$$

Diffusion of Bicoid mRNA

$$\begin{cases} \frac{\partial bcd}{\partial t} = -a_{bcd}bcd(x) + D_{bcd}\frac{\partial^2 bcd}{\partial x^2} \\ \frac{\partial BCD}{\partial t} = a_{bcd}bcd(x) \\ \frac{\partial cad}{\partial t} = -a_{cad}cad(x) - rBCD \cdot cad + D_{cad}\frac{\partial^2 cad}{\partial x^2} \\ \frac{\partial CAD}{\partial t} = a_{cad}cad(x) \end{cases}$$



How to calibrate the parameters with the experimental data?

Pareto Optimization

Multi-objective optimization problem

- to find the set of parameters (\mathbf{N}) that minimizes the objective (fitness) function

$$f = (f_1, \dots, f_M) : X \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$$

Parameters space (compact set)

- the parameters are sampled in a compact search space

$$X = \{x \mid l_i \leq x_i \leq u_i, i = 1, \dots, N\}$$

Fitness functions

$$f_1(x) = \|BCD(x) - BCD_{exp}(x)\|_2$$

$$f_2(x) = \|CAD(x) - CAD_{exp}(x)\|_2$$

- we consider two objectives: the fitness of Bicoid and the fitness of Caudal

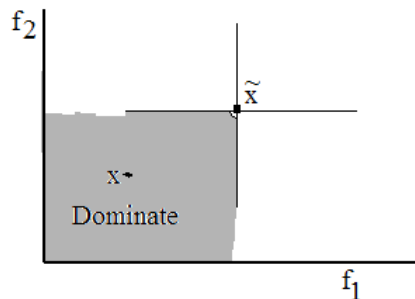
Comparison of solutions

- the solutions are selected according to the dominance criterium

$x \in X$ dominates $\bar{x} \in X$ ($x \prec \bar{x}$) if

$$\forall m \in \{1, \dots, M\} : f_m(x) \leq f_m(\bar{x}) \wedge$$

$$\exists m \in \{1, \dots, M\} : f_m(x) < f_m(\bar{x})$$



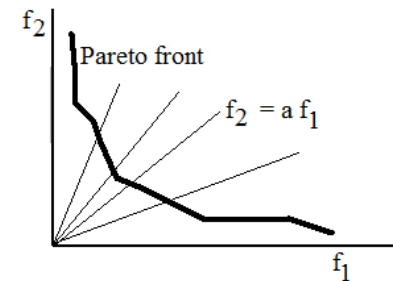
Goals

- to find a good approximation to the Pareto set

$$\text{Pareto set} := \{x | x \in X \wedge \nexists \bar{x} \in X : \bar{x} \prec x\}$$

- to distribute the solutions as uniformly as possible on the Pareto front

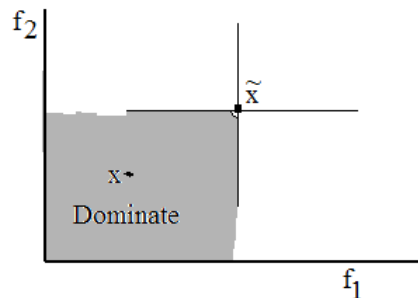
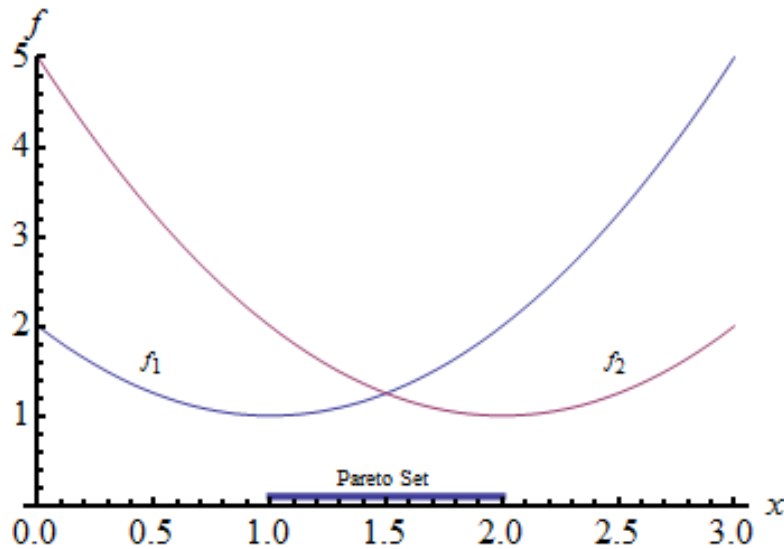
$$\text{Pareto front} := \{f(x) | x \in X \wedge \nexists \bar{x} \in X : \bar{x} \prec x\}$$



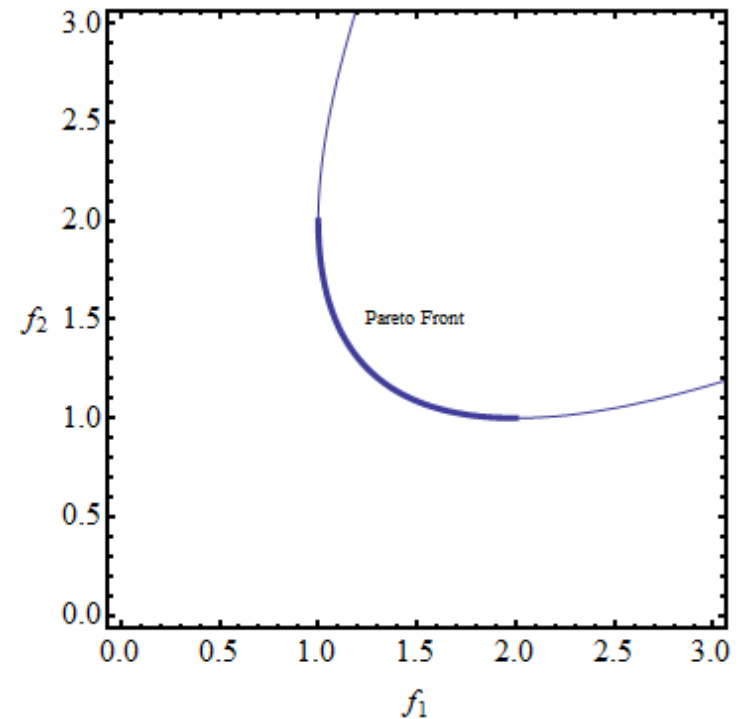
Fitness functions

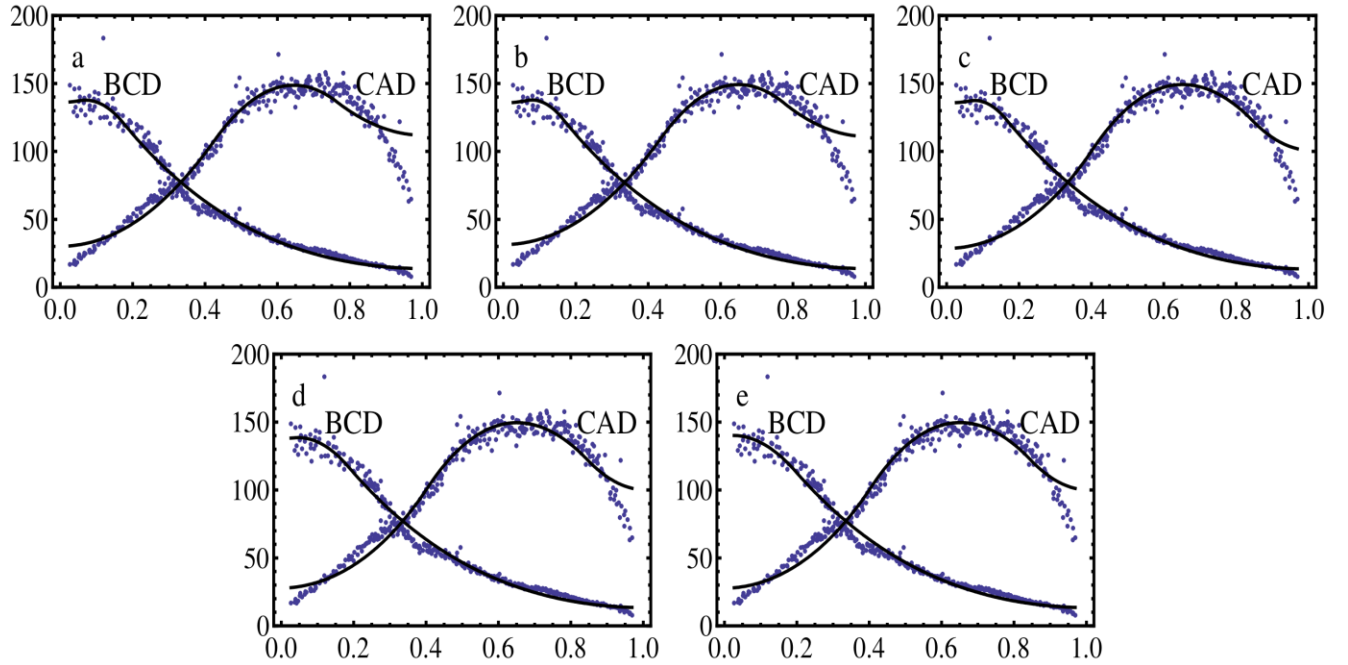
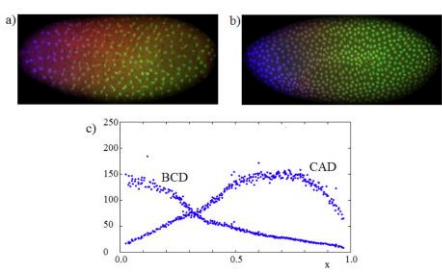
$$f_1(x) = (x-1)^2 + 1$$

$$f_2(x) = (x-2)^2 + 1$$



The solutions on the Pareto front are not dominated by any other solutions. So, from the parametric point of view, any solution of the Pareto set is admissible.



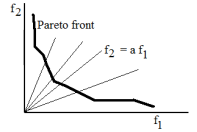


Non-unicity of parameter solutions

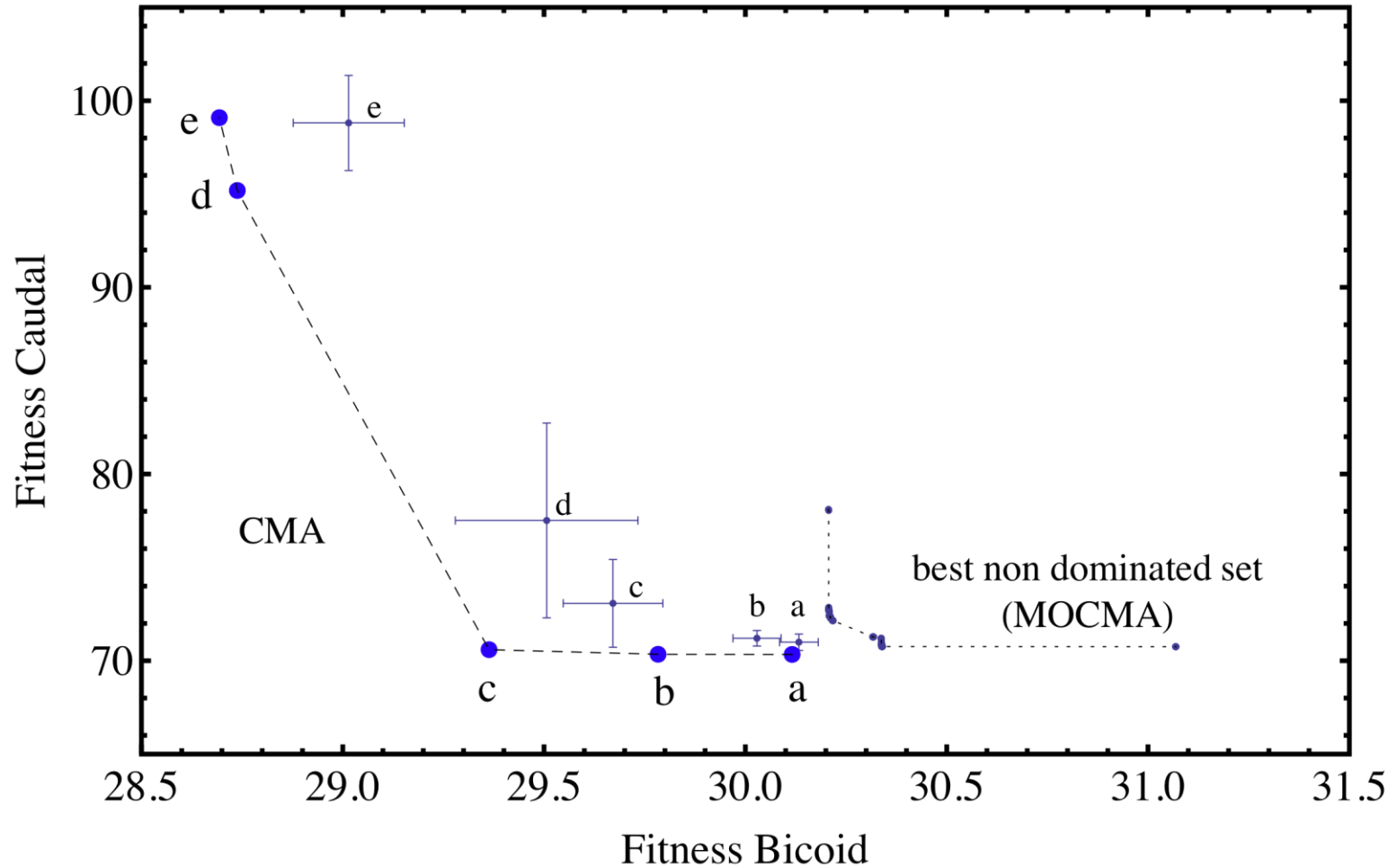
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	mean	σ
L_1	$5.68 \cdot 10^{-2}$	$6.72 \cdot 10^{-2}$	$6.25 \cdot 10^{-2}$	$3.29 \cdot 10^{-2}$	$1.43 \cdot 10^{-2}$	$4.67 \cdot 10^{-2}$	$2.24 \cdot 10^{-2}$
L_2	$1.73 \cdot 10^{-1}$	$1.68 \cdot 10^{-1}$	$1.62 \cdot 10^{-1}$	$1.84 \cdot 10^{-1}$	$1.94 \cdot 10^{-1}$	$1.76 \cdot 10^{-1}$	$0.12 \cdot 10^{-1}$
L_3	$4.28 \cdot 10^{-1}$	$4.35 \cdot 10^{-1}$	$4.04 \cdot 10^{-1}$	$4.07 \cdot 10^{-1}$	$4.04 \cdot 10^{-1}$	$4.16 \cdot 10^{-1}$	$0.14 \cdot 10^{-1}$
L_4	$7.63 \cdot 10^{-1}$	$7.74 \cdot 10^{-1}$	$8.45 \cdot 10^{-1}$	$8.45 \cdot 10^{-1}$	$8.48 \cdot 10^{-1}$	$8.15 \cdot 10^{-1}$	$0.42 \cdot 10^{-1}$
B	$1.53 \cdot 10^{+3}$	$1.98 \cdot 10^{+3}$	$3.47 \cdot 10^{+3}$	$2.36 \cdot 10^{+3}$	$1.98 \cdot 10^{+3}$	$2.26 \cdot 10^{+3}$	$0.73 \cdot 10^{+3}$
C	$1.06 \cdot 10^{+3}$	$1.08 \cdot 10^{+3}$	$1.26 \cdot 10^{+3}$	$1.28 \cdot 10^{+3}$	$1.28 \cdot 10^{+3}$	$1.19 \cdot 10^{+3}$	$0.11 \cdot 10^{+3}$
D_{bcd}	$1.00 \cdot 10^{-2}$	$1.09 \cdot 10^{-2}$	$1.99 \cdot 10^{-2}$	$2.03 \cdot 10^{-2}$	$2.04 \cdot 10^{-2}$	$1.63 \cdot 10^{-2}$	$0.53 \cdot 10^{-2}$
D_{cad}	$1.00 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$0.00 \cdot 10^{-2}$
a_{bcd}	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$1.31 \cdot 10^{+1}$
a_{cad}	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$9.99 \cdot 10^{+4}$	$3.96 \cdot 10^{+1}$
r	$8.64 \cdot 10^{+3}$	$6.74 \cdot 10^{+3}$	$3.34 \cdot 10^{+2}$	$5.74 \cdot 10^{+2}$	$6.71 \cdot 10^{+4}$	$3.07 \cdot 10^{+3}$	$4.26 \cdot 10^{+3}$
Iterations	$9.84 \cdot 10^{+3}$	$9.79 \cdot 10^{+3}$	$9.37 \cdot 10^{+3}$	$9.35 \cdot 10^{+3}$	$9.36 \cdot 10^{+3}$	$9.54 \cdot 10^{+3}$	$0.25 \cdot 10^{+3}$

Table 1. Parameter values for the five best non-dominated solutions of model equations (3), obtained with the CMA algorithm, for the experimental data set of Figure 1c). In Figure 5, we show this data set together with the solutions of equations (3) for the parameter values a-e. All the different choices of these parameter values are calibrated candidates of the experimental data set. We also show, for each parameter, the mean value (mean) and the standard deviation (σ) taken on the Pareto front.

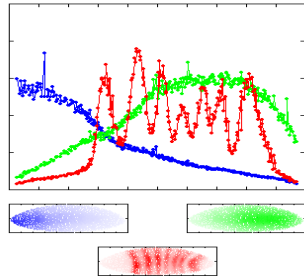
Pareto front for the Bicoid-Caudal multiobjective optimization



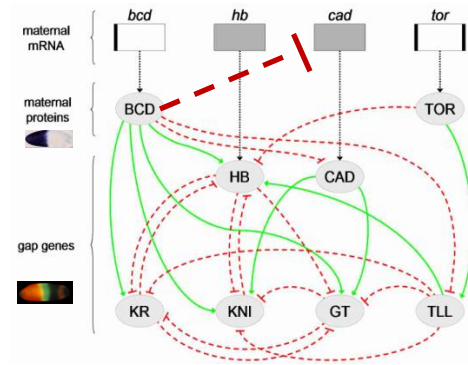
CMA and MOCMA solutions



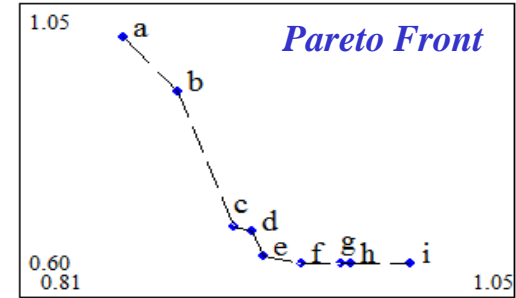
ad13



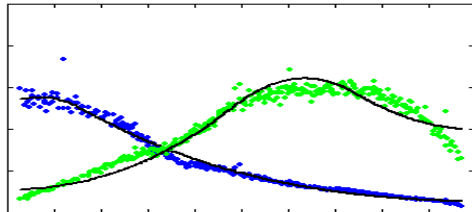
ad13 14A 4 even-skipped caudal bicoid



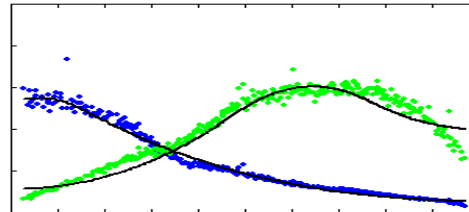
Fitness of Caudal



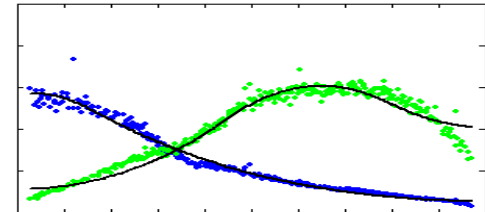
Fitness of Bicoid



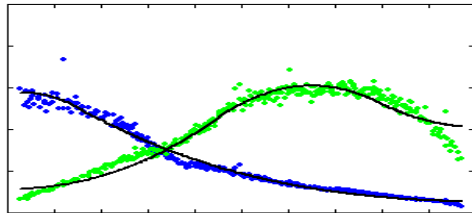
Fitness values: a



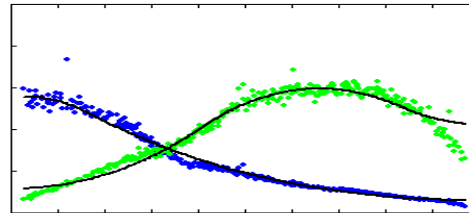
Fitness values: b



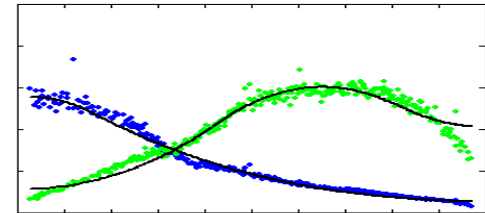
Fitness values: c



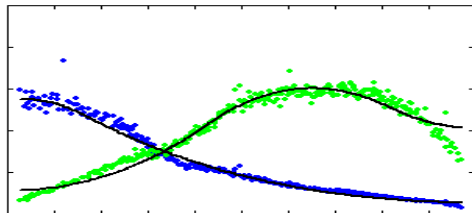
Fitness values: d



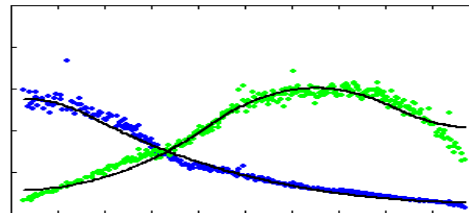
Fitness values: e



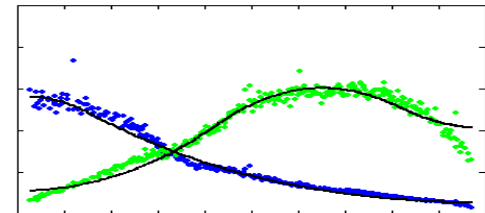
Fitness values: f



Fitness values: g



Fitness values: h

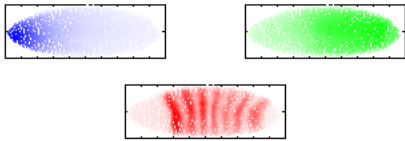
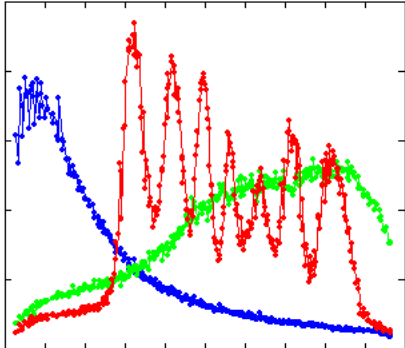


Fitness values: i

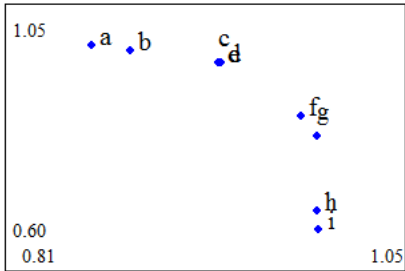
Picture a: best fit for Bicoid and worst for Caudal.

Picture i: best fit for Caudal and worst for Bicoid.

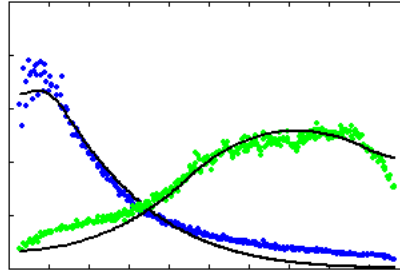
ab8



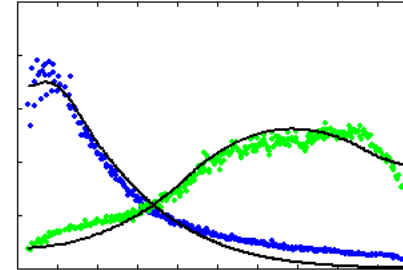
ab8 14A 5 even-skipped caudal bicoid



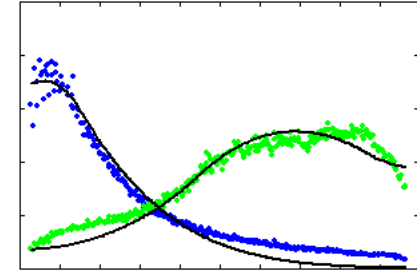
Pareto front
x-axis: fitness of bicoid
y-axis: fitness of caudal



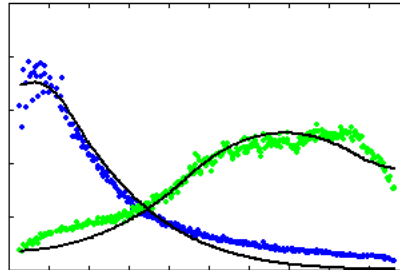
Fitness values: a



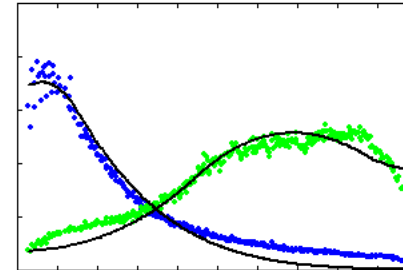
Fitness values: b



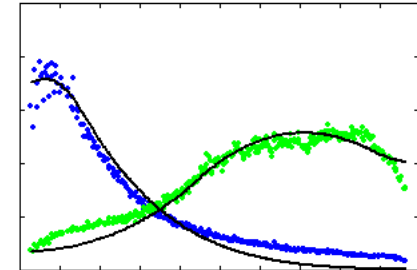
Fitness values: c



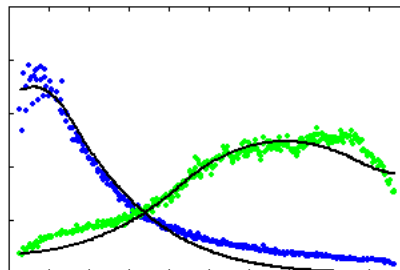
Fitness values: d



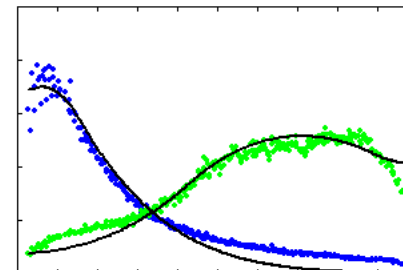
Fitness values: e



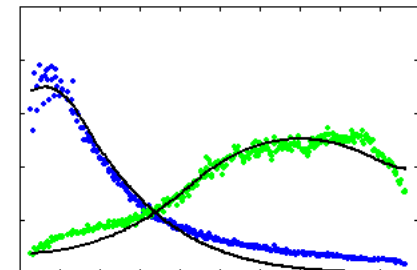
Fitness values: f



Fitness values: g



Fitness values: h



Fitness values: i

Fits with the Pareto set of parameters

Calibration and validation of a genetic regulatory network model describing the production of the protein Hunchback in *Drosophila* early development.

R. Dilão and D. Muraro, 2010, submitted.

--- We fit the parameters of a differential equations model describing the production of gap gene proteins Hunchback and Knirps along the antero-posterior axis of the embryo of *Drosophila*.

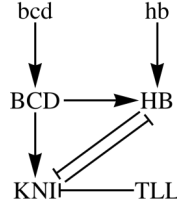
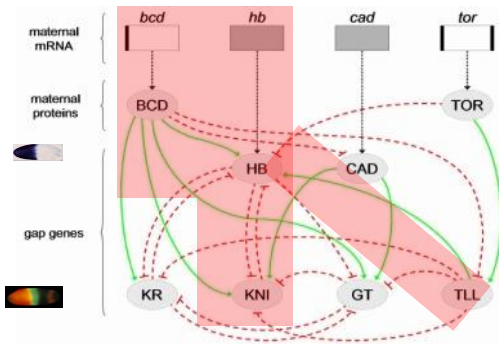
As initial data for the differential equations model, we take the antero-posterior distribution of the proteins Bicoid, Hunchback and Tailless at the beginning of cleavage cycle 14.

--- We calibrate and validate the model with experimental data using **single-** and **multi-objective** evolutionary optimization techniques. In the multi-objective optimization technique, we **compute the associated Pareto fronts**.

--- We analyze the **cross regulation mechanism** between the gap-genes protein pair **Hunchback-Knirps** and we show that the posterior distribution of Hunchback follow the experimental data if **Hunchback is negatively regulated by the Huchebein protein**. (Experimentaly supported).

!!! This approach enables to **predict the posterior localization** on the embryo of **the protein Huchebein**, and we validate with the experimental data the genetic regulatory network responsible for the antero-posterior distribution of the gap gene protein Hunchback.

!!! We discuss **the importance of Pareto multi-objective optimization techniques** in the calibration and validation of biological models (**evolutionary selection**).



$$I_{bcd}(x) = \begin{cases} \frac{a_1}{2} (e^{-a_2(L_1-x)/L} - e^{-a_2(L_2-x)/L}), & \text{if } x < L_1 \\ a_1 - \frac{a_1}{2} (e^{-a_2(x-L_1)/L} + e^{-a_2(L_2-x)/L}), & \text{if } L_1 \leq x \leq L_2 \\ \frac{a_1}{2} (e^{-a_2(x-L_2)/L} - e^{-a_2(x-L_1)/L}), & \text{if } x > L_2 \end{cases}$$

$$I_{hb}(x) = \begin{cases} \frac{a_3}{2} (e^{-a_4(M_1-x)/L} - e^{-a_4(M_2-x)/L}), & \text{if } x < M_1 \\ a_3 - \frac{a_3}{2} (e^{-a_4(x-M_1)/L} + e^{-a_4(M_2-x)/L}), & \text{if } M_1 \leq x \leq M_2 \\ \frac{a_3}{2} (e^{-a_4(x-M_2)/L} - e^{-a_4(x-M_1)/L}), & \text{if } x > M_2 \end{cases}$$

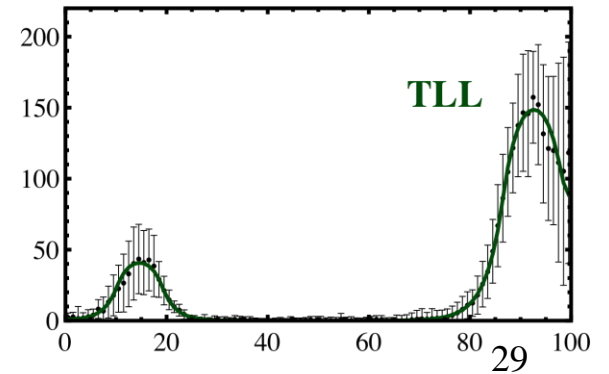
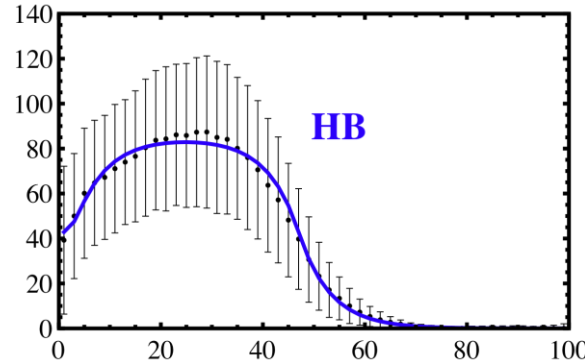
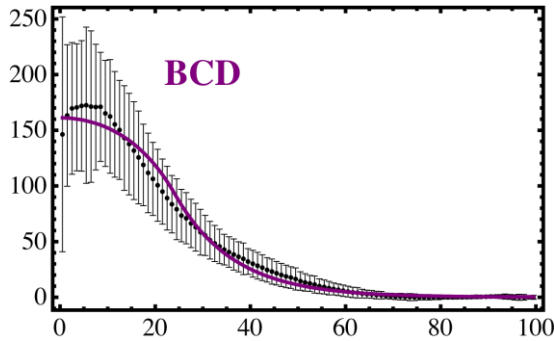
$$I_{1tll}(x) = \begin{cases} \frac{a_5}{2} (e^{-a_6(N_1-x)/L} - e^{-a_6(N_2-x)/L}), & \text{if } x < N_1 \\ a_5 - \frac{a_5}{2} (e^{-a_6(x-N_1)/L} + e^{-a_6(N_2-x)/L}), & \text{if } N_1 \leq x \leq N_2 \\ \frac{a_5}{2} (e^{-a_6(x-N_2)/L} - e^{-a_6(x-N_1)/L}), & \text{if } x > N_2 \end{cases}$$

$$I_{2tll}(x) = \begin{cases} \frac{a_7}{2} (e^{-a_8(N_3-x)/L} - e^{-a_8(N_4-x)/L}), & \text{if } x < N_3 \\ a_7 - \frac{a_7}{2} (e^{-a_8(x-N_3)/L} + e^{-a_8(N_4-x)/L}), & \text{if } N_3 \leq x \leq N_4 \\ \frac{a_7}{2} (e^{-a_8(x-N_4)/L} - e^{-a_8(x-N_3)/L}), & \text{if } x > N_4 \end{cases}$$

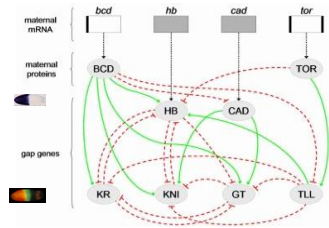
$$\text{BCD}_{eq}(x) = 2 \frac{a_1}{e^{2a_2/L} - 1} \cosh\left(a_2 \frac{x}{L}\right) \left(\sinh\left(a_2 \frac{L_2}{L}\right) - \sinh\left(a_2 \frac{L_1}{L}\right) \right) + \frac{a_1}{2} (e^{-a_2(x+L_1)/L} - e^{-a_2(x+L_2)/L}) + I_{bcd}(x)$$

$$\text{HB}_{eq}(x) = 2 \frac{a_3}{e^{2a_4/L} - 1} \cosh\left(a_4 \frac{x}{L}\right) \left(\sinh\left(a_4 \frac{M_2}{L}\right) - \sinh\left(a_4 \frac{M_1}{L}\right) \right) + \frac{a_3}{2} (e^{-a_4(x+M_1)/L} - e^{-a_4(x+M_2)/L}) + I_{hb}(x)$$

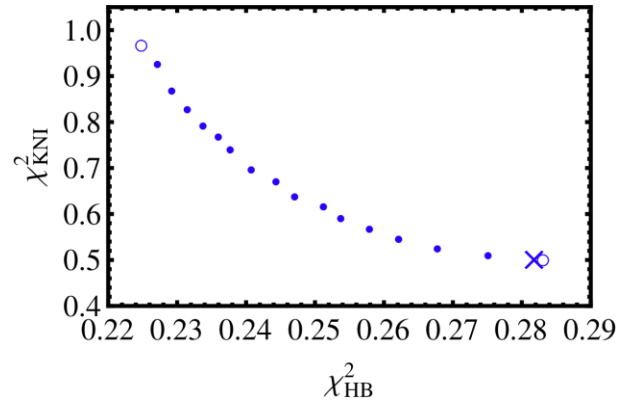
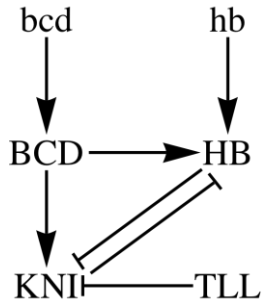
$$\text{TLL}_{eq}(x) = 2 \frac{a_5}{e^{2a_6/L} - 1} \cosh\left(a_6 \frac{x}{L}\right) \left(\sinh\left(a_6 \frac{N_2}{L}\right) - \sinh\left(a_6 \frac{N_1}{L}\right) \right) + \frac{a_5}{2} (e^{-a_6(x+N_1)/L} - e^{-a_6(x+N_2)/L}) + I_{1tll}(x) + 2 \frac{a_7}{e^{2a_8/L} - 1} \cosh\left(a_8 \frac{x}{L}\right) \left(\sinh\left(a_8 \frac{N_4}{L}\right) - \sinh\left(a_8 \frac{N_3}{L}\right) \right) + \frac{a_7}{2} (e^{-a_8(x+N_3)/L} - e^{-a_8(x+N_4)/L}) + I_{2tll}(x)$$



Calibration of the Hunchback-Knirps gap gene proteins with the mRNA diffusion hypothesis



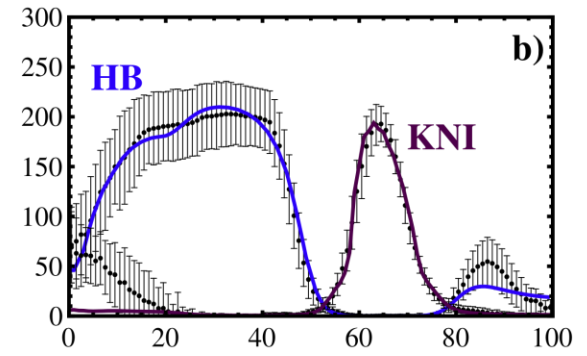
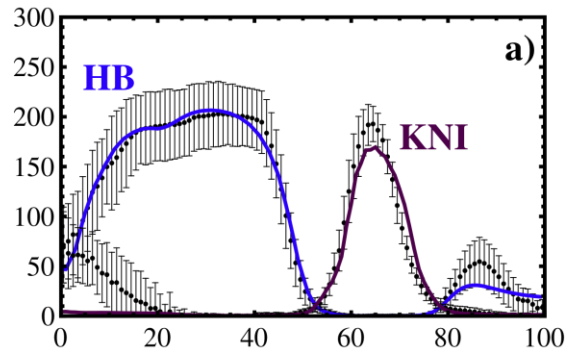
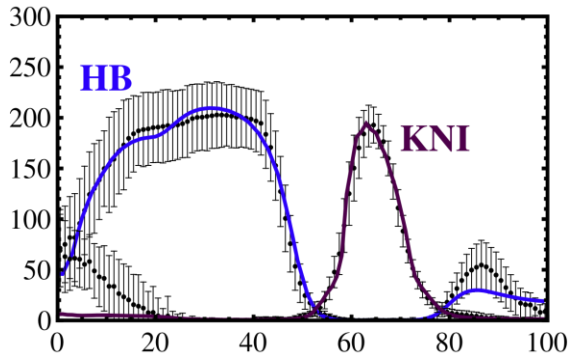
1 instantiation of the Pareto front



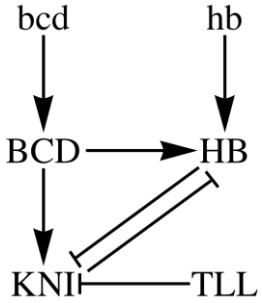
gap gene phase
NO diffusion
 pure mass action
 approach
 emergent thresholds

Maternal phase
mRNA diffusion

Pareto front for the pair HB-KNI

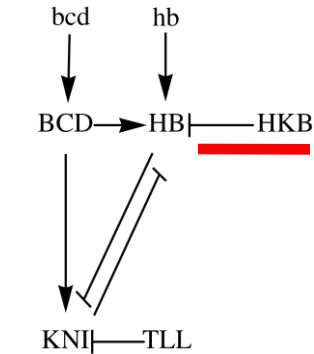
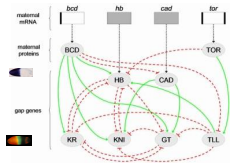


Embryo length, L



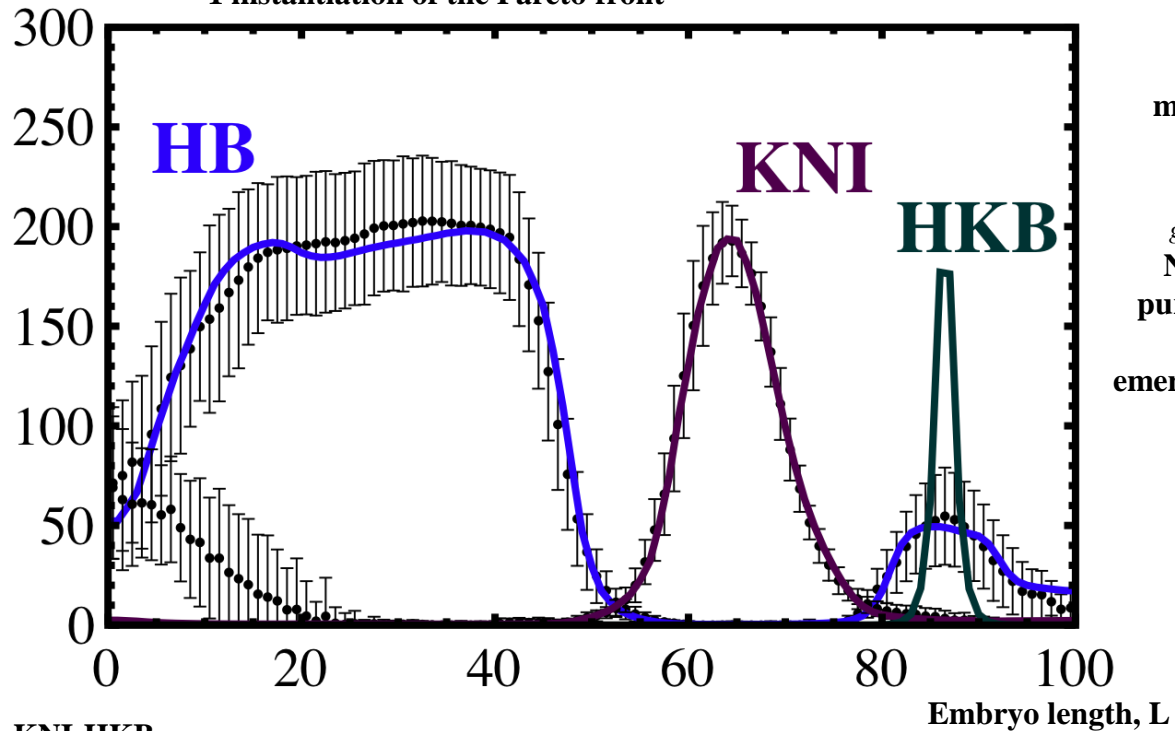
$$\begin{aligned}
HB' [t] &= -HB [t] d_{HB} + (P_{HB}^{BCD})_{HB} HB_0^{BCD} [t] + (P_{KNI}^{BCD})_{HB} HB_{KNI}^{BCD} [t] - HB [t] kni_{hb} KNI_{0,0}^0 [t] - \\
&HB [t] kni_{hb} KNI_{0,0}^{BCD} [t] - HB [t] kni_{hb} KNI_{0,TLL}^0 [t] - HB [t] kni_{hb} KNI_{0,TLL}^{BCD} [t] + \\
&kni_{-hb} KNI_{HB,0}^0 [t] + kni_{-hb} KNI_{HB,0}^{BCD} [t] + kni_{-hb} KNI_{HB,TLL}^0 [t] + kni_{-hb} KNI_{HB,TLL}^{BCD} [t] \\
KNI' [t] &= -KNI [t] d_{KNI} - KNI [t] hb_{kni} HB_0^0 [t] - KNI [t] hb_{kni} HB_0^{BCD} [t] + hb_{-kni} HB_{KNI}^0 [t] + hb_{-kni} HB_{KNI}^{BCD} [t] + \\
&(P_{0,0}^{BCD})_{KNI} KNI_{0,0}^{BCD} [t] + (P_{0,TLL}^{BCD})_{KNI} KNI_{0,TLL}^{BCD} [t] + (P_{HB,0}^{BCD})_{KNI} KNI_{HB,0}^{BCD} [t] + (P_{HB,TLL}^{BCD})_{KNI} KNI_{HB,TLL}^{BCD} [t] \\
(HB_0^0)' [t] &= -BCD hb_{bcd} HB_0^0 [t] - KNI [t] hb_{kni} HB_0^0 [t] + hb_{-bcd} HB_0^{BCD} [t] + hb_{-kni} HB_{KNI}^0 [t] \\
(HB_0^{BCD})' [t] &= BCD hb_{bcd} HB_0^0 [t] - hb_{-bcd} HB_0^{BCD} [t] - KNI [t] hb_{kni} HB_0^{BCD} [t] + hb_{-kni} HB_{KNI}^{BCD} [t] \\
(HB_{KNI}^0)' [t] &= KNI [t] hb_{kni} HB_0^0 [t] - BCD hb_{bcd} HB_{KNI}^0 [t] - hb_{-kni} HB_{KNI}^0 [t] + hb_{-bcd} HB_{KNI}^{BCD} [t] \\
(HB_{KNI}^{BCD})' [t] &= KNI [t] hb_{kni} HB_0^{BCD} [t] + BCD hb_{bcd} HB_{KNI}^0 [t] - hb_{-bcd} HB_{KNI}^{BCD} [t] - hb_{-kni} HB_{KNI}^{BCD} [t] \\
(KNI_{0,0}^0)' [t] &= -BCD kni_{bcd} KNI_{0,0}^0 [t] - HB [t] kni_{hb} KNI_{0,0}^0 [t] - \\
&TLL kni_{t11} KNI_{0,0}^0 [t] + kni_{-bcd} KNI_{0,0}^{BCD} [t] + kni_{-t11} KNI_{0,TLL}^0 [t] + kni_{-hb} KNI_{HB,0}^0 [t] \\
(KNI_{0,0}^{BCD})' [t] &= BCD kni_{bcd} KNI_{0,0}^0 [t] - kni_{-bcd} KNI_{0,0}^{BCD} [t] - \\
&HB [t] kni_{hb} KNI_{0,0}^{BCD} [t] - TLL kni_{t11} KNI_{0,0}^{BCD} [t] + kni_{-t11} KNI_{0,TLL}^{BCD} [t] + kni_{-hb} KNI_{HB,0}^{BCD} [t] \\
(KNI_{0,TLL}^0)' [t] &= TLL kni_{t11} KNI_{0,0}^0 [t] - BCD kni_{bcd} KNI_{0,TLL}^0 [t] - \\
&HB [t] kni_{hb} KNI_{0,TLL}^0 [t] - kni_{-t11} KNI_{0,TLL}^0 [t] + kni_{-bcd} KNI_{0,TLL}^{BCD} [t] + kni_{-hb} KNI_{HB,TLL}^0 [t] \\
(KNI_{0,TLL}^{BCD})' [t] &= TLL kni_{t11} KNI_{0,0}^{BCD} [t] + BCD kni_{bcd} KNI_{0,TLL}^0 [t] - \\
&kni_{-bcd} KNI_{0,TLL}^{BCD} [t] - HB [t] kni_{hb} KNI_{0,TLL}^{BCD} [t] - kni_{-t11} KNI_{0,TLL}^{BCD} [t] + kni_{-hb} KNI_{HB,TLL}^{BCD} [t] \\
(KNI_{HB,0}^0)' [t] &= HB [t] kni_{hb} KNI_{0,0}^0 [t] - BCD kni_{bcd} KNI_{HB,0}^0 [t] - \\
&kni_{-hb} KNI_{HB,0}^0 [t] - TLL kni_{t11} KNI_{HB,0}^0 [t] + kni_{-bcd} KNI_{HB,0}^{BCD} [t] + kni_{-t11} KNI_{HB,TLL}^0 [t] \\
(KNI_{HB,0}^{BCD})' [t] &= HB [t] kni_{hb} KNI_{0,0}^{BCD} [t] + BCD kni_{bcd} KNI_{HB,0}^0 [t] - \\
&kni_{-bcd} KNI_{HB,0}^{BCD} [t] - kni_{-hb} KNI_{HB,0}^{BCD} [t] - TLL kni_{t11} KNI_{HB,0}^{BCD} [t] + kni_{-t11} KNI_{HB,TLL}^{BCD} [t] \\
(KNI_{HB,TLL}^0)' [t] &= HB [t] kni_{hb} KNI_{0,TLL}^0 [t] + TLL kni_{t11} KNI_{HB,0}^0 [t] - \\
&BCD kni_{bcd} KNI_{HB,TLL}^0 [t] - kni_{-hb} KNI_{HB,TLL}^0 [t] - kni_{-t11} KNI_{HB,TLL}^0 [t] + kni_{-bcd} KNI_{HB,TLL}^{BCD} [t] \\
(KNI_{HB,TLL}^{BCD})' [t] &= HB [t] kni_{hb} KNI_{0,TLL}^{BCD} [t] + TLL kni_{t11} KNI_{HB,0}^{BCD} [t] + \\
&BCD kni_{bcd} KNI_{HB,TLL}^0 [t] - kni_{-bcd} KNI_{HB,TLL}^{BCD} [t] - kni_{-hb} KNI_{HB,TLL}^{BCD} [t] - kni_{-t11} KNI_{HB,TLL}^{BCD} [t]
\end{aligned}$$

Calibration of the Hunchback-Knirps and Huckebein gap gene proteins with the mRNA diffusion hypothesis



14 equations
 23 (42) free parameters
 Dynamic pattern: 29 seconds after 14A.

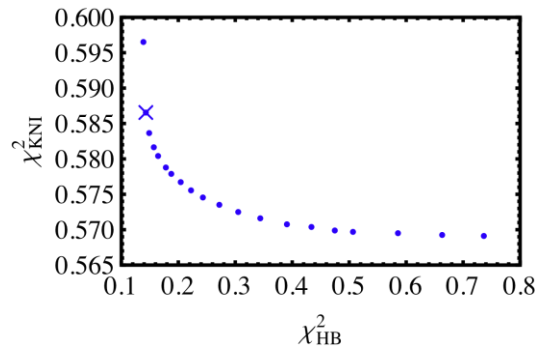
1 instantiation of the Pareto front



Maternal phase
 mRNA diffusion

gap gene phase
 NO diffusion
 pure mass action
 approach
 emergent thresholds

Pareto front for the pair HB-KNI-HKB



Prediction of the HKB distribution

Thank you