Modelling and inference for biological systems: from auxin dynamics in plants to protein evolution

Silvia Grigolon
silvia.grigolon@gmail.com
Former Ph.D. student
at LPTMS, Orsay, France

Currently Post-Doc at
The Francis Crick Institute, LIF, London, UK
Modelling

Modelling polarity patterns emerging for the intercellular transport of auxin in plants.

Inference

Inferring relevant sites from proteins Multiple Sequence Alignments by Critical Variable Selection (CVS).
S. Grigolon, S. Franz & M. Marsili (ICTP, Trieste), arXiv:1503.03815
Modelling
What is auxin?

Auxin is a **phytohormone** involved in **plant morphogenesis**, i.e., the set of processes driving organs’ formation and then growth. The **organs** allow the **fulfilment** of the **vital functions**.

**In plants:** flowers, roots, leaves, etc.
Auxin role in plant morphogenesis

Organs formation is allowed by the **initiation of cells differentiation**, in plants **driven** by auxin.

**How?**

Auxin concentration profiles set differences among the cells.

D. Reinhardt et al., *The Plant Cell*, 2000

**Root and Cambium Meristems**

***Shoot Apical Meristem***

[Images of root and shoot meristems]

de Reuille et al., *PNAS*, 2006

Band et al., *The Plant Cell*, 2014

**Meristem = Plant Tissue**
Auxin spatial distributions throughout tissues play a major role in **morphogenesis and cell differentiation**.

D. Reinhardt et al., *The Plant Cell*, 2000

How are these distributions set?

In each cell: **Auxin production and degradation + Transport among cells**

**Auxin routes**

PASSIVE TRANSPORT (diffusion)

ACTIVE TRANSPORT (carriers-mediated):

- **PIN formed proteins**: out-going flux
- **AUX1 (members of AUX-LAX)**: ingoing flux

D. Reinhardt et al., *Nature*, 2003
E. Zazimalovà et al., *Cold Spring Harb Perspect Biol*, 2010
Patterns of PIN intracellular localisation

During organs primordia settlement, PINs redistribute in a non-trivial fashion on cells’ membranes.

PINs pump auxin outside the cell inducing a FLUX.

If they are localised, they define a preferential direction for the flux of auxin:

they set a polarisation for cells.

How can these patterns be set spontaneously?

Bilou et al., Nature, 2005

de Reuille et al., PNAS, 2006
Why another model?

Complex Polarity Patterns:
how do the cells collectively generate organised polarisations?

Our main ingredients will be:
- auxin
- auxin transport
- PINs dynamics

We propose a deterministic and a stochastic model where:
- auxin gradient is not required as a main ingredient;
- spontaneous PIN polarisation arises if and only if cooperative effects in PIN recycling are strong enough;
- PIN polarity patterns are favoured by molecular noise.
Let us consider a tissue of a monolayer of cells modelled as a **2D lattice** of **cubic cells and apoplasts** (the space between two adjacent cells).

Auxin concentration in cell P at time $t$ $\rightarrow A_C(P, t)$

Auxin concentration between cell P and cell P' at time $t$ $\rightarrow A_a(P, P', t)$
We define a reaction-diffusion discrete model for auxin dynamics:

\[
\frac{dA_C(P, t)}{dt} = \beta - \rho A_C(P, t) + \sum_{P'} \Lambda^{-1} \frac{D}{\epsilon} [A_a(P, P', t) - A_C(P, t)]
\]

**Reactions**
- Production
- Degradation

**Diffusion through the cell wall**
- \(D\) diffusion constant
- \(\epsilon\) wall thickness

**Active Transport Terms**
- \(\phi_{PIN}(P, P')\)
- \(\phi_{AUX1}(P, P')\)
- \(\phi_{PIN}(P', P)\)
- \(\phi_{AUX1}(P', P)\)

**Auxin in cells:**

\[
\frac{dA_C(P, t)}{dt} = \beta - \rho A_C(P, t)
\]
We define a **reaction-diffusion discrete model** for auxin dynamics:

**Auxin in apoplasts:**

\[
\frac{dA_a(P, P', t)}{dt} = \lambda^{-1}D/\epsilon[A_C(P, t) + A_C(P', t) - 2A_a(P, P', t)]
\]

**Diffusion through the cell wall**

\[
+ \lambda^{-1}[\phi_{PIN}(P, P', t) - \phi_{AUX1}(P, P', t) + \phi_{PIN}(P', P, t) - \phi_{AUX1}(P', P, t)]
\]

**Active Transport Terms**
But what is the form of $\phi_{AUX1}$ and $\phi_{PIN}$?

We can imagine these pumps to work as enzymes

Enzyme $\rightarrow$ PIN/AUX1

Substrate $\rightarrow$ Cell Auxin/Apo Auxin

We assume the fluxes to have a Michaelis-Menten form, i.e.:

\[
\Lambda^2 \phi_{AUX1}(P, P', t) = N^{AUX1} \cdot \alpha \cdot \frac{A_a(P, P', t)}{1 + \frac{A_a(P, P', t)}{A^*} + \frac{A_c(P, t)}{A^{**}}}, \\
\Lambda^2 \phi_{PIN}(P, P', t) = N^{PIN} \cdot \gamma \cdot \frac{A_c(P, t)}{1 + \frac{A_a(P, P', t)}{A^*} + \frac{A_c(P, t)}{A^{**}}}
\]
PINs can rearrange on cells membranes in non-trivial patterns.

What affects PINs dynamics from one face of the cell to the other? We here assume that it is auxin flux to tune PINs dynamics.

\[ f = E, N, W, S \]

Total number of PINs conserved.

\[
\tau \frac{d N_{f}^{PIN}}{dt} = -\frac{3}{4} N_{f}^{PIN} \frac{1}{1 + \left(\frac{\phi_{f}^{OUT}}{\phi^{*}}\right)^{h}} + \frac{1}{4} \sum_{f'} N_{f'}^{PIN} \frac{1}{1 + \left(\frac{\phi_{f'}^{OUT}}{\phi^{*}}\right)^{h}}
\]

We introduce the phenomenological exponent \( h \) that mathematically tunes the non-linearity in the dynamics and biologically might be connected to cooperative effects.

\( h \) and \( D \) will be the key parameters in this model.
It is more convenient to work in terms of one single quantity:

**CELL POLARISATION**

What are the solution of these equations for PINs and consequently for the polarisation varying the **exponent** $h$ and the **diffusion constant** $D$?

Solving the deterministic equations we find **two regimes** for the polarisation.
Molecular noise enhances ordering


PINs non-linearity \( h = 2 \)

Varying the diffusion constant \( D \) only

Two phases robust even with noise.
Conclusions and further developments

We proposed a model for the spontaneous emergence of PIN polarity patterns.

Main Findings:

- **auxin gradient** is **not required** as a main ingredient;
- spontaneous PIN polarisation arises if and only if **cooperative effects in PIN recycling** are strong enough;
- PIN polarity patterns are **favoured by molecular noise**.

Further developments:

- clarifying the **biological meaning** of \( h \);
- full morphogenesis model.
Main principle:
The **structure** and **function** a protein carries out inside a cell are **encoded** in its amino acid sequence (**primary structure**)

However

**Living matter** undergoes **evolution**

for the same protein or protein domain remarkable differences across different organisms

Constraint:
structure and function must be somehow preserved across evolution
The ideas behind Critical Variable Selection

Marsili et al., *J. Stat.*, 2013

The number of times the sequence $s$ is sampled $M$.

Number of times the sequence $s$ is sampled:

$$K_s = \sum_{i=1}^{N} \delta_{s_i, s}$$

$$P(s_i = s) = \frac{K_s}{N}$$

$$P(K_{s_i} = k) = \frac{k m_k}{N}$$

$$m_k = \sum_s \delta_{K_s, k}$$
The ideas behind Critical Variable Selection

M. Marsili et al., *J. Stat.*, 2013

Entropy with respect to the sequences

\[ H[s] = - \sum_k m_k \frac{k}{N} \log\left(\frac{k}{N}\right) \]

Two limits:

1. \( s_1 \neq s_2 \neq \ldots \neq s_N \)
   \[ H[s] = \log N \]
   \[ k = 1 \]
   \[ m_k = N \]

2. \( s_1 = s_2 = \ldots = s_N \)
   \[ H[s] = 0 \]
   \[ k = N \]
   \[ m_k = 1 \]

Entropy with respect to the “clustering” (*coarse graining*)

\[ H[K] = - \sum_k \frac{k m_k}{N} \log\left(\frac{k m_k}{N}\right) \]

IDEA

Informative variables can be extracted by maximising \( H[K] \).
The aim is to find the subset of variables **maximising** $H[K]$.

After one run we get a sequence of **accepted** (1) and **unaccepted** (0) sites.
"Brain" ↓ "Conserved" ↓ "Noise"

Normal entropy max \( c_d(L_{\text{subseq}}) \)

H[K] max \( c_b(L_{\text{subseq}}) \)

\( L = 64 \)

\( L_{\text{subseq}} = 10 \)
\( L_{\text{subseq}} = 20 \)
\( L_{\text{subseq}} = 30 \)
The voltage sensor domain of ion channel proteins


Top 15

Top 40

Blue circles:

Lee et al., PNAS, 2005
lein et al., J. Gen. Physiol., 2014

\[ c_i = \sum_{L_{subseq}=\{10,20,30\}} c_i(L_{subseq}) \]

N=6652
L=114

site entropy

Total Count

23/24
Conclusions and further developments

We proposed a method aimed at identifying functional and structural positions out of a MSA.

Main Findings:

- CVS is able to afford a **sharp division between relevant and irrelevant sites**;
- CVS is able to capture **functional sites** out of the analysed MSAs;

Further developments:

- extending the analysis on **more proteins datasets**
  (with A. Haimovici too, University of Buenos Aires);
- performing more **insightful analysis** on biological properties such as SAS, intradomains analysis, etc.;
- new version soon coming!
Thank you for your attention