Interplay between the organization and function of cellular interaction networks

Réka Albert
Department of Physics & Huck Institutes for the Life Sciences
Pennsylvania State University

Course on Graphs and Networks in Systems Biology

http://www.phys.psu.edu/~ralbert/teaching.htm
Cells are complex systems

- functionally diverse elements
- diverse interactions that form networks
  - signal transduction-, gene regulatory-, metabolic-
- have a function that needs to be performed
  - sense and respond to the environment
  - maintain homeostasis
  - replicate
- need certain dynamical features
  - sensitive to some changes, insensitive/adaptable to others
  - robust to unwanted perturbations
  - evolvable, shaped by evolution
- Which network topological features can ensure reliable and robust dynamics?

Dynamics: state of the nodes and changes in the state
protein-gene interactions

protein-protein interactions

pCOLOR

proTEOMe

GENOME

citrate cycle

metabolism

Biochemical reactions
Cellular processes form networks on many levels

**Protein interaction networks**
- Nodes: proteins
- Edges: protein-protein interactions (binding)

**Reaction networks**
- Nodes: substrates, enzymes
- Edges: chemical reactions, catalysis
Cellular processes form networks on many levels

**Regulatory networks**
- Nodes: mRNAs, proteins
- Edges: translation \( \rightarrow \)
or regulation \( \rightarrow \) \( \rightarrow \) \( \rightarrow \)\( \rightarrow \)
- activating or inhibiting

**Signal transduction networks**
- Nodes: proteins, molecules
- Edges: reactions and processes reflecting information transfer (e.g. ligand/receptor binding, protein conformational changes)
Protein interaction maps now contain thousands of nodes and edges

- Although usually tested in a given bait/prey setting, protein interactions are considered symmetrical.

- All networks have giant connected components.

- The topological properties of diverse protein interaction networks are similar.


Ito et al (yeast): 8868 interactions between 3280 proteins
Uetz et al (yeast): 4480 interactions, 2115 proteins
Giot et al (Drosophila): 4780 interactions among 4679 proteins
Li et al (C. elegans): 5534 interactions, 3024 proteins
Rual et al (human): 2800 interactions
Biological networks are highly heterogeneous (scale – free)
Many nodes have only a few edges, but highly interactive (hub) nodes are also possible.

This suggests robustness to random mutations, but vulnerability to mutations in highly-connected components.


Comparison of yeast interaction networks

Degree distribution

\[ P(k) \sim k^{-2.5} \]

Clustering coefficient

\[ C(k) \sim k^{-2} \]

Connected components

Yook, Oltvai and Barabási, Proteomics 4, 928 (2004)
Not all interactions are simultaneously active

Calculate the correlation between the expression time-course of genes encoding the first neighbors of hub proteins.

Two peaks – two different types of hubs. Party hubs are inside connected modules that interact simultaneously. Date hubs connect different modules.

Networks of chemical reactions

Metabolism: Sum of chemical processes by which energy is stored or released.
Tri-partite representation of metabolic network

- Node types:
  - Metabolites (substrates or products), open rectangles
  - No distinction between metabolites and coenzymes
  - Metabolite-enzyme complexes, black rectangles
  - Enzymes, open ovals

- Edges:
  - Substrate to complex or complex to product
  - Symmetrical edges between enzyme and complex
Reaction Stoichiometry

\[ \begin{align*}
A + B & \rightarrow C + D \quad (1) \\
A + D & \rightarrow E \quad (2) \\
B + C & \rightarrow F \quad (3)
\end{align*} \]

Stoichiometric Matrix (S)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( S_{ij} \) = Number of molecules of substrate \( i \) participating in reaction \( j \)

- \( S_{ij} < 0 \) if substrate \( i \) is a reactant in reaction \( j \)
- \( S_{ij} > 0 \) if substrate \( i \) is a product in reaction \( j \)

\[ \begin{align*}
i &= 1, 2, \ldots, N = \# \text{ of substrates} = \# \text{ rows} \\
j &= 1, 2, \ldots, M = \# \text{ of reactions} = \# \text{ columns}
\]
Network Representation – Bi-partite Graph

Reaction

\[ A + B \rightarrow C + D \] (1)
\[ A + D \rightarrow E \] (2)
\[ B + C \rightarrow F \] (3)

Bi-partite Graph ("S-Graph")

- Two types of nodes:
  - Substrate Node
  - Reaction Node
- Directed edges
- No direct arcs between nodes of the same type
Network Representation – Substrate Graph

Reaction Pathway

A + B → C + D \hspace{1cm} (1)
A + D → E \hspace{1cm} (2)
B + C → F \hspace{1cm} (3)

- One type of node: Substrate Node
- Un-directed edges
- Each reaction represented as a clique

Network Representation – Reaction Graph

Reaction

A + B → C + D  \( (1) \)
A + D → E  \( (2) \)
B + C → F  \( (3) \)

Reaction Graph

- One type of node: Reaction Node
- Un-directed edges
- An edge between two reactions if they share at least one substrate in common

Three alternate network representations for the same reaction pathway!
Key Properties of Metabolic Networks

- Metabolic networks are scale-free

\[ P(k) = \text{Probability that a given substrate participates in } k \text{ reactions} \approx k^{-\gamma} \]

- In- and out-degree of substrate nodes in the bi-partite representation

- Existence of “hub” substrates such as ATP, ADP, NADP, NADPH (Carrier Metabolites)

Degree distributions in metabolite and reaction networks

Construct non-directed projections to metabolite and reaction networks

Rank vs. degree plot, similar to $P(k > K)$. The degree exponent $\gamma = |\text{slope}| + 1$

Undirected substrate network  Undirected reaction network

Distances in Metabolic Networks

Paths defined to connect educts to products, the average is calculated on the reachable pairs only

Distance distribution

Average degree

Relatively small and constant network diameter across organisms

Clustering-degree relation in metabolic networks

Average clustering coefficient of nodes with degree $k$

Open symbols: model with the same degree distribution

Straight line: $C(k) \sim k^{-1}$

Suggests hierarchical modularity

Ravasz et al., Science 297, 1551 (2002)
Gene regulatory networks

- nodes: genes (circle) mRNAs (ovals), proteins (boxes)
- edges: mass flow (continuous) or regulation (dashed)
  - regulatory edges acting on edges – similar to catalysis
  - edges can be activating or inhibiting
Simplified representation for gene regulatory networks

- Direct modulation towards product, i.e. TF – mRNA edge or protein- modified protein edge
- General meaning of edges: (positive or negative) information propagation
- Need rules for combining several regulatory effects
Genome-wide transcription networks

- Contract mRNA and protein into a single node, describe transcriptional regulation as a directed gene-gene edge, thick – activation, thin - inhibition
- Sources: TFs that are not regulated at the transcriptional level
- Sinks: non-TF genes, others are both regulators and regulated
Out-degree distribution long-tailed, in-degree distribution more limited

Guelzim et al, Nature Genetics 31, 60 (2002)

S. cerevisiae
Abundant regulatory motifs

- Feedforward loop: convergent direct and indirect regulation possible role: noise filter
- Single input module: one TF regulates several genes possible role: temporal program
- Dense overlapping regulons: groups of genes regulated combinatorially

Shen – Orr et al., Nature Genetics (2002)
More detailed regulatory motifs

- Regulators (TFs), blue circles
- Genes, red rectangles
- Dashed edges mean translation

Regulatory themes

R: transcrip. reg
P: prot. Interaction
H: seq. homology

Feed-forward

Co-pointing

Co-regulation

Protein complex

Zhang et al, J. Biol 4, 6 (2005)
Condition-dependent transcription sub-networks

Representation of chemical reactions and regulation

irreversible reaction
\[ x_1 \rightarrow x_2 \]

reversible reaction
\[ x_1 \rightleftharpoons x_2 \]

divergence
\[ x_1 \rightarrow x_2 \]
\[ x_3 \rightarrow x_1 \]

convergence
\[ x_2 \rightarrow x_1 \]
\[ x_3 \rightarrow x_1 \]

two reactant reactions
\[ x_2 \rightarrow x_1 \]
\[ x_3 \rightarrow x_1 \]

coenzyme
\[ x_3 \rightarrow x_4 \]
\[ x_1 \rightarrow x_3 \]
\[ x_2 \rightarrow x_4 \]

positive modulation
\[ x_1 \rightarrow x_2 \]
\[ x_3 \rightarrow x_2 \]

negative modulation
\[ x_1 \rightarrow x_2 \]
\[ x_3 \rightarrow x_2 \]

autokinase
\[ x_1 \rightarrow x_2 \]

E. O. Voit, Computational Analysis of Biochemical Systems
Signal transduction network example

Red: enzymes
Blue: transport
Orange: small molecules
Green: sign. transd. proteins
Black points: unknown intermediary nodes

Song Li, PSU
Signal transduction network of the hippocampal CA1 neuron

Data (binary interactions) collected form the experimental literature
System of interacting cellular components involved in phenotypic behavior
Edges can be directed or undirected (neutral)
Directed edges are activating or inhibitory

Motif abundance, homeostasis, and plasticity

Looked at three key regulators of plasticity

1. Motif counts increase linearly with steps for all regulators – preferential paths to key effectors;
2. Positive and negative motifs are balanced for glutamate and BDNF - homeostasis;
3. More positive than negative FBL and FFL in NE – long-term info storage

Rapid-change ligands engage more motifs in fewer steps;
At early steps, more FFL than expected; at later steps, more +FBL than expected

Looked at three key regulators of plasticity

1. Motif counts increase linearly with steps for all regulators – preferential paths to key effectors;
2. Positive and negative motifs are balanced for glutamate and BDNF - homeostasis;
3. More positive than negative FBL and FFL in NE – long-term info storage
Importance of a dynamical understanding

Only subsets of the genome-wide interaction networks are active in a given external condition

Han et al. 2004 – dynamical modularity of protein interaction networks
Luscombe et al. 2004 – endogeneus and exogeneus transcriptional subnetworks

Network topology needs to be complemented by a description of network dynamics – states of the nodes and changes in the state
First step - pseudo-dynamics: propagation of reactions in chemical (interaction) space, starting from a source (signal)

Complete dynamical description is only feasible on smaller networks (modules):
Signal transduction in bacterial chemotaxis, NF-kB signaling module, the yeast cell cycle, Drosophila embryonic segmentation
Access dynamics through modeling

First step: define the system; collect known states or behavior
Input: components; states of components
Hypotheses: interactions; kinetics (rates, parameters).
Validation: capture known behavior.
Explore: study cases that are not accessible experimentally
change parameters, change assumptions

Tyson 1991 – cell cycle
Barkai & Leibler 1997, Spiro et al. 1997 – chemotaxis
Bhalla & Iyengar 1999 - EGF pathway
Kholodenko 2000 – MAPK signaling module
G. von Dassow et al. 2000 – segment polarity gene network
Hoffman et al. 2002 - NF-kB signaling
Types of models

1. **Continuous** - similar to chemical kinetics
   - differential equations
2. **Discrete** - assume a small set of qualitative states
   - e.g. active or inactive; basal, intermediate, high
   - the changes in state are given by discrete (logical) rules

1. **Deterministic** - no randomness is involved in the development of future states of the system
2. **Stochastic** - non-deterministic in that the next state of is not fully determined by the previous state.
   - can take into account the fluctuations in mRNA/protein numbers and external noise
Enzyme-catalyzed reactions

Most reactions in biological systems would not take place at perceptible rates in the absence of enzymes. In enzyme-catalyzed reactions the rate of product synthesis depends non-linearly on the concentration of the substrate.

\[
\frac{dP}{dt} = k_2 E_T \frac{S}{K_M + S}
\]

\(K_M\) is equal to the substrate concentration at which the reaction rate is half its maximal value. 

Limits:

\[
\frac{dP}{dt} \approx k_2 E_T \quad \frac{dP}{dt} \approx \frac{k_2}{K_M} E_T S
\]

\(k_2 E_T\) is the number of substrate molecules converted in a unit time when the enzyme is fully saturated with substrate.
Chemical kinetics-like models of cellular processes

Assumption: cellular synthesis and degradation processes can be described as simple or enzyme-catalyzed reactions

Ex.: receptor - ligand binding
    methylation reactions – catalyzed by methylating enzymes,
    phosphorylation - catalyzed by kinases
    dephosphorylation – spontaneous or catalyzed by phosphatases
    protein synthesis – catalyzed by mRNA,
    protein degradation – spontaneous or catalyzed

Kinetics of protein synthesis and degradation

Protein synthesis: mRNA \(\rightarrow\) protein (sufficient supply of amino-acids)

Protein degradation: protein \(\rightarrow\)

\[
\frac{dR}{dt} = k_1 S - k_2 R \quad \text{Steady state:} \quad R_{ss} = \frac{k_1 S}{k_2}
\]
Kinetics of phosphotransfer

Phosphorylation:  protein → phospho-protein
Dephosphorylation:  phospho-protein → protein

The first reaction is catalyzed by a kinase, assume first-order kinetics

\[ \frac{dR_P}{dt} = k_1 SR - k_2 R_P \]

Steady state:

\[ R_{P,ss} = R_T \frac{S}{k_2/k_1 + S} \]

production

degradation
Phosphotransfer with Michaelis-Menten kinetics

Assume that the phosphorylation and dephosphorylation reactions follow Michaelis-Menten kinetics

\[ k_1 SR \rightarrow k_1 S \frac{R}{K_{M1} + R} \quad k_2 R_P \rightarrow k_2 \frac{R_P}{K_{M2} + R_P} \]

Steady state:

\[ R_{Pss} = R_T G\left( k_1 S, k_2, \frac{K_{M1}}{R_T}, \frac{K_{M2}}{R_T} \right) \]

G - Goldbeter-Koshland function
Negative feed-forward loop

The signal acts on R both directly, and through an intermediary.

\[
\frac{dR}{dt} = k_1S - k_2X R
\]

\[
\frac{dX}{dt} = k_3S - k_4X
\]

Steady state: \( R_{SS} = \frac{k_1 k_4}{k_2 k_3} \)

Assume that S has several step changes. R responds to the change, but comes back to the resting level when S stabilizes

- adaptation
Dose-response curves for regulated processes

- $Y$ – regulator (e.g. transcriptional activator)
- $X$ – target (e.g. mRNA)
- Synthesis is a nonlinear function of activator
- Decay is un-catalyzed
- Parameters:
  - maximum rate $T_{\text{max}} \rho_X$
  - Half-maximal activity $K_Y$
  - Hill coefficient $\nu$
  - Half-life $H_X$

Combinatorial regulation of synthesis is approximated with similar sigmoidal curves.

\[
\frac{dX_j}{dt} = T_{j}^{\text{max}} g\left(\sum_i R_{ij} X_i\right) - \frac{X_j}{H_j}
\]
From dose-response curves to switches

If $v$ is large, the dose-response curve becomes a switch

If $Y>K_Y$ \( \frac{dX}{dt} > 0 \)
If $Y<K_Y$ \( \frac{dX}{dt} < 0 \)

The activation threshold is $K_Y$
If activation is weak, mRNA can decay.

$X$ – mRNA
$Y$ – transcriptional activator

Boolean simplification:
$Y>K_Y \implies Y=\text{ON}$
$Y<K_Y \implies Y=\text{OFF}$

$X(t+1) = Y(t)$

Activation:
If $Y(t)=\text{ON}$ \( X \) produced
$X(t+1)=\text{ON}$

Decay:
If $Y(t)=\text{OFF}$ \( X \) decays
$X(t+1)=\text{OFF}$
Hybrid models: Boolean activation combined with continuous decay

- Each node is characterized by both a continuous ($x_i$) and a Boolean ($X_i$) variable.

$$\frac{dx_i}{dt} = B(X_1, X_2..) - x_i$$

- $X_i$ is defined by the threshold rule

$$X_i = \begin{cases} 
0, & \text{if } x_i < 0.5 \\
1, & \text{if } x_i > 0.5
\end{cases}$$

- Compared to

$$\frac{dX}{dt} = T_{max} \rho_x \left( \frac{Y^\nu}{K^\nu_y + Y^\nu} \right) - \frac{X}{H_x}$$

his assumes constant activation threshold=0.5, maximal synthesis rate = decay rate= 1

Modeling the segment polarity gene network

First: System is biologically defined; known expression patterns
Input: segment polarity genes
Hypotheses:
  continuous model: transcription factors act as enzymes
  Boolean model: mRNA and protein activity is switch-like
Validation: reproduces known gene expression patterns.
Explored: changes in kinetic parameters
  mutations
  changes in initial conditions
Insight: topology is a main source of robustness.

Segmentation is governed by a cascade of genes. Transient gene products, initiate the next step then disappear.
Evolution of gene expression patterns

*en*  
early stages  
2:50 h

*hh*  
pre-pattern  
3:00-3:30 h

*wg*  
stable pattern  
4:20-7:20 h

3:30 h
Wild type, stable gene patterns

• *en* is expressed in the anterior part of the parasegment.
• *wg* is expressed in the posterior part of the parasegment.
• Parasegmental grooves form between the *wg* and *en* stripes.

• *two ptc* stripes in each parasegment.
• *ci* pattern is complementary to that of *en*. 
Second reconstruction of the segment polarity gene interaction network

- mRNA
- PROTEIN
- PROT
- COMPL

- repression
- translation, activation, modification

EN \(\rightarrow\) en \(\rightarrow\) SLP \(\rightarrow\) wg \(\rightarrow\) ptc \(\rightarrow\) smo \(\rightarrow\) CI \(\rightarrow\) ci \(\rightarrow\) hh

FZ \(\rightarrow\) WG \(\rightarrow\) PTC \(\rightarrow\) SMO \(\rightarrow\) PH

cell \(\rightarrow\) neighbor cell
Transcripts and proteins are either \textbf{ON} (1) or \textbf{OFF}(0).

- Transcription depends on transcription factors; inhibitors are dominant.
- Translation depends on the presence of the transcript.
- Transcripts and most proteins decay if not produced.
- Synchronous update: transcription, translation, mRNA/protein decay on the same timescale, protein binding faster


- Asynchronous update & hybrid model: post-translational processes faster than pre-translational


Updating rules

\[ hh_i^* = EN_i \text{ and not CIR}_i \]

\[ en_i^* = (WG_{i-1} \text{ or } WG_{i+1}) \text{ and not SLP}_i \]

\[ ptc_i^* = CIA_i \text{ and not EN}_i \text{ and not CIR}_i \]

\[ ci_i^* = \text{not } EN_i \]

\[ EN_i^* = en_i \]

\[ WG_i^* = wg_i \]

\[ CI_i^* = ci_i \]

\[ HH_i^* = hh_i \]

Synchronous update: the next state is the state in the next timestep

\[ hh_i^{t+1} = EN_i^t \text{ and not CIR}_i^t \]
Ingredients of the model

Components: mRNAs and proteins
State: expressed or not in a certain cell
Expression pattern:

Initial state - updating rules– steady state
The updating rules are determined by the network of interactions and could depend on the assumed interaction durations.
The steady state repertoire of the model is independent of durations.
Start with the synchronous model, then explore whether conclusions change by asynchronicity.
The model reproduces the wild type steady state

The interaction network and the net effect of the interactions (with reasonable assumptions on timing) is enough to capture the functioning of the network.
What happens if the components are perturbed?

The most severe perturbation is caused by gene mutations. To model a null mutation, we assume that the mRNA is kept OFF, thus the protein cannot be translated. The effects of the mutation propagate throughout the network.
wg, en or hh mutations are lethal

No wg, en and hh stripes, no segmentation, regardless of initial state or interaction durations.
*ptc* mutation broadens the stripes

The *wg*, *en* and *hh* stripes broaden, regardless of initial state or interaction durations.
ci mutation can preserve the prepattern

The effect of ci mutation depends on the initial state. For wild type prepattern, the wg, en, hh stripes remain, independent of durations.
Model correctly reproduces experimental results on knock-out mutants

Tabata, Eaton, Kornberg, Genes & Development 6, 2635 (1992)
Gallet et al., Development 127, 5509 (2000)
Dynamic repertoire: four steady states

- **Wild type**
- **Broad**
- **Displaced**
- **Ectopic furrow**
- **Lethal**
- **No segmentation**
Divergence from wild-type development in asynchronous model

The concept of timestep (or round of update) is still maintained, but a state change can now affect other nodes’ states within the same step.

Assume that posttranslational processes are always faster than pre-translational ones

Start from the wild type initial state.
Randomize the durations/order in each step.

The steady states of the model are the same as the synchronous model’s, but now oscillations are also possible.
Asynchronous update causes that the WT initial state can lead to one of two steady states

<table>
<thead>
<tr>
<th>Std. state</th>
<th>Incid.</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>87.5%</td>
</tr>
<tr>
<td>broad</td>
<td>12.5%</td>
</tr>
</tbody>
</table>

Hybrid model

- Each node is characterized by both a continuous \( (x_i) \) and a Boolean \( (X_i) \) variable.

\[
\frac{dx_i}{dt} = \alpha_i \left( B(X_1, X_2, \ldots) - x_i \right)
\]

\[
X_i = \begin{cases} 
0, & \text{if } x_i < \theta_i \\
1, & \text{if } x_i > \theta_i 
\end{cases}
\]

- Individual activation threshold, maximal synthesis rate = decay rate = \( \alpha_i \)

Hybrid model more robust than asynchronous Boolean model

- Scale-separation: choose the scale factors from $A_{mRNA}$ and $A_{prot}$
- Faster protein synthesis/decay
- Start from WT initial condition, calculate incidence of steady states

If protein scale factors disjoint from mRNA scale factors, the only possible steady state of the hybrid model is the WT.
Regulation of post-translational modifications crucial for correct dynamic behavior

If a perturbation leads to a transient imbalance between CIA and CIR, the wild type steady state becomes unreachable.

Only CIA - broad stripes; Only CIR - no segmentation

The condition of CIA/CIR complementarity is that PTC be initiated before SMO – true
Interplay between topology and function

- The network contains two activating clusters that inhibit each other in each cell, *en*, *hh* and *ci*, *wg*, *ptc*

- At the same time *en* and *wg* enforce each other in neighboring cells through the secreted proteins HH and WG

- SLP is a regulatory source that maintains asymmetry and limits *en* and *wg* to different halves of the parasegment.
Acknowledgements

Topology of biological networks:
Albert- László Barabási
Hawoong Jeong, Zoltán Oltvai

Drosophila segmentation:
Hans G. Othmer, Eduardo Sontag
Madalena Chaves

Graph inference:
Bhaskar Dasgupta, Riccardo Dondi
Eduardo Sontag

Gene regulation:
Anshuman Gupta, Claire Christensen
Costas Maranas, Yu-wen Chiu,
Anton Nekrutenko, Kateryna Makova

Protein interaction prediction:
István Albert

ABA signaling in plants:
Sarah Assmann, Song Li

Pathogen-immune system interactions:
Eric Harvill, Jaewook Joo, Juilee Thakar

Alfred P. Sloan Foundation
National Science Foundation

Course on Graphs and Networks in Systems Biology

http://www.phys.psu.edu/~ralbert/teaching.htm
Modeling abscisic acid (ABA) signaling in plants

First step: Process is biologically defined: stomatal closure
Input: proteins and molecules affecting ABA-induced stomatal closure
Hypotheses: network synthesis, Boolean kinetics
Validation: reproduce known behavior in wild type and mutant plants.
Explored: changes in initial state
    single and multiple component disruptions

Insight: redundant topology is a main source of robust signaling.

S. Li, S. Assmann and R. Albert 2006.
Stomata open in the morning and close during the night. The immediate cause is a change in the turgor (fullness) of the guard cells.

90% of the water taken up by a plant is lost in transpiration, while the stomata are open.

During drought conditions the hormone abscisic acid (ABA) triggers the closing of the stomata.

More than 20 proteins and molecules participate in ABA-induced closure, but their interaction network has not been synthesized yet.
Mediators of ABA-induced stomatal closure

Inference methods: genetic & pharmacological perturbations
biochemical evidence

NO, cADPR,
cGMP, S1P, IP3,
IP6 etc…

Ca^{2+} increase/
oscillation

pH increase
K^+ efflux
anion efflux

ABA → Closure

ABI1(PP2C), ABI2(PP2C), RCN(PP2A), ERA1-2, etc..
Database construction

- Literature mining & curation - Song Li
- Define network
  - nodes: proteins, chemical messengers, ion channels, concepts
    Examples: ABA, SphK, K efflux, pH, depolarization, closure
  - edges: interactions, activating or inhibiting effects on nodes or other edges
  - classify biological information into activation or inhibition
    Examples: ABA → SphK, SphK → (ABA → closure)

<table>
<thead>
<tr>
<th>Node A</th>
<th>Node/Process B</th>
<th>interaction</th>
<th>species</th>
<th>ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS</td>
<td>ABA → closure</td>
<td>promotes</td>
<td>Vicia faba</td>
<td>(1)</td>
</tr>
<tr>
<td>PLC</td>
<td>ABA → closure</td>
<td>promotes</td>
<td>Commelina communis</td>
<td>(3)</td>
</tr>
<tr>
<td>SphK</td>
<td>ABA → AnionEM</td>
<td>partially promotes</td>
<td>Arabidopsis</td>
<td>(4)</td>
</tr>
<tr>
<td>ABA</td>
<td>SphK</td>
<td>promotes</td>
<td>Arabidopsis</td>
<td>(4)</td>
</tr>
</tbody>
</table>
Network construction

Need to synthesize experimental inferences into the simplest network that incorporates all effects.

Edges should connect pairs of nodes: introduce intermediary nodes (1,3)

Limit redundancy to minimal supported: contract intermediary nodes (2)

The full algorithm is an example of a binary transitive reduction problem.

enzymes
sign. trans. proteins
transport
small molecules
- interm. node
inf. edges
Two pathways of Ca\(^{2+}\) activation

At least two separate ABA-closure pathways, one through Ca\(^{2+}\), the other through pH\(_c\).

Pathway redundancy suggests robustness to perturbations.

Actin reorganization, pH\(_c\) increase, malate breakdown, membrane depolarization need to be simultaneously disrupted to block all ABA-closure paths.
Boolean model of network dynamics

• Each node has two states: 1 (active) and 0 (inactive)
• “closure=1” does not mean “stomata are closed” because “open” and “closed” stomatal apertures are both distributions
• The next state of each node is determined by a function of the state of its regulators
• The regulatory rules are based on the reconstructed interactions.
• Synergy -- AND; independence -- OR; inhibitors -- NOT.

\[
\begin{align*}
\text{pH}^* &= \text{ABA} \\
\text{CaATPase}^* &= \text{Ca}^{2+}_c \\
\text{Ca}^{2+}_c^* &= (\text{CaIM or CIS}) \text{ and } (\text{not CaATPase}) \\
\text{CIS}^* &= (\text{cGMP and cADPR}) \text{ or } (\text{Ca}^{2+}_c \text{ and InsP3}) \text{ or InsP6} \\
\text{AnionEM}^* &= ((\text{Ca}^{2+}_c \text{ or } \text{pH}_c) \text{ and } \text{not ABI1}) \text{ or } (\text{Ca}^{2+}_c \text{ and } \text{pH}_c) \\
\text{Closure}^* &= (\text{KOUT or KAP}) \text{ and AnionEM and Actin and not Malate}
\end{align*}
\]
Random asynchronous updates

- It is probably not realistic to assume equal interaction timescales, but no kinetic information
- Asynchronous algorithm with randomly selected timing/order.
- The update order is changed after each round

Randomize the initial states of all the nodes to mimic the noise in the internal environment of the guard cell.

- We do 10,000 simulation runs and interpret the number of different runs having achieved closure at a certain timestep as the probability of closure.

- We simulate node deactivation (e.g. pharmacological inhibition) by setting and maintaining the state of the node to 0. This means that although the node is not completely absent, it cannot reach the activated state and cannot transmit the signal.
Signal transduction is resilient to perturbations

- Normal response to ABA stimulus.
- No stimulus
- ABI1 knockout mutants respond faster (hypersensitivity).
- Ca\(^{2+}\) clamping leads to slower response (hyposensitivity)

Perturbations in □ anion efflux or depolarization cause ABA insensitivity.

Perturbations in ■ SphK or S1P, ○ GPA1, PLD or PA, or △ pH\(_c\) lead to decreased sensitivity.

Prediction: pH disruption more severe than Ca\(^{2+}\) disruption.
Model predicts remarkable robustness

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65%</td>
<td>5%</td>
<td>5%</td>
<td>17.5%</td>
<td>7.5%</td>
</tr>
<tr>
<td>2</td>
<td>46%</td>
<td>8.1%</td>
<td>8.1%</td>
<td>26.5%</td>
<td>15.8%</td>
</tr>
<tr>
<td>3</td>
<td>32%</td>
<td>9.8%</td>
<td>10.5%</td>
<td>29.6%</td>
<td>24.8%</td>
</tr>
</tbody>
</table>

Continuum of close-to normal sensitivity

Cumulative prob. of closure: the sum of PC over 12 steps
Experimental validation: disruption of Ca\textsuperscript{2+} versus pH

Normal: “open” and “closed” state distinguishable

Ca\textsuperscript{2+} disrupted: “open” and “closed” state distinguishable

pH disrupted: “open” and “closed” state indistinguishable

Qualitative agreement with theoretical prediction.