The Structure, Function, and Evolution of Biological Systems

Instructor: Van Savage
Spring 2016 Quarter
4/13/2016
Alon Book
Other important ways to construct gene networks: Gene regulation and motifs

- Organism must be able to respond to environment and gene expression is one way to do this

- One gene can create a protein that either *activates* or *represses* another gene

![Diagram showing activation and repression]

- promoter region (easily evolvable)
Self loops
E coli network has been measured

- N=424 genes, E=519 total edges, and 40 of those are self edges

- Is this a lot or a little? What do we compare with?

$E_p = E/N$ is expected number of self edges and mean connectivity

For network below N=6 and E=8 edges, so $E_p = 4/3$

Variance is approximately $E/N$ (Poisson), so standard deviation is $2/Sqrt(3)$, so for our example The one self edge is expected by random chance

For E Coli $E_p = 1.2 +/-.11$, so many more self edges that expected
Compare with unregulated

\[ \frac{dX}{dt} = \beta - \alpha X \]

unregulated

\[ X^{st} = \frac{\beta}{\alpha} \quad \text{and} \quad T_{1/2} = \frac{X^{st}}{2\beta} \]

regulated

\[ \frac{dX}{dt} = \beta \left( \frac{1}{1 + \left( \frac{X}{K} \right)^n} \right) - \alpha X \]

Hill function

\[ X^{st} = \left( \frac{\beta}{\alpha K} \right)^{1/n+1} K \rightarrow K \]

Can now adjust speed, \( \beta \), without affecting asymptote/steady state, which is fixed physiologically. Allows faster response times.
What about activation instead of repression?

\[ \frac{dX}{dt} = \beta \left( \frac{X}{K + X^n} \right) - \alpha X \]

Hill function

Slows down response time and allows for a type of memory
What other types of “motifs” might be possible to look for?

Consider all possible networks with 3 nodes.

Only 13 possibilities.

How do we look to see if a given motif is overrepresented in the data?
All 3-node subgraphs

Feed-forward loop

3-node feedback loop (cycle)

(a)

(b)
Expected number of motifs in random network

There are $N^2$ possible edges between $N$ genes, so chance of picking specific edge is $1/N^2$. If there are $E$ edges in network, you have $E$ chances to pick it, so $p=1 - (1 - 1/N^2)^E \sim E/N^2$ is chance of picking specific edge.

For a given subgraph, $\sim N^n$ ways to pick a subgraph with $n$ nodes and $p^g$ chance of picking correct $g$ edges in subgraph.

Expected number of subgraphs with $n$ nodes and $g$ edges is

$$\frac{N^n p^g}{a} = \frac{\lambda^g N^{n-g}}{a}$$

where $a$ is symmetry number and $\lambda=E/N$ is mean connectivity.
Keep connectivity fixed while increasing network size

For $n=3$ subgraphs

Two edge patterns become more common

Three edge patterns stay constant number

Four and higher edge patterns become vanishingly small

Only one $n=3$ subgraph is overrepresented: Feed Forward Loop

42 FFLs exist. We would expect on the order of $\lambda^{3-1.7+/-1.1}$

0 FBLs exist. May even be selected against.
FFLs in E Coli
Possible types of FFLs

Coherent FFL

<table>
<thead>
<tr>
<th>Coherent type 1</th>
<th>Coherent type 2</th>
<th>Coherent type 3</th>
<th>Coherent type 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
</tbody>
</table>

Incoherent FFL

<table>
<thead>
<tr>
<th>Incoherent type 1</th>
<th>Incoherent type 2</th>
<th>Incoherent type 3</th>
<th>Incoherent type 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Types of FFLs in Yeast and E Coli
What are function of FFLs?

Now, can add detail of whether arrows are activating or repressing

One is coherent and one is incoherent

Coherent allows us to integrate two signals with an AND or an OR gate.

AND creates delays in response to make sure signal is really there

OR creates delays in fall to make sure signal is really gone

Incoherent allows integration of two contradictory signals

Other types cause interference or cross signals
Problems with this approach

All subgraphs are taken out of context.

Is that reasonable or appropriate?

How useful is it?