Lecture 5: Auto-regulation – a network motif

Chapter 3 of Alon

3.1 Introduction

1. Define a way (based on statistical significance) to detect network motifs
2. Auto-regulation motif
3. Auto-regulation has useful functions: speed-up of response, stabilizer
3.2 Patterns, randomized networks, and network motifs

- To define statistical significance, compare the network to an ensemble of randomized networks
- Patterns that occur in the real network significantly more often than in randomized networks with the same characteristics (number of nodes, number of edges) are called network motifs

- Edges are easily lost in a transcription network: a mutation that changes a single DNA letter in a promoter can abolish (or create) a binding of a transcription factor and cause the loss or addition of an edge
- See the example in the book (pp 28-29): a change of any DNA letter of the genome can be reached many times very rapidly (within less than a day) in bacterial populations

=> Edges in the network motifs must be constantly selected in order to survive randomization forces in unexpected high amounts
=> Motifs must give some advantages to the organism

3.2.1 Detecting motifs by comparison to randomized networks

- Erdős-Renyi (ER) model of random graphs: directed edges are assigned at random between pairs of nodes
  - N nodes & E edges
  => There are N^2 possible directed edges (includes the self-edges)

  - In the ER model, the E edges are placed at random in the N^2 possible positions
  => each possible edge is present with probability p = E/N^2 (Explain why!)
Fig 3.1a. Self-regulating genes in a network of transcription interactions in E. coli. Nodes that correspond to genes (operons) which encode transcription factor proteins that regulate their own promoters (self-regulating genes, represented by self-edges) are shown in blue. This network, which we will use as an example in the coming chapters, has N=420 nodes, E=520 edges and Es=40 self-edges.

Fig 3.1b

‘Real’ Network

Randomized network (Erdos – Renyi model)

N=10 nodes  
E= 14 edges  
Es=4 self-edges

N=10 nodes  
E= 14 edges  
Es=1 self-edge
3.3 Auto-regulation - a network motif

- **Self-edge**: originates and ends at the same node
  - *E. Coli* network: 40 self-edges (Fig 3.1 a)
- Self-edges ↔ auto-regulation
- **Negative auto-regulation**: repressor proteins that repress their own transcription
  - *E. Coli* network: 34 cases of negative auto-regulation

Is negative auto-regulation significantly more frequent in the real network than in a random graph with the same number of nodes and edges?

Fig 3.2a: Gene X is simply regulated by A.

Fig 3.2b: Gene X is **negatively auto-regulated**, and simply regulated by A. Repressor X binds a site in its own promoter and thus acts to repress its own transcription. The symbol $\rightarrow{\text{--|}}\rightarrow$ stands for repression. The repression threshold is $K$ (defined as the concentration of X needed to repress the promoter activity by 50%).
Auto-regulation (cont.)

- The selection probability of a self-edge
  \[ p_{\text{self}} = \frac{1}{N} \]

- Probability of having k self-edges is binomially distributed (throwing a coin \( E \) times and getting k heads):
  \[ P(k) = \binom{E}{k} p_{\text{self}}^k (1-p_{\text{self}})^{E-k} \]

- Average number \(<\cdot>\) of self-edges in a random graph (from Poisson approximation of binomial distribution):
  \[ <N_{\text{self}}>_{\text{rand}} \sim \frac{E p_{\text{self}}}{N} = \frac{E}{N} \]

- Standard deviation of the number of self-edges (Poisson approximation)
  \[ \sigma_{\text{rand}} = \frac{E}{N^{1/2}} \]

- Def. (Z-score): The Z-score \( Z(a) \) of a value \( a \) of a random variable \( x \) is
  \[ Z(a) = \frac{|a - \text{mean}(x)|}{\sigma(x)} \]
  where \( \sigma(x) \) is the standard deviation of \( x \). Score \( Z(a) \) is the deviation from the mean measured in standard deviations.

Example: self-edges of E. coli network are a motif

- \( E. coli \)'s network (Fig 3.1) has \( N = 424, E = 519 \). Then
  \[ <N_{\text{self}}>_{\text{rand}} \sim \frac{E}{N} \sim 1.2 \]
  \[ \sigma_{\text{rand}} \sim \sqrt{1.2} \sim 1.1 \]
  \[ <N_{\text{self}}>_{\text{real}} = 40 \]

  \[ Z(40) = (40 - 1.2)/1.1 \sim 35 \]

- 35 standard deviations mark a very high significance
  => self-edges (and also the 34 negatively auto-regulating self-edges for which \( Z \sim 30 \)) are a network motif
3.4 Negative auto-regulation speeds the response time

- Let protein X be negatively auto-regulated: $X \rightarrow X$
- Recall that the dynamics of X is (from the dynamics of $X \rightarrow Y$ in Lect 4 when $X=Y$):
  \[
  \frac{dX}{dt} = f(X) - \alpha X
  \]
- where $f(X) = \frac{\beta}{1 + \left(\frac{X}{K}\right)^n}$
- If X is much smaller than the repression coefficient K, then the production rate of X reaches its maximal value $\beta$
- If X is high, then no transcription occurs and hence $f(X) \sim 0$.

Speed-up of response time (cont.)

- We solve the dynamics using logic approximation where $f(X) = 0$, if $X > K$, and $f(X) = \beta$ if $X < K$:
  \[
  f(X) = \beta \Theta(X < K)
  \]
- Let X be initially absent ($X=0$) and its production starts at $t=0$. Then
  \[
  \frac{dX}{dt} = \beta - \alpha X \quad \text{while } X<K
  \]
- At early times we have $\alpha X << \beta$. Hence we can neglect degradation $\alpha$ and have
  \[
  X(t) \sim \beta t \quad \text{while } X<K \text{ and } X << \beta/\alpha
  \]
- However, when X levels reach the self-repression threshold $X = K$, the production of X stops (small oscillations will occur around $X=K$ if there are delays in the system)
  \[
  \Rightarrow X \text{ effectively locks itself into a steady-state level equal to the repression coefficient of its own promoter (Fig. 3.3)}
  \]
  \[
  X_{st} = K
  \]
Fig 3.3: Dynamics of a negatively auto-regulated gene product. Production starts at t=0. Full line: Negatively auto-regulated gene with maximal production rate $\beta = 5$, auto-repression threshold $K=1$, and degradation/dilution rate $\alpha = 1$. Dashed line: Dynamics of the same gene if auto-regulation is removed, resulting in simple regulation that approaches a higher, unrepressed steady-state $X_{ss} = \frac{\beta}{\alpha} = 5$.

Speed-up of response time (cont.)

- Response time: $X(T_{1/2}) = \frac{X_{ss}}{2}$
- Using linear approximation $X = \beta t$, we obtain
  $$T_{1/2}^{(n.a.r.)} = \frac{X_{ss}}{2\beta} = \frac{K}{2\beta}$$
  $(n.a.r.) = \text{negative auto-regulation}$

- Note: Evolutionary selection can tune parameters $\beta$ and $K$ independently
  - $K$ modified, for example, by mutations in the binding site of $X$ in the promoter
  - $\beta$ tuned by mutations in the binding site of RNAp (RNA polymerase) in the promoter

- Response time of simply regulated vs negatively auto-regulated genes?
  $$\frac{T_{1/2}^{(n.a.r.)}}{T_{1/2}^{(n.a.r.)}} = \frac{\beta_{\text{simple}}}{\beta} \frac{1}{2 \ln 2}$$

  $\Rightarrow$ the larger is $\beta$, the smaller is the n.a.r. response time as compared with the simple response time (Figs 3.4 & 3.6)
Fig 3.4 Dynamics of negatively auto-regulated gene product (full line) and simply regulated gene product (dashed line) which reach the same steady-state level and have equal degradation/dilution rates \( \alpha \). The response time is the time that the protein level reaches 50% of the steady state, denoted \( T_{1/2}^{(\text{nar})} \) and \( T_{1/2}^{(\text{simple})} \) for the negatively auto-regulated and simply regulated gene products. The parameters \( \beta=5, \alpha=1, \beta_{\text{simple}}=1 \) were used.

Fig 3.6: Experiment on negatively auto-regulated and simply-regulated genes. The experiment used green-fluorescent protein fused to the TetR repressor as a reporter and automated fluorescence measurements on growing E. coli cells. Protein concentration is normalized to its steady-state level. Shown also are the analytical solutions for a simply auto-regulated gene and for a negatively auto-regulated gene with a Hill input function with \( n=1 \) in the limit of strong auto-repression (solved exercise 3.1). Source: Rosenfeld, Elowitz, Alon, JMB 323:785 2002
3.5 Negative auto-regulation promotes robustness to fluctuations in production rate

- Simple gene regulation is affected quite strongly by fluctuations in production rate $\beta$, as $X_{st} = \beta/\alpha$ and hence a change in $\beta$ leads to proportional change in $X_{st}$.

- In contrast, negative auto-regulation can buffer such fluctuations, as the steady-state level depends only on the repression threshold of $X$ for its own promoter: $X_{st} = K$.

- **Positive auto-regulation** slows the response time relative to simple regulation (see Fig 3.5). The dynamics are initially slow but with a growing level of $X$, its production rate increases due to positive autoregulation loop. This results in a concave curve that reaches 50% of its steady-state value at a delay relative to simple regulation.

![Fig 3.5: Dynamics of negatively auto-regulated gene, a simply regulated gene and a positively auto-regulated gene. The negatively and positively auto-regulated genes have a Hill-input function with Hill coefficient $n=1$. Shown is protein concentration normalized by its steady-state value, $X/X_{st}$, following an increase in production rate. Time is in cell-generations, or for actively degraded proteins, $\log(2)/\alpha$, where alpha is the protein degradation/dilution rate. Note that the response-time is $T_{1/2} = \log(2)/\alpha = 1$ for simple regulation, $T_{1/2} = 0.21$ for negative auto-regulation, and $T_{1/2} = 2$ for positive auto-regulation with the present parameters. The response-time is constructed by the intersect of the dynamics with horizontal line at $X/X_{st} = 0.5$.](image-url)