

Algo de dinámica en redes biológicas

¿Por qué nos interesa?

Entre otras cosas porque se observó que, en distintas circunstancias, el ingreso de los factores de transcripción al núcleo (en células eucariotas) es pulsátil.

En ese caso, ¿importa la amplitud de los pulsos o importa su frecuencia?

Por otro lado uno de los aspectos importantes del desarrollo embrionario es la coordinación de las decisiones regulatorias a lo largo del tiempo.

Otro ejemplo de variación en el tiempo es el de las baterías de genes que se expresan en diferentes momentos durante el ciclo celular.

También hablamos de switches que pueden variar al cambiar las condiciones del medio.

Para entender estos aspectos varios es que abordaremos brevemente la cuestión de la expresión génica dependiente del tiempo.

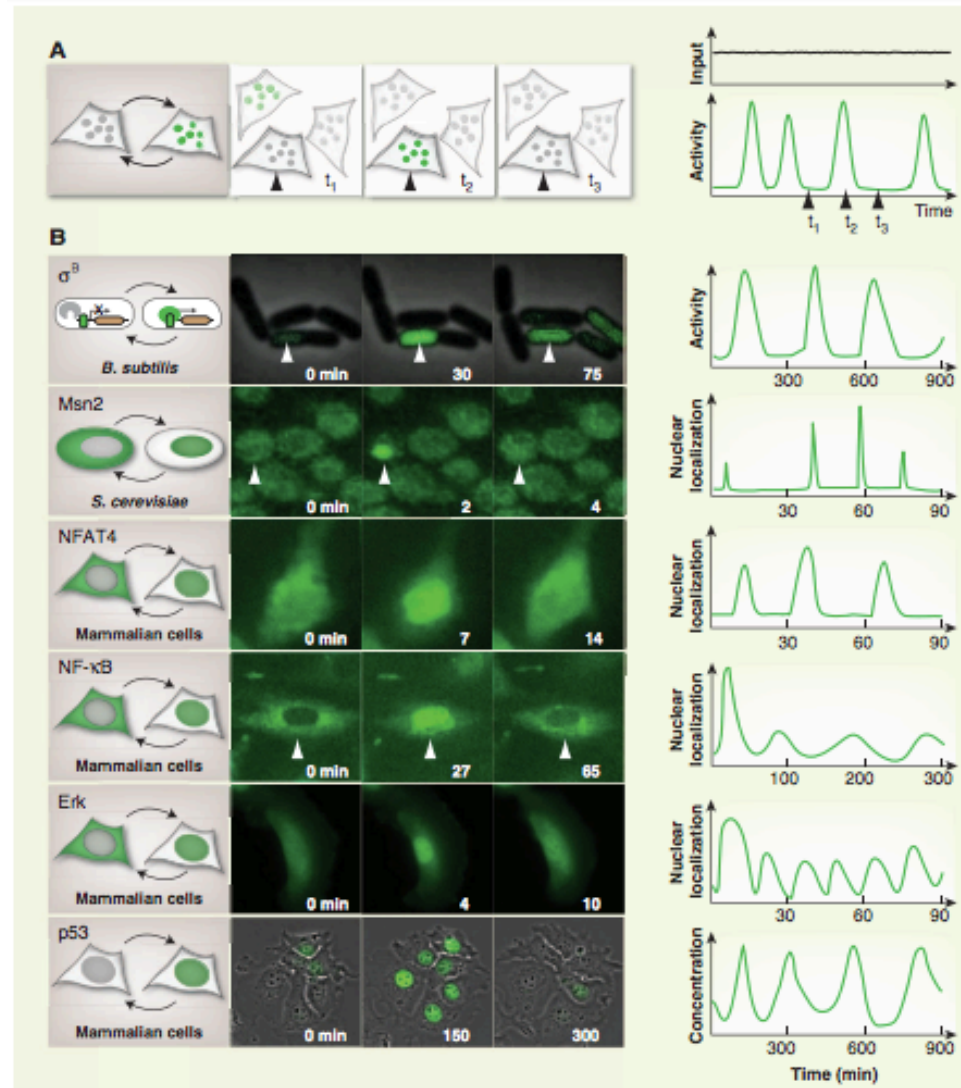


Fig. 1. Pulsing is ubiquitous in cellular regulation. (A) Pulsatile dynamics involve the transient, simultaneous activation of many molecules of a given type (circles), even under constant environmental conditions. Cells pulse asynchronously, making pulsing difficult to detect with static snapshots and necessitating tracking of cell lineages over time (right, schematic). (B) Pulsing occurs in a diverse array of pathways, molecular types, organisms, and time scales (1–6). For each example, a schematic of the type of regulation is shown at left, a typical filmstrip is shown at center, and a qualitative schematic plot of typical dynamics is shown at right.

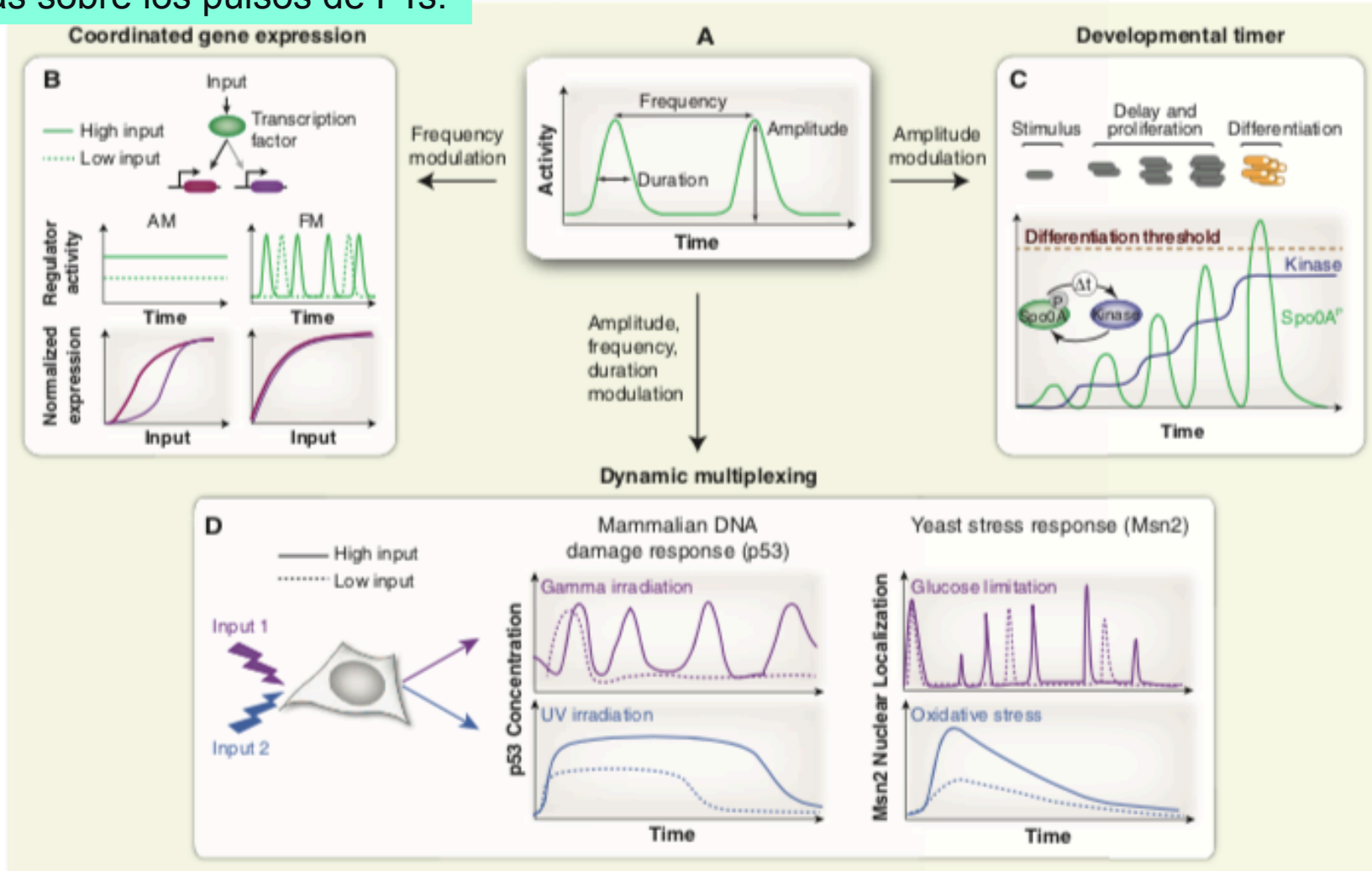


Fig. 2. Pulsing enables diverse cellular functions. (A) Cells modulate pulse characteristics, including amplitude, frequency, and duration, to implement diverse regulatory functions. (B) A transcription factor (green) may activate different target promoters at different thresholds or with different affinities (light and dark arrows). Concentration-based regulation (amplitude modulation, AM) would therefore lead to different, nonproportional, response profiles (bottom left). In contrast, frequency-modulated (FM) pulsing, by effectively controlling the fraction of time that all target genes are expressed, leads to expression of targets in fixed proportions (bottom right), indicated by overlap of expression curves (each is normalized to its own maximum) (11). (C) Pulsed regulation functions in

a developmental timer. *B. subtilis* respond to sudden nutrient limitation by proliferating for multiple cell cycles before sporulating (schematic). A model of the underlying circuit (inset) is based on a positive-feedback loop (arrows) with a hypothesized time delay (Δt). This circuit can generate progressive growth in pulses of phosphorylation of the sporulation master regulator Spo0A (green trace), via steplike growth in the kinase concentration (blue trace). The timer terminates when a threshold level of Spo0A is reached (dashed line) (10). (D) Examples in which dynamic multiplexing enables a single pathway to transmit multiple signals (2, 13). In each case, distinct types and levels of inputs generate distinct dynamic activation patterns for the indicated regulatory protein.

Fig. de Levine et al, 2013

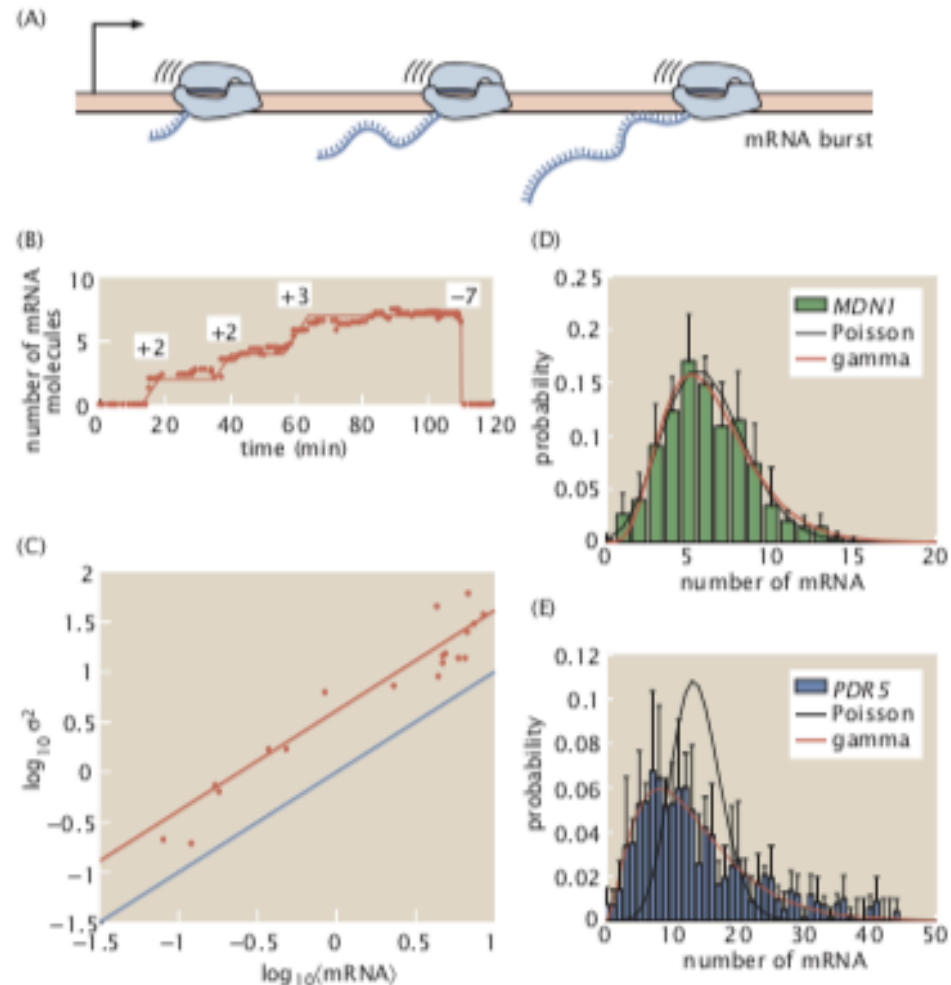
Modelando la dinámica de la transcripción más mecánicamente

La idea es escribir ecuaciones dinámicas para las concentraciones y/o probabilidades relevantes en el problema regulatorio de interés. Para ello podemos llevar adelante una descripción en términos de la cinética de acción de masas. Una descripción alternativa es usando variables aleatorias (en lugar de ecuaciones para probabilidades o concentraciones).

El escenario más simple es considerar solo las ecuaciones para las concentraciones de las proteínas relevantes. Dependiendo de la pregunta que se quiera responder, en ocasiones es también necesario caracterizar la evolución temporal de las transcripciones de ARNm.

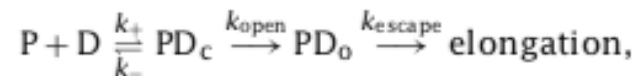
En particular, esto último es cada vez más relevante dadas las observaciones experimentales con las que se cuenta actualmente donde se ven, e.g., “bursts” en la transcripción (ver figura).

Figure 19.33: Time-dependent dynamics of transcriptional networks. (A) A burst of mRNA production. (B) Time history of the number of mRNA molecules in a given *E. coli* cell, revealing periods of no production punctuated by bursts of production with the size of the burst indicated by the numbers in the white boxes. (C) Noise in *E. coli* gene expression (measured by the variance) as a function of the mean level of gene expression. (D) Distribution of mRNA in yeast for the *MDJ1* gene. (E) Distribution of mRNA in yeast for the *PDR5* gene. The distributions in (D) and (E) are fitted to various models considered in the section. (B, C, adapted from I. Colding et al., *Cell* 123:1025, 2005; D, E adapted from D. Zenklusen et al., *Nat. Struct. Mol. Biol.* 15:1263, 2008.)



Vamos a hacer primero unas estimaciones para ver hasta qué punto son válidos los cálculos de equilibrio que vinimos haciendo. Esta suposición de equilibrio se justifica cuando la unión de la polimerasa al promotor se produce en una escala de tiempo mucho más rápida que el tiempo que tarda la polimerasa en iniciar la transcripción desde el momento en que se une. En ese caso, la polimerasa está en un pre-equilibrio rápido con el ADN y la cantidad de transcripción es proporcional a la fracción de tiempo que la polimerasa está unida.

Pensemos en los pasos clave a lo largo del proceso de transcripción: la unión de la ARN polimerasa, P, al promotor, D, que regula el gen de interés formando un complejo cerrado con el ADN, PD_c ; formación de un complejo abierto, PD_o , en el que las dos hebras de ADN se separan permitiendo que la ARN polimerasa lea una de ellas; partida del promotor, cuando la ARN polimerasa comienza a transcribir el gen. Lo podemos modelar como:



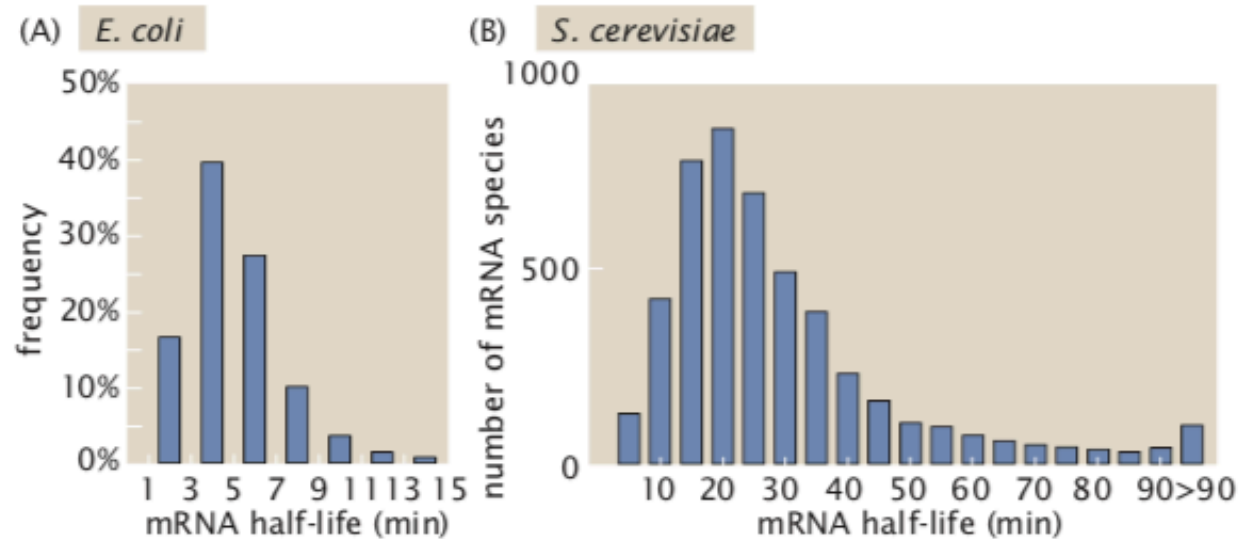
donde la última transición deja al promotor libre para que se pueda pegar otra polimerasa.

Si la ARN polimerasa tiene tiempo para unirse y separarse del promotor varias veces antes de que se forme el complejo abierto, entonces podemos pensar que el primer paso alcanza el equilibrio que está caracterizado por la constante de equilibrio $K_P = k_+/k_-$. En particular, existe una separación de escalas temporales que avala esta aproximación: en el caso del operón *lac*, por ejemplo, la unión de la ARN polimerasa al promotor *lacUV5 in vitro* es tan rápida que las velocidades k_{\pm} ni siquiera se miden en experimentos típicos. En cambio, se mide la constante de equilibrio de $K_P \approx 200 \mu\text{M}^{-1}$, y $k_{open} \approx 0.1 \text{s}^{-1}$.

Una vez que la ARN polimerasa inicia la transcripción va produciendo el ARNm a una velocidad de aproximadamente 50 nucleótidos por segundo, lo que significa que un gen típico, de aproximadamente 1000 nucleótidos, se transcribirá en aproximadamente 20s, es decir, una polimerasa produce 3 ARNm por minuto.

La producción de ARNm a partir de un gen típico en *E. coli* se produce a una velocidad de aproximadamente 10 por minuto.

Figure 3.14: mRNA lifetimes. (A) Measurements of the lifetimes of thousands of different mRNA transcripts in *E. coli* using microarrays. The mean lifetime is slightly in excess of 5 minutes. (B) Measurements of the lifetimes of thousands of different mRNA transcripts in the yeast *S. cerevisiae* using microarrays. (A, adapted from J. A. Bernstein et al., *Proc. Natl Acad. Sci. USA* 99:9697, 2002; B, adapted from Y. Wang et al., *Proc. Natl Acad. Sci. USA* 99:5860, 2002.)



El cálculo que acabamos de hacer implica entonces que hay alrededor de 3 ARN polimerasas transcribiendo simultáneamente un gen en cada momento.

Estos números son típicos de la producción de ARN mensajero en *E. coli*. El ARN ribosómico, que no se traduce y es un componente clave de los ribosomas, se produce a velocidades de aproximadamente 1 ARNm por segundo, o sea, mucho más rápido.

Más números: la vida media de un ARNm como el del operón lac es de un poco más de 1 minuto. En estado estacionario, el número de ARNm creados en la célula durante cualquier intervalo de tiempo debe, en promedio, equilibrar el número degradado. Dado que las moléculas de ARNm se crean a 10 por minuto, la misma cantidad de moléculas debe degradarse cada minuto, y concluimos que, en promedio, habrá 10 moléculas de ARNm por célula.

En el caso de un promotor regulado como *lacUV5*, también deberíamos considerar las velocidades a las que los factores de transcripción como el represor Lac se pegan y despegan de la región regulatoria del ADN.

La tasa de unión limitada por difusión del represor Lac al ADN del operador es de aproximadamente $0.003\text{s}^{-1}\text{nM}^{-1}$. Suponiendo que hay 10 moléculas represoras en la célula y que una sola molécula dentro del volumen típico de una *E. Coli* corresponde a una concentración de 1 nM, la probabilidad por unidad de tiempo de que el represor se pegue al operador es 0.03s^{-1} .

La tasa de disociación del ADN del operador varía con la fuerza del operador. Para el represor Lac, puede ir entre 2s^{-1} para O_3 y 0.002s^{-1} para O_{id} . Estas tasas son del orden e incluso más lentas que la tasa de inicio de la transcripción, lo que sugiere que la suposición de equilibrio para este promotor podría no ser válida. Sin embargo, si agregamos una simple represión al esquema de reacción que vimos antes se llega a que el razonamiento del equilibrio también es válido en este caso.

Veamos ahora brevemente cómo describir la dinámica de la producción del ARNm.

La pintura más sencilla está esquematizada en esta figura:

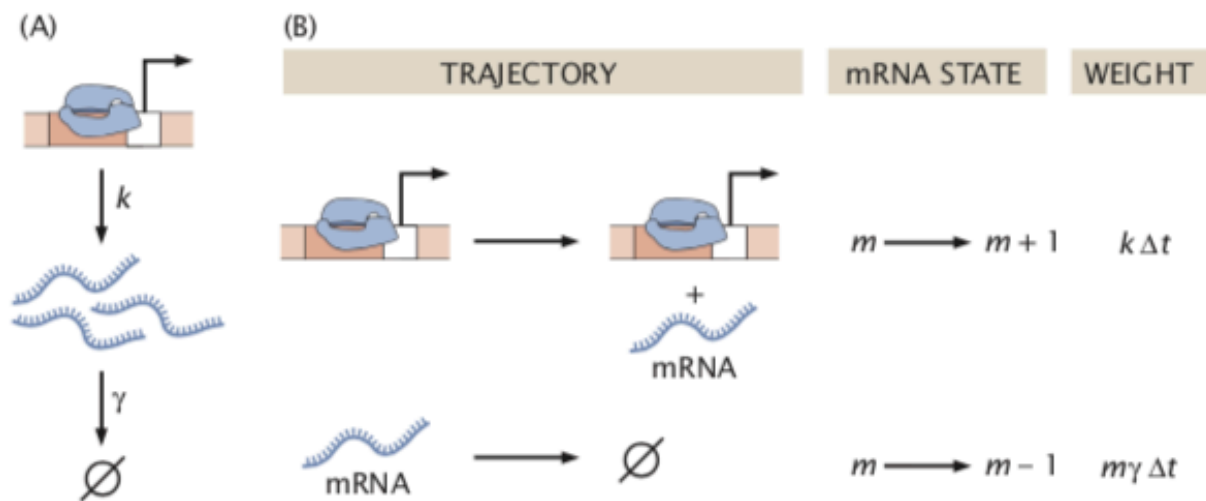


Figure 19.34: Trajectories and weights for the case of an unregulated promoter. (A) Schematic of the processes that can occur for the unregulated promoter, namely, mRNA production with rate k and mRNA decay with rate γ . (B) The individual processes that can occur in a time step of length Δt and their corresponding statistical weights.

Acá k es la tasa a la que se produce el ARNm y γ es la tasa a la que se degrada. La ecuación que satisface la probabilidad, $p(m,t)$, de tener m ARNMs al tiempo t es entonces:

$$\frac{dp(m, t)}{dt} = \underbrace{-kp(m, t)}_{m \rightarrow m+1} + \underbrace{kp(m-1, t)}_{m-1 \rightarrow m} - \underbrace{\gamma mp(m, t)}_{m \rightarrow m-1} + \underbrace{\gamma(m+1)p(m+1, t)}_{m+1 \rightarrow m}$$

En el caso $m = 0$ el segundo término de la ecuación no está (ya que los $m < 0$ no tienen sentido). Una forma alternativa de trabajar es permitir $m < 0$ con la condición $p(m < 0, t) = 0$.

Se puede calcular la ecuación para el valor medio, $\langle m \rangle$. Como era de esperar, resulta:

$$\frac{d\langle m(t) \rangle}{dt} = k - \gamma \langle m(t) \rangle.$$

cuya solución estacionaria es: $\langle m \rangle = k/\gamma$

Por otro lado, la solución estacionaria de la ecuación para p (ec. maestra) es:

$$p(m) = \frac{\lambda^m e^{-\lambda}}{m!}, \quad \text{Con } \lambda = k/\gamma$$

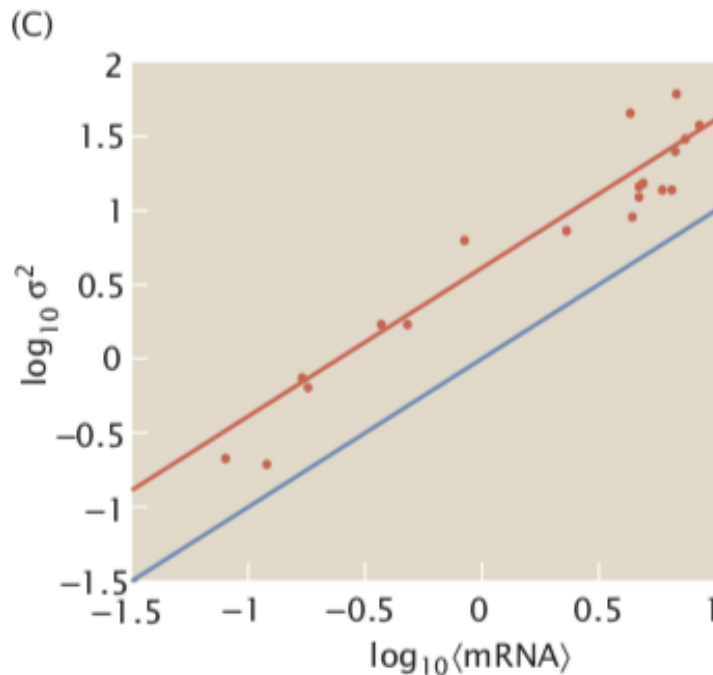
Como se desprende de los temas que abordan algunos de los papers que podían elegir contar, un área que viene siendo estudiada desde hace varios años es la del ruido en la expresión génica.

Una forma de cuantificar el ruido es mediante el factor de Fano:

$$\text{Fano factor} = \frac{\langle m^2 \rangle - \langle m \rangle^2}{\langle m \rangle}.$$

Este factor es 1 para una distribución como la de Poisson que acabamos de obtener en nuestro modelo sencillo.

Este factor se puede calcular a partir de experimentos. En una figura ya lo mostramos y lo repetimos acá.



En este caso (E. coli) el factor es ~4 en lugar de 1!

En los estudios de equilibrio que hicimos vimos que los promotores pueden existir en muchos estados diferentes: reprimido, vacío, ocupado por ARN polimerasa, activado, etc.

El modelo de un estado que exploramos antes corresponde a la situación más simple en la que siempre hay una ARN polimerasa en el promotor: cuando una de ellas comienza la transcripción y deja el sitio, su lugar es inmediatamente ocupado por otra polimerasa.

Un modelo un poco más sofisticado como el de la figura puede explicar un factor de Fano > 1

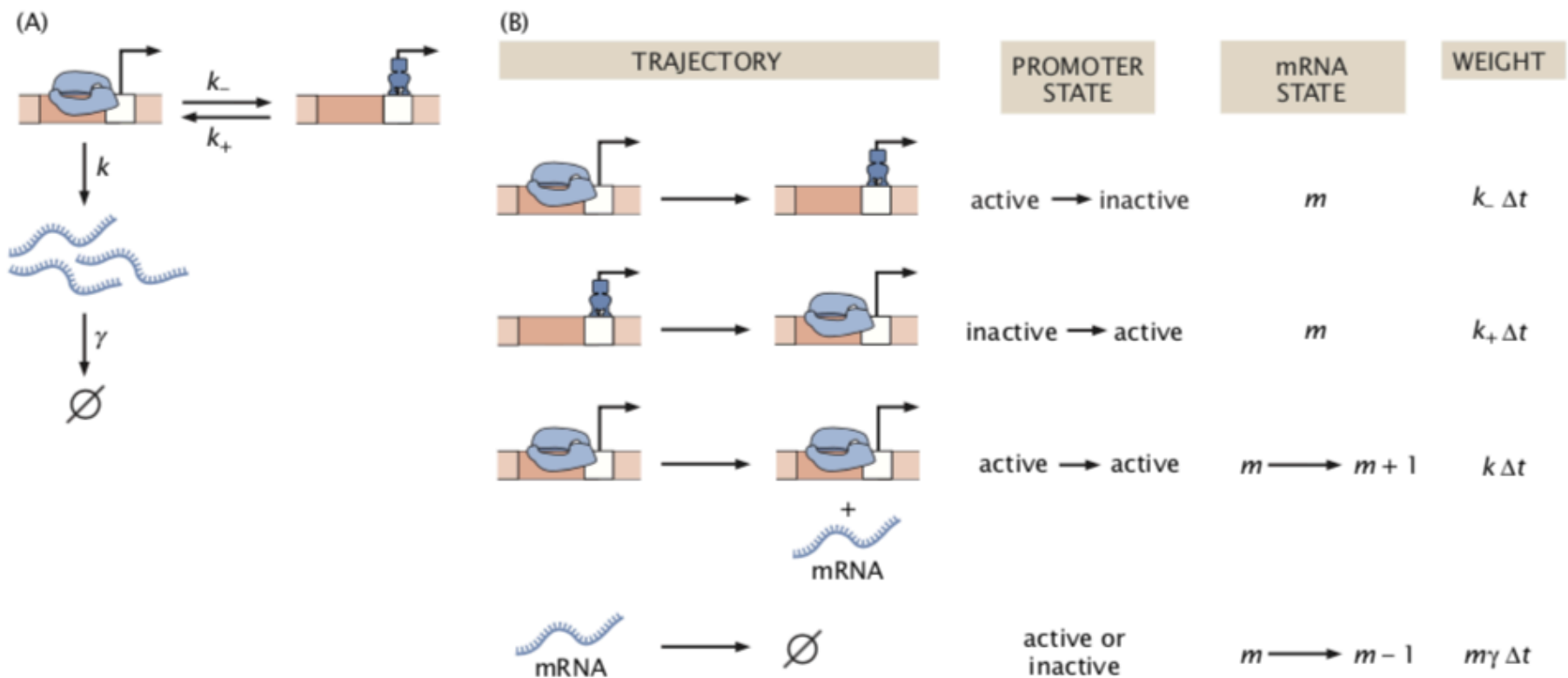
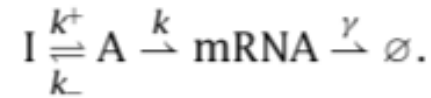


Figure 19.36: Trajectories and weights for the case of simple repression. (A) In an increment of time Δt , the system can suffer several different fates, including switching between the active and inactive states, degradation of an individual mRNA molecule, and production of a mRNA molecule while in the active state. (B) The individual trajectories available to the system in time Δt and their corresponding weights.

En este modelo el promotor puede estar en dos estados: activado, A, e inhibido, I. Las transiciones se pueden representar como:



En este modelo cinético hay que seguir dos variables: el estado del promotor (I o A) y el número de moléculas de ARNm, m . Para ello, se definen las siguientes probabilidades: la probabilidad, $p_I(m, t)$, de tener el promotor en estado inactivo y m moléculas de ARNm; la probabilidad, $p_A(m, t)$, de tener el promotor en estado activo y m moléculas de ARNm.

Las ecuaciones que rigen las variaciones de p_I y p_A son:

$$\begin{aligned} \frac{dp_I(m, t)}{dt} &= \underbrace{-k^+ p_I(m, t)}_{I \rightarrow A} + \underbrace{k^- p_A(m, t)}_{A \rightarrow I} - \underbrace{\gamma m p_I(m, t)}_{m \rightarrow m-1} + \underbrace{\gamma(m+1) p_I(m+1, t)}_{m+1 \rightarrow m}. \\ \frac{dp_A(m, t)}{dt} &= \underbrace{k^+ p_I(m, t)}_{I \rightarrow A} - \underbrace{k^- p_A(m, t)}_{A \rightarrow I} - \underbrace{k p_A(m, t)}_{m \rightarrow m+1} \\ &\quad + \underbrace{k p_A(m-1, t)}_{m-1 \rightarrow m} - \underbrace{\gamma m p_A(m, t)}_{m \rightarrow m-1} + \underbrace{\gamma(m+1) p_A(m+1, t)}_{m+1 \rightarrow m}. \end{aligned} \quad (19.68)$$

Con estos modelos cinéticos se pueden derivar varios resultados analíticamente. En particular resulta $\frac{\sigma^2}{\langle m^1 \rangle} = 1 + \langle m^1 \rangle \frac{k^-}{k^+} \frac{\gamma}{k^+ + k^- + \gamma}$ y usando estimaciones derivadas de experimentos se obtiene:

$$\begin{aligned} \frac{\sigma^2}{\langle m^1 \rangle} &= 1 + 10 \times \frac{0.2 \text{ min}^{-1}}{0.03 \text{ min}^{-1}} \frac{0.014 \text{ min}^{-1}}{0.03 \text{ min}^{-1} + 0.2 \text{ min}^{-1} + 0.014 \text{ min}^{-1}} \\ &\approx 1 + 3.6 = 4.6. \end{aligned} \quad (19.94)$$

Distintas arquitecturas regulatorias tienen distintos factores de Fano. Algunos ejemplos:

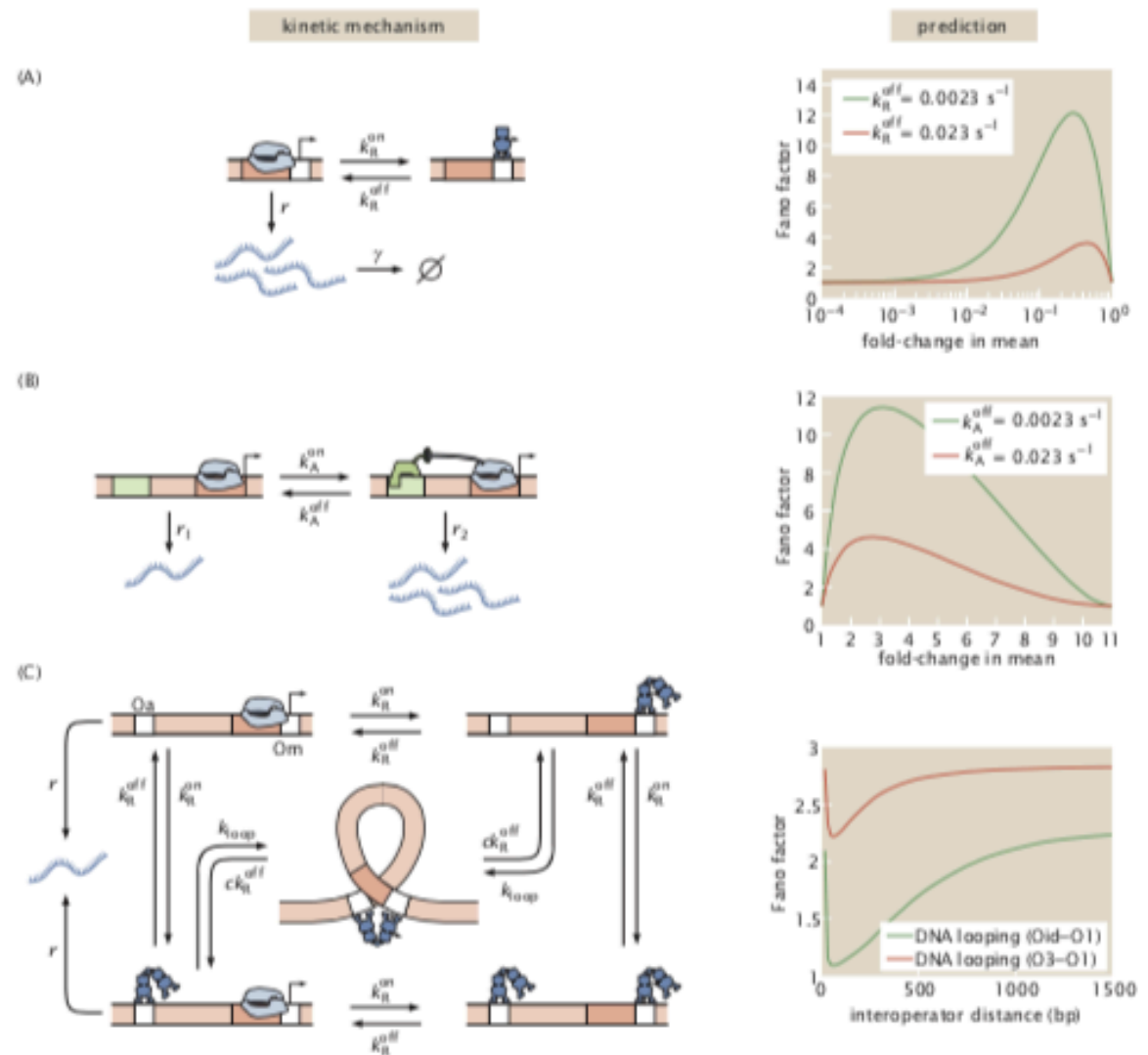


Figure 19.37: Stochastic models of transcriptional regulation for several regulatory architectures. The Fano factor is shown as a function of the fold-change in gene expression for (A) simple repression, (B) simple activation, and (C) repression by DNA looping. The kinetic models assumed for each architecture are shown on the left and their corresponding predictions are shown on the right. For all three figures, we use $r = 0.33 \text{ mRNA s}^{-1}$ and $\gamma = 0.011 \text{ s}^{-1}$. The specific parameters for each figure are (B) $r_1 = r$ and $r_2/r_1 = 11$ and (C) $c = 1$, $k_{\text{loop}} = [J]k_R^0$, with $[J] = (\ln M)e^{-\beta\Delta F_{\text{loop}}}$ the same as given in Figure 19.29 and $k_R^0 = 2.7 \times 10^{-3} \text{ s}^{-1} \text{ nM}^{-1}$, $k_R^{\text{off}}(\text{Oid}) = 1/(7 \text{ min})$, $k_R^{\text{off}}(\text{O1}) = 1/(2.4 \text{ min})$ and $k_R^{\text{off}}(\text{O3}) = 1/(0.47 \text{ s})$. The fold-change in mean gene expression is obtained by varying k_R^{on} and k_A^{on} and is given by (A) $(1 + k_R^{\text{on}}/k_R^{\text{off}})^{-1}$, (B) $[(k_A^{\text{on}}/k_A^{\text{off}}) r_1/r_2 + 1] / (k_A^{\text{on}}/k_A^{\text{off}} + 1)$, and (C) $[c k_R^{\text{off}} (k_R^{\text{off}} + k_R^{\text{on}})] / [k_{\text{loop}} k_R^{\text{on}} + c (k_R^{\text{off}} + k_R^{\text{on}})^2]$. The connection between fold-change in mean gene expression and Fano factor is explored in the problems at the end of the chapter. (Adapted from A. Sanchez et al., *PLoS Comput. Biol.* 7:e1001100, 2011.)

La dinámica y los motivos (o arquitecturas) desde el abordaje que usa ideas de teoría de redes.

Pensemos en un gen que está regulado por un factor de transcripción sin entradas adicionales (o con todas las demás entradas y modos de regulación postranscripcionales mantenidos constantes a lo largo del tiempo). Esta interacción de transcripción se describe en la red mediante $X \rightarrow Y$ que dice "el factor de transcripción X regula el gen de Y".

En ausencia de su señal de entrada, S_X , el factor de transcripción X está inactivo y no se produce Y. Cuando aparece la señal S_X , X pasa rápidamente a su forma activa X^* y se une al promotor del gen Y. El gen Y comienza a transcribirse y el ARNm se traduce, lo que da como resultado la acumulación de proteína Y. La célula produce la proteína Y a un tasa β .

La producción de Y se equilibra mediante dos procesos, la degradación de la proteína (su destrucción específica por proteínas especializadas en la célula) y la dilución (la reducción de la concentración debido al aumento del volumen celular durante el crecimiento). Llamemos α_{deg} y α_{dil} , a las tasas de degradación y dilución respectivamente. Definiendo: $\alpha = \alpha_{dil} + \alpha_{deg}$ podemos escribir la ecuación para Y como:

$$dY/dt = \beta - \alpha Y$$

donde β va a depender de X.

En este ejemplo sencillo la escala temporal de variación de Y (tanto para su crecimiento como para su degradación) es α , o sea, está determinado por la degradación que es independiente de X (de la regulación).

Veamos cómo se modifica si X regula su propia producción. Consideremos:

$$\frac{dX}{dt} = f(X) - \alpha X$$

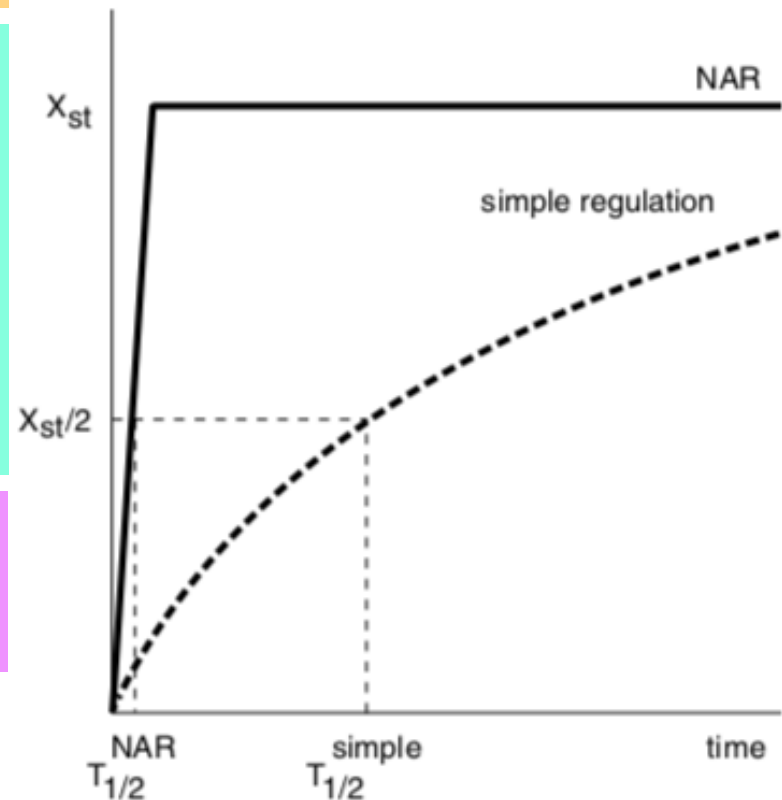
con $f(X)$ una función decreciente de X. Una posibilidad es: $f(X) = \frac{\beta K^n}{K^n + X^n}$

Es posible obtener soluciones numéricas de la ecuación para X con esta f de un modo sencillo, pero no es fácil obtenerlas analíticamente. Para ganar cierta intuición trabajemos con la aproximación “lógica” mencionada anteriormente: $f(X) = \beta$ si $X < K$ y 0 en caso contrario:

$$\frac{dX}{dt} = \beta - \alpha X \quad \text{while} \quad X < K$$

Si se empieza con $X=0$, esto implica $X = \beta/\alpha (1 - \exp(-\alpha t))$ hasta que $X=K$. Si β es grande y α es pequeño, esto ocurre para $K = \beta/\alpha (1 - 1 + \alpha t)$, o sea, $t = K/\beta$. A partir de ahí empieza a decaer, $X = K/(\alpha\beta) \exp(-\alpha t)$ hasta que $X < K$ cuando se vuelve a producir, etc. O sea, hay una oscilación alrededor de K cuya amplitud típicamente va decayendo, de modo que X se fija en $X=K$.

Si comparamos con una ecuación con el mismo estado estacionario, $X=K$, pero con β_{simple} y α_{simple} fijos, es decir: $K = \beta_{\text{simple}} / \alpha_{\text{simple}}$ la dinámica es diferente:



La respuesta acelerada de un circuito con autorregulación negativa en comparación con la regulación simple se demostró experimentalmente utilizando mediciones de expresión génica de alta resolución como se muestra en la figura del libro de U. Alon donde los experimentos están en líneas llenas y la teoría en líneas discontinuas.

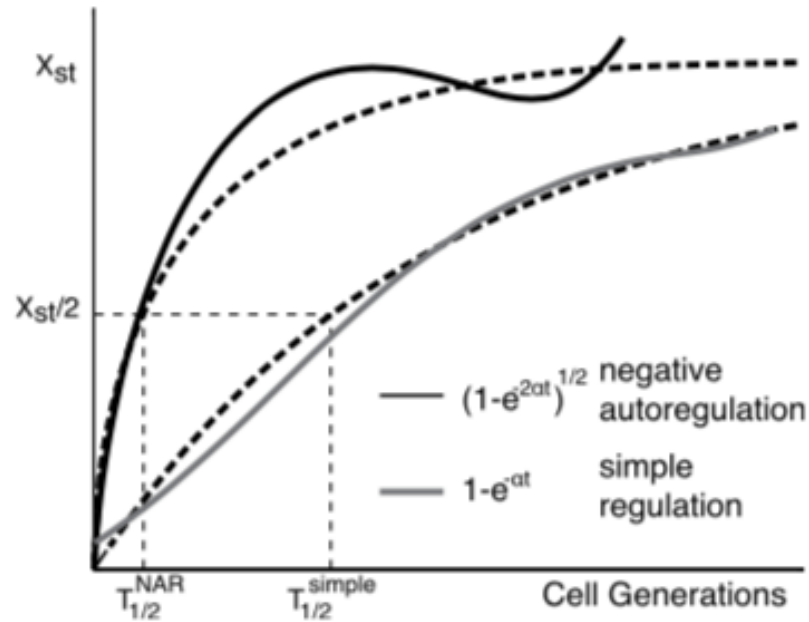
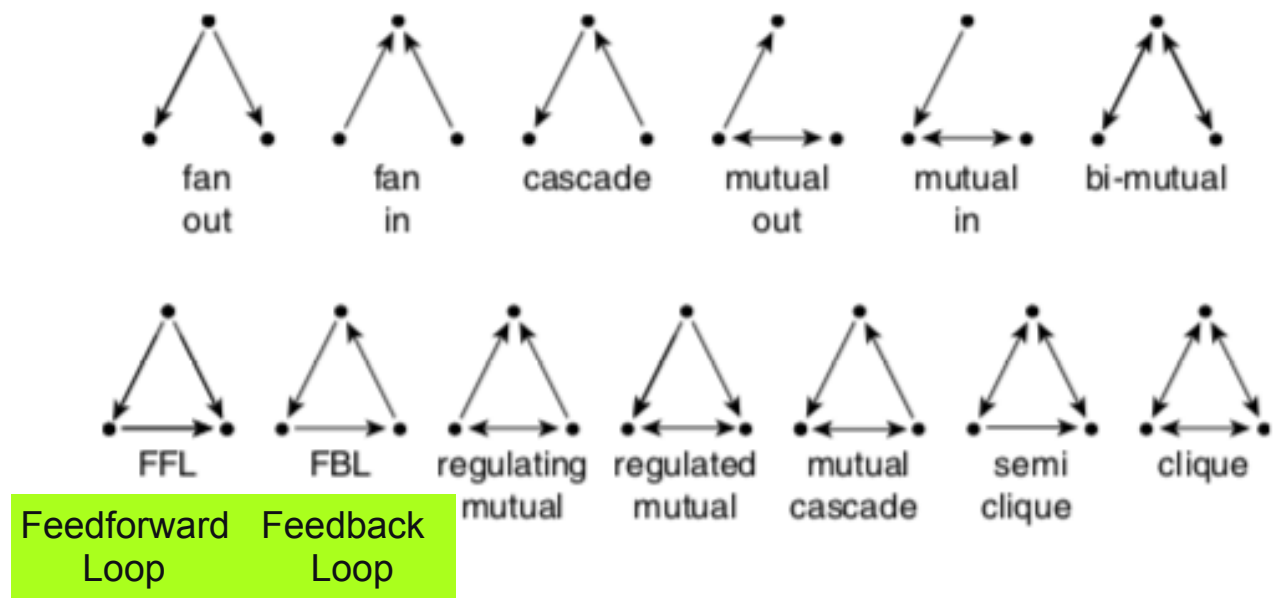


FIGURE 2.6 Adapted from (Rosenfeld, Elowitz and Alon, 2002).

Como ya mencionamos, distintas arquitecturas también resultan en distintos niveles de ruido transcripcional.

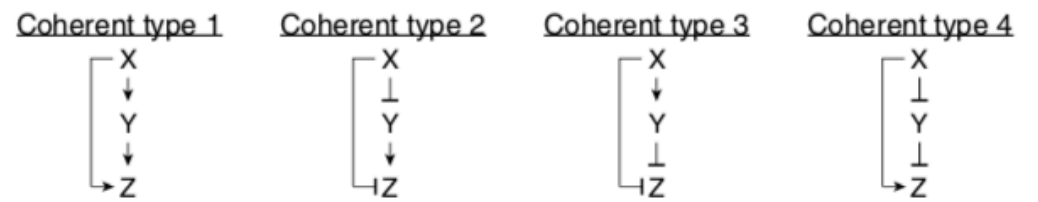
Algunos ejemplos de subgrafos de 3 nodos que pueden aparecer:



En E. coli hay 42 “feedforward loops” y ningún FBL de 3 nodos. En las redes aleatorias con la misma conectividad media $\lambda = 500/400 \sim 1.2$, solo hay alrededor de 2 “feedforward loops”.

Cada una de las tres flechas en el FFL puede corresponder a activación (signo más) o represión (signo menos). Por tanto, hay $2^3 = 8$ tipos posibles de FFL

Coherent FFL



Incoherent FFL

