

El transporte y la lectura precisa de concentraciones

Con cosas tomadas de papers varios: Berg y Purcell, 1977; Kaizu et al, 2015; Gregor et al, 2007a y b; Zhang & Dudko, 2016; R. Metzler, 2019 y algo del libro de Bill Bialek.

El transporte de moléculas es muy importante para muchos procesos, en particular, para que los FTs encuentren los promotores a lo largo del ADN donde ejercer su efecto o para que distintas sustancias "efectoras" se ligen a ciertos receptores y ejerzan su efecto.

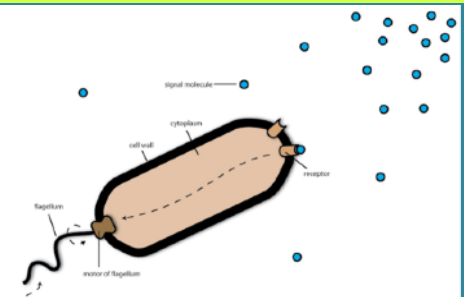
La información que las células reciben, intercambian y sobre la que actúan está muchas veces codificada en la concentración de diferentes sustancias.

En particular, en muchas circunstancias depende de la "amplitud" de dicha concentración. Decodificar la información significa entonces que las células (o componentes celulares) puedan "leer" concentraciones, detectar cambios de concentración y responder de alguna manera a esos cambios.

La concentración se "lee" ligando moléculas de la sustancia que representa la información. Al ligarse en mayor o menor medida a receptores (e.g., en la membrana plasmática), se generan distintos tipos de respuestas.

La cantidad de sitios de unión que participan de esta "lectura", la cantidad de moléculas efectoras presentes y la rapidez con que llegan a la vecindad de los sitios de unión determina la rapidez con la que la célula puede obtener una lectura precisa de la concentración de la sustancia efectora más allá de su inmediata vecindad.

Berg y Purcell escribieron un trabajo muy interesante y pionero hace muchos años en el que estudiaron cómo pueden las bacterias "sensar" la concentración de nutrientes en el medio que las rodea para poder desplazarse hacia la zona donde hay mayor cantidad de ellos.



Bacterial chemotaxis

En el caso de la quimiotaxis, la mayor o menor concentración de sustancias capaces de ligarse a algún receptor en la membrana plasmática va a determinar la fracción de tiempo que los sitios de ligaduras (receptores) estén ocupados.

La fracción de tiempo que están ocupados depende de la concentración, $[P]$, de la sustancia

Berg & Purcell, 1977, scheme for bacterial chemotaxis:

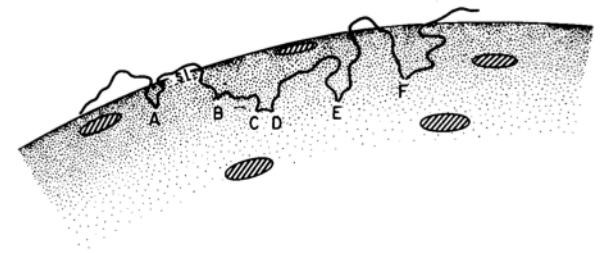


FIGURE 1 The path of a diffusing molecule that has touched the surface of a cell of radius a at a sequence of points A, B, \dots, F . The receptor patches, shown shaded, are of radius s . A and B constitute independent tries at hitting a patch, but C and D do not. Note between A and B the excursion of distance s perpendicular to the surface of the sphere.

¿Con cuánta precisión puede $[P]$ ser leída? ¿Cuál es la mínima diferencia que puede ser detectada?

En este ejemplo la precisión aumenta promediando a lo largo del tiempo. ¿Cuánto tiempo hace falta para alcanzar cierta precisión?

Pasemos un poco al pizarrón

Berg and Purcell (resumen)

PHYSICS OF CHEMORECEPTION

HOWARD C. BERG AND EDWARD M. PURCELL, *Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309 and the Department of Physics, Harvard University, Cambridge, Massachusetts 02138 U.S.A.*

BIOPHYSICAL JOURNAL VOLUME 20 1977

Suppose that one binding site can interact with effector molecules that are contained inside a spherical volume of radius a . The ideal situation would correspond to a case in which the site “senses” all molecules $m(t)$ in this volume at any given time and can compute its average over time.

$$m_T = \frac{1}{T} \int_{t_1}^{t_1+T} m(t) dt.$$

For large enough T , the average becomes a good approximation of the mean (which would correspond to the bulk concentration). The variance of m_T gives an estimate of the distance² between mean and average.

The variance, $\langle m_T^2 \rangle - \langle m \rangle^2$ is related to the autocorrelation function G :

$$G(\tau) = \langle m(t)m(t + \tau) \rangle,$$

by:

$$\langle m_T^2 \rangle = \frac{1}{T^2} \int_0^T dt' \int_0^T G(t' - t) dt,$$

Defining the mean correlation time as:

$$\tau_0 = \int_0^{\infty} u(\tau) d\tau,$$

where $u(\tau)$ is the probability that if a particle is within the sensing volume at time t , it is still inside this volume at time $t+\tau$, we obtain:

$$\langle m_T^2 \rangle = \langle m_T \rangle^2 + (2\tau_0/T) \langle m_T \rangle$$

for T much larger than the mean correlation time.

Now, for a spherical volume of radius, a , and molecules that diffuse with coefficient D it is:

$$\tau_0 = \frac{2}{5} a^2 / D.$$

Then: $\langle \Delta m_T^2 \rangle = (4a^2/5DT) \bar{m},$

The mean, $\langle m \rangle$ and the bulk concentration, c , are related by:

$$c = \langle m \rangle / (4\pi a^3/3)$$

Thus: $\Delta c_{\text{rms}} / \bar{c} = (5\pi TD\bar{c}a/3)^{-1/2}.$

For this “perfect device”, the time it takes to estimate the concentration in the bulk, c , with a certain precision scales as $(a c D)^{-1}$.

But the site generates a response only when bound. So, what matters is the occupancy, $p(t)$, of the site (1 when it's bound and 0 when not).

As before we compute the average, p_T , over an observation time, T , the variance (which is related to the autocorrelation function, G , of p) and relate the concentration in the bulk, c , to $\langle p \rangle$.

$$p_T = (1/T) \int_{t_1}^{t_1+T} p(t) dt, \quad G(\tau) = \langle p(t)p(t+\tau) \rangle, \quad c/c_{1/2} = \langle p \rangle / (1 - \langle p \rangle)$$

Assumptions:

- the site is a circular "sink" of radius s that, if empty, binds an incoming molecule with probability 1 (# of molecules that reach the site per unit time is $4 D s c$, with c the molecules concentration).
- the molecule unbinds from the site with rate $1/\tau_b$.

$$\bar{p}/\tau_b = 4(1 - \bar{p}) D s c$$

Then: $G(\tau) = \bar{p}^2 + \bar{p}(1 - \bar{p}) \exp[-|\tau|/(1 - \bar{p})\tau_b]$, for large enough T &

$$\langle p_T^2 \rangle - \langle p_T \rangle^2 = (2/T)\bar{p}(1 - \bar{p})^2\tau_b$$

So that: $\Delta c_{\text{rms}}/\bar{c} = (\nu/2)^{-1/2}$, with $\nu = 4 D s \bar{c} (1 - \bar{p}) T$.

The time to achieve a certain precision again scales as $(\text{size } c D)^{-1}$.

El trabajo de Berg y Purcell fue revisitado en distintas oportunidades.

976

Biophysical Journal Volume 106 February 2014 976–985

The Berg-Purcell Limit Revisited

Kazunari Kaizu,[†] Wiet de Ronde,[‡] Joris Pajmans,[‡] Koichi Takahashi,[†] Filipe Tostevin,[‡]
and Pieter Rein ten Wolde^{†*}

[†]RIKEN Quantitative Biology Center (QBIC), RIKEN, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan; and [‡]FOM Institute AMOLF, Science Park 104, 1098 XG, Amsterdam, The Netherlands

La interpretación del resultado de Berg y Purcell dada en la intro de este trabajo:

$$\delta c = \frac{dc}{d\bar{n}} \delta \bar{n},$$

$$\frac{\delta n}{\bar{n}} = \sqrt{\frac{(1 - \bar{n})}{2D\sigma cT}},$$

Acá c es la concentración de ligando y n el estado de ocupación de EL receptor ($n=1$, ocupado y $n=0$, vacío). La bacteria infiere c a partir de promediar n . Si promedia durante un tiempo, T , mucho más largo que el tiempo de correlación del estado del receptor, el error relativo con el que se infiere c está dado por:

$$\frac{\delta c}{c} = \sqrt{\frac{2}{4D\sigma c(1 - \bar{n})T}}.$$

$4D\sigma c$ es el flujo de moléculas de ligando que llegan a un receptor circular de radio, σ (lo vimos en las clases de difusión) y $1 - \bar{n}$ es la probabilidad de que el receptor esté libre. Por lo tanto, $4D\sigma c(1 - \bar{n})$ es la tasa efectiva de unión ligando-receptor si cada colisión da lugar a una unión exitosa. Berg y Purcell argumentan que su resultado es válido para reacciones no limitadas por difusión, que después de un encuentro infructuoso receptor-ligando, el ligando chocará rápidamente con el receptor una y otra vez hasta unirse, y estas rondas de encuentros pueden capturarse mediante la renormalización.

El trabajo de Berg y Purcell fue revisitado en distintas oportunidades.

976

Biophysical Journal Volume 106 February 2014 976–985

The Berg-Purcell Limit Revisited

Kazunari Kaizu,[†] Wiet de Ronde,[‡] Joris Pajmans,[‡] Koichi Takahashi,[†] Filipe Tostevin,[‡]
and Pieter Rein ten Wolde^{†*}

[†]RIKEN Quantitative Biology Center (QBIC), RIKEN, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan; and [‡]FOM Institute AMOLF, Science Park 104, 1098 XG, Amsterdam, The Netherlands

La interpretación del resultado de Berg y Purcell dada en la intro de este trabajo:

$$\delta c = \frac{dc}{d\bar{n}} \delta \bar{n},$$

$$\frac{\delta n}{\bar{n}} = \sqrt{\frac{(1 - \bar{n})}{2D\sigma cT}},$$

Acá c es la concentración de ligando y n el estado de ocupación de EL receptor ($n=1$, ocupado y $n=0$, vacío). La bacteria infiere c a partir de promediar n . Si promedia durante un tiempo, T , mucho más largo que el tiempo de correlación del estado del receptor, el error relativo con el que se infiere c está dado por:

$$\frac{\delta c}{c} = \sqrt{\frac{2}{4D\sigma c(1 - \bar{n})T}}.$$

$4D\sigma c$ es el flujo de moléculas de ligando que llegan a un receptor circular de radio, σ (lo vimos en las clases de difusión) y $1 - \bar{n}$ es la probabilidad de que el receptor esté libre. Por lo tanto, $4D\sigma c(1 - \bar{n})$ es la tasa efectiva de unión ligando-receptor si cada colisión da lugar a una unión exitosa. Berg y Purcell argumentan que su resultado es válido para reacciones no limitadas por difusión, que después de un encuentro infructuoso receptor-ligando, el ligando chocará rápidamente con el receptor una y otra vez hasta unirse, y estas rondas de encuentros pueden capturarse mediante la renormalización.

Sigue la intro de Kaizu et al, 2014:

Bialek y Setayeshgar (PNAS2005) generalizaron el resultado de Berg y Purcell incluyendo la dinámica de unión ligando-receptor. Consideraron un modelo en el que las moléculas de ligando pueden difundir, unirse al receptor al estar en su vecindad con una tasa de asociación k_a y despegarse con una tasa, k_d . Invocando el teorema de fluctuación-disipación, linealizaron la ecuación de reacción-difusión no lineal, y obtuvieron:

$$\frac{\delta c}{c} = \sqrt{\frac{1}{\pi D \sigma c T} + \frac{2}{k_a c (1 - \bar{n}) T}}$$

en lugar de:

$$\frac{\delta c}{c} = \sqrt{\frac{2}{4 D \sigma c (1 - \bar{n}) T}}$$

El primer término es debido a la llegada estocástica de las moléculas de ligando al receptor por difusión, mientras que el segundo se debe a la estocasticidad intrínseca de la cinética de unión del receptor.

Si la reacción está (muy) limitada por difusión, $k_a \rightarrow \infty$ ambas expresiones deberían coincidir (a menos de factores geométricos). Sin embargo, las expresiones difieren en un factor $1-n$. La expresión de Bialek-Setayeshgar predice que la incertidumbre permanece finita si $n \rightarrow 1$ mientras que la de Berg-Purcell diverge en este límite.

Kaizu et al rederivan este error relativo y obtienen:

$$\frac{\delta c}{c} = \sqrt{\frac{1}{2\pi\sigma Dc(1-\bar{n})T} + \frac{2}{k_a c(1-\bar{n})T}}$$

Esta expresión se puede obtener trabajando “a la Bialek”, partiendo de otra versión linealizada de las ecuaciones de reacción-difusión.

Más allá de la discrepancia sobre el 1-n, en ambas expresiones el error relativo decae en el tiempo como $T^{-1/2}$. Este decaimiento es característico de estimaciones que provienen de promediados temporales. Vayamos al pizarrón.

Supongamos una secuencia temporal de una variable aleatoria que corresponde al número de partículas de una especie, s

$$\{N_\ell^{(s)} \equiv N^{(s)}(t_\ell)\}_{\ell=0}^{n-1} \quad t_\ell = \ell dt$$

Supongamos que dt es suficientemente “grande” como para que $N_i(t+dt)$ y $N_i(t)$ no estén correlacionadas ($dt \sim \tau_{\text{corr}}$). El promedio temporal

$$\bar{N}^{(s)}(T_{\text{obs}}) = \frac{\sum_i N_\ell^{(s)}}{n}$$

es una mejor aproximación del valor medio a medida que se promedia a lo largo de un tiempo más largo. Por otro lado, suponemos que el valor medio está relacionado con la concentración:

$$\langle N^{(s)} \rangle = c^{(s)} V. \quad \text{Éste es el valor “verdadero” que se quiere estimar.}$$

El error relativo de estimar la media con el promedio sobre un tiempo de observación, T_{obs} es:

$$\Delta_r(\bar{N}^{(s)}) \equiv \frac{\text{var}(\bar{N}^{(s)}(T_{\text{obs}}))^{1/2}}{\langle N^{(s)} \rangle}$$

La varianza de la suma de n variables aleatorias equidistribuidas, independientes entre sí, es n veces la varianza de cada una de ellas. Por lo tanto, la varianza del promedio es

$$\text{var}(\bar{N}^{(s)}(T_{\text{obs}})) = \frac{\text{var}(N^{(s)}(t))}{n}$$

Pero $n = T_{\text{obs}}/dt$, por lo tanto, el error relativo del promedio decae como $T_{\text{obs}}^{-1/2}$.

Esta cuenta se puede hacer también sin recurrir a la independencia de las variables aleatorias y da, a tiempos largos, esa dependencia.

Entre otras cosas este cálculo está diciendo: si bien un receptor va pegando moléculas de ligando de a una y pega solo a aquéllas que llegan a su vecindad, al hacerlo a lo largo del tiempo (mientras la concentración de ligando permanezca constante) promedia esa información y puede sacar una buena estimación de la concentración del ligando en el “bulk” (es decir, fuera de la vecindad del receptor), al menos después de un tiempo.

En el caso de bacterias y nutrientes cuantificar concentraciones correctamente no es tan importante. Solo importa saber dónde hay más nutriente! Pero en morfogénesis es MUY relevante!

In this case there are substances (“morphogens”) that are distributed non-uniformly inside the embryo. the fate of otherwise undifferentiated cells depending on their position in the embryo. To this end, morphogen concentrations need to be “read”.

This uneven distribution determines the fate of otherwise undifferentiated cells depending on their position.

Being able to “read” the morphogen’s concentration accurately allows the “faithful” determination of the position in the embryo.

With what accuracy can position be determined from reading the concentration of a “morphogen”? In what time?

A paradigmatic example that is (still) being studied in great detail: *Drosophila* development (como ya vimos antes!)

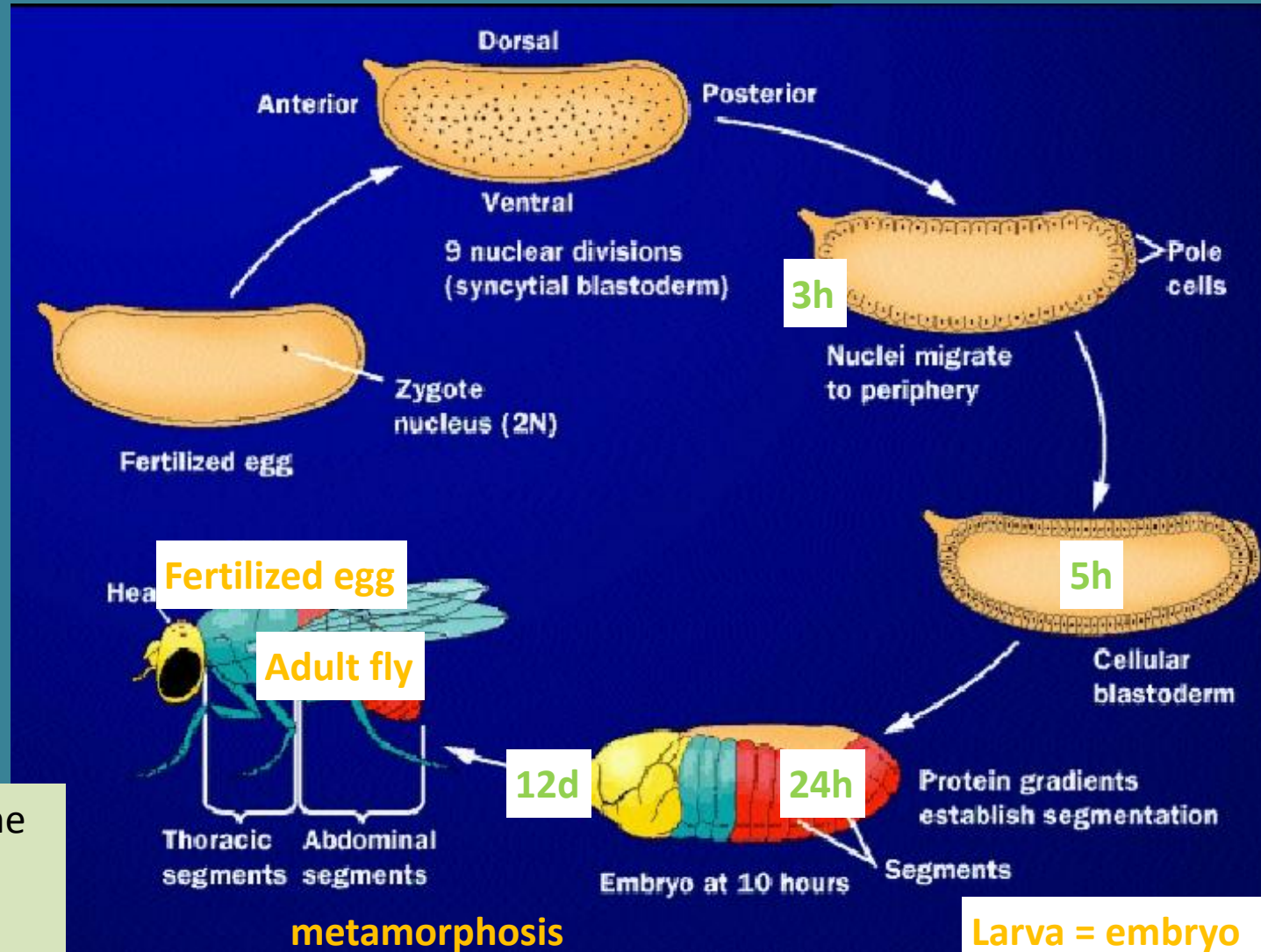
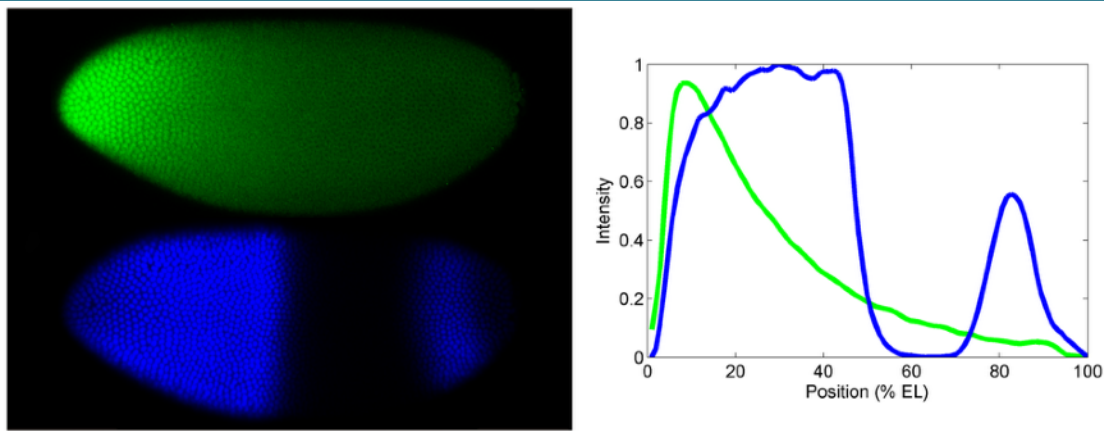


Fig. by K. Nowick, Trome and Transcriptome Evolution. Leipzig

Particular attention has been paid to the establishment of the anterior-posterior axis by Maternal Effect Genes during early development

Here we focus on the morphogen (transcription factor) Bicoid.



Concentration profiles of the proteins Bicoid (Bcd) and Hunchback (Hb), Holloway et al, PLoS Biol 2011

A cascade of events:

Bicoid is unevenly distributed along the embryo

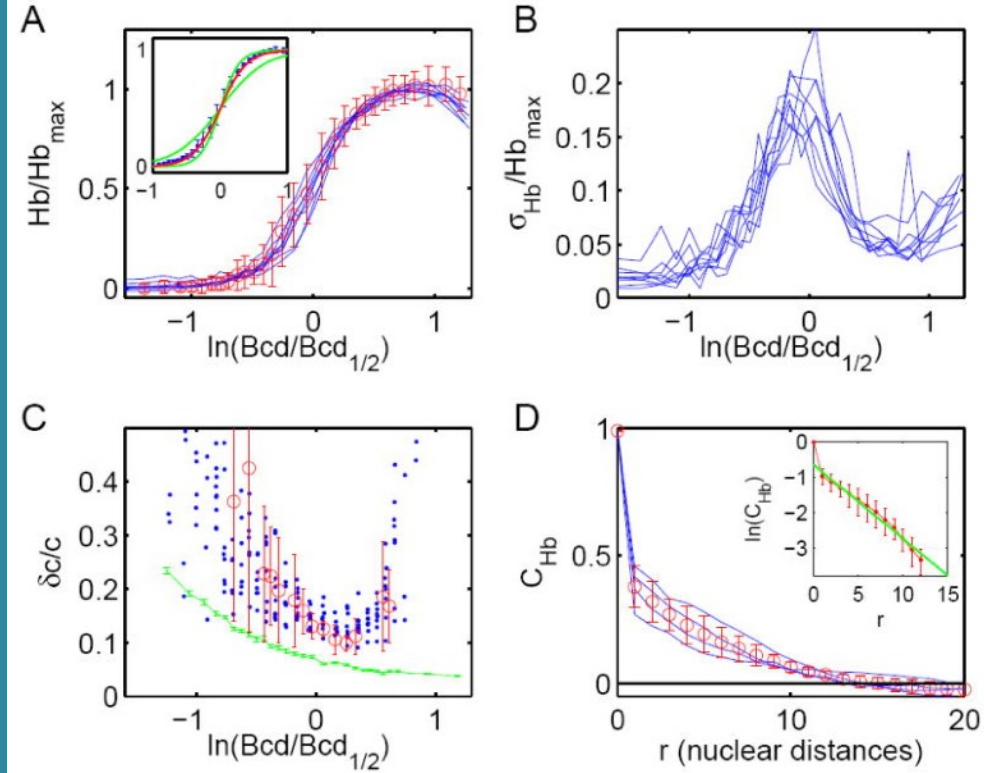
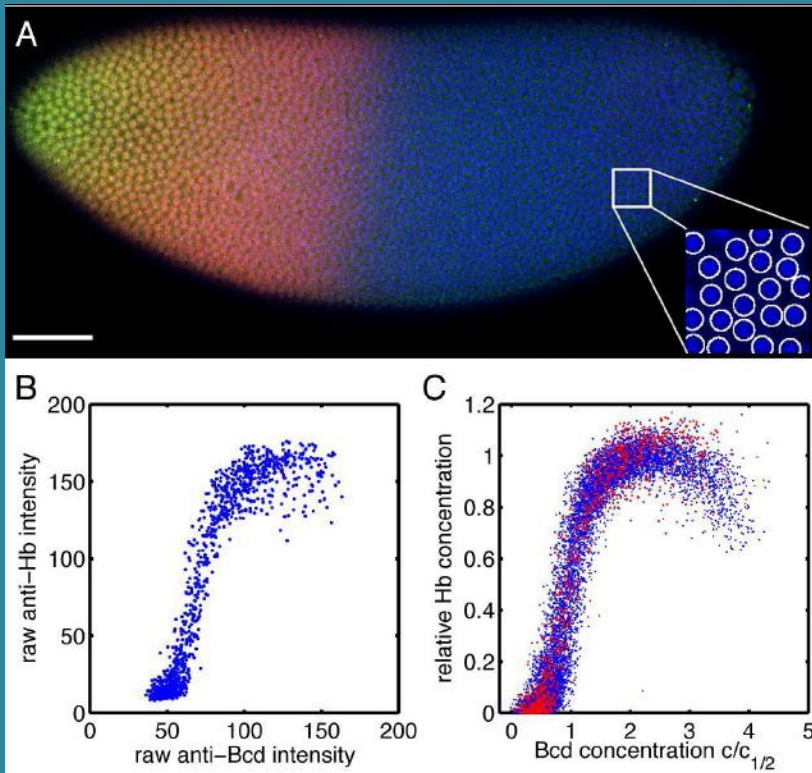
Bicoid regulates *hunchback* transcription

The transcription of *hb* depends on [Bcd], i.e., it does not occur at a uniform pace along the embryo.

The spatial changes in the transcription of *hunchback* translate into changes in downstream genes.

This cascade of events generates a very precise spatial pattern where neighboring nuclei have readily distinguishable levels of expression for several genes. These patterns are reproducible from embryo to embryo.

Quantifying the [Bcd], [Hb] relationship (Gregor et al, 2007).



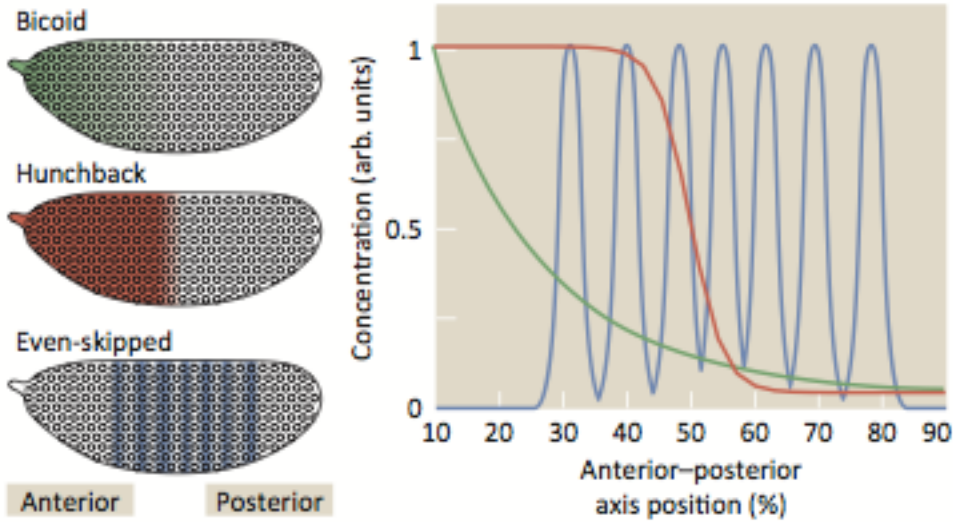
[Hb] vs. [Bcd] from fixed and stained embryos. **A:** Drosophila embryo in early nuclear cycle 14, stained for DNA (blue), Hb (red) and Bcd (green); scale bar 50 μm. **B:** Hb vs Bcd immunofluorescent staining levels from 1299 nuclei of a single embryo. **C:** [Hb] vs [Bcd] from 13,366 nuclei in 9 embryos. **Observe the sharp change in [Hb] at a particular [Bcd] along its gradient.**

A: Input/output relations between Bcd and Hb from 9 embryos. Fitted well by with $n=5$.

$$Hb = Hb_{max} \frac{Bcd^n}{Bcd^n + Bcd_{1/2}^n}$$

B: standard deviation in Hb levels as a function of [Bcd]. Output fluctuations are below 10% when [Bcd] is high.

C: Relative errors in [Bcd] that lead to the observed deviations in [Hb] (minimum value ~ 0.1). Green curve: detection error.

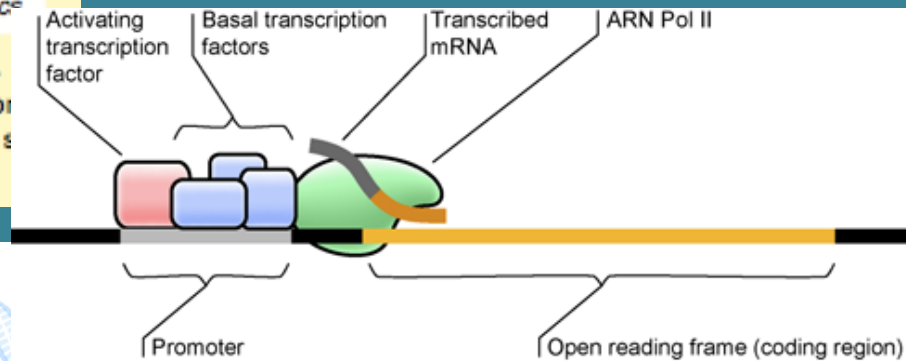


TRENDS in Genetics

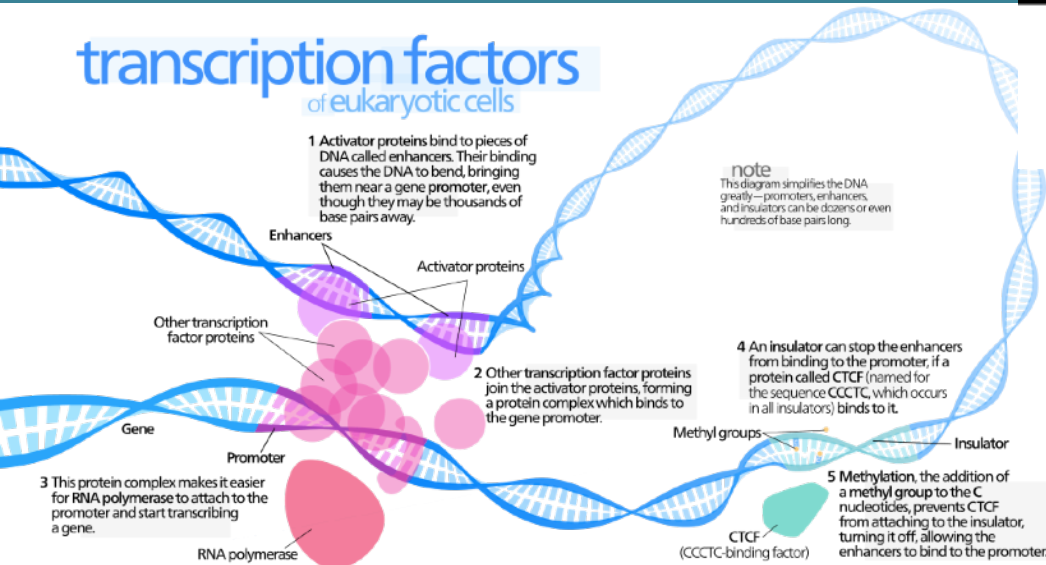
Figure 1. Hierarchy of patterning genes in the early *Drosophila* embryo. maternal factor Bicoid activates *hunchback* and various other gap genes in broad domains, which all work together to regulate the activity of pair-rule genes such as *even-skipped*. Gregor et al, 2014

But, who “reads” the concentration of Bcd to produce Hb?

Individual binding sites along DNA that can sense Bcd molecules within a very small volume.



transcription factors of eukaryotic cells



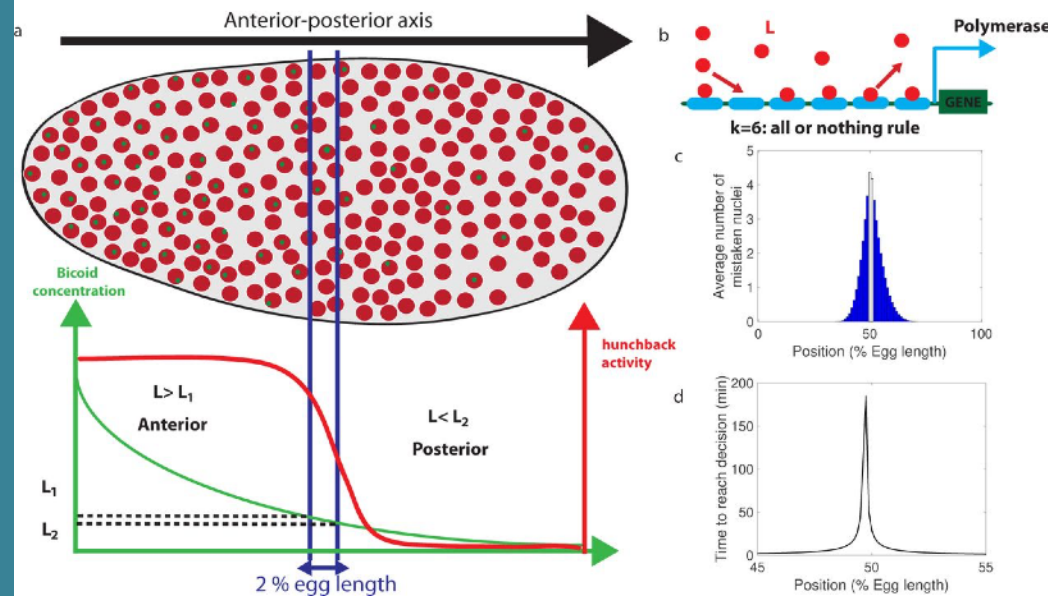
How long does it take for each of the binding sites to “detect” the bulk concentration? (B. Bialek & others have worked on this)

From J Desponds et al, BioRxiv, 2019:

Hunchback spatial profiles are sharp and the variance in hunchback expression of nuclei at similar positions along the AP axis is small.

Mitosis resets the binding of transcription factors (TF) during the phase of synchronous divisions, suggesting that the decision about the nuclei's position is made by using information accessible within one nuclear cycle.

Experiments show that during the nuclear cycles 10-13 the positional information encoded by the Bicoid gradient is read out by hunchback promoters precisely and **within 3 minutes**.



a. Projection on one plane of the embryo at nuclear cycles 10-13, when nuclei (red dots) have migrated to the surface of the embryo. The green dots represent active transcription loci of the *hunchback* gene. *hunchback* undergoes a sharp drop around the boundary region (between the blue lines) which width δx is 2% of the egg length.

The *hunchback* readout defines the cell fate whether each nucleus will follow an anterior or the posterior gene expression program. c. The average number of nuclei making a mistake in the decision process as a function of the egg length position at cycle 11. For a fixed-time decision process completed within the interphase duration during cycle 11, $T = 270s$.

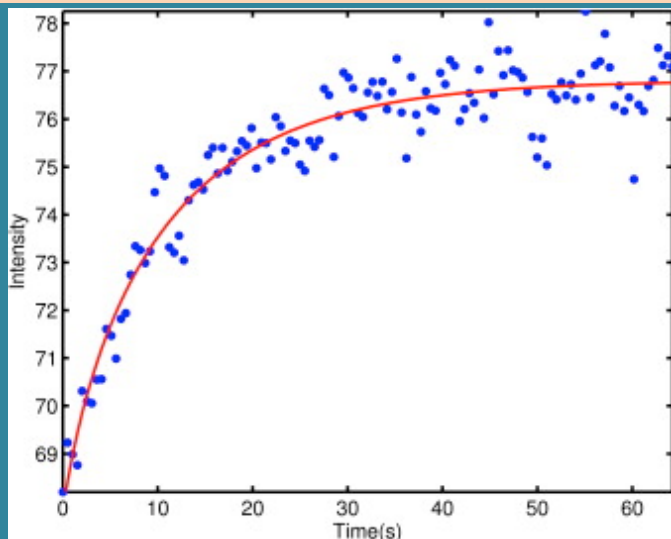
Let's go back to the work of Berg and Purcell.

The formula of their perfect device was used by Gregor et al to determine the time, T , it takes for a binding site of length a to estimate $c=[Bcd]$ (at each nucleus) with a certain (e.g. 10%) error:

$$\frac{\delta c}{c} \sim \frac{1}{\sqrt{DacT}}$$

Setting $\delta c/c=0.1$, with D the diffusion coefficient of Bicoid they could obtain T . For that, they needed an estimate of D .

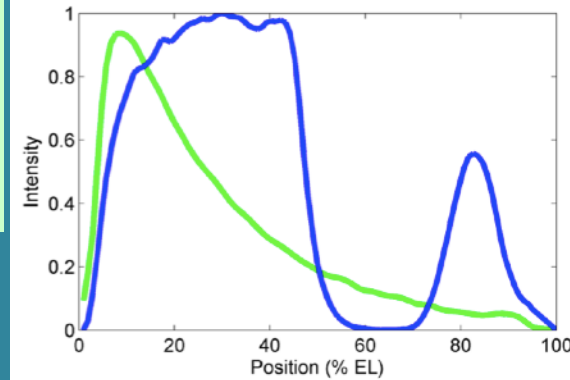
Gregor et al, (Cell, 2007), also did Fluorescence Recovery After Photobleaching (FRAP) experiments to determine D .



They obtained: $D = 0.27 \pm 0.07 \mu\text{m}^2/\text{s}$, way too small!

Doing this for the transition region where Hb switches from high to low concentration and $[Bcd]=c= 6/\mu\text{m}^3$,

Green: Bcd, Blue: Hb



using $a=3\text{nm}$ and $D = 0.27\mu\text{m}^2/\text{s}$, they determined that a ratio $\delta c/c \sim ([Bcd]aDT)^{-1/2} = 0.1$ is achieved only after $T \sim 340\text{min}!!$

Solution proposed: spatial averaging of mRNA

But spatial averaging will wipe out the transition zone.

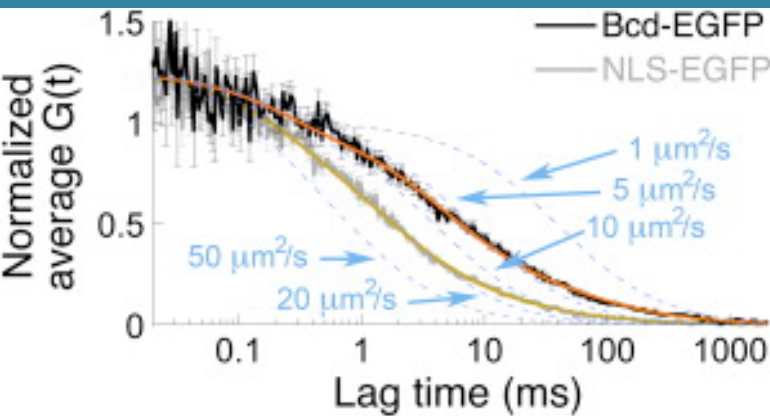
Various things to challenge about the calculation:

Diffusion coefficient, D (Abu-Arish et al, 2010; Sigaut et al, 2014)

What is the correct asymptotic expression for the relative error?
(Bialek & others 2005 vs Kaizu et al 2013)

Asymptopia (Perez Ipina & SPD, PLoS ONE, 2016).

Abu-Arish et al, 2010 did FCS experiments to determine the diffusion coefficient of Bcd.



Normalized ACF's for **Bcd-EGFP** and **NLS-EGFP** (EGFP that enters the nuclei) in the cortical cytoplasm during the interphase of cycles 12-14. Blue: Expected functions in the case of a single diffusing species for different values of D .

In the case of NLS-EGFP they determined 1 single diffusing component with $D \sim 20 \mu\text{m}^2/\text{s}$

In the case of Bcd-EGFP they determined various diffusing components with coefficients: 14, 1.6, $0.095 \mu\text{m}^2/\text{s}$

They averaged over these values (with the weights derived from the fitting) and obtained:

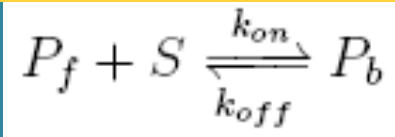
$$D_{Bcd}^{FCS} \approx [7 - 9] \mu\text{m}^2 / \text{s}$$

Repeating the previous calculation with this D , the ratio $\delta c / c \sim ([Bcd]aDT)^{-1/2} = 0.1$ is achieved after $T \sim 10-13$ min. Still off.

Our interpretation of the results obtained on the diffusion of Bicoid using FRAP and FCS (Pando et al PNAS06, Sigaut et al PRE10, PLoS Comp Biol 2014)

Bicoid binds to numerous sites. Thus, it does not diffuse freely.

Simplest model: Free Bicoid, P_f , diffuses freely with coefficient D_f & binds to slowly moving or immobile binding sites, S , becoming P_b (which diffuses at the same rate as S , D_s):



$$K_D = k_{off}/k_{on}$$

$$S_{eq} = \frac{K_D S_T}{K_D + P_{feq}}$$

All concentrations at equilibrium, S_T : total S concentration

We identified two (main) types of effective diffusion coefficients (Pando06):

$$D^t = \frac{D_f + \frac{S_{eq}}{K_D} D_s}{1 + \frac{S_{eq}}{K_D}},$$

$$D^u = \frac{D_f + \frac{S_{eq}^2}{S_T k_D} D_s}{1 + \frac{S_{eq}^2}{S_T k_D}},$$

FRAP (applied to Bcd-EGFP)
Gives: D^t

FCS, applied to Bcd-EGFP,
gives: D^u , D^t , and D_s (Sigaut2010);

Applied to NLS-EGFP, gives D_f

$$D_{FRAP} = D^t = 0.3-1; D_{FCS}(Bcd) = 14, 1.6, 0.095; D_{FCS}(NLS) = 19 \text{ (um}^2/\text{s)}$$

Everything is compatible if: $D_f = 20, D_u = 14, D_t = 1.6, D_s = 0.095 \text{ (um}^2/\text{s)}$

Now, what diffusion coefficient should we use to estimate the time it takes for $\delta c/c \sim 0.1$? (remember that $\delta c/c \sim ([Bcd]aDT)^{-1/2}$)

We expected it would be the collective diffusion coefficient, $D^u \sim 14 \mu\text{m}^2/\text{s}$, but our studies showed that if the binding that slows down the diffusion is with the sites that regulate the expression of hb, then it is $D_f \sim 20 \mu\text{m}^2/\text{s}$.

The Berg-Purcell expression gives: $T \sim 4.6\text{-}6\text{min}$ depending on whether we use D_u or D_f . Much better than before, but not “perfect”.

The expression used holds in the asymptotic limit. How good is the asymptotic expression? Can we do better?

Resumen sobre lo de Bicoid

The information that cells and organisms need to “make decisions” is many times encoded in the concentration of “ligands”.

These concentrations are read by binding sites that have access to a very small volume and the system not always averages over many binding sites, the “reading” process is very fluctuating.

The information is transmitted reliably because noise is eventually reduced due to (time) averaging.

In the case of Bcd, a transcription factor that regulates the expression of the gene *hb* and which is fundamental for the establishment of the anteroposterior axis in flies, how fast the initial noise is reduced also depends on how fast Bcd diffuses.

Interpreting correctly the experiments with which diffusion coefficients are estimated in environments where the molecules of interest diffuse and bind to sites is key for quantification purposes.

Understanding the nature of the fluctuations is key to determine the time within which they decay below a certain level.

Este tipo de cálculo, sobre el tiempo que les lleva a las partículas encontrar los sitios a los que se tienen que ligar son ejemplos de cálculos de “first passage”

Supongamos una partícula que difunde en una dimensión espacial, que parte de $x_0 > 0$ al tiempo $t = 0$, se mueve en $x > 0$ y que es absorbida al tocar $x=0$. Resolviendo la ecuación de evolución de la densidad de probabilidad con esta condición de contorno da en realidad la probabilidad de supervivencia y menos su derivada es la densidad del “first passage time” (o sea, el tiempo que hay que esperar para que una partícula que partió de x_0 y hace un paseo al azar llegue a $x=0$).

Los trabajos detallados en este tema involucran mucha matemática (subí algunos papers como algunos de Ralf Metzler u otro de Benichou y Voituriez, pero tal vez son demasiado sofisticados). Hay un trabajo, sin embargo, que puede ser interesante para contar (Zhang y Dudko) ya que conecta muy directamente con la biología.

First-Passage Processes in the Genome

Yaojun Zhang^{1,2} and Olga K. Dudko¹

¹Department of Physics, University of California at San Diego, La Jolla, California 92093;
email: dudko@physics.ucsd.edu

²Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544

Este trabajo menciona que muchos procesos esenciales en biología comparten un paso fundamental común: establecer contacto físico entre segmentos distantes de ADN. La rapidez con que se logra este paso establece el "límite de velocidad" para los procesos a gran escala que pueden ocurrir. Ejemplos de este tipo de procesos pueden ser la producción de anticuerpos por parte del sistema inmunológico o la diferenciación de tejidos en un embrión en desarrollo.

El paper pregunta, entonces: ¿Cuánto tiempo tardan los segmentos de ADN que están repartidos sobre millones de pares de bases a lo largo de la fibra de cromatina para encontrarse entre sí en la célula que, además está llena de cosas?

Esta pregunta, fundamental para la biología, puede reconocerse como el problema de la física del tiempo de primer paso, o el tiempo de espera para el primer encuentro.

El paper revisa, entonces, una serie de enfoques para revelar los principios físicos mediante los cuales las células resuelven, con asombrosa eficiencia, el problema del primer paso para las interacciones genómicas remotas.

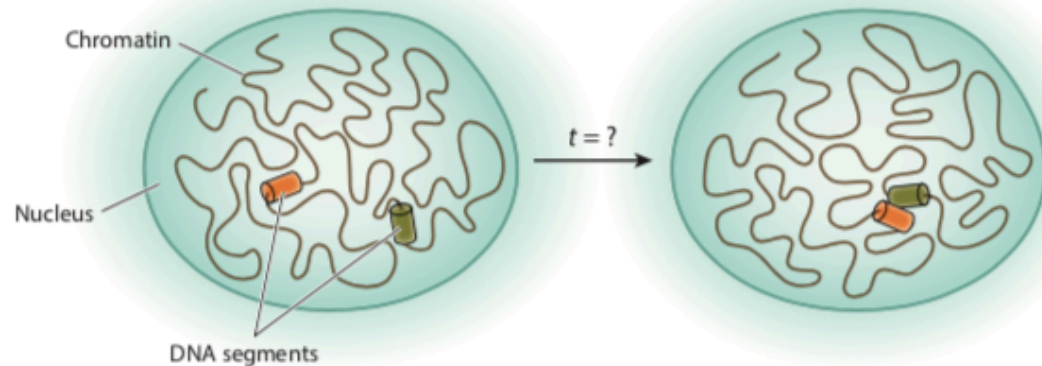


Figure 1

A genomic event in which distant parts of chromosomal DNA have to find each other—and have to do so in a timely fashion—is a precursor to a wide range of essential processes of life. This review discusses a variety of approaches to revealing the physical principles that enable cells to efficiently solve the first-passage time problem for its genome.

Les doy alguna idea del contenido del paper mirando sus figuras, pero estaría bueno que alguien contara al menos algunas cosas con un poco más de detalle.

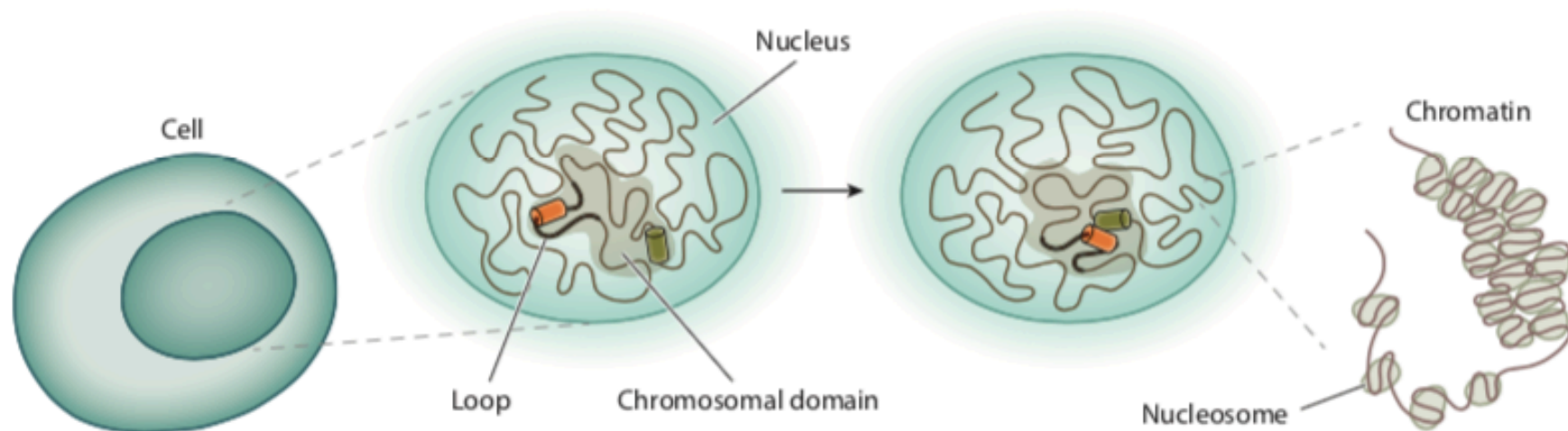


Figure 2

Motion and interactions of genomic segments are strongly affected by the genome architecture. Interphase chromatin is folded into loops, which are further compacted into chromosomal domains. Genomic interactions mostly occur within the same domain.

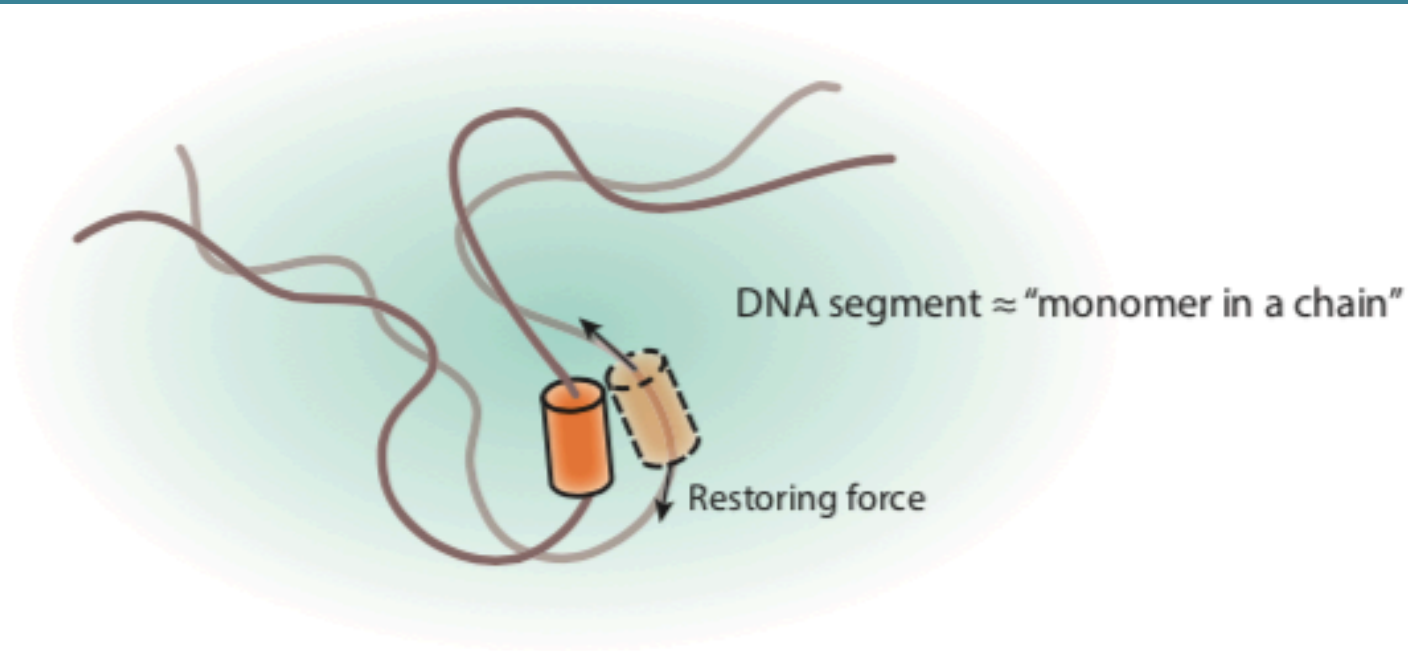


Figure 3

Unlike the case of a free particle, any movement of a DNA segment results in restoring forces on this segment exerted by the neighboring segments of the chain.

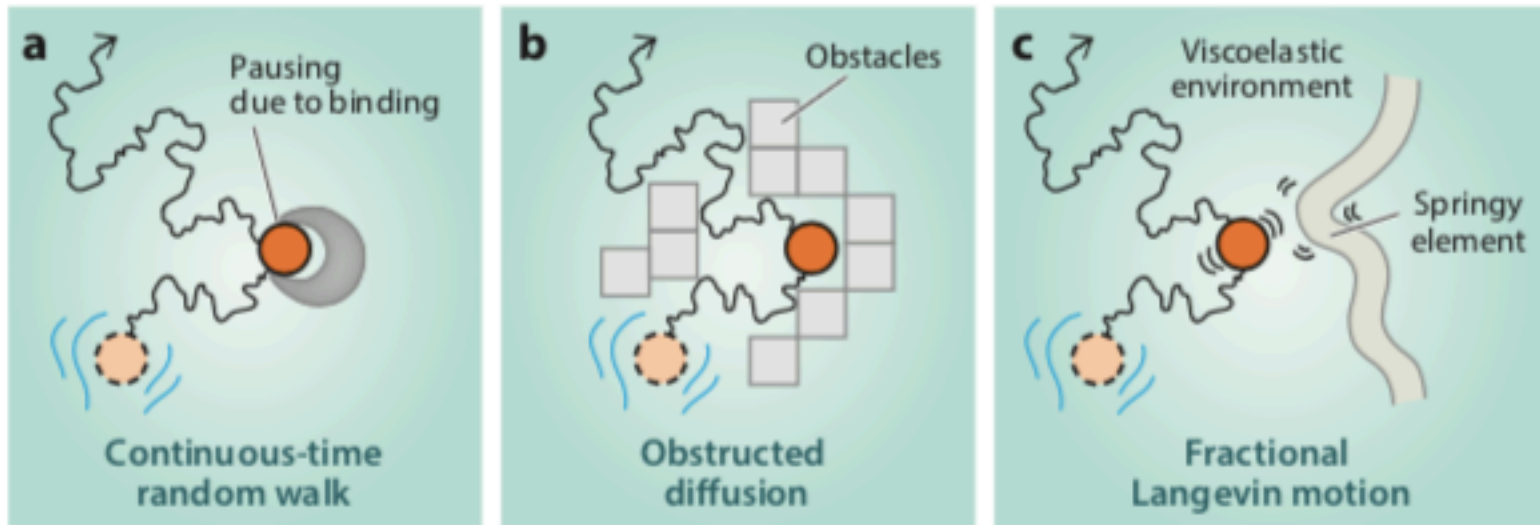
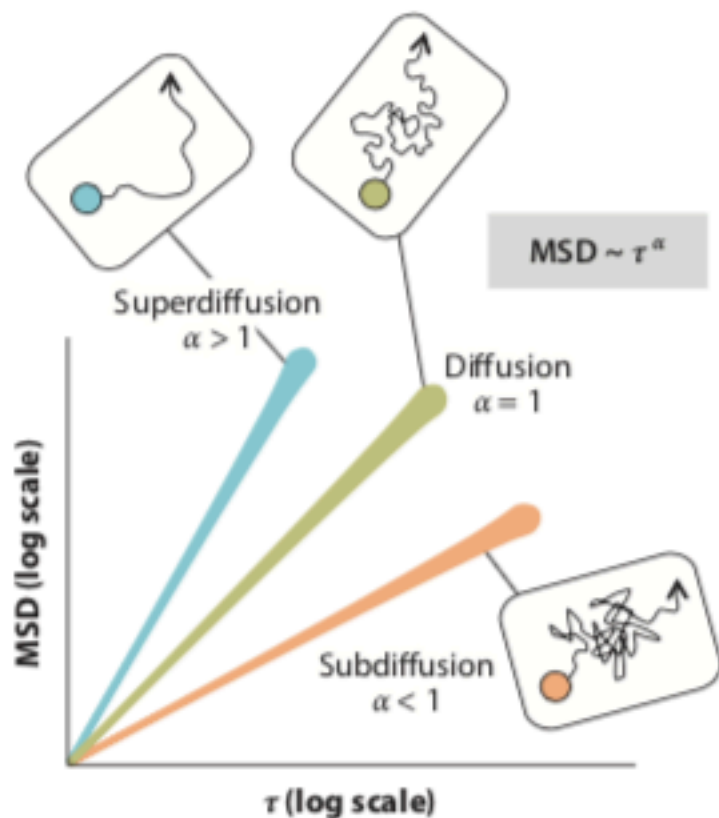


Figure 5

Diffusive motion of a particle in different types of the crowded environment. (a) The continuous-time random walk model describes diffusive motion interspersed by pauses resulting from binding events. (b) The obstructed diffusion model describes the motion of a particle in a high concentration of hard-wall-like obstacles. (c) The fractional Langevin motion model describes the thermal motion of a particle that bounces back and forth against a springy network.

No vimos nada sobre el caso c), se trata del caso en que en la solución hay macromoléculas semiflexibles. Una partícula que se mueve a través de un entorno de este tipo experimenta efectos tanto viscosos como elásticos, de modo que no solo difunde, sino que también rebota hacia adelante y hacia atrás contra la red macromolecular elástica (Figura 5c) y eso cambia su tasa de transporte.

a Mean squared displacement



b Velocity autocorrelation function

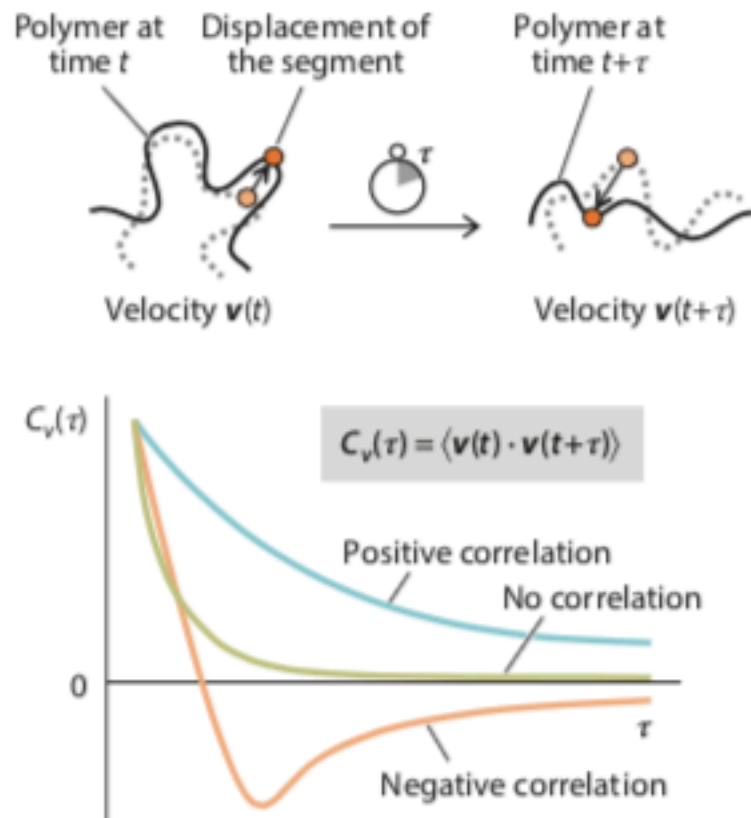


Figure 6

Quantitative tools for establishing the mechanism of motion from trajectories: (a) mean squared displacement (MSD) and (b) velocity autocorrelation function (VAF). MSD is a measure of the extent of excursions of the random walker during time τ . For many systems, MSD scales with the elapsed time as $\text{MSD} \sim \tau^\alpha$, resulting in a linear graph of slope α on a log-log plot. The scaling exponent α determines the type of diffusion (normal diffusion versus anomalous diffusion). The VAF, $C_v(\tau)$, indicates the degree to which the velocities of the particle are correlated when separated by time τ . Normal diffusion exhibits no correlations, whereas subdiffusion and superdiffusion lead to negative and positive correlations, respectively.

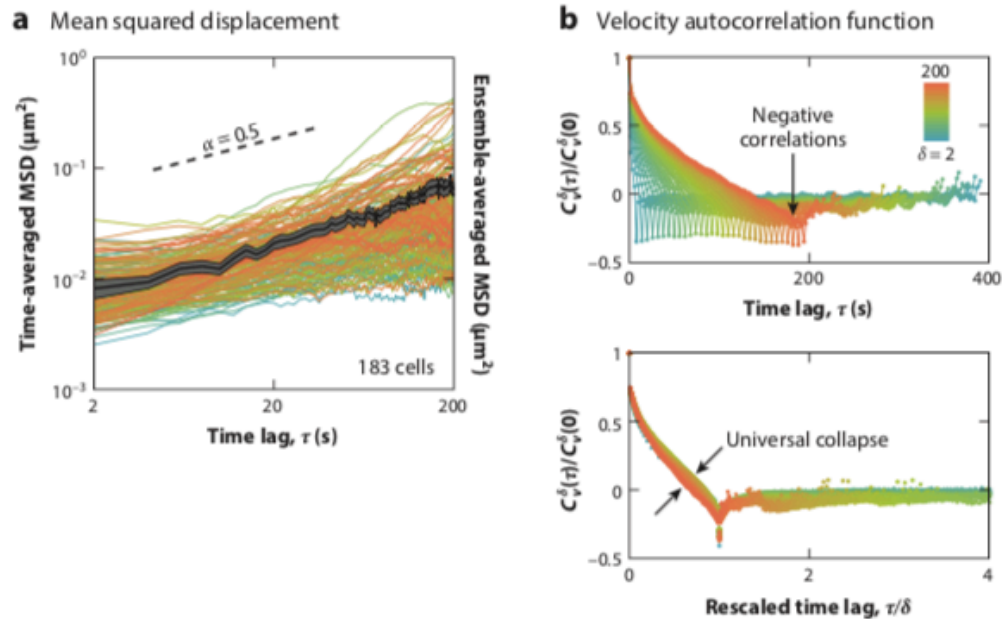


Figure 7

Mean squared displacement (MSD) and velocity autocorrelation function (VAF) as diagnostic tools for DNA motion: an illustration using real DNA trajectories. The data are for D_{H1H} segments in recombination-activating gene-deficient pro-B cells (32). (a) Time-averaged MSD (colored curves) and ensemble-averaged MSD (black curve; shadowed region is the standard error of the mean) as functions of the time lag τ . MSD is calculated from the individual DNA trajectories recorded in 183 cells. On the log-log scale, the slope of MSD yields the scaling exponent α , whereas the y intercept is proportional to the anomalous diffusion coefficient D . Both time- and ensemble-averaged MSD yield similar values of α , a sign of ergodicity. $\alpha \approx 0.5$ indicates that the motion is subdiffusive. (b, top) VAF curves for different values of the discretization intervals (δ), plotted as a function of the time lag τ . A dip into the negative values indicates the presence of negative correlation. (b, bottom) VAF curves as a function of the rescaled time lag (τ/δ). A universal collapse is an indication of self-similarity of motion.

Esta información se usa para inferir el tipo de movimiento dominante en el ADN (contrastando con distintos modelos posibles sobre movimientos de polímeros). A partir del análisis concluyen que el mecanismo viscoelástico es el que gobierna el movimiento de los segmentos de ADN en el abarrotado interior del núcleo.

Después ya entra en los cálculos de first passage.

Busca responder, dado que sabe cuál es el tipo de movimiento dominante, ¿cuánto tiempo tardan dos segmentos distantes de ADN en encontrarse en el interior nuclear?

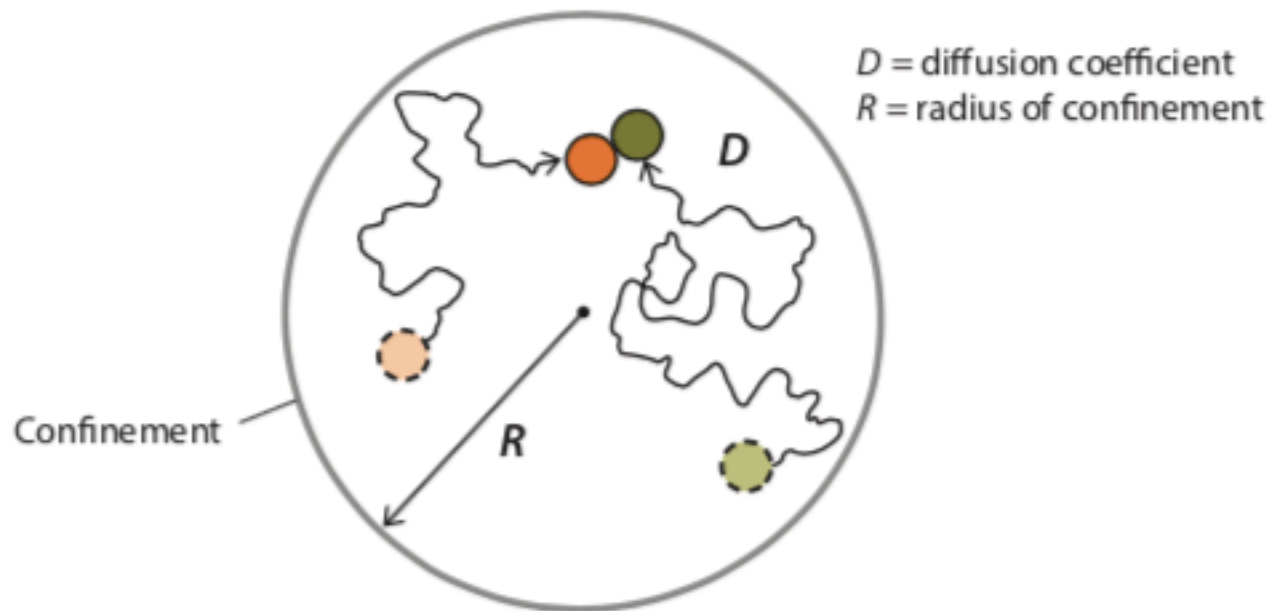


Figure 8

The problem of first-passage times for two genomic elements can be simulated as the anomalous diffusion of two particles in a confinement of radius R (representing a chromosomal domain; see **Figure 2**) with the anomalous diffusion coefficient D . The motion of the particles is governed by the fractional Langevin equation, Equation 2.

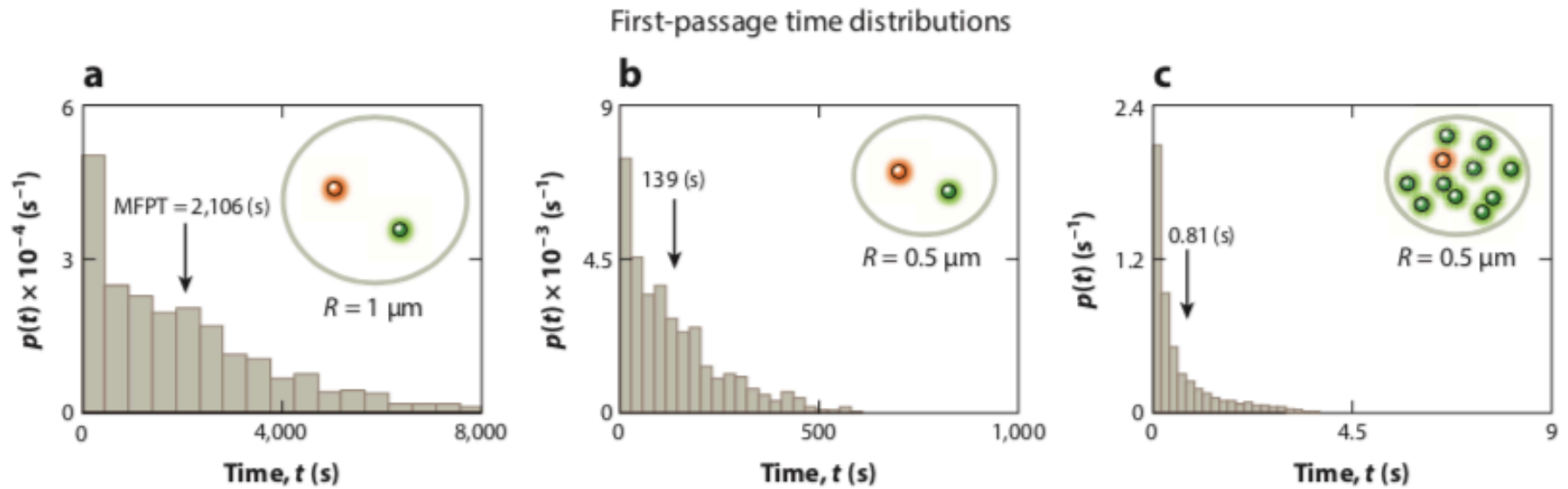


Figure 9

Distributions of the first-passage times obtained from the simulations of distant DNA segments undergoing fractional Langevin motion in a confining sphere of radius R . (a) A single pair of DNA segments diffusing in a spatial confinement of radius $1 \text{ } \mu\text{m}$ and with initial positions randomly chosen from the equilibrium distribution. The arrow indicates the mean first-passage time (MFPT). (b) Simulation as performed in panel a but with a confinement of radius $0.5 \text{ } \mu\text{m}$. Note that decreasing the confinement size by a factor of 2 accelerates the first-passage time by a factor of 16, in agreement with the prediction in Equation 4. (c) Simulation as performed in panel b but with 100 copies of one of the DNA segments. Note that the MFPT decreases by approximately a factor of 100, in accord with the prediction by Equation 4. Figure adapted from Reference 32 with permission.

En 9a), el MFPT para dos segmentos de ADN en un confinamiento de radio $R = 1 \text{ } \mu\text{m}$ es de aproximadamente 30 min. En 9b) el radio de confinamiento se reduce a la mitad y el MFPT es solo de aproximadamente 2 min. Cuando uno de los segmentos de ADN está presente en 100 copias, la MFPT disminuye aún más, a aproximadamente un segundo (Figura 9c).