

Modelo de Hardy-Weinberg

- Variantes ya introducidas en modelo Fisher-Wright

Organismos diploides ●

Reproducción sexual ●

Generaciones no solapantes

Apareamientos al azar

Población de tamaño infinito ●

No hay migración

No hay mutación ●

No actúa la selección ●

Las frecuencias alélicas no difieren entre sexos ●

- La selección (sexual o no sexual) puede modificar los patrones dentro y entre poblaciones

Frecuencias absolutas

Incorporando la endocria

Muchas poblaciones reales se apartan de la panmixia, de modo que los apareamientos ocurren con una mayor probabilidad entre individuos emparentados. Aunque con frecuencia la reproducción entre individuos fuertemente emparentados se evitan, efectos de vecindario, organización social, y otros tienen a hacer que los apareamientos entre individuos con un parentesco mayor que el promedio de la población sean más frecuentes de lo esperado por azar.

Para incorporar de manera sencilla este fenómeno general (sin obligarnos por ello a estudiar pedigrís de manera directa), introducimos F , el coeficiente de endocria (o endogamia) de la población. Se trata de un único parámetro adicional que procura capturar el efecto neto de la endocria sobre las frecuencias genotípicas esperadas.

Si los alelos se aparean con sus similares (A_1 con A_1 , y A_2 con A_2) con probabilidad F , entonces aumentarán las frecuencias de homocigotas A_1A_1 y A_2A_2 a expensas de las frecuencias de heterocigotas (A_1A_2).

Planteamos ahora lo razonado más arriba de manera explícita:

1. Una fracción $1-F$ de las combinaciones gaméticas se realizan al azar, produciendo los siguientes resultados parciales:

$$\text{frec. } (A_1A_1) = p^2(1 - F)$$

$$\text{frec. } (A_1A_2) = 2pq(1 - F)$$

$$\text{frec. } (A_2A_2) = q^2(1 - F)$$

2. Por otra parte, la restante fracción F de las combinaciones gaméticas corresponden a la endocria, es decir que no son al azar, sino que combinan gaméticos idénticos. Por tanto, estas combinaciones estarán en proporción a las frecuencias de los alelos, y serán pF y qF para los genotipos AA y aa , respectivamente. En combinación con el resultado anterior, obtenemos:

$$frec. (A_1A_1) = p^2(1 - F) + pF$$

$$frec. (A_1A_2) = 2pq(1 - F)$$

$$frec. (A_2A_2) = q^2(1 - F) + qF$$

Sumando las frecuencias, verificamos el resultado:

$$p^2(1 - F) + 2pq(1 - F) + q^2(1 - F) + pF + qF = (1 - F)(p^2 + 2pq + q^2) + F(p + q) = 1$$

Notamos, de paso, que la expresión de frecuencias esperadas obtenida en (2) es también válida cuando $F = 0$, en cuyo caso se simplifica a la expresión en (1).

Podemos hacer una simulación de muestreo de genotipos como la de más arriba, solamente que introduciendo a F para ponderar las probabilidades asociadas a cada genotipo.

Coefficiente de endocría observado

La frecuencia observada de heterocigotas H_o nos permite estimar F , puesto que:

$$f(A_1A_2) = 2pq(1 - \hat{F}) = H_o = H_e(1 - \hat{F})$$

(Notar que usamos \hat{F} para indicar que vamos a obtener una estimación de "F", cuyo verdadero valor desconocemos).

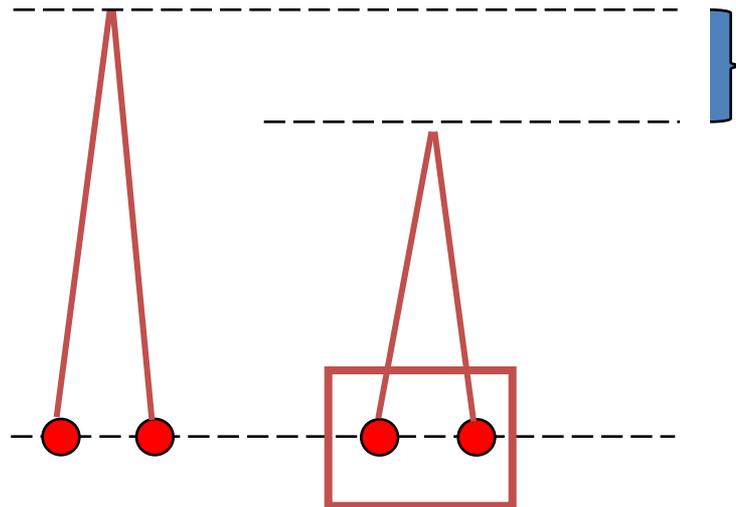
Despejando, tenemos que

$$\hat{F} = (H_o - H_e)/H_e$$

A continuación agregaremos las estimaciones de F para cada una de las muestras de genotipos obtenidas más arriba: una de ellas corresponde a muestreos realizados bajo el modelo de Hardy-Weinberg, y la otra a muestreos en los que las frecuencias esperadas están ajustadas asumiendo un valor de $F \neq 0$.

Estimación de la endogamia: una aproximación genealógica

Parentesco medio entre pares de alelos tomados al azar en la población: reflejado en antigüedad del ancestro común

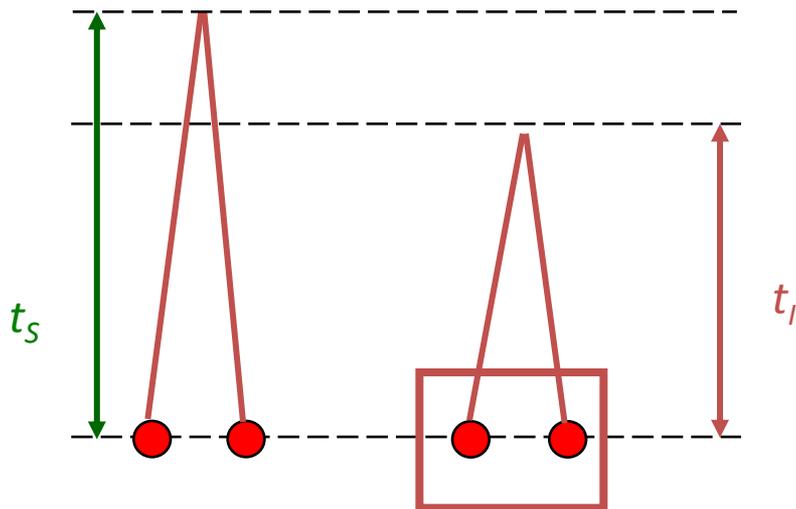


Efecto de la endogamia

Parentesco medio entre pares de alelos de **individuos** tomados al azar: reflejado en ancestro común más reciente

Endogamia: para **estimar** la diferencia, uso datos genéticos (heterocigosidad, número de diferencias entre alelos... o medidas equivalentes aptas para el tipo de locus que estoy estudiando).

Estimación de la endogamia: una aproximación genealógica



$$F_{IS} = \frac{t_S - t_I}{t_S}$$

Normalmente no tenemos estimaciones directas de tiempos, pero las aproximamos por sus consecuencias genéticas:

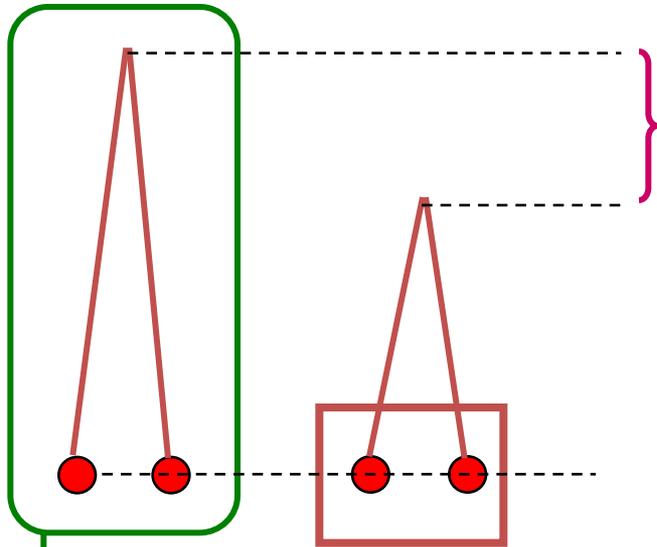
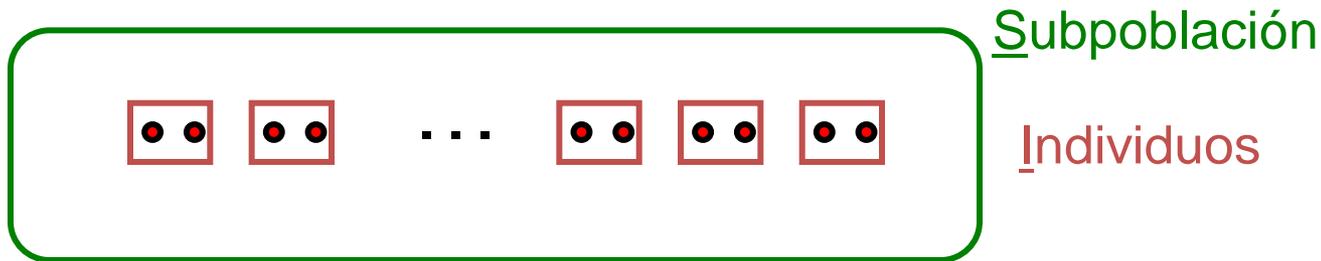
a) Heterocigocidad

$$F_{IS} = \frac{H_S - H_I}{H_S}$$

b) Número de diferencias entre alelos

$$F_{IS} = \frac{\pi_S - \pi_I}{\pi_S}$$

Endogamia local (en las subpoblaciones) en base a H



Coeficiente de endogamia:

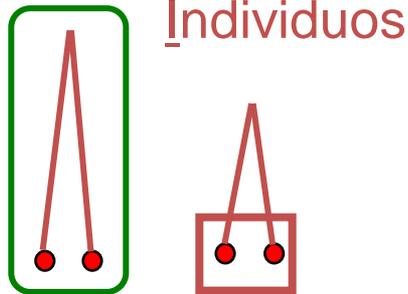
$$F_{IS} = \frac{H_S - H_I}{H_S} = \frac{H_e - H_o}{H_e}$$

$H_I = H_o$: heterocigosidad observada en los individuos

$H_S = H_e$: heterocigosidad esperada en las subpoblaciones, si las mismas son panmícticas (HW).

Endogamia local (en las subpoblaciones)

Subpoblación



Coefficiente de endogamia:

$$F_{IS} = \frac{H_S - H_I}{H_S} = \frac{H_e - H_o}{H_e}$$

Fis = -1	frec heterocigotas > He	bajo neutralidad Exogamia	si hay selección, favorece heterocigotas
Fis = 0	= He	Panmixia	--
Fis = 1	< He	Endogamia	homocigotas

Interpretación:

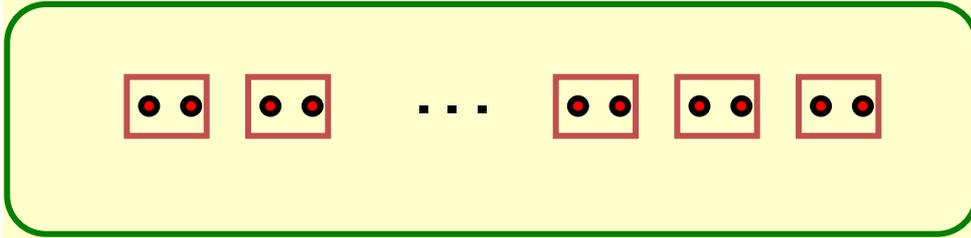
Bajo neutralidad, los apartamientos de lo esperado bajo panmixia deben atribuirse al régimen de apareamiento de los individuos en las subpoblaciones.

La selección puede producir un exceso o un déficit de heterocigotas.

Subdivisión geográfica

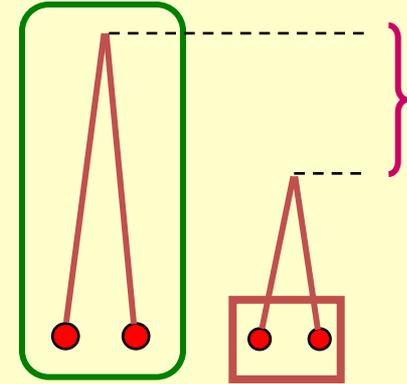
(¿podemos aplicar el mismo enfoque!)

$$F_{ST} = \frac{t_T - t_S}{t_T}$$

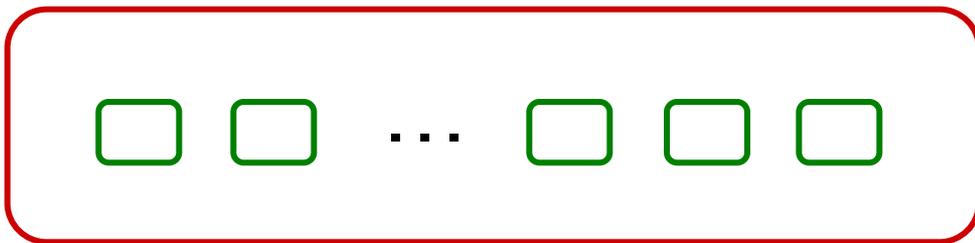


$$F_{IS} = \frac{H_S - H_I}{H_S} \quad F_{IS} = \frac{\pi_S - \pi_I}{\pi_S}$$

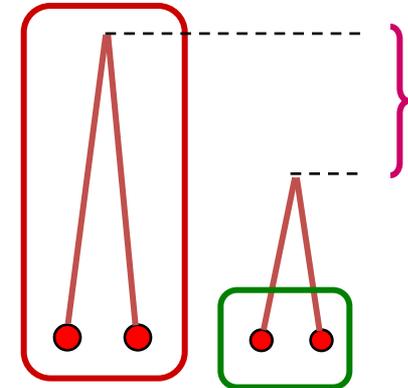
Subpoblación Individuos



Población Total Subpoblaciones



$$F_{ST} = \frac{H_T - H_S}{H_T} \quad F_{ST} = \frac{\pi_T - \pi_S}{\pi_T}$$



$$-1 < F_{IS} < 1$$

	frec heterocigotas	bajo neutralidad	si hay selección, favorece
F _{IS} = -1	> H _e	Exogamia	heterocigotas
F _{IS} = 0	= H _e	Panmixia	--
F _{IS} = 1	< H _e	Endogamia	homocigotas

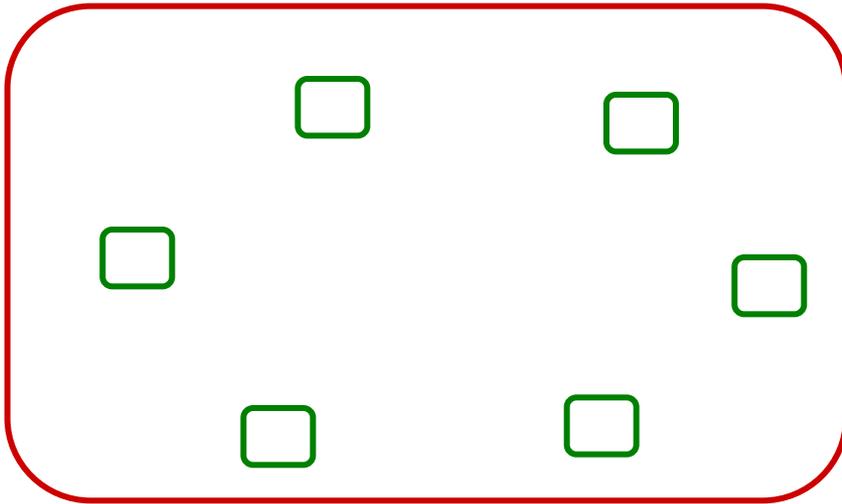
$$0 < F_{ST} < 1$$

F_{ST} = 0 Subpoblaciones idénticas
F_{ST} = 1 Fijación de alelos alternativos
efectos de la selección y/o de la divergencia neutral

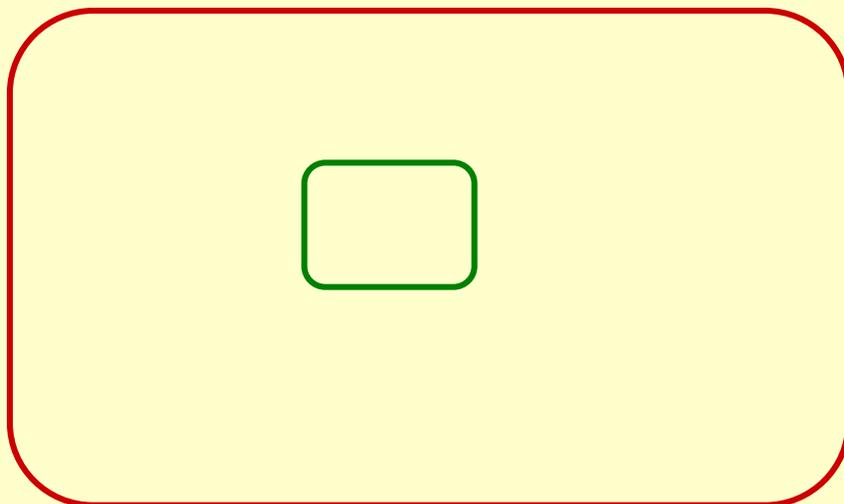
$$-1 < F_{IT} < 1$$

Combinación de efectos intra e interpoblacionales

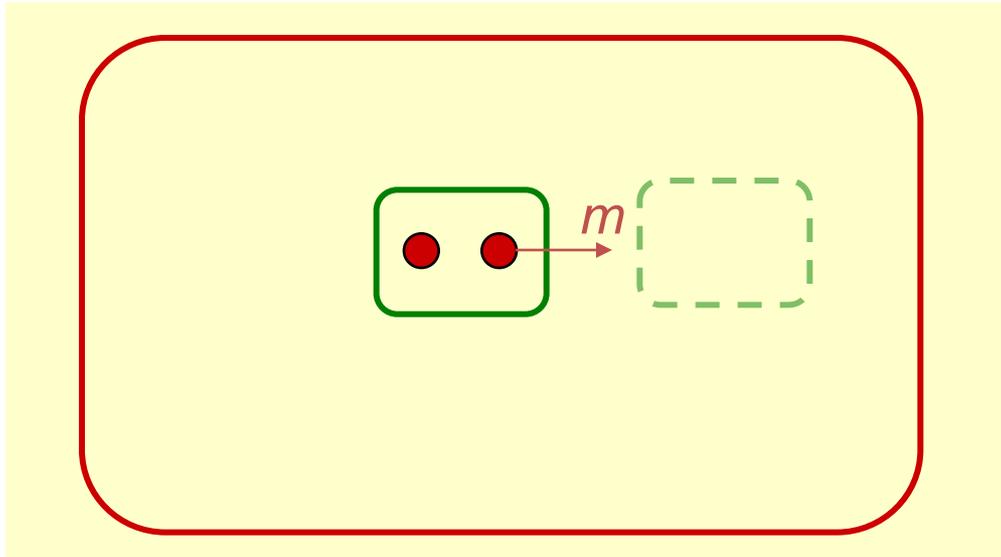
Subdivisión y flujo génico: modelo de islas (neutral)



- numerosas (en el límite, infinitas) subpoblaciones (islas), cada una de ellas con $2N$ alelos
- conectadas por flujo génico o migración de manera simétrica:
- Prob de migración por alelo y por generación = m .



- por ser un modelo simétrico, considero el caso de una isla o subpoblación con respecto al total

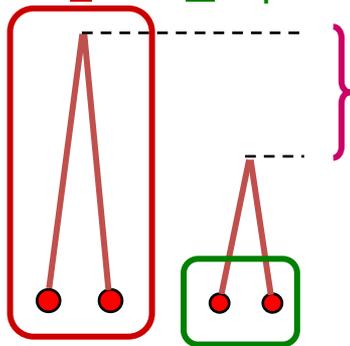


- F_{ST} es el exceso de homocigosis en las subpoblaciones con respecto al total.
 - Para calcular su valor de equilibrio, tomo dos alelos al azar dentro de la subpoblación, y considero su historia hacia el pasado:
 - Prob. de coalescencia = $1/2N$
 - Prob. de migración $\approx 2m$
- Aprox: eventos excluyentes en una generación*

Notar que F_{ST} estima la probabilidad de que ocurra un coalescente **antes (hacia el pasado)** que una migración, puesto que, una vez que un alelo migró, la probabilidad de coalescencia pasa a ser la de un par de alelos tomados al azar de toda la población (en el límite, tienden a ser de islas diferentes).

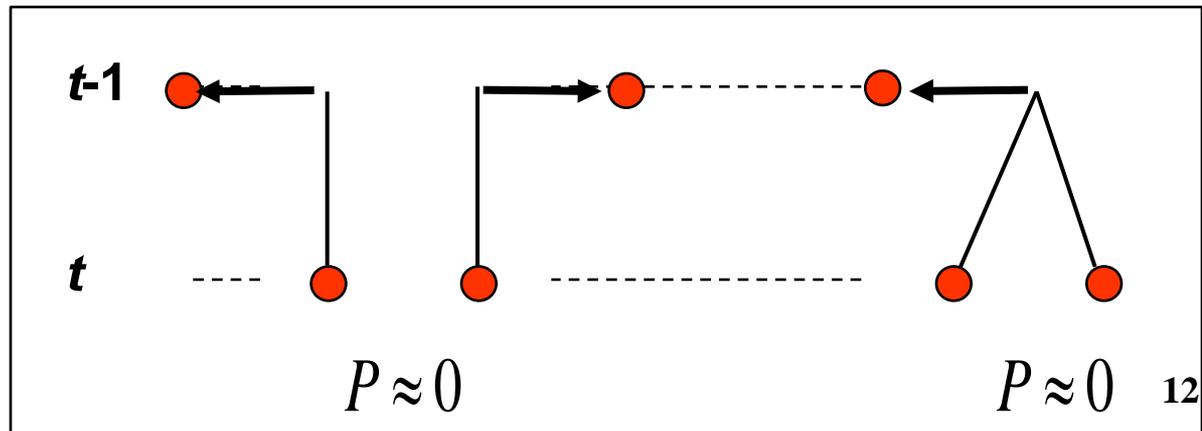
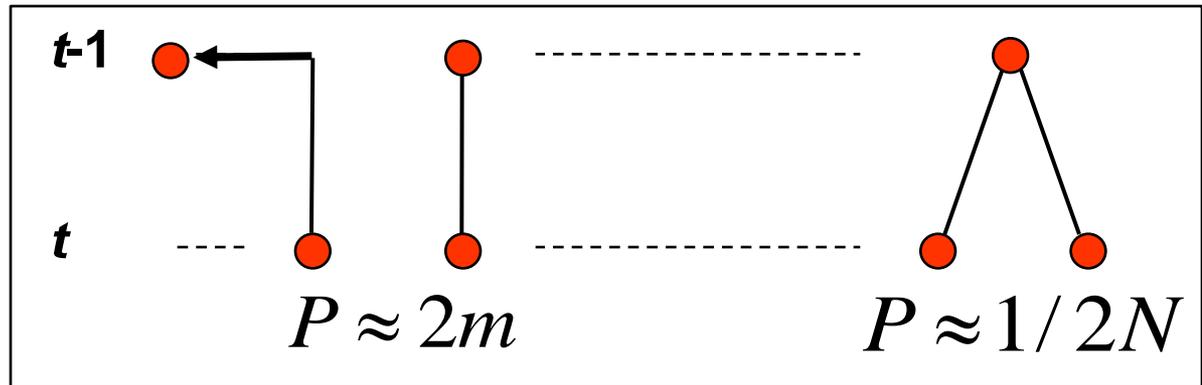
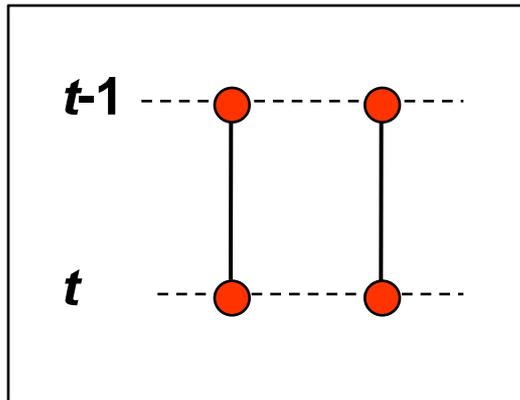
Contando casos favorables vs. casos adversos:

Población Total Subpoblaciones



$$E(F_{ST}) \approx \frac{1/2N}{2m + 1/2N} = \frac{1}{4Nm + 1}$$

Acción simultánea de deriva local y migración (modelo de islas)



Habíamos visto que el equilibrio entre mutación y deriva en una población panmíctica era:

$$E(F_{DM}) \approx \frac{1}{4N\mu + 1} = \frac{1}{\theta + 1}$$

Remplazando el efecto de la mutación por el del flujo génico

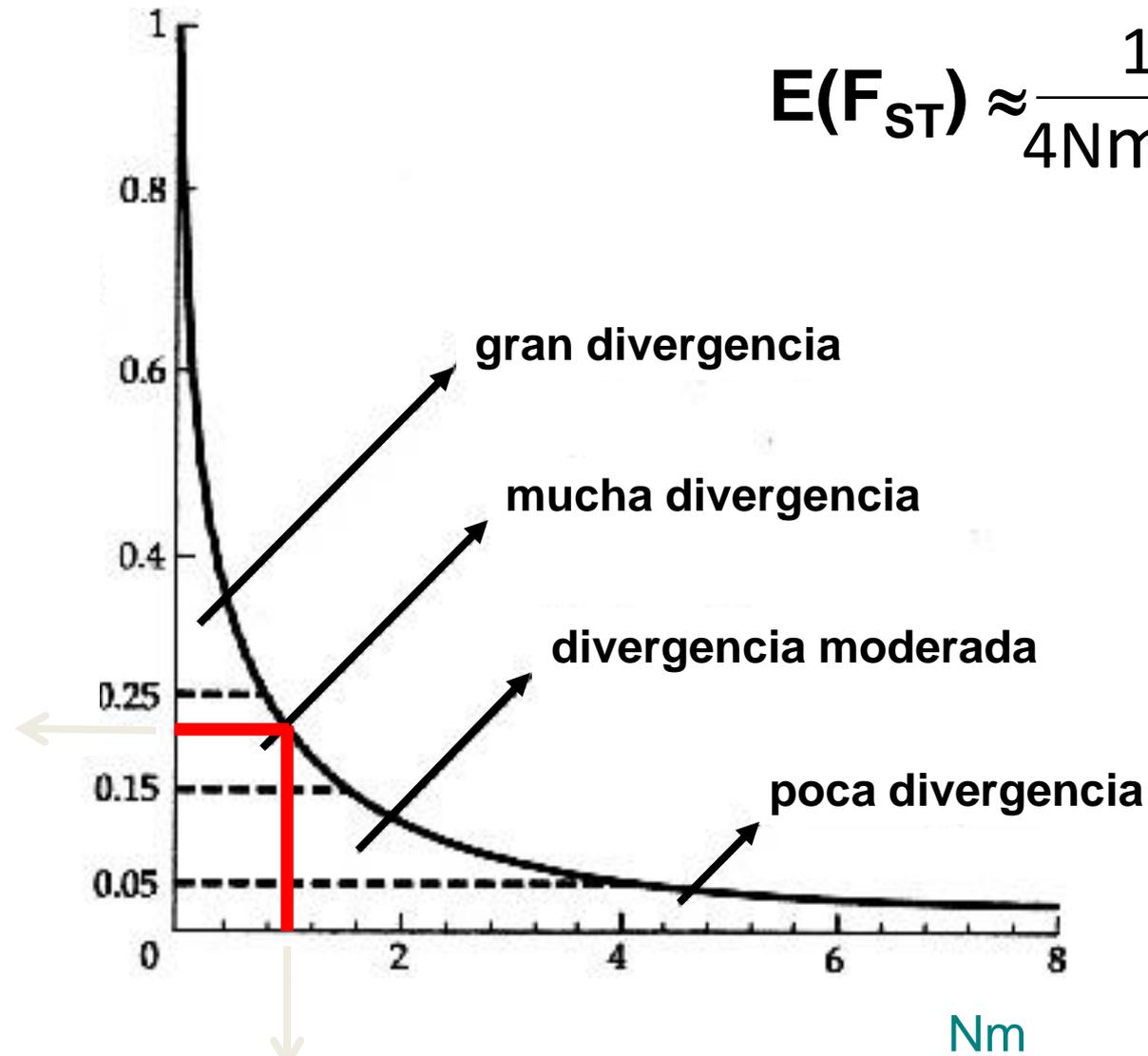
$$E(F_{ST}) \approx \frac{1}{4Nm + 1}$$

Este es el F_{ST} esperado **en equilibrio** entre deriva local y flujo génico en un modelo de islas

$E(F_{ST})$ en equilibrio

$$E(F_{ST}) \approx \frac{1}{4Nm + 1}$$

$F_{ST} = 0.2$



$Nm = 1$

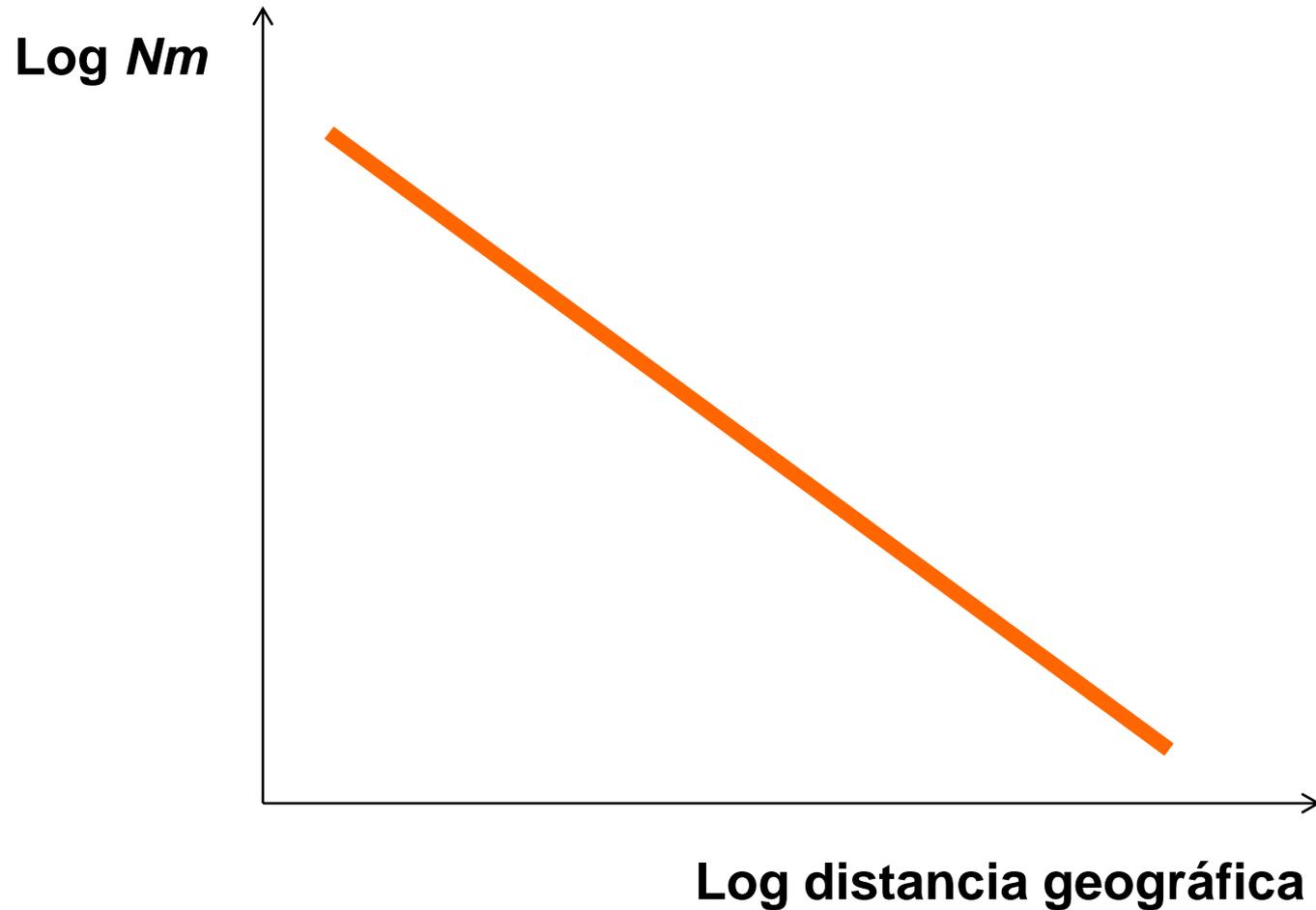
**$Nm \ll 1$ Predomina la deriva genética
(divergencia al azar de subpoblaciones)**

**$Nm \gg 1$ Predomina el flujo génico
(homogeneidad geográfica)**

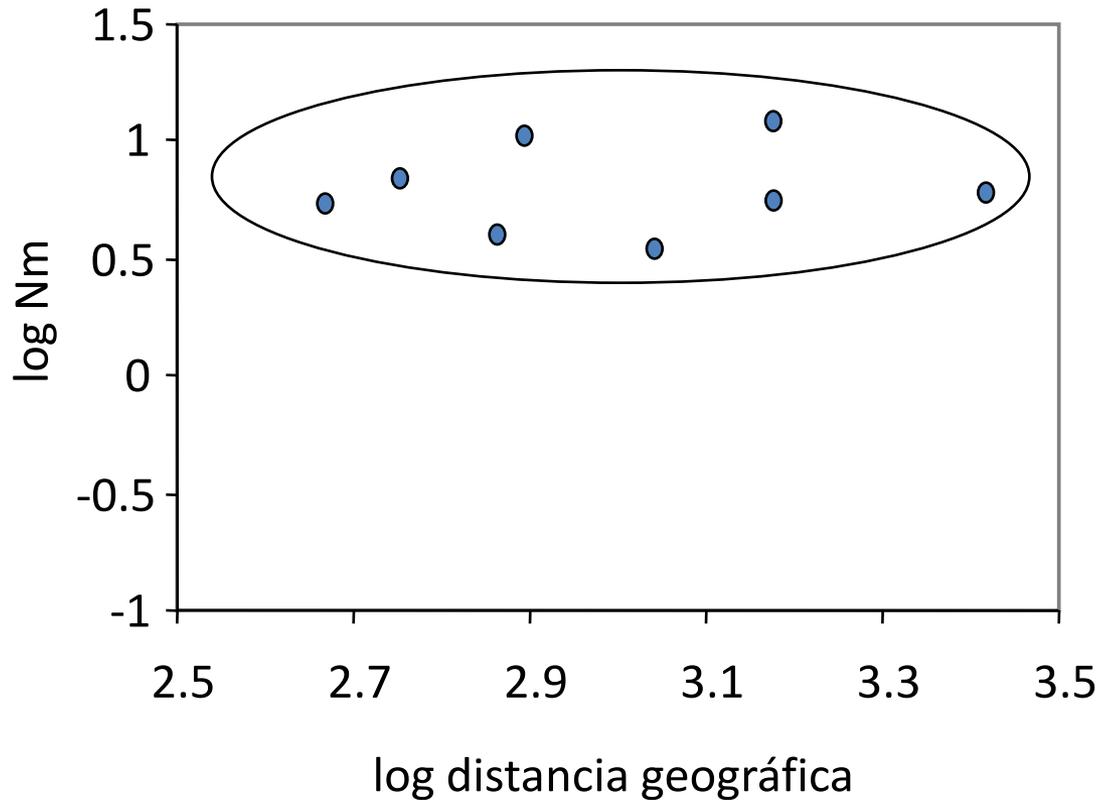
- Alcanza con **sólo unos pocos migrantes en promedio por generación** para impedir la diferenciación por deriva
- Esto es **INDEPENDIENTE** del tamaño poblacional

Aislamiento por distancia

Para pares de poblaciones:



Expansión reciente (ausencia de equilibrio)



- Nm mayor que el flujo actual

- Nm aparece como independiente de la distancia