

**Curso de Evolución 2020**

**Facultad de Ciencias**

**Montevideo, Uruguay**

<http://evolucion.fcien.edu.uy/>

<http://eva.fcien.udelar.edu.uy/>



Sewall Wright

4a. **Deriva genética.** Mutación. Equilibrios entre mutación y deriva.

# Modelo de Hardy-Weinberg

Organismos diploides

Reproducción sexual

Generaciones no solapantes

Apareamientos al azar

→ Población de tamaño infinito → N individuos (2N alelos)

No hay migración

No hay mutación

No actúa la selección

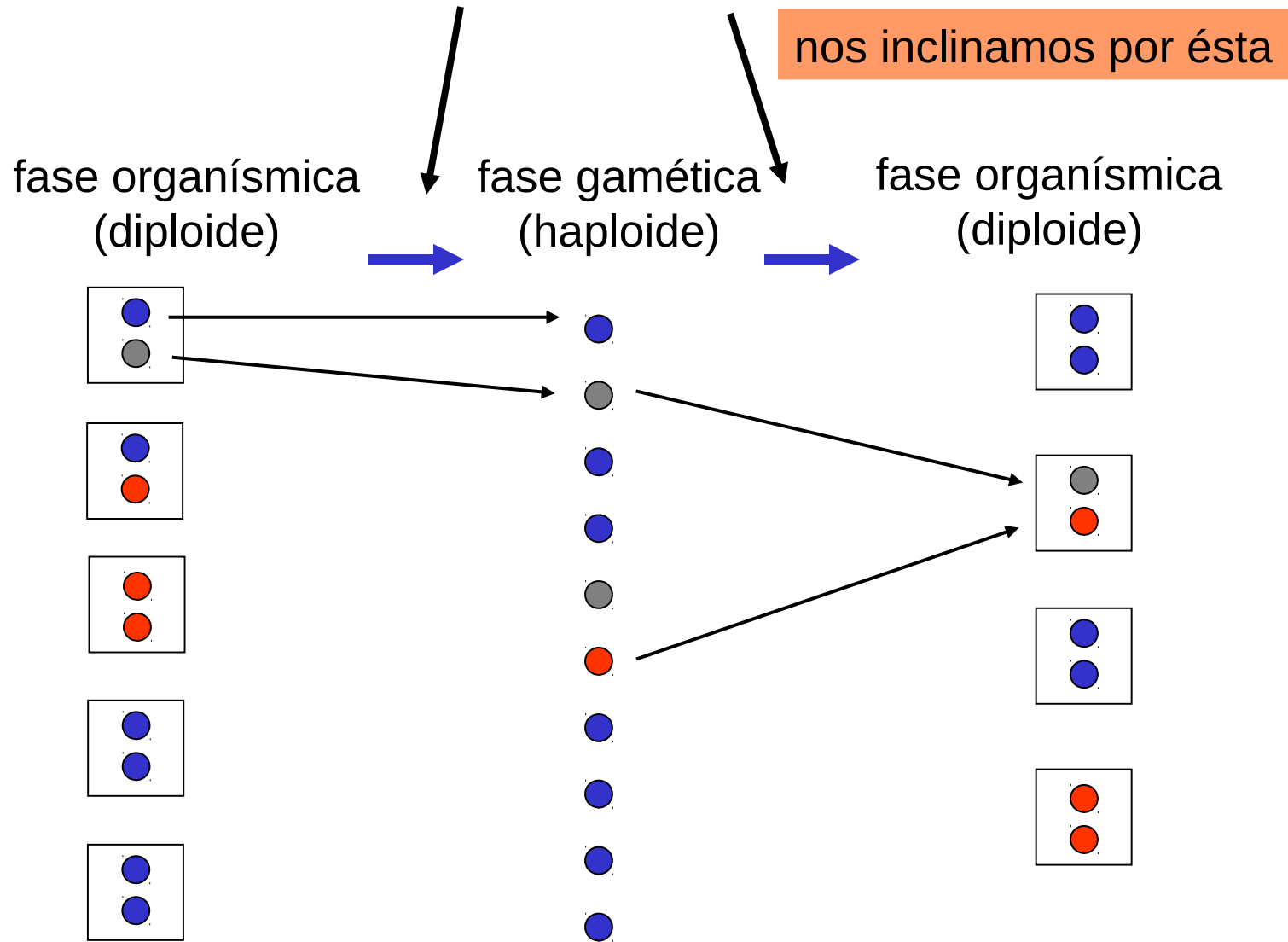
Las frecuencias alélicas no difieren entre sexos

**En estas condiciones hay un equilibrio estable  
(¡no hay evolución!)**

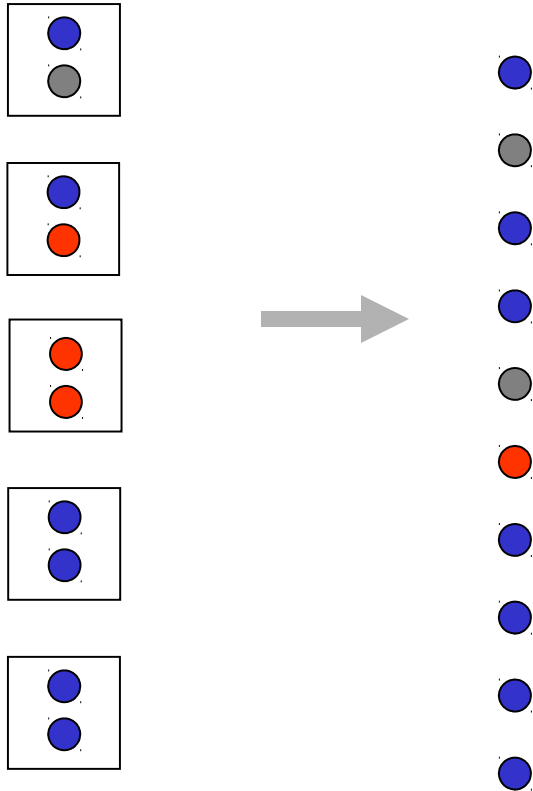
## Abandonando el supuesto del tamaño poblacional infinito

- aparecen efectos de muestreo de una generación a la siguiente
- ocurre evolución (cambio en la composición genética de la población)
- pero al no haber selección, ésta ocurre al azar

# Dos oportunidades para que ocurran efectos de muestreo



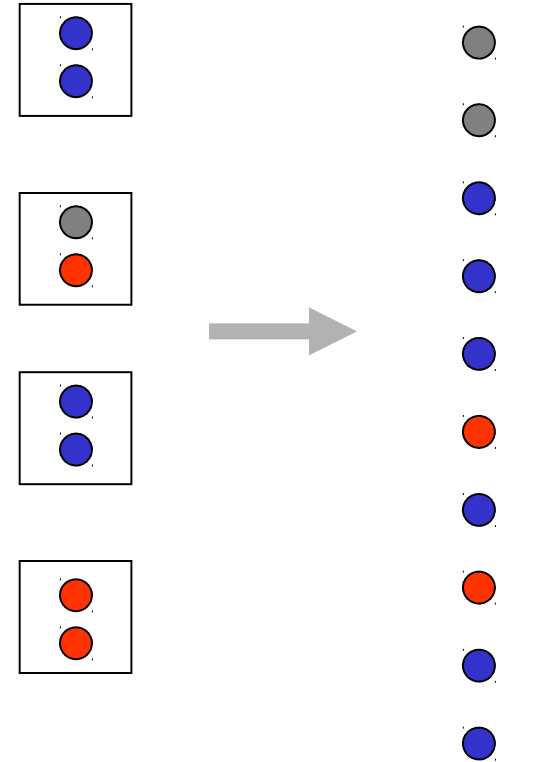
Generación 1



efectos de muestreo



Generación 2



## efectos de muestreo



## Modelo de Fisher-Wright

- muestreo al azar con reposición



- es un modelo haploide (la fase diploide no es necesaria, ni cambia nada)\*



\* Detalle importante: nuestro conteo de alelos, sin embargo, es de 2 por locus diploide y por individuo.

## Modelo de Fisher-Wright

- Incorpora el efecto de la deriva, mediante un **tamaño poblacional finito**
- Se parte de una población de **N individuos diploides**.
- Gen autosómico: **2N alelos**
- Cada generación se **construye** mediante un **muestreo con reposición** a partir de la generación previa.
- El tamaño poblacional se mantiene **constante**.

## Deriva genética

- Fluctuación aleatoria de las frecuencias alélicas como consecuencia del tamaño poblacional finito.
- Esto se da porque en cada generación, hay un muestreo al azar de los gametos para formar la generación siguiente.
- La deriva puede conducir a la **evolución sin selección** (en concreto, sin selección positiva), ya que los cambios que promueve son simplemente producto del azar.

**Sewall Wright, Motoo Kimura**



## Cómo puede modelarse el efecto de la deriva...

Poblacion tipo Fisher-Wright que consta de 4 alelos de dos tipos diferentes:

en la primera generación **A** (frec.  $A=0.75$ )  
**a** (frec.  $a=0.25$ )

Se puede escoger el alelo A con una probabilidad  $p=0.75$

Se puede escoger el alelo a con una probabilidad  $q=0.25$

La segunda generación puede entonces conformarse únicamente con las siguientes combinaciones:

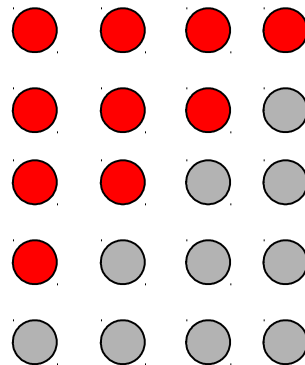
4A, 0a

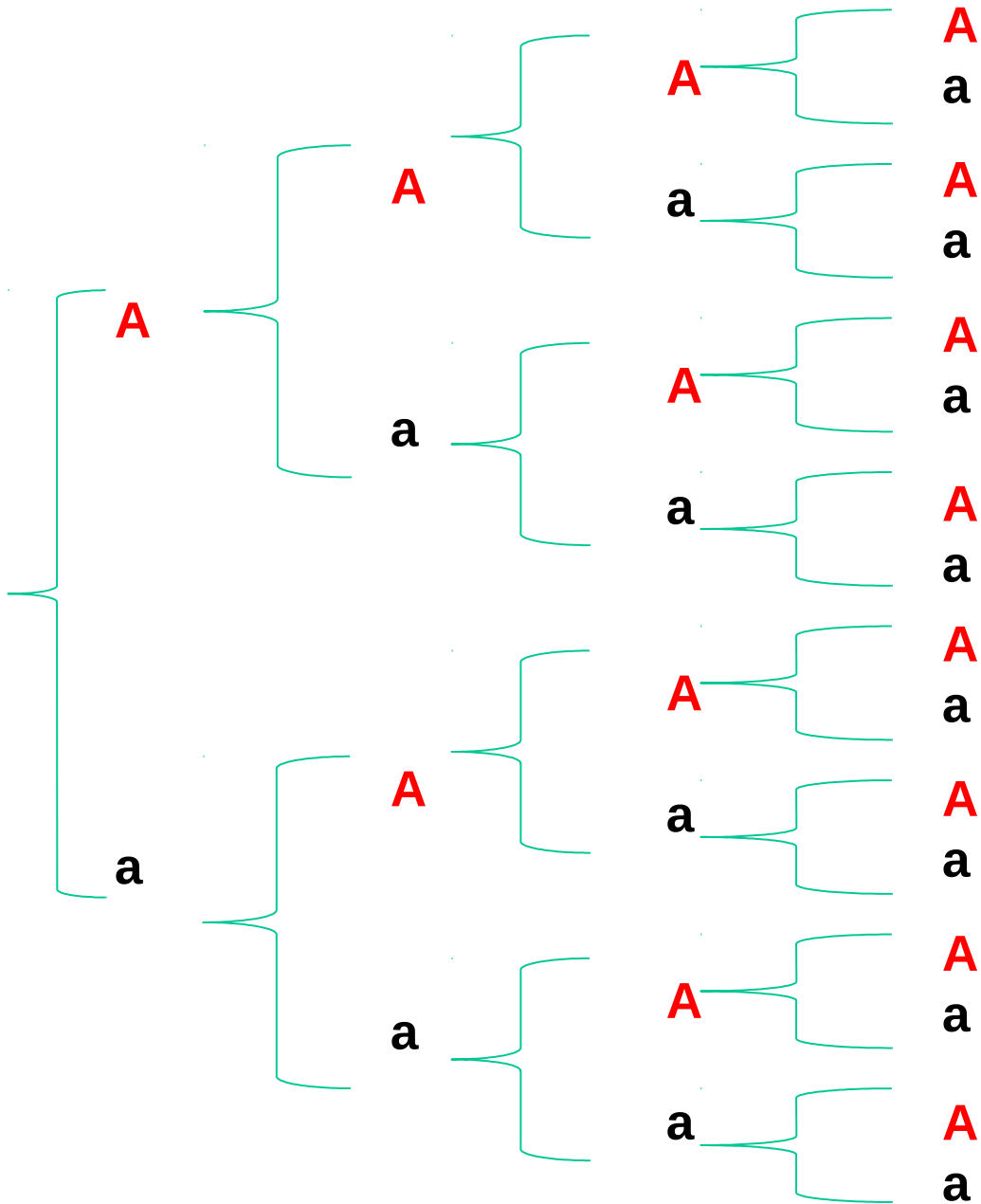
3A, 1a

2A, 2a

1A, 3a

0A, 4a





**AAAA**  
**AAAa**  
 ... etc.

De los 16 arreglos, hay:

- AAAA 1
- AAAa 4
- AAaa 6
- Aaaa 4
- aaaa 1

4A, 0a



3A, 1a



2A, 2a



1A, 3a



0A, 4a



- 4A y 0a con una probabilidad =  $p^4$  = 0,31640625

- 3A y 1a “ =  $4p^3q$  =  $4p^3(1-p)$  = 0,421875

- 2A y 2a “ =  $6p^2q^2$  =  $6p^2(1-p)^2$  = 0,2109375

- 1A y 3a “ =  $4pq^3$  =  $4p(1-p)^3$  = 0,046875

- 4a y 0a “ =  $q^4$  = 0,00390625

Total = 1

Vemos que existe **una sola forma de obtener 4A o 4a:**

$$p \times p \times p \times p = \mathbf{p^4} = 0.316$$

$$q \times q \times q \times q = \mathbf{q^4} = (1-p)^4 = 0.004$$

...pero hay **varias posibilidades de obtener el resto** de las combinaciones. Por ejemplo:

Un arreglo de 3A y 1a, tiene una probabilidad

$$p \times p \times p \times (1-p) = p^3(1-p)$$

Y además hay 4 arreglos de ese tipo:

AAAa, AAaA, AaAA y aAAA

La probabilidad acumulada para esas formas es  $4 p^3 q$

A partir de las frecuencias iniciales de los dos alelos y del número total de alelos (tamaño poblacional), se pueden calcular las probabilidades de todos los posibles resultados del proceso en una generación.

Esto se puede generalizar mediante la distribución binomial:

Se define una variable aleatoria que cuenta el número de casos ( $i$ ) que es muestreado el alelo A en una serie de  $n$  ensayos (en nuestro caso  $n=2N$  alelos):

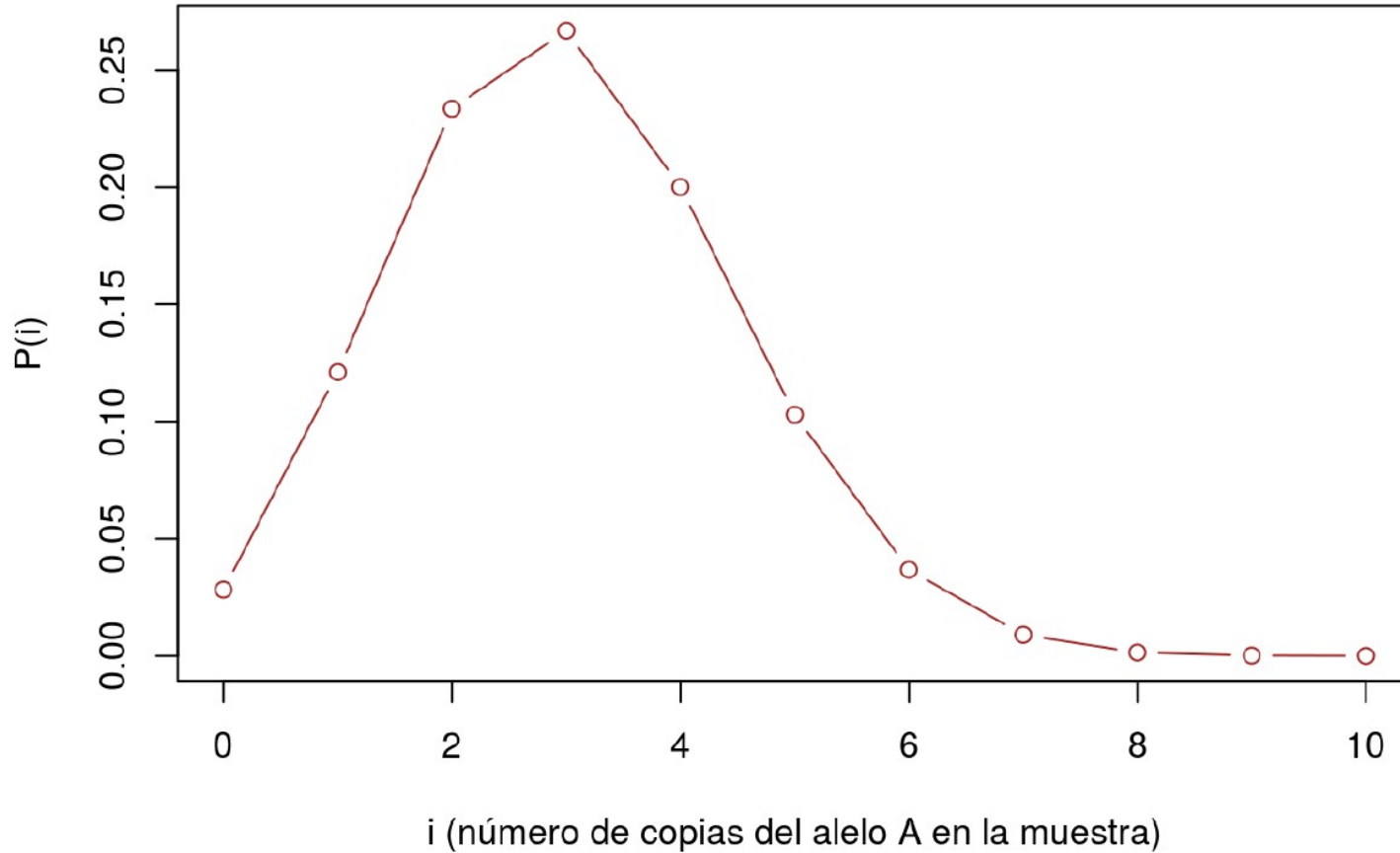
$$P(i) = \binom{2N}{i} p^i (1-p)^{2N-i}, \quad i = 0, 1, \dots, 2N$$

donde  $\binom{2N}{i} = \frac{2N!}{(2N-i)!i!}$

Este es el llamado coeficiente binomial, calcula el número de formas de obtener  $i$  copias de A y  $2N-i$  copias de a.

# Ejemplo

$$p = 0.3, 2N = 10$$



```
## [1] 0.0282475249 0.1210608210 0.2334744405 0.2668279320 0.2001209490
## [6] 0.1029193452 0.0367569090 0.0090016920 0.0014467005 0.0001377810
## [11] 0.0000059049
```

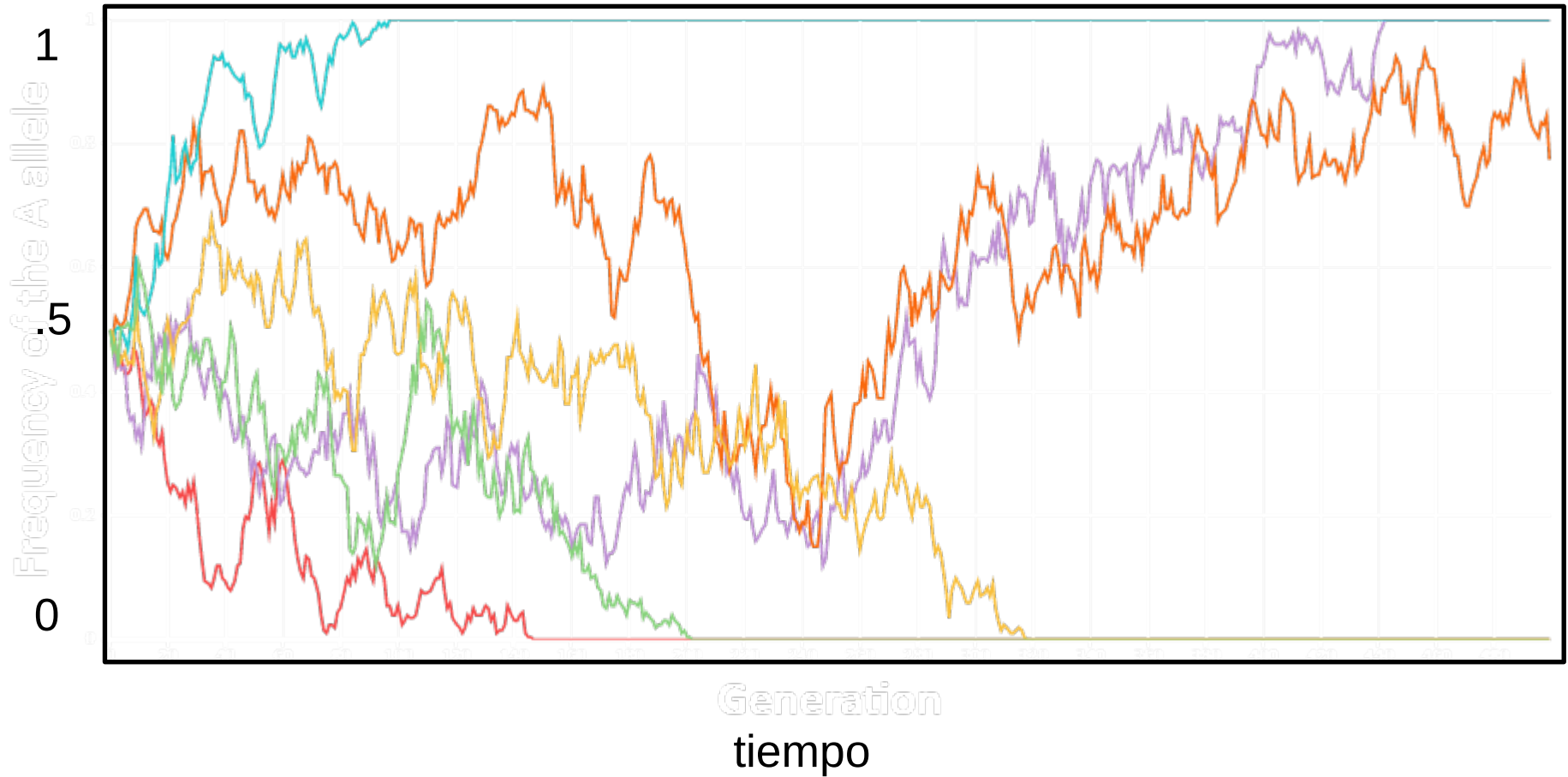
# Simulaciones

Usaremos dos herramientas para realizar simulaciones del proceso de deriva genética:

- 1) Interfase web de “learnPopGen” (<http://www.phytools.org/PopGen/>), un paquete armado para R (también disponible como tal) por Liam Revell (2018).
- 2) Código R casero, usando Rstudio cloud, también a través de internet.
- 3) Sitio <http://popgensimulator.pitt.edu/graphs/allele>

# Simulaciones en

<http://popgensimulator.pitt.edu/graphs/allele>



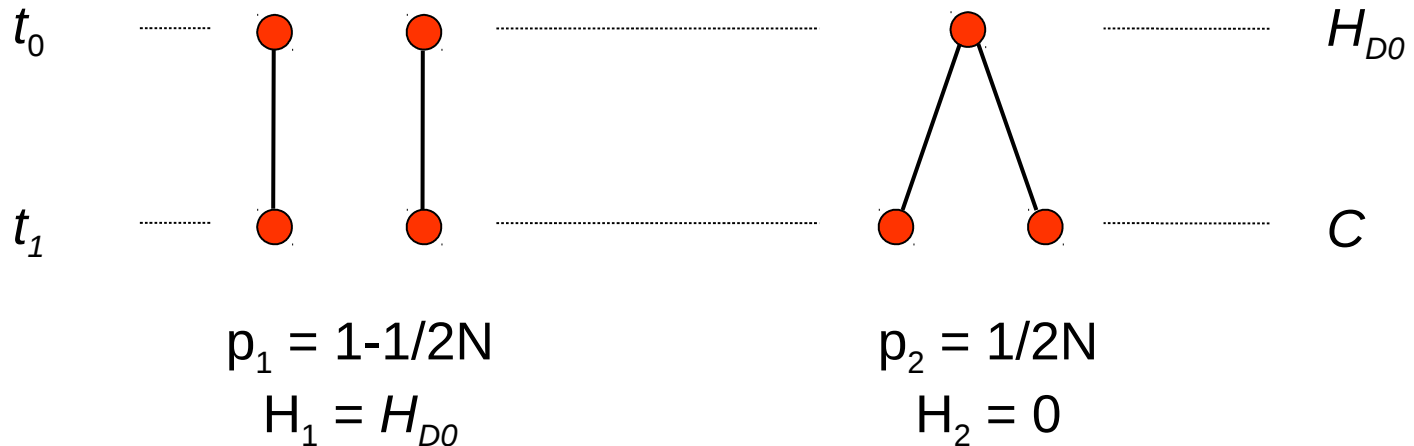


## OBSERVACIONES

- Cambios en las frecuencias: aleatorios en magnitud y signo (siguiendo las propiedades de la binomial).
- Efecto acumulado a corto plazo: errático.
- Efecto acumulado a largo plazo: reducción de la variabilidad (hasta la fijación de alguna variante).
- Las pequeñas poblaciones tienen fluctuaciones más marcadas que las grandes (también determinado por la binomial).
- Tienen que pasar un cierto número de generaciones, del orden de  $N$ , para observar un cambio apreciable en las frecuencias alélicas.

## Efectos de la deriva genética: reducción de heterocigosidad ( $H_D$ )

- heterocigosidad: probabilidad de que dos alelos cualesquiera en la población tengan estados diferentes



- el valor esperado de  $H_{D1}$  es la suma ponderada de los valores de la variable:

$$E(H_{D1}) = \sum_{i=1}^2 H_i p_i = H_{D0} \left( 1 - \frac{1}{2N} \right) + 0 \cdot \left( \frac{1}{2N} \right)$$

$$E(H_{D1}) = H_{D0} \left( 1 - \frac{1}{2N} \right) = H_{D0} - H_{D0} \left( \frac{1}{2N} \right)$$

## Efectos de la deriva genética: reducción de heterocigosidad ( $H_D$ )

$$E(H_{D1}) = H_{D0} (1 - 1/2N)$$

$$E(H_{D2}) = H_{D1} (1 - 1/2N)$$

$$E(H_{D2}) = H_{D0} (1 - 1/2N) (1 - 1/2N)$$

$$E(H_{Dt}) = H_{D0} (1 - 1/2N)^t$$

- $1/2N$  es la tasa de reducción de heterocigosidad por deriva genética (“coeficiente de deriva”  $d$ )

# Usos del término heterocigosidad

1. Modelo de Hardy-Weinberg, y análisis de frecuencias genotípicas:

Frecuencia esperada de heterocigotas ( $H_e = H_{HW}$ ).

Frecuencia observada de heterocigotas ( $H_o$ ).

Válido para sistemas diploides.

Modelo de Fisher-Wright

Probabilidad de que dos alelos cualesquiera en la población tengan estados diferentes ( $H_D$ , y próximamente, cuando introduzcamos la mutación,  $H_{DM}$ ).

Válido para cualquier sistema. Es una medida de variación genética en las poblaciones.

Los subíndices se usan en forma irregular, y a veces deben inferirse según el contexto.

Discutir relación y diferencias entre estos conceptos para un sistema diploide autosómico.

## Simulaciones vs. experimentos

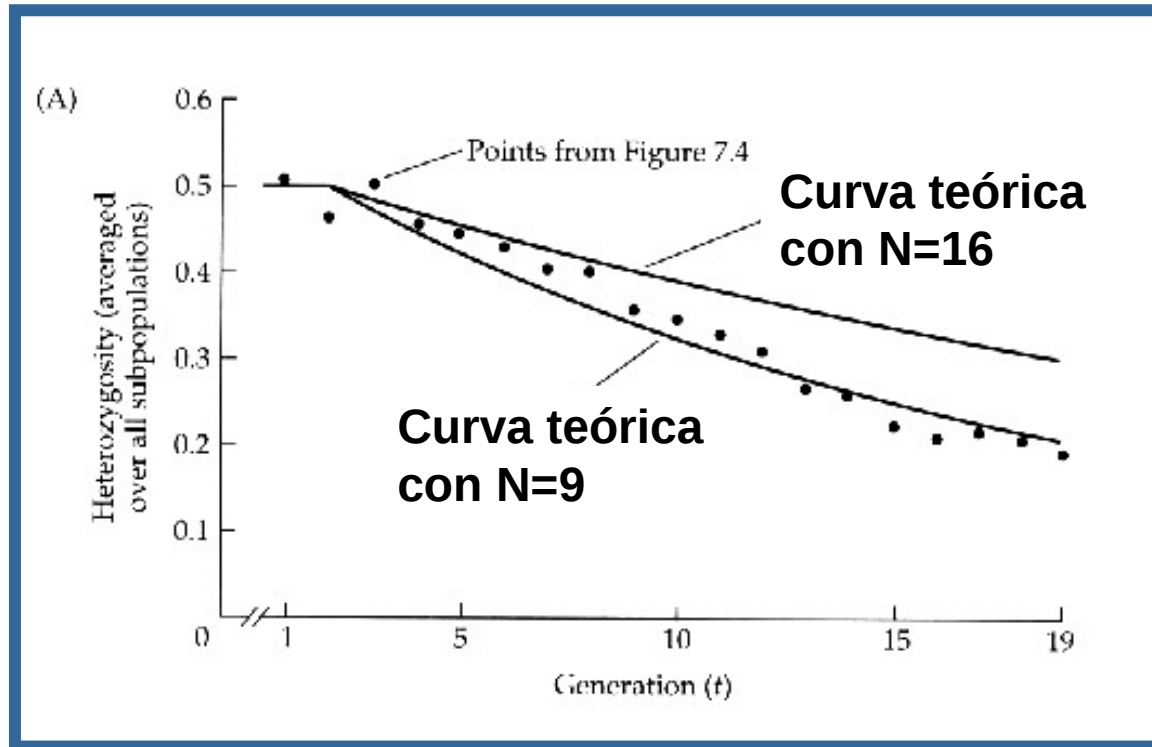
Simulaciones: 107 simulaciones para 16 individuos diploides (32 alelos), comenzando con  $p = 0.5$  para el alelo de referencia. Cada simulación se corrió durante 20 generaciones y se acumularon los resultados.

Experimentos: lo mismo, pero con 107 poblaciones reales de 16 (8 machos, 8 hembras) moscas de la fruta (*Drosophila*). Se sigue la variación en frecuencia de un alelo (*bw75*), cuya frecuencia inicial fue  $p = 0.5$ .

En los dos casos se trazó la pérdida de heterocigosidad por generación, promediando entre las distintas réplicas.

# Concepto de tamaño poblacional efectivo

tamaño de una población que, en el modelo Fisher-Wright, pierde heterocigosidad al mismo ritmo de la población real



Vemos que las poblaciones reales de 16 moscas pierden heterocigosidad a un ritmo comparable al de una población ideal (Fisher-Wright) de 9 moscas. Su “tamaño efectivo”  $N_e \cong 9$

# Tamaño poblacional efectivo

Típicamente menor que el tamaño censal

$$N_e \leq N$$

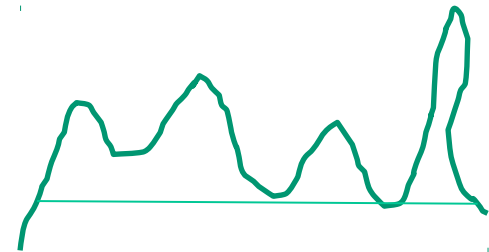
En particular, cuando:

- hay distinto N° de machos y hembras reproductoras

$$\frac{1}{N_e} = \frac{1}{4N_f} + \frac{1}{4N_m} \quad N_e = \frac{4N_f N_m}{N_f + N_m}$$

- fluctuaciones demográficas

$$\frac{1}{N_e} = \frac{1}{t} \left( \frac{1}{N_1} + \frac{1}{N_2} + \dots + \frac{1}{N_t} \right)$$



## Síntesis

- la deriva genética opera en todas las poblaciones
- su efecto es más marcado en poblaciones pequeñas (ya que depende del tamaño efectivo)
- produce fluctuaciones en las frecuencias génicas
- a largo plazo resulta en la pérdida de variación (reducción de heterocigosidad, pérdida de alelos) hasta la fijación de alguna variante



# Problema

- $N_e$  estima cuál debería ser el tamaño de una población ideal, que siga el modelo Fisher-Wright, para que la reducción esperada de la heterocigosis por generación en dicha población ideal y en nuestra población de interés sean iguales
- podemos valernos de un “coeficiente de deriva”  $d = 1/2N$  en una población F-W para razonar sobre el problema
- en el ejemplo anterior,  $N_e$  en una población que cambia de tamaño resulta de promediar los valores de  $d$  a lo largo del tiempo
- Problema: aplicar estos conceptos para mostrar que, cuando el número de hembras ( $N_f$ ) y machos ( $N_m$ ) difiere en una población,

$$N_e = \frac{4N_f N_m}{N_f + N_m}$$