

**Curso de Evolución 2024**  
**Facultad de Ciencias**  
**Montevideo, Uruguay**

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

<https://www.youtube.com/@CursoEvolucion/videos>

10. Evolución molecular. Tasas y patrones de evolución a nivel proteico y nucleotídico. Relojes moleculares. Selecciónismo y neutralismo. Análisis de la selección a nivel molecular.



# Comparación esquemática de modelos nulos (de referencia)

	Hardy-Weinberg	Wright-Fisher sin mutación	Wright-Fisher + mutación neutral
parámetros	$\{p_i\}$	$\{p_i\}, N$	$\{p_i\}, N, \mu$
escala de tiempo	$\sim$ generación	$\sim N$	$\sim N$
			Modelo neutral estándar
	Selección natural en tiempo generacional		Selección natural en tiempo evolutivo

# Incorporando la selección natural

Selección natural en tiempo generacional

Selección natural en tiempo evolutivo

parámetros

$\{p_i\}$ ,  $N$ ,  $\{s_i\}$

$\{p_i\}$ ,  $N$ ,  $\mu$ ,  $\{s_i\}$

Modelo nulo

HW

WF

Comentarios

incorporamos selección sobre los genotipos, pero desestimamos mutación

incorporamos selección sobre los alelos, pero desestimamos el régimen de apareamiento

Ejemplo de aplicación (“cargando los datos”)

$$\begin{aligned}P(A_1A_1) &= p_1^2 w_{11} = p_1^2(1-s_{11}) \\P(A_1A_2) &= 2p_1p_2 w_{12} = \\ & 2p_1p_2(1-s_{12}) \\ & \dots\end{aligned}$$

$$\begin{aligned}P(A_1) &= p_1^2 w_1 = p_1^2(1-s_1) \\P(A_2) &= p_2^2 w_2 = p_2^2(1-s_2) \\ & \dots\end{aligned}$$

# Deriva y mutaciones neutrales: el modelo neutral estándar

	Modelo Wright-Fisher	
Tamaño poblacional	N (finito, constante)	
Reproducción	Los alelos de t son muestreados al azar con reposición de t-1	
Mutación	Sí	Modelo de sitios infinitos
Recombinación	No	
Selección	No	
Migración (en particular inmigración)	No	

# Teoría neutral de evolución molecular

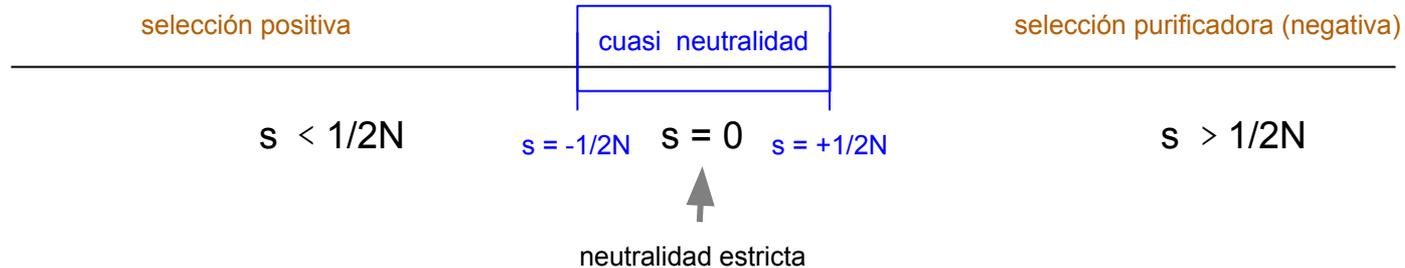
(Kimura 1968, King & Jukes 1969)

1. Hipótesis nula (contra la cual buscamos, por ejemplo, evidencia de selección natural).

2. Explicación de (al menos parte de) la realidad:

- polimorfismos (variación genética) dentro de las poblaciones.
- sustituciones (diferencias fijas) entre especies.

Selección natural (s en contra de una nueva mutación):



# Teoría neutral: sustituciones

Probabilidad de fijación = frecuencia del alelo en la población

Reloj molecular: tasa de sustitución  $k$  = tasa de mutación neutral  $\mu$   
(¡independientemente del tamaño poblacional!)

Prob. sustitución = prob. mutación neutral x prob. fijación

$$k = 2N\mu \times 1/2N$$

$$k = \mu$$

(notamos que el reloj molecular predicho es en tiempo generacional)

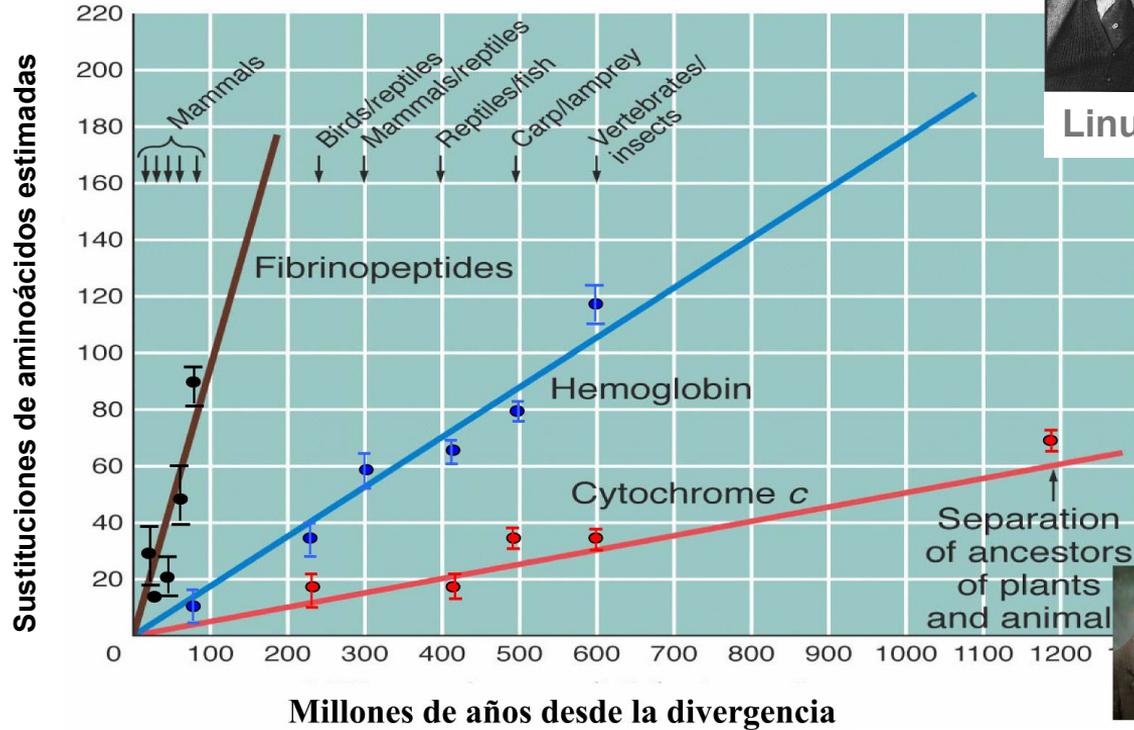
# El Reloj Molecular

Zuckerkandl y Pauling 1965

(“las moléculas como documentos históricos”)



Linus Pauling



Emile Zuckerkandl

# Reloj molecular

Cuasi neutralidad: específicamente selección contra variantes ligeramente deletéreas.

- Tamaño corporal
  - pequeño : poblaciones grandes, tiempo generacional corto
  - grande: poblaciones pequeñas, tiempo generacional largo

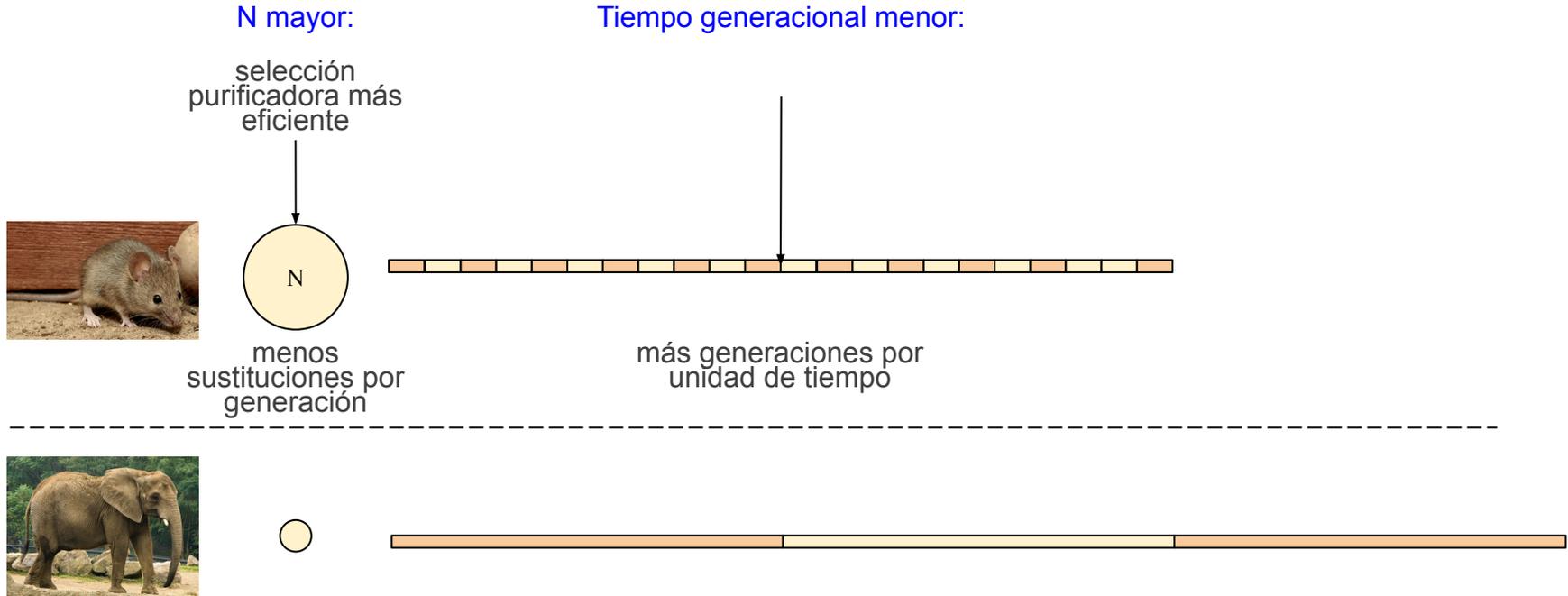
Tamaño pequeño:

- selección purificadora más eficiente: menos sustituciones por generación
- tiempo generacional menor: habrá más generaciones por unidad de tiempo

Si los dos factores se compensan, la tasa neta de sustituciones por millón de años será parecida en organismos de tamaños corporales diferentes. En otras palabras, el reloj molecular empírico podría explicarse en un modelo cuasi neutral que incluya selección (débil) contra variantes ligeramente deletéreas.

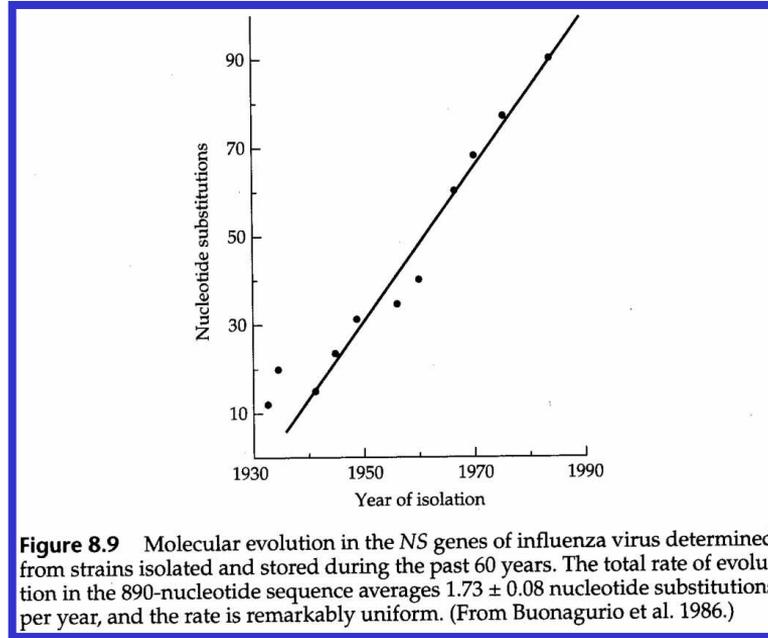
# Reloj molecular

Empírico: tasas constantes por unidad de tiempo (millones de años)



Se conocen otros factores que afectan el reloj molecular: por ejemplo, tasas metabólicas.

# Reloj molecular en virus: la evolución en tiempo real



<https://nextstrain.org/ncov/gisaid/global/6m>

# Teoría neutral: sustituciones

Probabilidad de fijación = frecuencia del alelo en la población

Reloj molecular:  $k = \mu$  (neutralidad estricta)

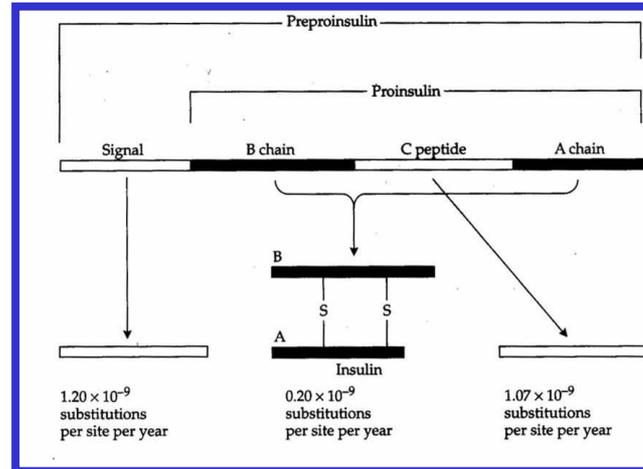
- reloj empírico más cercano a cuasi neutralidad
- variaciones por diversas causas (tasa metabólica, sel. natural, etc)
- el tiempo medio entre eventos de fijación es  $1/\mu$  (consecuencia del punto anterior y, una vez más, independiente del tamaño poblacional)
- el tiempo esperado de fijación de un alelo desde su aparición por mutación es  $4N$

# Patrones de evolución molecular

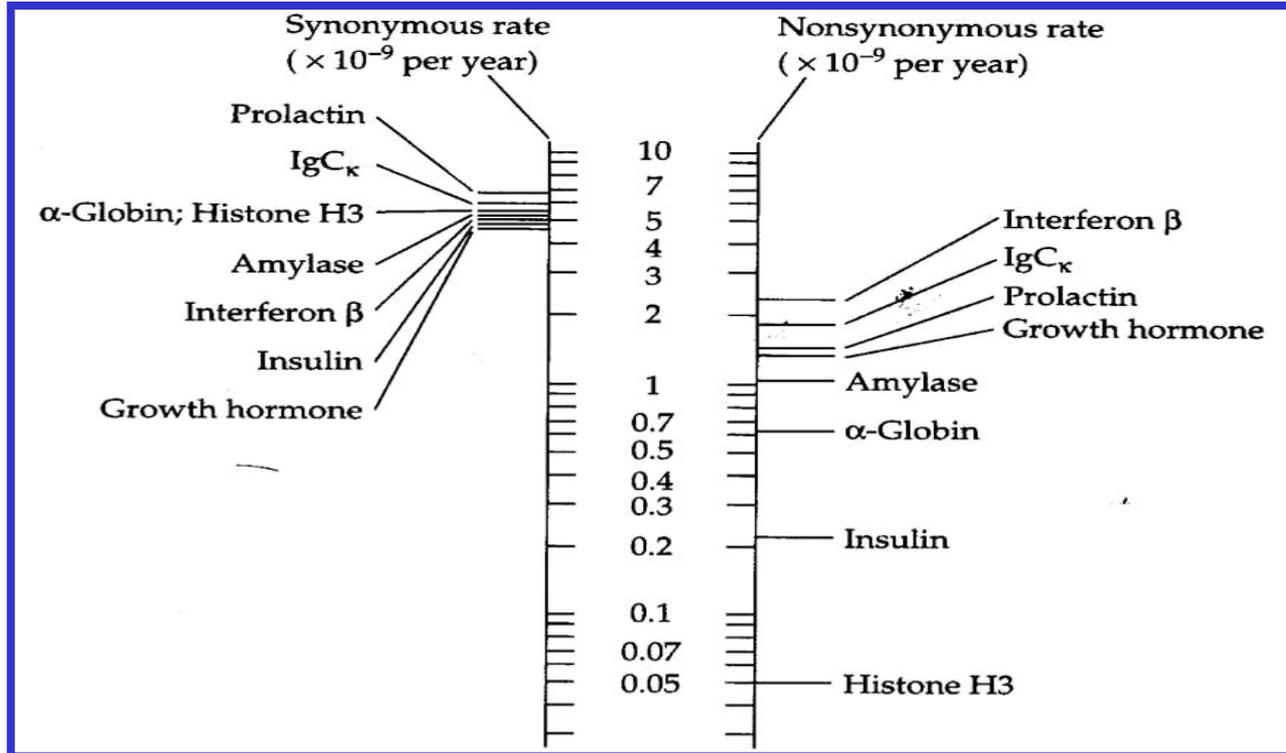
En general, reflejan distintos niveles de selección purificadora:

- sustituciones sinónimas > sustituciones no sinónimas
- sustituciones de AA similares > sustituciones de AA de propiedades más divergentes

Podemos verificarlo en filogenias y, en algunos casos, vincular niveles de conservación de AA con los efectos funcionales de los cambios por mutagénesis dirigida.



# Patrones de evolución molecular



# Selección a nivel molecular: escala filogenética

En general, reflejan distintos niveles de selección purificadora:

- sustituciones sinónimas > sustituciones no sinónimas
- sustituciones de AA similares > sustituciones de AA de propiedades más divergentes
- ... y muchos casos más

Podemos verificarlo en filogenias y, en algunos casos, vincular niveles de conservación de AA con los efectos funcionales de los cambios por mutagénesis dirigida.

# Evolución adaptativa a nivel molecular

Un ejemplo de selección direccional:  
lisina espermática en moluscos

genes implicados en el  
reconocimiento óvulo-espermio y el  
aislamiento reproductivo

Evaluar  $\omega = dN/dS$



# Evolución de regiones codificantes

$q_{ij}$  - probabilidad de pasar del codón  $i$  al codón  $j$ :

$$q_{ij} = \begin{cases} 0, & \text{if the two codons differ at more than} \\ & \text{one position,} \\ \pi_j, & \text{for synonymous transversion,} \\ \kappa\pi_j, & \text{for synonymous transition,} \\ \omega\pi_j, & \text{for nonsynonymous transversion,} \\ \omega\kappa\pi_j, & \text{for nonsynonymous transition,} \end{cases}$$

$\pi_j$  frecuencia de equilibrio del codón  $j$

$\kappa$ : razón transición:transversión

$\omega = dN/dS$ : razón de cambios no sinónimos:sinónimos

## Selección natural:

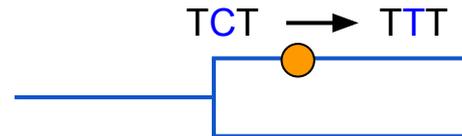
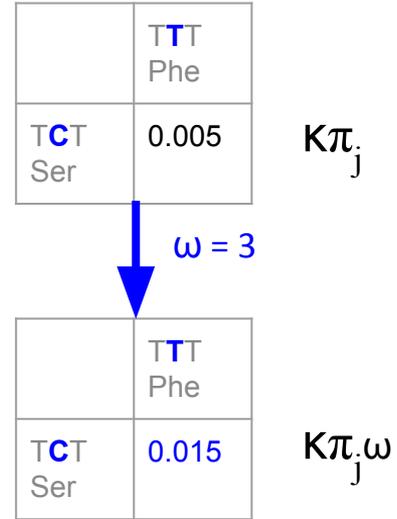
$\omega < 1$  purificadora (negativa)

$\omega = 1$  neutralidad

$\omega > 1$  **positiva**

# Esquema

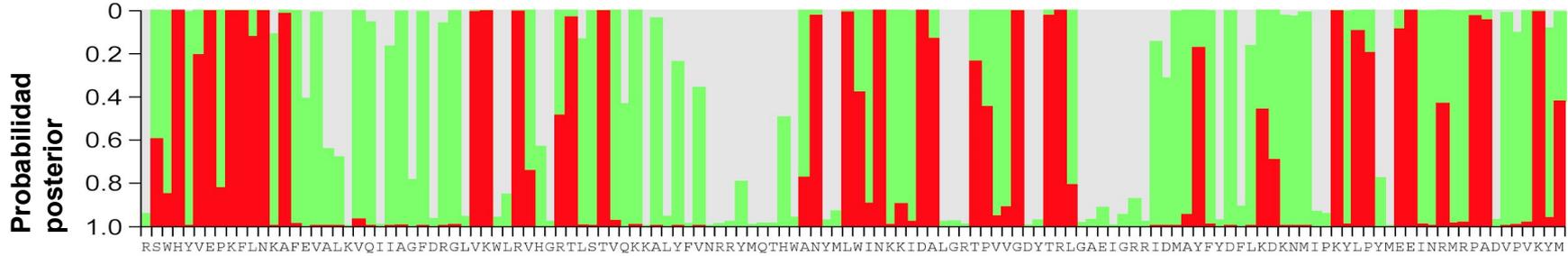
- Línea de base:
  - Modelo general de sustituciones (61x61)
  - Excluyendo codones Stop
- Parámetros adicionales para detectar selección:
  - $\omega = dN/dS$
  - $\omega < 1$ : purificadora
  - $\omega = 1$ : neutralidad (no hay selección)
  - $\omega > 1$ : selección direccional positiva





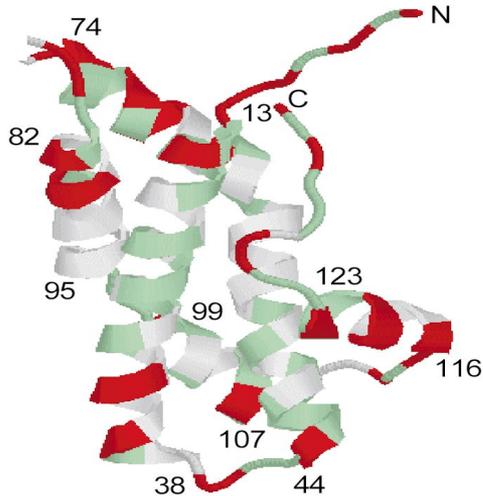
# LISINA ESPERMÁTICA EN ABALONES

Yang y Bielawski 2000



Sitios aminoacídicos (codones) en la lisina espermática

*ds in Ecology & Evolution*



$$\omega = dN/dS$$

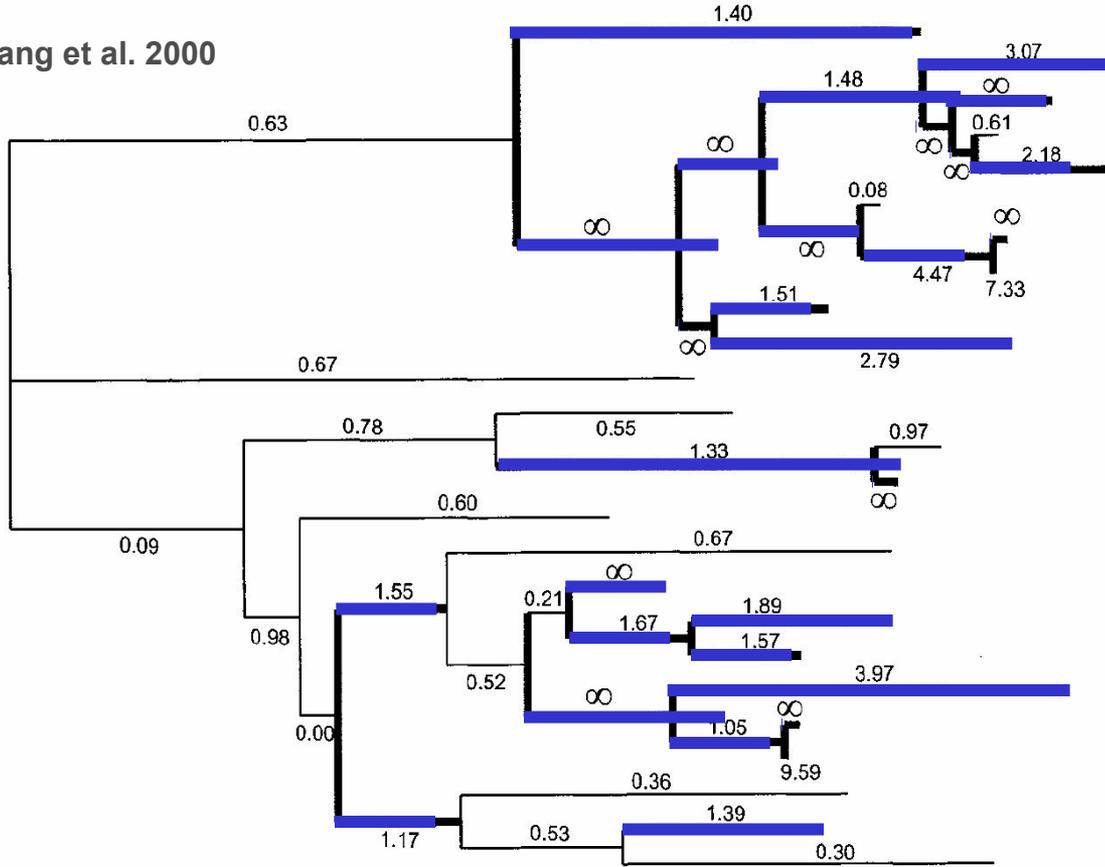
- 0,085
- 0,911
- 3,065

Los sitios seleccionados positivamente se agrupan en los extremos de la estructura tridimensional

tal structure. *Reproduced, with permission, from Ref. 39.*

# LISINA ESPERMÁTICA EN ABALONES

Yang et al. 2000



$\omega = dN/dS > 1$   
(selección positiva)

diferente entre  
linajes y entre  
sitios...

# Coevolución de genes de la vía de la insulina en primates neotropicales

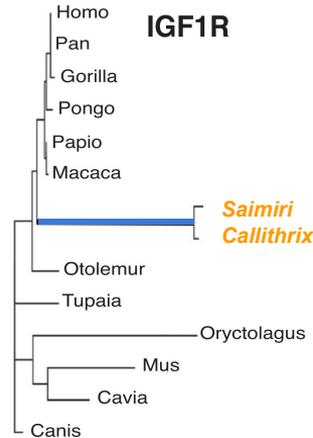
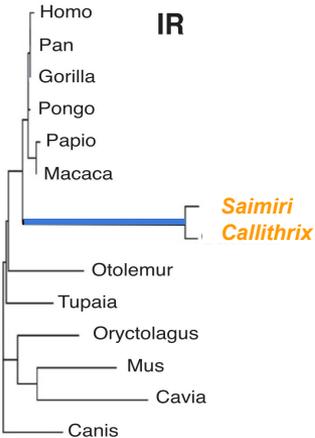
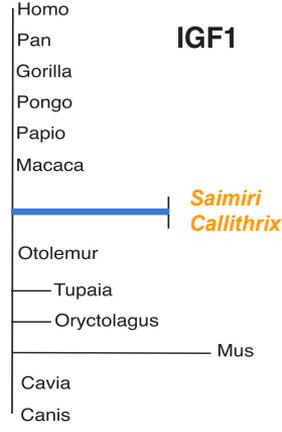
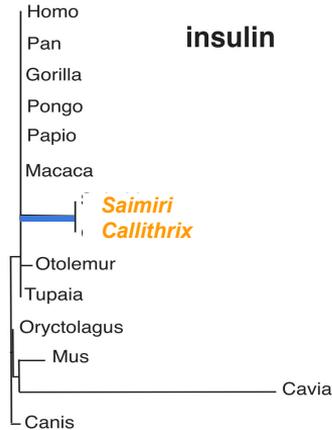
Dos grupos de mamíferos neotropicales tienen insulinas fisiológicamente divergentes (por ejemplo, menor afinidad con el receptor):

- primates (monos del Nuevo Mundo)
- roedores caviomorfos (cobayo, carpincho, chinchilla... )

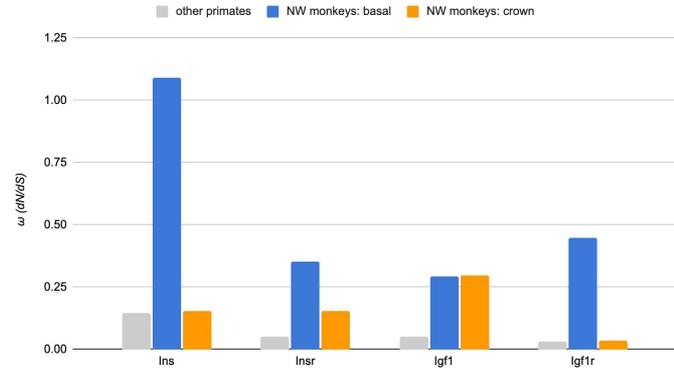
**Wallis (2015)** estudió la coevolución del gen de la insulina, su receptor y genes funcionalmente vinculados a la insulina extrayendo las secuencias de dos genomas del grupo y combinándolas con las de otros mamíferos.



# Cuatro de los genes estudiados por Wallis (2015)



Variation of  $\omega$  (dN/dS) in insulin and related primate genes



$\omega$  por ramas, usando datos de más genomas

Selección episódica (rama basal):

- Insulina (Ins)
  - Receptor de insulina (InsR, IR)
  - Factor de crecimiento parecido a la insulina (Igf1)
  - Receptor de Igf1 (Igf1r)
- 
- Igf1: la selección positiva también involucra al “grupo corona”)

# Resumen: evolución molecular a escala filogenética

La teoría neutral ofrece un modelo nulo. Sirve de hipótesis nula y explica algunos patrones generales.

Un modelo cuasi neutral mejora mucho la calidad de las predicciones.

Aunque no existe un reloj molecular universal e invariante, los cambios a nivel molecular son mucho más regulares que a otros niveles (por ejemplo, morfológico). “Relojes locales” o ajustables.

Podemos buscar evidencia de selección natural a escala filogenética usando

- secuencias (en especial de regiones codificantes para proteínas)
- filogenias
- modelos que ajustan  $\omega = dN/dS$  para detectar variaciones
  - entre ramas del árbol
  - entre sitios en la secuencia
  
- un corolario importante de la teoría neutral es que, aún cuando la evolución de un gen esté dictada por la selección positiva, una clase particular de cambios (aquellos estrictamente neutros) se fijará con una tasa igual a la correspondiente tasa de mutación

# Estudio de la adaptación al nivel molecular

- el debate entre seleccionismo y neutralismo se ha transformado en una discusión sobre la **importancia relativa** de distintas clases de selección
- existen métodos diversos para poner a prueba hipótesis adaptativas a nivel molecular
- los patrones generales de cambio molecular tienen una “forma general” compatible con el neutralismo (y la cuasi neutralidad)
- recordemos que los cambios neutrales pueden acompañar la evolución adaptativa (“haciendo dedo”)
- estudios genómicos recientes sugieren que, en *Drosophila*, hasta un 35 % de los codones pueden estar evolucionando al influjo de la selección positiva, en humanos un 9%, en *A.thaliana* aún menos.

# Teoría neutral de evolución molecular

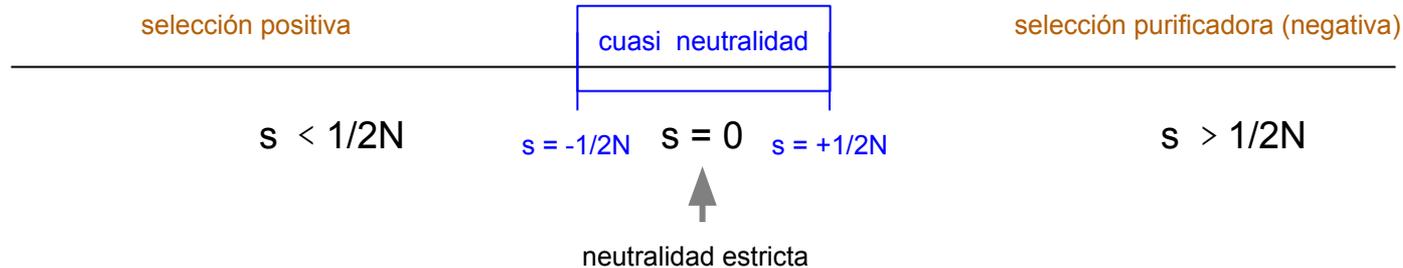
(Kimura 1968, King & Jukes 1969)

1. Hipótesis nula (contra la cual buscamos, por ejemplo, evidencia de selección natural).

2. Explicación de (al menos parte de) la realidad:

- polimorfismos (variación genética) dentro de las poblaciones.
- sustituciones (diferencias fijas) entre especies.

Selección natural (s en contra de una nueva mutación):



# Polimorfismos neutrales

Ya obtuvimos un resultado que nos dice la variación esperada como equilibrio entre deriva y mutaciones estrictamente neutras:

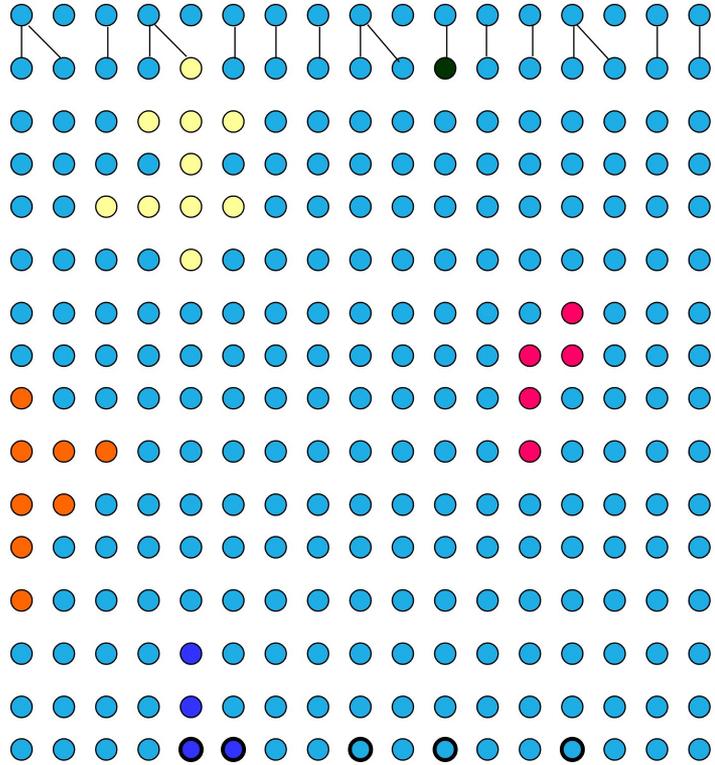
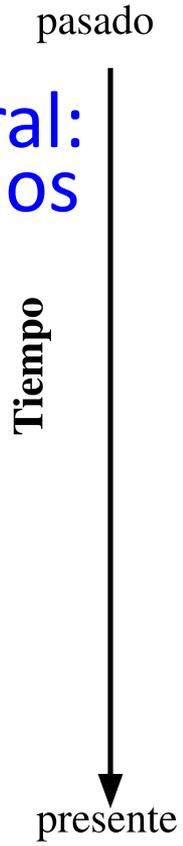
Población (demografía): WF

Modelo mutacional: sitios infinitos

$$E(H_{DM}) = 1 - E(F) = \frac{\theta}{1 + \theta}$$

$\theta$ , el parámetro poblacional neutral, combina los efectos de la deriva (N) y la mutación ( $\mu$ ). Bajo neutralidad, todas las medidas de variación genética dependen de  $\theta$ .

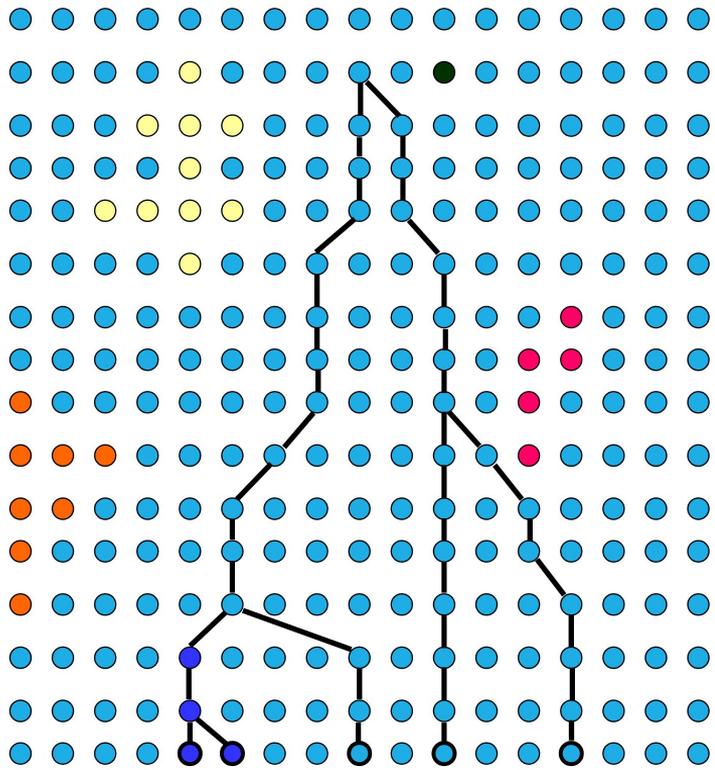
# Teoría neutral: polimorfismos



pasado

Tiempo

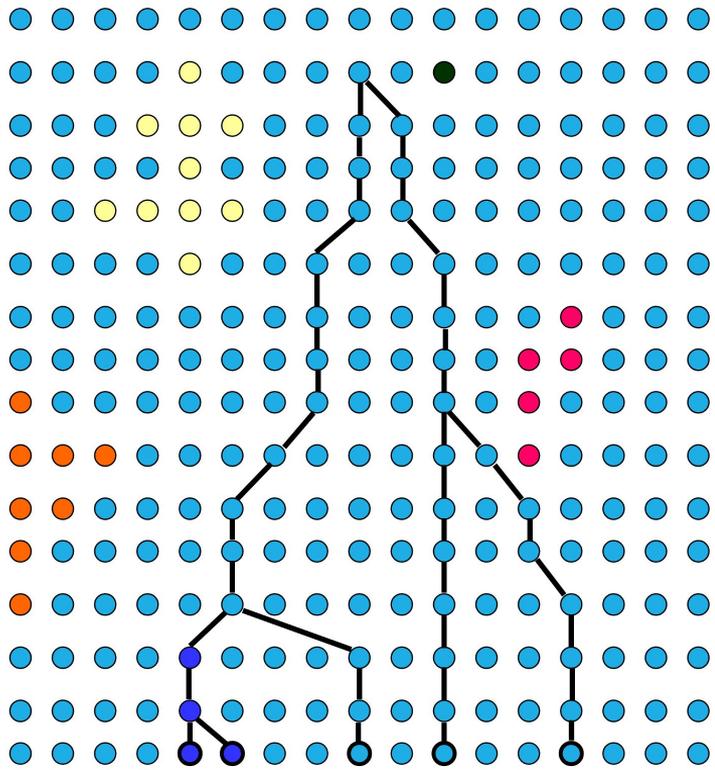
presente



pasado

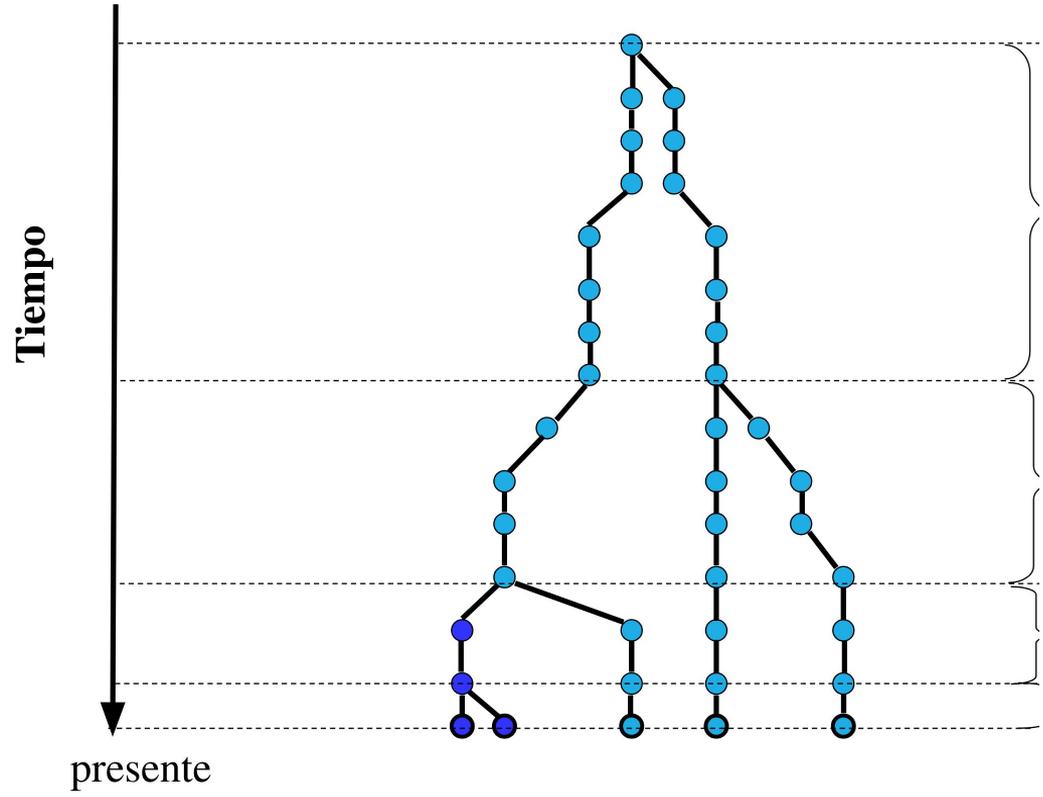
Tiempo

presente



pasado

To exit full screen, press `esc`



# El coalescente

Estudio de las propiedades de una muestra de alelos.

En nuestro análisis: comenzamos con el modelo neutral estándar

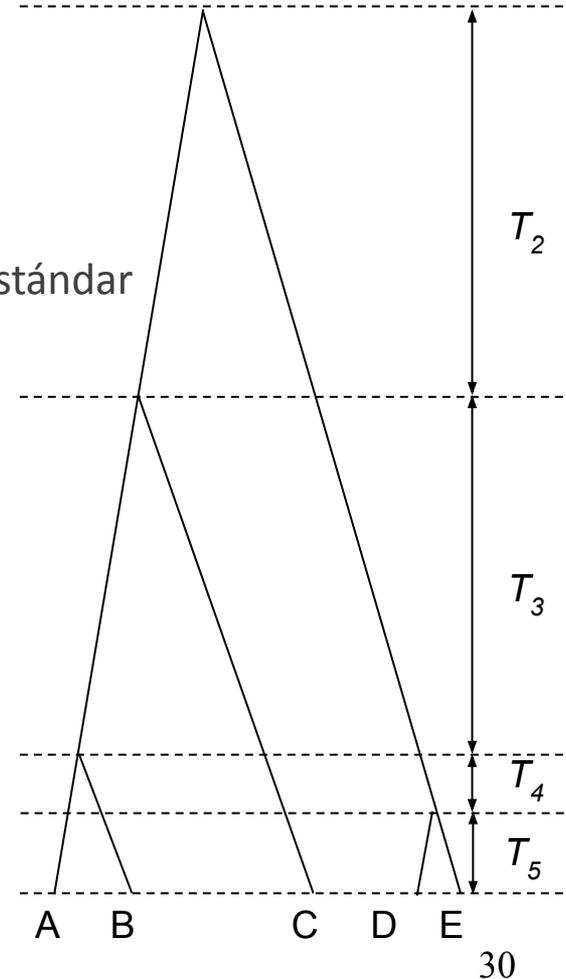
- Población de tipo Wright-Fisher
- Variación neutral en un locus (sin recombinación).
- Modelo mutacional de sitios infinitos.

Aproximación: tamaño de la muestra  $n \ll 2N$ .

Una realización:

**Componentes genealógicos:**

- Topología
- Tiempos entre eventos de coalescencia  $\{T_i\}$

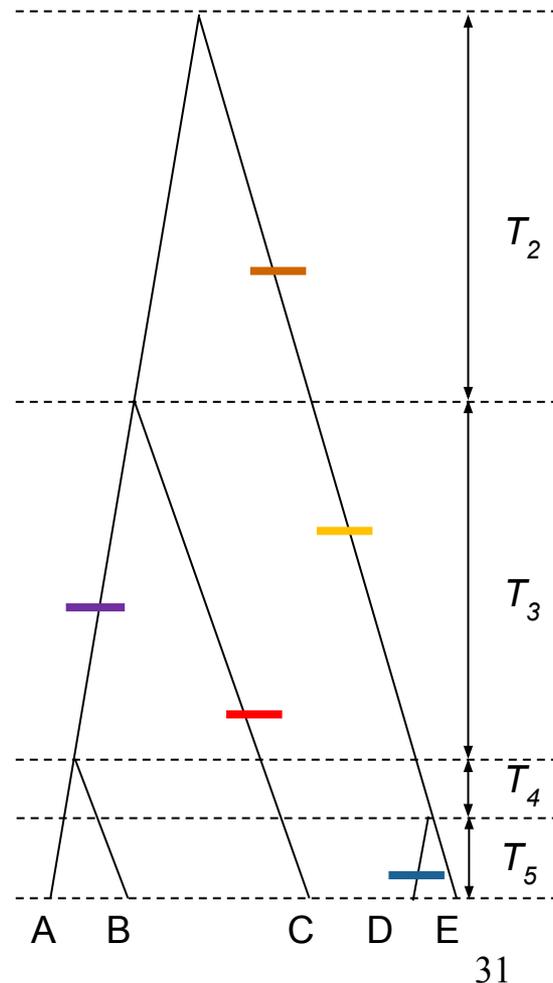


# El coalescente

## Aspectos mutacionales:

- Mutaciones a lo largo de las ramas. La genealogía (topología,  $\{T_i\}$ ) condiciona
- Bajo el modelo de sitios infinitos:
  - El “tamaño” (número de copias en la muestra) de cada mutación es igual al número de alelos terminales en la muestra.
  - El número de sitios variables ( $S$ , “segregantes” en sentido mendeliano) en la muestra es igual al número de mutaciones ( $M$ ) que ocurrieron a lo largo de la genealogía (no hay “homoplasia”: convergencia, paralelismo, reversión). Nota: la situación más simple ( $S=M$ ) de las posibles ( $S \leq M$ ).

Sitio	3	6	9	12	15	18
ancestro	gag	gtc	tac	att	gct	act
A	.g.	...	...	...	...	...
B	.g.	...	...	...	...	...
C	...	a.	...	...	...	...
D	...	...	.g.	.c.	.t.	...
E	...	...	.g.	.c.	...	...



# Teoría neutral: polimorfismos

## Coalescente para pares de alelos ( $n = 2$ ):

- no hay variación topológica
- demografía: variación en  $T_2$
- mutación:  $\mu$

Número esperado de diferencias entre pares de alelos ( $\pi$ )

Probabilidad de coalescencia (por generación):  $1/2N$

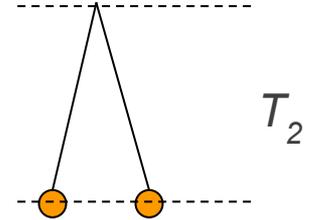
$$E(T_2) = 2N \text{ generaciones}$$

## Modelo neutral estándar

Heterocigosidad esperada:

$$E(H) = \frac{\theta}{\theta + 1}$$

( $2 \times \text{“g”} \times \mu$ ), donde “g” es el número (efectivo) de genes



# Teoría neutral: polimorfismos

## Realizaciones del proceso poblacional en el modelo:

- Genealogía:  $T_2$  tiene una distribución geométrica, de la que se sortea al azar un valor

$$P(T_2 = t) = \left(1 - \frac{1}{2N}\right)^{t-1} \frac{1}{2N}$$

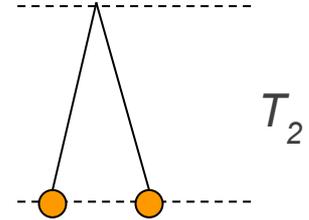
$$E(T_2) = 2N$$

- Mutación: dado  $T_2$ , el número de mutaciones que se asocian a las ramas tiene una distribución binomial ( $2N, \mu$ ), de la que se sortea al azar un valor.

$$E(\pi) = 2 \cdot E(T_2) \cdot \mu$$

$$E(\pi) = 2 \cdot 2N \cdot \mu = 4N\mu$$

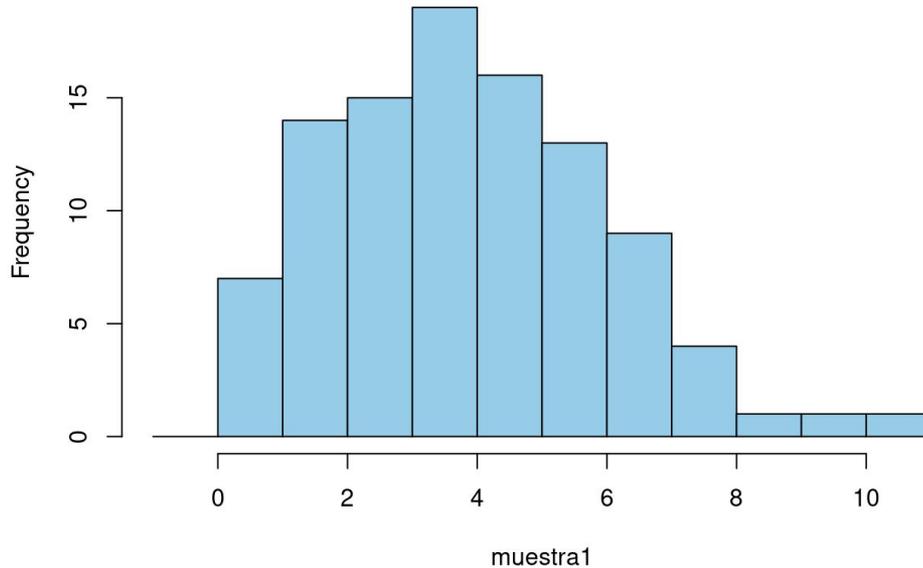
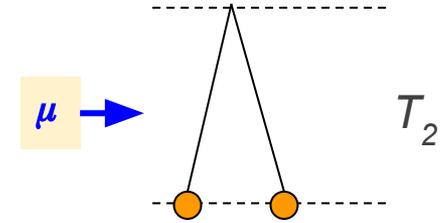
$$E(\pi) = \theta$$



# Realizaciones del proceso

Parámetros:  $N = 10\,000$   $\mu = 1/10\,000$

fijamos  $T_2$  en su valor esperado y muestreamos  $\pi_{ij}$  en 100 réplicas (rbinom)



```
print(mean(muestra1))
```

```
## [1] 4.39
```

```
print(range(muestra1))
```

```
## [1] 1 11
```

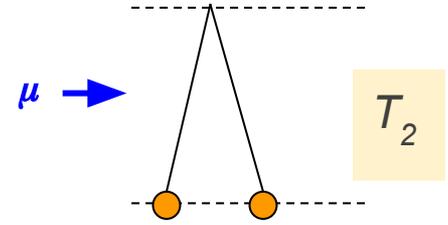
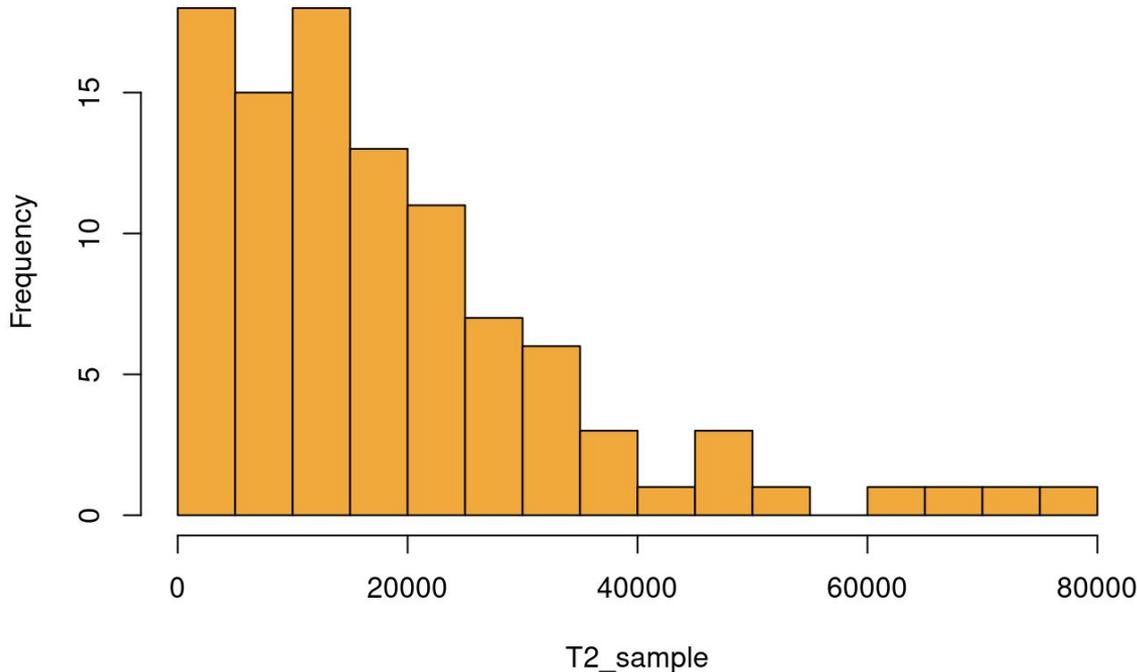
```
print("E(pi) = 4")
```

```
## [1] "E(pi) = 4"
```

# Realizaciones del proceso

Parámetros:  $N = 10\ 000$   $\mu = 1/10\ 000$

100 realizaciones al azar de  $T_2$  (rgeom)



```
print(mean(T2_sample))
```

```
## [1] 18800.86
```

```
print(range(T2_sample))
```

```
## [1] 228 77551
```

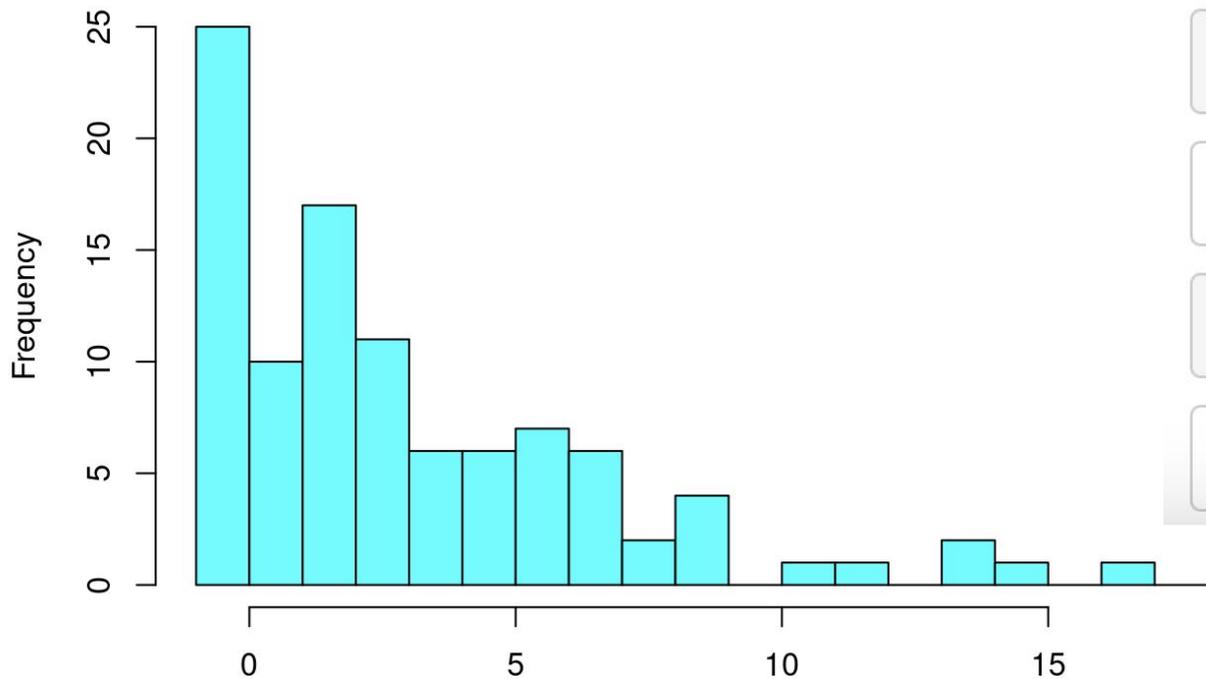
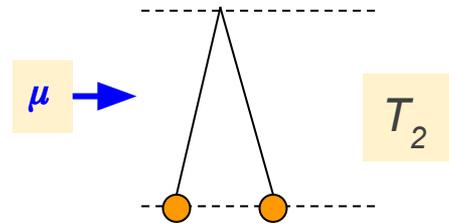
```
print("E(T2) = 20 000")
```

```
## [1] "E(T2) = 20 000"
```

# Realizaciones del proceso

Parámetros:  $N = 10\,000$   $\mu = 1/10\,000$

100 realizaciones al azar de  $T_2$ , mutaciones al azar para  $c/u$



```
print(mean(muestra3))
```

```
## [1] 3.5
```

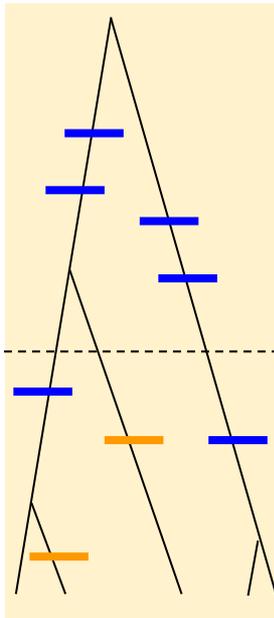
```
print(range(muestra3))
```

```
## [1] 0 17
```

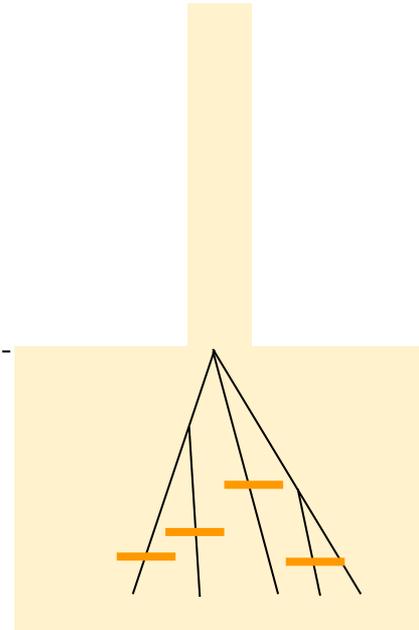
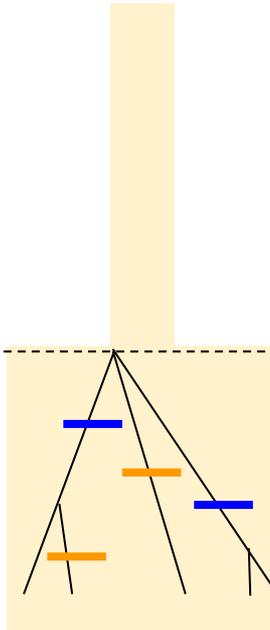
# Desviaciones del modelo neutral estándar

## Cambio demográfico

Tamaño constante

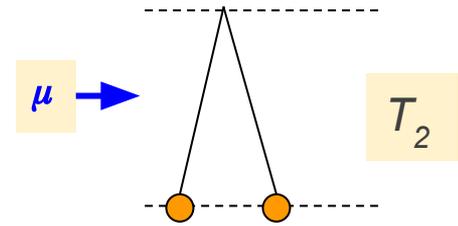


Expansión reciente

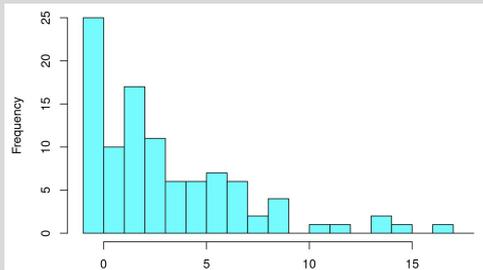
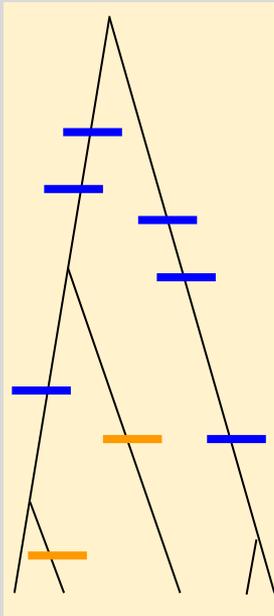


# Desviaciones del modelo neutral estándar

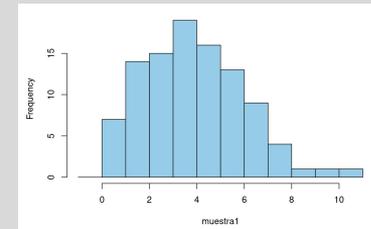
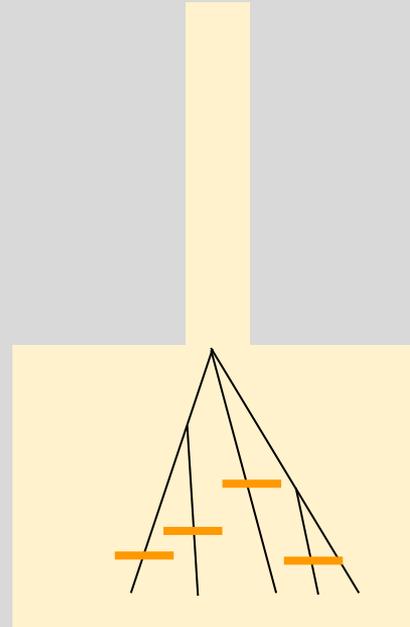
## Cambio demográfico



### Tamaño constante



### Expansión reciente



# ¿Expansión postglacial?

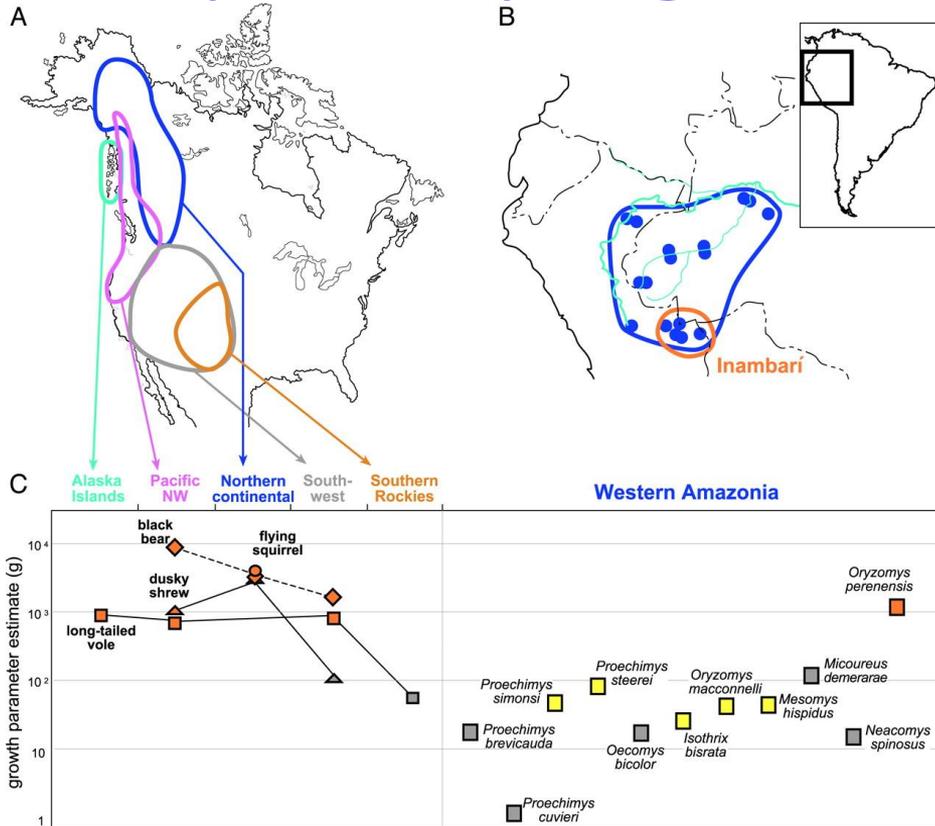


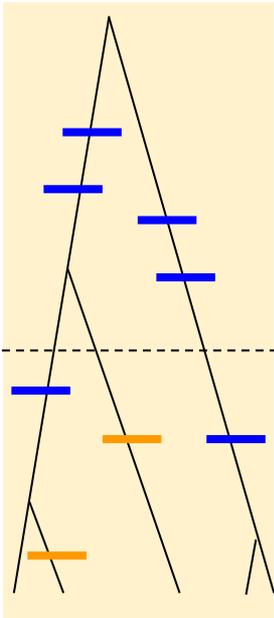
Table 1. Results of Fu's  $F_s$  test (28) for mitochondrial DNA sequences of North American (clades correspond to areas shown in Fig. 1) and Western Amazonian mammals

Area/taxon	Clade	$F_s$	$P$
North America			
	Black bears		
	Pacific Northwest	-2.592	0.022
	Northern Continental	-3.242	0.008
	Southwest	-0.475	0.223
Long-tailed voles	Alaska Islands	-12.600	0.001
	Pacific Northwest	-8.842	0.001
	Southwest	-0.444	0.333
Dusky shrews	Southern Rockies	0.788	0.570
	Pacific Northwest	-9.915	0.001
	Northern Continental	-6.576	0.001
	Southwest	-0.003	0.531
Flying squirrels	Northern Continental	-3.796	0.001
Western Amazonia			
	<i>Proechimys breviceauda</i>	-4.608	0.071
	<i>P. cuvieri</i>	0.865	0.649
	<i>P. simonsi</i>	-7.231	0.081
	<i>P. steerei</i>	-14.326	0.004
	<i>P. bicolor</i>	-2.722	0.161
	<i>Isothrix bisrata</i>	-1.276	0.235
	<i>Oryzomys macconnelli</i>	-2.478	0.075
	<i>Mesomys hispidus</i>	-6.457	0.033
	<i>Micoureus demerarae</i>	-1.364	0.181
	<i>Neacomys spinosus</i>	-6.083	0.008
	<i>Oryzomys perenensis</i>	-21.496	0.001

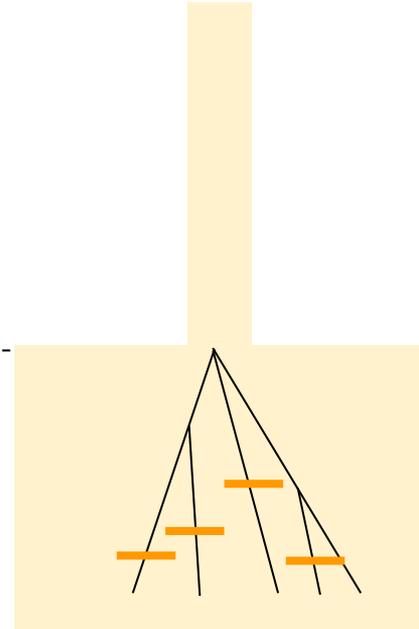
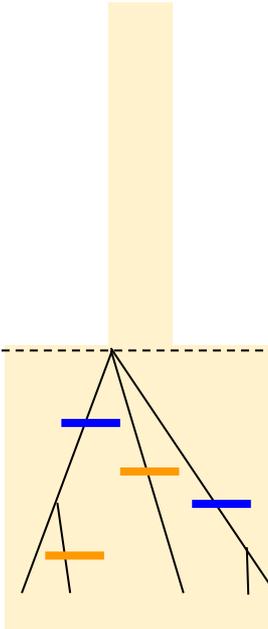
# Desviaciones del modelo neutral estándar

## Barrida selectiva

Locus neutro



Barrida selectiva



# Pruebas (tests) de selección

**Las que usan cambios en la distribución de frecuencias de polimorfismos son sensibles al cambio demográfico y a la selección natural.**

Cambio demográfico: afecta a todos los genes

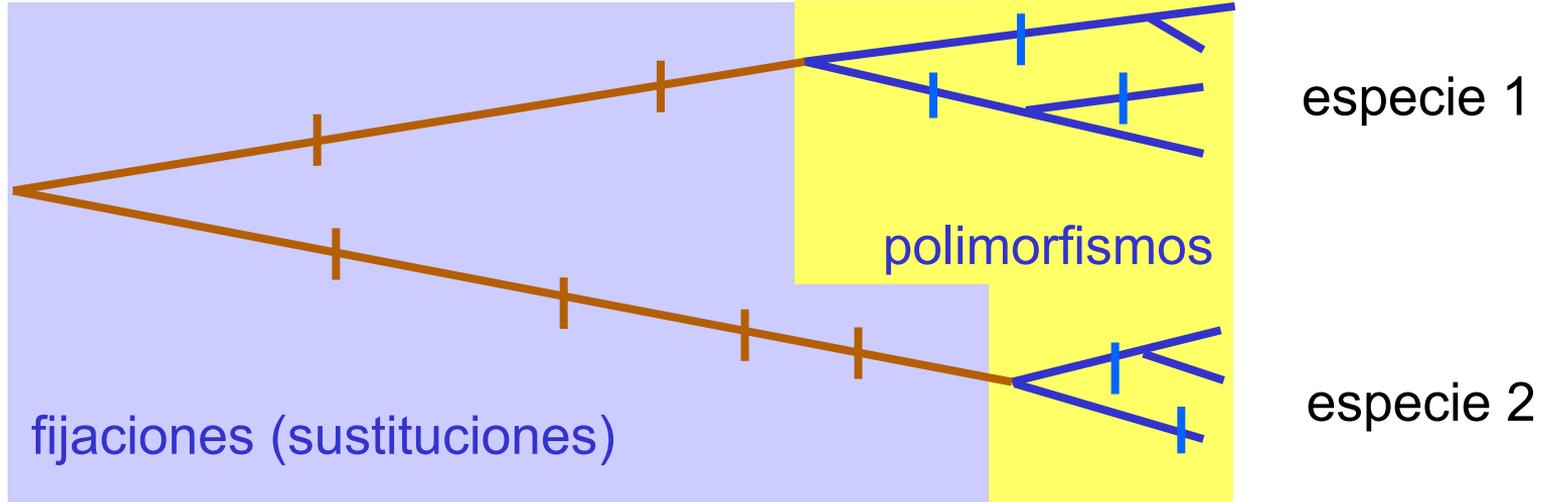
Selección natural: gen a gen

**Las que usan cambios sinónimos y no sinónimos son más sensibles a la selección natural.**

Ejemplos a continuación.

# Adaptación a nivel molecular:

Prueba (test) de McDonald-Kreitman (1991)



g	g	a	c	t a t	especie 1
.	.	.	.	c g .	
a	t	g	t	. . c	especie 2
a	t	g	t	. . .	
a	t	g	t	. . .	

# Adaptación a nivel molecular:

## Prueba de McDonald-Kreitman (1991)

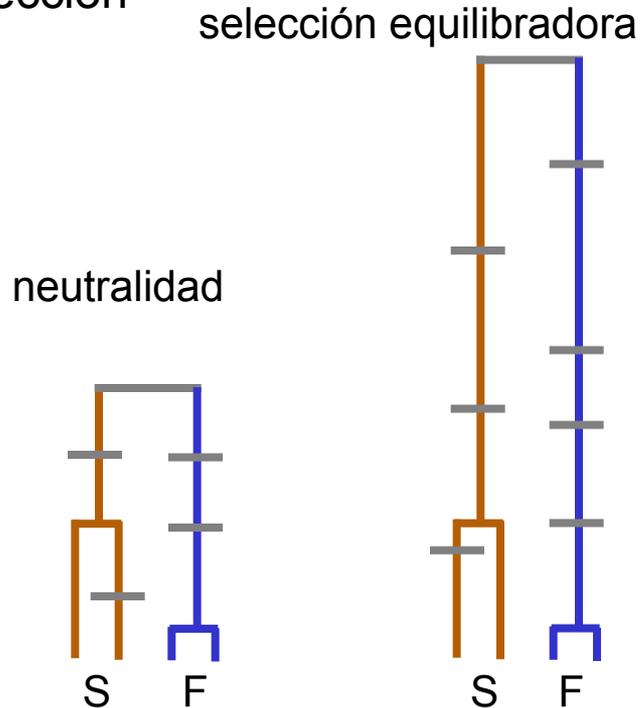
Trabajo original: 3 especies de *Drosophila*  
gen de la enzima ADH

Cambios	Sustituciones	Polimorfismos
No sinónimos	7	2
Sinónimos	17	42

sugieren un exceso de fijaciones no sinónimas  
(selección direccional en la divergencia entre especies)

## Polimorfismos balanceados

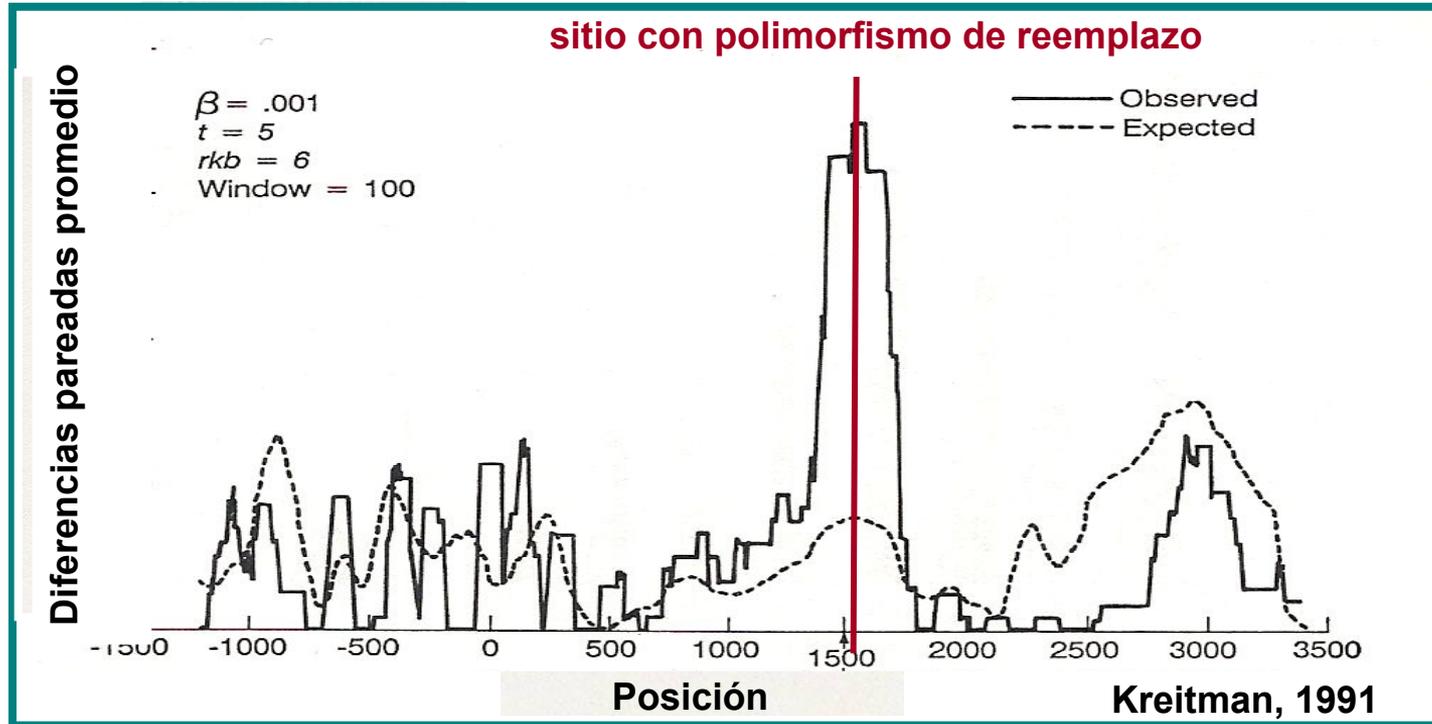
- si son de larga data, se extiende  $T_2$  más allá de lo esperado por azar
- los polimorfismos neutros se asocian a estas ramas largas
- pero la recombinación “borra” ese efecto si nos alejamos del sitio sometido a selección



- exceso de cambios silenciosos
- asociados a diferencias entre las 2 clases alélicas ( $T_2$ )
- en el entorno del codón bajo selección

# Adaptación a nivel molecular

## Polimorfismo en la enzima ADH



11 alelos de la enzima ADH en *D. melanogaster*

# Pruebas (tests) de selección

## Desvíos del equilibrio HWE

Un solo gen no podemos discriminar otras violaciones de modelo HWE

$\omega = dN/dS$ , relación entre cambios de reemplazo y sinónimos

**1) Nivel filogenético.** Si  $\omega > 1$ , es evidencia de selección positiva

Puede ser por rama, por sitio o por rama y sitio

**2) Nivel poblacional / filogenético.** Test de **McDonald-Kreitman**: si  $\omega$  se mantiene constante dentro y entre poblaciones (polimorfismos y sustituciones respectivamente)

## Polimorfismo balanceado

Exceso de mutaciones neutras en torno al polimorfismo de reemplazo

# Síntesis

- como modelo deductivo, la teoría neutralista representa un aporte fundamental para entender cómo ocurre la evolución por mutación, selección purificadora, y deriva
- unifica en un mismo modelo los polimorfismos poblacionales y las sustituciones filogenéticas
- es la base de las hipótesis nulas que, por contraste, permiten detectar desviaciones del neutralismo estricto (evolución cuasi-neutra, selección positiva)
- aún las regiones sobre las que actúa la selección positiva, las mutaciones neutras (o cuasi-neutras) “hacen dedo” y se acumulan a las tasas esperadas

# Síntesis

- existe evidencia creciente del papel de la selección positiva en los polimorfismos y las sustituciones
- los modelos cuasi-neutros involucran regímenes selectivos más diversos (y por ello son más realistas que el neutralismo estricto)
- una visión evolutiva resulta esencial para entender la composición del genoma, la diversidad de distintas regiones, la evolución viral, la evolución de resistencias a herbicidas, antibióticos.....
- ... las aplicaciones en el campo crecen, al igual que se vuelven más evidentes los desafíos que presenta