

Capítulo 6: FUNDAMENTOS DE GENÉTICA DE POBLACIONES

Josefa Cabrero y Juan Pedro M. Camacho

DEPARTAMENTO DE GENÉTICA, FACULTAD DE CIENCIAS. UNIVERSIDAD DE GRANADA.
18071-Granada. E-mails: jcabrero@ugr.es y jpmcamac@ugr.es

La evolución biológica consiste en el cambio en las características hereditarias de grupos de organismos a través de las generaciones. Ocurre como consecuencia de varios procesos fundamentales, tanto aleatorios como no aleatorios. La variación en las características de los organismos de una población se origina a través de la mutación al azar de secuencias de ADN (los genes) que las determinan. Además, la variación genética aumenta por recombinación durante la reproducción sexual, que produce nuevas combinaciones de genes, y también por el flujo génico, es decir, la entrada de nuevos genes desde otras poblaciones. El cambio evolutivo dentro de una población consiste en un cambio en las frecuencias génicas y genotípicas. Los dos principales procesos impulsores del cambio evolutivo son la selección natural y la deriva genética. La selección natural resulta de cualquier diferencia heredable en la tasa de supervivencia o reproducción entre organismos portadores de diferentes alelos o genotipos (diferencias en eficacia biológica). En la mayoría de los casos, las circunstancias ambientales determinan qué variante tiene la mayor eficacia biológica, y una consecuencia común de la selección natural es la adaptación, una mejora en la habilidad media de los miembros de la población para sobrevivir y reproducirse en su ambiente. La deriva genética resulta de la variación al azar en la supervivencia y reproducción de los diferentes genotipos. En la deriva genética, las frecuencias de los alelos fluctúan por puro azar. Ocasionalmente, un alelo reemplazará a los otros (es decir, se fijará en la población). La deriva genética es la fuerza predominante cuando los alelos de un gen son neutros, esto es, cuando no difieren substancialmente en sus efectos sobre la supervivencia o la reproducción, y actúa más intensamente cuanto más pequeña es la población. La deriva genética resulta en cambio evolutivo, pero no en adaptación, aunque en combinación con la selección puede favorecerla.

Introducción

Genes, individuos y poblaciones

El concepto darwiniano de evolución por selección natural es aplicable a entidades que se replican como, por ejemplo, un ser vivo o un algoritmo genético. En la evolución biológica, los genes son los “replicadores” y los individuos son los “vehículos” que utilizan los genes para su transmisión. Los genes, con las modificaciones que son toleradas por la selección natural, tienen continuidad evolutiva, pero los individuos son efímeros.

El estudio de la evolución ha sido facilitado enormemente por la genética de poblaciones, una disciplina que ha desarrollado herramientas matemáticas que permiten predecir cómo evolucionan los genes en las poblaciones de individuos. Las poblaciones, como los genes, tienen continuidad de generación en generación y su constitución genética puede cambiar (evolucionar) a través de las generaciones. Aunque la selección natural actúa con mayor intensidad sobre los niveles jerárquicos inferiores (genes, células y, sobre todo, individuos), los cambios evolutivos son más visibles en los niveles superiores (poblaciones, especies y clados). Por ello, los mecanismos genéticos de evolución alcanzan su máximo sentido cuando son enfocados sobre estos últimos. El nivel poblacional (la genética de poblaciones) es, por ahora, el que ha desa-

rollado las herramientas matemáticas más sofisticadas, mientras que en el nivel macroevolutivo (evolución por encima del nivel de especie) quedan aun muchas incógnitas por resolver (ver Capítulo 5).

Concepto de población

Una población es un grupo local de individuos que pertenece a una especie, dentro de la que ocurren apareamientos real o potencialmente. La especie es la mayor unidad evolutiva. La población es una entidad genética abierta (que puede intercambiar genes con otras poblaciones de la misma especie), mientras que la especie (y las categorías taxonómicas superiores) es una entidad cerrada (que no puede intercambiar genes con otras entidades). El conjunto de informaciones genéticas que llevan todos los miembros de una población, se denomina acervo génico. Para un locus dado, este acervo incluye todos los alelos de dicho gen que están presentes en la población. Puesto que la especie es una entidad genética cerrada, puede decirse que su sino evolutivo está escrito en su acervo, mientras que no ocurre lo mismo con la población.

En genética de poblaciones la atención se centra en la cuantificación de las “frecuencias alélicas y genotípicas” en generaciones sucesivas. Los gametos producidos en una generación dan lugar a los cigotos de la genera-

ción siguiente. Esta nueva generación tiene un acervo génico reconstruido, que puede diferir del de la generación anterior.

Las poblaciones son dinámicas; pueden crecer y expansionarse o disminuir y contraerse mediante cambios en las tasas de nacimiento o mortalidad, o por migración o fusión con otras poblaciones. Esto tiene consecuencias importantes y, con el tiempo, puede dar lugar a cambios en la estructura genética de la población.

Del examen fenotípico de los individuos de una población se deduce la existencia de variaciones fenotípicas, muchas de las cuales provienen de la diversidad genética subyacente. Cuando existe correspondencia biunívoca e inequívoca entre fenotipo y genotipo, podemos calcular las "frecuencias alélicas". Por ello, primero es necesario conocer la naturaleza genética de los caracteres variables, mediante el análisis de los resultados de determinados apareamientos, y entonces podremos definir la población por el número relativo de genotipos de cada clase en los adultos (frecuencias genotípicas) y el número relativo de alelos de cada tipo en los gametos (frecuencias alélicas).

Supongamos un gen con los alelos A y a , y sean n_{AA} , n_{Aa} y n_{aa} los números de individuos con los genotipos AA , Aa y aa , respectivamente, de modo que $n_{AA} + n_{Aa} + n_{aa} = N$, siendo N el nº total de individuos de la población. Si representamos por X , Y y Z las proporciones de los genotipos AA , Aa y aa en la población, las frecuencias genotípicas serán:

$$X_{AA} = \frac{n_{AA}}{N}, \quad Y_{Aa} = \frac{n_{Aa}}{N} \quad \text{y} \quad Z_{aa} = \frac{n_{aa}}{N}, \quad \text{de forma que}$$

$$X_{AA} + Y_{Aa} + Z_{aa} = 1.$$

Las frecuencias alélicas (denominemos p a la frecuencia de A y q a la de a) se calculan a partir del número de individuos de cada genotipo:

$$p = \frac{2n_{AA} + n_{Aa}}{2N} = \frac{n_{AA} + \frac{1}{2}n_{Aa}}{N}, \quad \text{y} \quad q = \frac{2n_{aa} + n_{Aa}}{2N} = \frac{n_{aa} + \frac{1}{2}n_{Aa}}{N},$$

o bien a partir de las frecuencias genotípicas:

$$p = \frac{n_{AA} + \frac{1}{2}n_{Aa}}{N} = \frac{n_{AA}}{N} + \frac{1}{2} \frac{n_{Aa}}{N} = X_{AA} + \frac{1}{2}Y_{Aa}$$

$$q = \frac{n_{aa} + \frac{1}{2}n_{Aa}}{N} = \frac{n_{aa}}{N} + \frac{1}{2} \frac{n_{Aa}}{N} = Z_{aa} + \frac{1}{2}Y_{Aa}$$

Como $p+q=1$, basta calcular p para conocer q : $q=1-p$.

Es decir, "la frecuencia del alelo A se determina sumando la mitad de la frecuencia de heterocigotos a la frecuencia de homocigotos AA ", y "la frecuencia del alelo a es igual a la suma de la mitad de la frecuencia de heterocigotos y la proporción de homocigotos aa " (Apéndice 1).

A nivel genético, el cambio evolutivo en una población puede describirse como cambio en las frecuencias de los diferentes alelos y en las frecuencias genotípicas. Si todos los individuos de una población son genéticamente idénticos para un cierto locus, es decir, son homocigotos para el mismo alelo ($p=1$), no puede haber evolu-

ción en ese locus, ya que las frecuencias alélicas no pueden cambiar de una generación a otra. Por el contrario, si en otra población hay dos alelos en ese locus, sí puede haber cambio evolutivo en ese locus porque la frecuencia de un alelo puede incrementar a expensas de la del otro. Esto supone que cuanto mayor sea el número de loci variables y más alelos haya en cada locus variable, mayor será la probabilidad de que las frecuencias alélicas cambien en algún locus.

Equilibrio Hardy-Weinberg

En 1908, G.H. Hardy y W. Weinberg demostraron, por separado, que "en una población panmíctica (es decir, donde los individuos se aparean al azar), de gran tamaño y donde todos los individuos son igualmente viables y fecundos, el proceso de la herencia, por sí mismo, no cambia las frecuencias alélicas ni genotípicas de un determinado locus" (Apéndice 2). En esencia, el principio de Hardy-Weinberg enuncia que, en ausencia de fuerzas, la descripción del sistema no cambia en el tiempo una vez alcanzado el equilibrio, y que la consecución de éste puede llevar una o más generaciones, dependiendo de las restricciones físicas impuestas por la organización del genoma.

La relación general entre frecuencias alélicas y genotípicas puede describirse en términos algebraicos: si p es la frecuencia de A_1 y q es la de A_2 , se cumple que $p+q=1$ si no existen más que esos dos alelos. Las frecuencias genotípicas de equilibrio vienen dadas por: p^2 (A_1A_1), $2pq$ (A_1A_2), q^2 (A_2A_2). Por ejemplo, si $p=0.6$ y $q=0.4$, las frecuencias genotípicas son: $p^2=0.36$ (A_1A_1), $2pq=0.48$ (A_1A_2), $q^2=0.16$ (A_2A_2). Obsérvese que las frecuencias genotípicas resultan del desarrollo de $(p+q)(p+q)=(p+q)^2=p^2+2pq+q^2$. Con valores cualesquiera de p y q y con apareamiento aleatorio, una generación es suficiente para alcanzar el equilibrio en las frecuencias alélicas y genotípicas (Tabla 1). La Figura 1 muestra las frecuencias genotípicas de equilibrio para distintos valores de las frecuencias alélicas. Obsérvese que la frecuencia de heterocigotos no puede superar el valor 0.5.

En una población en equilibrio, la frecuencia de heterocigotos es relativamente más alta cuanto más raro sea el fenotipo recesivo (Tabla 2). Por ejemplo, una de cada 20,000 personas son albinas, por lo que, suponiendo equilibrio para el carácter, $q^2=1/20,000=0.00005$, y $q=0.007$, y $p=0.993$. En ese caso, $2pq=0.014$. Puesto que $0.014/0.00005=290$, en las poblaciones humanas hay 290 veces más heterocigotos que homocigotos recesivos para el albinismo. Esto es una muestra de la dificultad de eliminar de las poblaciones los caracteres deletéreos recesivos, ya que la mayoría se encuentran en estado heterocigótico inexpressado y contra ellos no puede actuar la selección.

Equilibrio Hardy-Weinberg para alelos múltiples

Para los genes con más de dos alelos se pueden adoptar tres soluciones:

1) Si únicamente estamos interesados en las frecuencias genotípicas determinadas por uno de los alelos, por ejemplo A_1 , podemos considerar las frecuencias de A_1 como p y reunir las frecuencias de los demás alelos del locus (A_2, A_3, \dots, A_n) en una frecuencia q . Las frecuencias de equilibrio serán: p^2 (A_1A_1), $2pq$ ($A_1A_2, A_1A_3, \dots, A_1A_n$), q^2 (A_2A_2, A_3A_3, \dots).

2) Si estamos interesados en los valores de equilibrio de dos alelos, por ejemplo, A_1 y A_2 con frecuencias q_1 y q_2 , las frecuencias genotípicas en el equilibrio serán: $q_1^2 A_1A_1$, $2q_1q_2 A_1A_2$, $q_2^2 A_2A_2$. Como $q_1 + q_2 < 1$, entonces $q_1^2 + 2q_1q_2 + q_2^2 < 1$.

3) Si queremos considerar todos los alelos, las frecuencias genotípicas de equilibrio se obtienen al desarrollar el multinomio correspondiente. Por ejemplo, si en un locus hay 3 alelos posibles, A_1, A_2 y A_3 , con frecuencias p, q y r , de modo que $p+q+r=1$, las frecuencias genotípicas de equilibrio quedan determinadas por el desarrollo del polinomio $(p+q+r)^2$:

$$p^2(A_1A_1) + 2pq(A_1A_2) + 2pr(A_1A_3) + q^2 + (A_2A_2) + 2qr(A_2A_3) + r^2(A_3A_3)$$

Equilibrio Hardy-Weinberg para un gen ligado al sexo

Si el determinismo cromosómico del sexo es XX/XY, el número de genotipos posibles aumenta a 5: A_1A_1, A_1A_2 y A_2A_2 en hembras y A_1 y A_2 en machos. Si p es la frecuencia de A_1 y q la de A_2 , las frecuencias genotípicas de equilibrio son $p^2(A_1A_1)$, $2pq(A_1A_2)$ y $q^2(A_2A_2)$ en las hembras, y $p(A_1)$ y $q(A_2)$ en los machos. En éstos, las frecuencias alélicas coinciden con las frecuencias genotípicas.

Si existe apareamiento aleatorio, las frecuencias genotípicas serán estables tanto en machos como en hembras (Tabla 3), siempre que las frecuencias alélicas (p y q) sean idénticas en ambos sexos. Si hubiese diferencia en las frecuencias alélicas de machos y hembras, la población no estaría en equilibrio. Por ejemplo, en una población con las proporciones $0.2A_1 : 0.8A_2$ en los machos y $0.2A_1A_1 : 0.6A_1A_2 : 0.2A_2A_2$ en las hembras, la frecuencia de A_1 es 0.2 en los machos y 0.5 en las hembras. ¿Cuáles serían las frecuencias de equilibrio de los 5 genotipos? ¿Cuánto se tardaría en alcanzar esas frecuencias?

Como en los machos sólo hay un cromosoma X y en las hembras dos, los machos llevarán 1/3 de todos los alelos de genes ligados al sexo, y las hembras llevarán 2/3. En consecuencia, la frecuencia media de un alelo liga-

Tabla 1

Cálculo de las frecuencias genotípicas en la primera generación de una población donde se dan las condiciones para el equilibrio Hardy-Weinberg y las frecuencias genotípicas parentales son $p^2 A_1A_1, 2pq A_1A_2$ y $q^2 A_2A_2$

Tipo de cruce	Nº de cruces por sexo	Frecuencia relativa	Proporción de descendientes		
			A_1A_1	A_1A_2	A_2A_2
$A_1A_1 \times A_1A_1$	1	$p^2 \times p^2 = p^4$	p^4		
$A_1A_1 \times A_1A_2$	2	$2 \times p^2 \times 2pq = 4p^3q$	$2p^3q$	$2p^3q$	
$A_1A_1 \times A_2A_2$	2	$2 \times p^2 \times q^2 = 2p^2q^2$		$2p^2q^2$	
$A_1A_2 \times A_1A_2$	1	$2pq \times 2pq = 4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
$A_1A_2 \times A_2A_2$	2	$2 \times 2pq \times q^2 = 4pq^3$		$2pq^3$	$2pq^3$
$A_2A_2 \times A_2A_2$	1	$q^2 \times q^2 = q^4$			q^4
TOTAL	9	I= 1	II=p^2	III=$2pq$	IV=q^2

Total I: $p^4 + 4p^3q + 2p^2q^2 + 4p^2q^2 + 4pq^3 + q^4 = (p^2 \text{ factor común de los tres primeros miembros}) = p^2(p^2 + 2pq + q^2) + 2p^2q + p^2q^2 + 4p^2q^2 + 4pq^3 + q^4 = (2pq \text{ factor común de los tres miembros subrayados}) = p^2(p^2 + 2pq + q^2) + 2pq(p^2 + 2pq + q^2) + p^2q^2 + 2pq^3 + q^4 = (q^2 \text{ factor común de los tres últimos miembros}) = p^2(p^2 + 2pq + q^2) + 2pq(p^2 + 2pq + q^2) + q^2(p^2 + 2pq + q^2) = (p^2 + 2pq + q^2) = 1$, Total I = $p^2 + 2pq + q^2 = 1$
 Total II: $p^4 + 2p^3q + p^2q^2 = p^2(p^2 + 2pq + q^2) = p^2$
 Total III: $2p^3q + 2p^2q^2 + 2p^2q^2 + 2pq^3 = 2p^3q + 4p^2q^2 + 2pq^3 = 2pq(p^2 + 2pq + q^2) = 2pq$
 Total IV: $p^2q^2 + 2pq^3 + q^4 = q^2(p^2 + 2pq + q^2) = q^2$

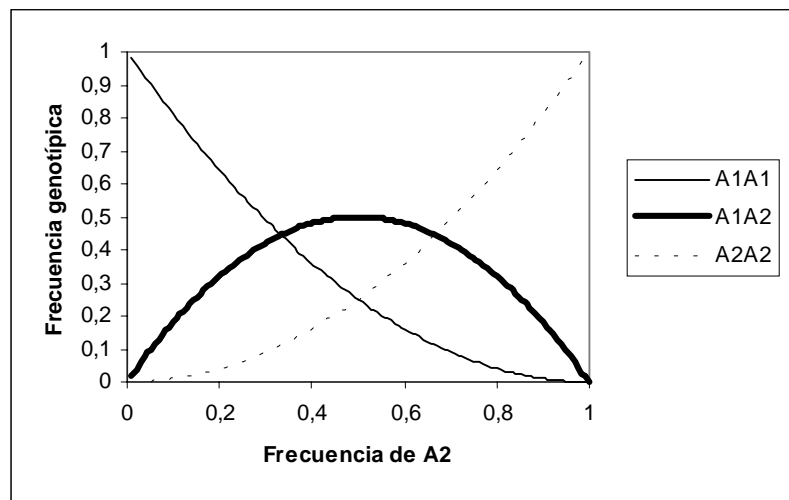


Figura 1. Frecuencias genotípicas predichas por el principio de Hardy-Weinberg, en función de las frecuencias alélicas. Obsérvese que la frecuencia de heterocigotos no puede sobrepasar el valor 0.5.

Tabla 2

Frecuencia relativa de alelos deletéreos enmascarados en los heterocigotos con respecto a la frecuencia de homocigotos recesivos

q	$2pq / q^2$	Observaciones
0.5	2	Hay 2 veces más heterocigotos que homocigotos recesivos.
0.1	18	Hay 18 veces más heterocigotos que homocigotos recesivos.
0.01	198	Hay 198 veces más heterocigotos que homocigotos recesivos.
0.001	1998	Hay 1998 veces más heterocigotos que homocigotos recesivos.

do al sexo en una población reproductora con el mismo número de machos que de hembras, será igual a 1/3 de la frecuencia en machos más 2/3 de la frecuencia en hembras: $p_{eq} = 1/3 p_m + 2/3 p_h = (p_m + 2p_h)/3$.

En el ejemplo anterior, la frecuencia de A_1 en el equilibrio es $p_{eq} = (0.2+2\cdot 0.5)/3 = 0.4$, y la frecuencia de A_2 es 0.6.

Por tanto, en el equilibrio, las frecuencias genotípicas esperadas son:

Hembras: $p^2(A_1A_1) = 0.16$, $2pq(A_1A_2) = 0.48$, $q^2(A_2A_2) = 0.36$

Machos: $p(A_1) = 0.4$, $q(A_2) = 0.6$.

Cuando las frecuencias alélicas difieren entre sexos, el equilibrio no se alcanza en una generación. Como los machos heredan sus cromosomas X únicamente de sus madres, la frecuencia en los machos de un gen ligado al sexo será igual a la frecuencia materna. Por tanto, si en la población inicial la frecuencia materna de A_1 era $p_h = 0.5$, en la siguiente generación $p_m = 0.5$. Las hembras, sin embargo, heredan un X materno y otro paterno, por lo que la frecuencia de un gen ligado al sexo, en ellas, será la media aritmética de las frecuencias parentales: $p_h = (p_m + p_h)/2$. Por ejemplo, si $p_m = 0.2$ y $p_h = 0.5$, entonces $p_h = (0.2+0.5)/2 = 0.35$. Por tanto, después de una generación de apareamiento aleatorio: $p_m = 0.5$ y $p_h = 0.35$.

En la siguiente generación: $p_m = 0.35$ y $p_h = (0.35+0.5)/2 = 0.425$.

En la siguiente: $p_m = 0.425$ y

$p_h = (0.35+0.425)/2 = 0.3875$.

En la 4ª: $p_m = 0.3875$ y $p_h = 0.40625$.

En la 5ª: $p_m = 0.40625$ y $p_h = 0.396875$.

En la 6ª: $p_m = 0.396875$ y

$p_h = 0.4015625$.

Los valores de p se irán aproximando al valor de equilibrio (0.4) a través de las generaciones (Fig. 2). El Apéndice 3 muestra al cabo de cuántas generaciones se alcanzan los valores de equilibrio.

Nótese que la relación entre la frecuencia del fenotipo recesivo ligado al sexo en los machos y en las hembras es q/q^2 , o sea, $1/q$. Si $q = 0.1$, la frecuencia de machos recesivos es 10 veces mayor que la frecuencia de hembras homocigóticas recesivas. Sin embargo, si $q = 0.0001$, $1/q = 10,000$, es decir, la frecuencia de machos recesivos es 10,000 veces mayor. De este modo, cuanto más pequeña es q mayor es, relativamente, la frecuencia del fenotipo recesivo en los machos. Por ejemplo, la hemofilia ligada al sexo tiene una frecuencia fenotípica en los varones de $1/10,000$. En contraste, la frecuencia de mujeres homocigóticas sería 0.0001^2 , o sea, una de cada 100 millones, un valor esperado que está de acuerdo con los poquísimos casos observados de mujeres hemofílicas.

El Apéndice 4 muestra algunos ejemplos prácticos para probar si las poblaciones naturales muestran equilibrio Hardy-Weinberg para loci autosómicos, ligados al sexo y con alelos múltiples.

Tabla 3
Equilibrio Hardy-Weinberg para un gen ligado al sexo

Tipo de cruce	Frecuencia relativa	Proporción de descendientes			
		Hembras		Machos	
		A_1A_1	A_1A_2	A_2A_1	A_2A_2
$A_1A_1 \times A_1A_1$	$p^2 \times p = p^3$	p^3			p^3
$A_1A_2 \times A_1A_1$	$2pq \times p = 2p^2q$	p^2q	p^2q		p^2q
$A_2A_2 \times A_1A_1$	$q^2 \times p = pq^2$		pq^2		pq^2
$A_1A_1 \times A_2A_2$	$p^2 \times q = p^2q$		p^2q		p^2q
$A_1A_2 \times A_2A_2$	$2pq \times q = 2pq^2$		pq^2	pq^2	pq^2
$A_2A_2 \times A_2A_2$	$q^2 \times q = q^3$			q^3	q^3
TOTAL	I=1	II= p^2	III= $2pq$	IV= q^2	V= p VI= q

Total I: $p^3 + 2p^2q + pq^2 + p^2q + q^3 = p(p^2 + 2pq + q^2) + q(p^2 + 2pq + q^2) = p+q=1$
 Total II: $p^3 + p^2q = p^2(p+q) = p^2$
 Total III: $2p^2q + 2pq^2 = 2pq(p+q) = 2pq$
 Total IV: $pq^2 + q^3 = q^2(p+q) = q^2$
 Total V: $p^3 + 2p^2q + pq^2 = p(p^2 + 2pq + q^2) = p$
 Total VI: $p^2q + 2pq^2 + q^3 = q(p^2 + 2pq + q^2) = q$

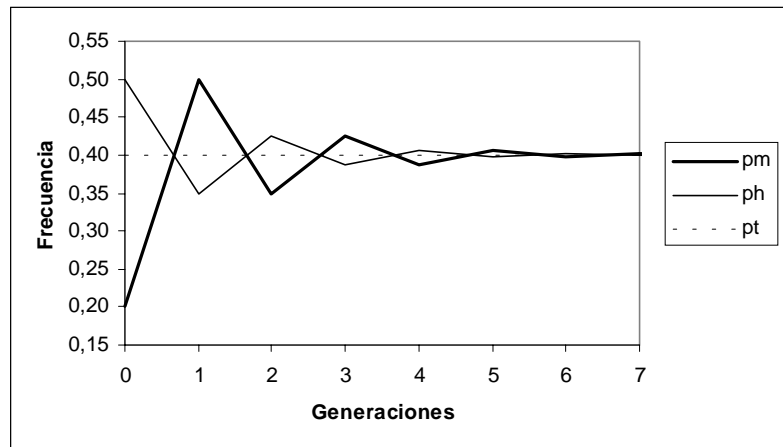


Figura 2. Evolución de frecuencias alélicas desiguales en machos (pm) y hembras (ph) para un gen ligado al sexo. Las diferencias se reducen a la mitad en cada generación y tienden a igualarse en la media ponderada de ambas ($pt = 1/3pm + 2/3ph$).

Equilibrio para dos loci

El alcanzar el equilibrio en las frecuencias genotípicas después de una generación de panmixia es cierto para todos los loci autosómicos al ser considerados separadamente, pero no es cierto cuando se consideran dos o más loci a la vez. Para dos loci, por ejemplo Aa y Bb , el número de genotipos diferentes es 3^2 (Tabla 4).

Tabla 4

Combinaciones genotípicas para dos loci dialélicos (A/a y B/b)			
	AA	Aa	aa
BB	AABB	AaBB	aaBB
Bb	AABb	AaBb	aaBb
bb	AAbb	Aabb	aabb

Supongamos que los dos loci son independientes y que las frecuencias de sus respectivos alelos son $p(A)$, $q(a)$, $x(B)$, $y(b)$. Las frecuencias genotípicas de equilibrio en cada locus, por separado, son:

$p^2(AA)$: $2pq(Aa)$: $q^2(aa)$ para el locus A/a , y $x^2(BB)$: $2xy(Bb)$: $y^2(bb)$ para el locus B/b .

Consideradas juntas, las proporciones de los genotipos bifactoriales en el equilibrio son:

$(p^2+2pq+q^2)(x^2+2xy+y^2) = p^2x^2 + p^22xy + p^2y^2 + 2pqx^2 + 2pq2xy + 2pqy^2 + q^2x^2 + q^22xy + q^2y^2$, que son las frecuencias de los genotipos $AABB$, $AABb$, $AAbb$, $AaBB$, $AaBb$, $Aabb$, $aaBB$, $aaBb$ y $aabb$, respectivamente (Tabla 5). El polinomio anterior corresponde al cuadrado de la suma de las frecuencias gaméticas: $(px+py+qx+qy)^2$. Por tanto, el equilibrio genotípico depende de las frecuencias de equilibrio de los gametos, que son $px(AB)$, $py(Ab)$, $qx(aB)$ y $qy(ab)$, y el problema de alcanzar el equilibrio se reduce a que las frecuencias gaméticas alcancen dichos valores.

Por ejemplo, si imaginamos una población constituida solamente por heterocigotos $AaBb$, las frecuencias alélicas serán $p=q=x=y=0.5$, los cuatro tipos de gametos (AB , Ab , aB y ab) se forman en las frecuencias de equilibrio ($px=py=qx=qy=0.25$) y se alcanza el equilibrio genotípico en la siguiente generación.

Tabla 5
Frecuencias genotípicas para dos loci considerados simultáneamente

	AA p^2	Aa $2pq$	aa q^2
BB x^2	$AABB$ p^2x^2	$AaBB$ $2pqx^2$	$aaBB$ q^2x^2
Bb $2xy$	$AABb$ p^22xy	$AaBb$ $2pq2xy$	$aaBb$ q^22xy
bb y^2	$AAbb$ p^2y^2	$Aabb$ $2pqy^2$	$aabb$ q^2y^2

Sin embargo, esta es la única situación en que se alcanza el equilibrio de forma tan rápida. En el extremo opuesto, supongamos una población constituida únicamente por los genotipos $AABB$ y $aabb$: sólo se formarán dos tipos de gametos (AB y ab) y no puede alcanzarse el equilibrio para todos los genotipos en la siguiente generación, ya que sólo habrá genotipos $AABB$, $AaBb$ y $aabb$, y faltarán 6 genotipos.

Generalmente, si las frecuencias iniciales de los distintos alelos no son iguales, el que los gametos y los genotipos alcancen el equilibrio quedará retrasado cierto número de generaciones. En tales casos, ¿con qué rapidez se alcanza el equilibrio?

Para una mayor generalidad, supongamos que los dos genes están ligados con una fracción de recombinación c (si no hubiera ligamiento bastaría hacer $c=0.5$). En una generación dada podemos admitir que los gametos AB , Ab , aB y ab se producen con unas frecuencias r , s , t y u , respectivamente, con lo que las frecuencias genotípicas, en la generación siguiente, serán las mostradas en la Tabla 6.

Al calcular las frecuencias gaméticas de la nueva generación hay que tener en cuenta la posible recombinación en los loci Aa y Bb , pero sólo tiene efecto en el caso de los

Tabla 6
Frecuencias genotípicas para dos loci tras una generación de apareamiento aleatorio

Ovulos	Espermatozoides			
	AB r	Ab s	aB t	ab u
AB r	$AABB$ r^2	$AABb$ rs	$AaBB$ rt	$AaBb$ ru
Ab s	$AABb$ rs	$AAbb$ s^2	$AaBb$ st	$Aabb$ su
aB t	$AaBB$ rt	$AaBb$ st	$aaBB$ t^2	$aaBb$ tu
ab u	$AaBb$ ru	$Aabb$ su	$aaBb$ tu	$aabb$ u^2

individuos heterocigotos $AaBb$, de los cuales $2ru$ tienen el ligamiento en fase de acoplamiento y $2st$ en fase de repulsión. Teniendo esto en cuenta, podemos calcular la frecuencia r' del gameto AB en la nueva generación:

$$r' = r^2 + rs + rt + 2ru \frac{1}{2}(1-c) + 2st \frac{1}{2}c = r^2 + rs + rt + ru - ruc + stc = r[r+s+t+u] - c(ru-st) = r - c(ru - st)$$

Si llamamos $ru-st=d_0$ y haciendo un cálculo similar para cada gameto, tenemos que las nuevas frecuencias gaméticas son:

$$\begin{aligned} r' &= r - cd_0 \\ s' &= s + cd_0 \\ t' &= t + cd_0 \\ u' &= u - cd_0 \end{aligned}$$

Para que la población esté en equilibrio se ha de cumplir que $c=0$ (ligamiento absoluto, en cuyo caso los dos loci se comportarían a estos efectos como uno solo) o bien que $d_0=0$; es decir, $ru=st$.

Cuando no se dan estas condiciones la población no está en equilibrio, pero tiende hacia él (Apéndice 5), alcanzando las frecuencias gaméticas de equilibrio ($r_\infty = r - d_0$, $s_\infty = s + d_0$, $t_\infty = t + d_0$ y $u_\infty = u - d_0$) que, como vemos, son independientes de c .

Los cálculos realizados anteriormente son válidos también para el caso de factores no ligados, pues basta con hacer $c=0.5$. Sin embargo, es interesante observar que la tendencia hacia el equilibrio es más lenta cuanto más intenso sea el ligamiento (sin llegar a ser absoluto), puesto que $(1-c)^n$ tiende a cero más rápidamente cuando el valor de c es más próximo a 0.5. Como $d_1 = d_0 - d_0c$, el valor de d disminuirá cada generación en una proporción que es precisamente la fracción de recombinación c entre los dos loci. Si los loci son independientes, $d_1 = d_0 - 0.5d_0$, el valor de d se reduce a la mitad en cada generación. Si están ligados y $c=0.1$, $d_1 = d_0 - 0.1d_0$, se reduce un 10% en cada generación. El cambio del valor de d es $\Delta d = d_1 - d_0 = d_0 - d_0c - d_0 = -d_0c$. Por tanto, d se reduce en cada generación una fracción c de su valor.

Cambio en las frecuencias alélicas y genotípicas

Hemos visto anteriormente que una población panmíctica y suficientemente grande se mantiene estable con respecto a las frecuencias alélicas y genotípicas. Sin embargo, existen algunos procesos que cambian estas frecuencias. Son de dos tipos:

- Procesos “direccionales”, que cambian las frecuencias de forma predecible en cantidad y dirección (las poblaciones se consideran de tamaño infinito). Analizaremos los efectos de la migración, la mutación y la selección.
- Procesos “dispersivos”, que se dan en poblaciones pequeñas por efecto de muestreo. Los cambios de frecuencia son predecibles en cantidad, pero no en dirección. Es el caso de la deriva genética y la consanguinidad, aunque esta última puede darse también en poblaciones grandes con apareamiento no aleatorio.

La migración (flujo génico)

Las poblaciones raramente son sistemas cerrados. Por lo regular, se produce cierta cantidad de transferencia de genes, lo cual es más probable cuando las poblaciones se hallan estrechamente relacionadas espacial y genéticamente. Entre las poblaciones adyacentes de una especie el flujo de genes puede ser grande, por lo que es de esperar que las poblaciones contiguas posean una composición génica más semejante que las que están más alejadas geográficamente. Por ello, la migración entre grupos geográficamente aislados es un suceso de gran importancia porque los complejos génicos de adaptación se rompen, las frecuencias alélicas se alteran y, en general, todas las diferencias genéticas entre las poblaciones se reducen.

La eficacia del intercambio de genes depende de la estructura de las dos poblaciones (emigrante y receptora) y, más específicamente, de la cantidad de migración (índice de migración, m) y de la magnitud de la diferencia en frecuencias génicas entre las dos poblaciones.

Para considerar lo anterior en términos cuantitativos, imaginemos un modelo más teórico. Para simplificar el problema, consideremos que la migración efectiva es unidireccional, es decir, que se da desde una población grande (p. ej. continental) a otra menor y parcialmente aislada (p. ej. insular). Supongamos un gen con dos alelos (A_1 y A_2) y un índice de migración m constante del continente a la isla. La frecuencia del alelo A_2 se representa por Q en el continente y por q en la isla.

Se considera constante el valor de Q , porque la fracción que se pierde en cada generación (m) es muy pequeña en comparación con el tamaño de la población continental. Sin embargo, el valor de q cambia en cada generación porque al ser pequeña la población insular, una parte importante de su acervo génico estará formada por inmigrantes y sus descendientes.

En toda generación, una fracción de la población insular es de inmigrantes (m) y el resto de los individuos son endémicos ($1-m$). El valor de q en cada generación es la media ponderada de Q y q , es decir, el promedio de las frecuencias Q y q ponderado por las fracciones relativas de la población insular que representan. Si denominamos q_0 a la frecuencia de A_2 en la generación inicial, la frecuencia en la primera generación (q_1) será:

$$q_1 = q_0(1-m) + Qm = q_0 - q_0m + Qm = q_0 - m(q_0 - Q)$$

El cambio en la frecuencia alélica, Δq , en una generación es:

$$\Delta q = q_1 - q_0 = q_0 - m(q_0 - Q) - q_0 = -m(q_0 - Q)$$

Por tanto, el efecto de la migración sobre las frecuencias alélicas (Δq) depende del índice de migración (m) y de la diferencia genética entre las poblaciones ($q - Q$).

Después de una generación de migración, la diferencia entre las frecuencias alélicas de las dos poblaciones será:

$$q_1 - Q = q_0(1-m) + Qm - Q = q_0(1-m) - Q(1-m) = (1-m)(q_0 - Q)$$

En la siguiente generación:

$$q_2 = q_1(1-m) + Qm = (q_0 - q_0m + Qm)(1-m) + Qm$$

y la diferencia genética entre las dos poblaciones:

$$q_2 - Q = (q_0 - q_0m + Qm)(1-m) + Qm - Q = q_0 - q_0m + Qm - q_0m + q_0m^2 - Qm^2 + Qm - Q = q_0 - 2q_0m + q_0m^2 + 2Qm - Qm^2 - Q = q_0(1 - 2m + m^2) - Q(1 - 2m + m^2);$$

Por tanto,

$$q_2 - Q = q_0(1-m)^2 - Q(1-m)^2 = (1-m)^2(q_0 - Q)$$

Al cabo de n generaciones de migración, la frecuencia alélica vale:

$$q_n = q_{n-1}(1-m) + Qm$$

y la diferencia genética entre las poblaciones es:

$$q_n - Q = (1-m)^n(q_0 - Q)$$

Para conocer otros modelos sobre el flujo génico, consultar Fontdevila y Moya (1999).

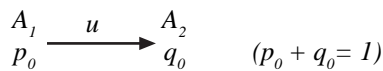
La mutación

Las mutaciones pueden ocurrir una sola vez (mutaciones “no recurrentes”) o bien darse con una cierta tasa (“recurrentes”). Estas últimas pueden ser “reversibles” o “irreversibles”. Las mutaciones génicas y cromosómicas muestran una diferencia fundamental: las primeras suelen ser recurrentes, mientras que las reordenaciones cromosómicas pueden considerarse como un suceso muy raro y probablemente único. Las reordenaciones cromosómicas provienen de dos fracturas que pueden producirse casi al azar a lo largo de los cromosomas, por lo que es muy poco probable que se repita una reordenación en dos ocasiones diferentes y por fracturas precisamente en los mismos puntos. Por eso, las reordenaciones cromosómicas se suelen considerar como sucesos únicos.

Como demostró Fisher (1930), las mutaciones no recurrentes tienen una probabilidad de permanecer en las poblaciones dependiendo, en principio, del número de descendientes que tenga el individuo mutante: a mayor número de descendientes la probabilidad de permanecer será mayor. No obstante, la probabilidad de permanecer disminuye constantemente generación tras generación, lo que indica que una mutación única, sin ventaja selectiva,

no puede producir un cambio permanente en la población. Existen otros mecanismos, además de la selección, que pueden ayudar a las mutaciones no recurrentes a aumentar en frecuencia en las poblaciones. Son los casos del impulso meiótico, que ocurre cuando uno de los alelos de un heterocigoto es transmitido a más del 50% de los gametos, y la deriva genética.

Cuando las mutaciones son recurrentes, podemos modelar los cambios en las frecuencias génicas como sigue. Sea un gen A_1 (con frecuencia p_0) que muta a A_2 (cuya frecuencia representaremos por q_0) con una tasa de mutación u . Suponiendo que no se produce mutación en sentido inverso:



En la generación siguiente:

$$p_1 = p_0 - p_0 u = p_0 (1 - u);$$

$$p_2 = p_1 - p_1 u = p_1 (1 - u); \text{ sustituyendo } p_1 \text{ por su valor:}$$

$$p_2 = p_0 (1 - u) (1 - u) = p_0 (1 - u)^2;$$

En la n -ésima generación:

$p_n = p_0 (1 - u)^n$; tomando logaritmos en esta ecuación, podemos despejar n :

$$L p_n = L p_0 + n L (1 - u); \quad n = \frac{L p_n - L p_0}{L (1 - u)}; \quad n = \frac{L \frac{p_n}{p_0}}{L (1 - u)}$$

Cuando u es muy pequeña, inferior a 10^{-5} , se cumple que el valor de $(1-u)^n$ es aproximadamente igual a e^{-nu} , por lo que:

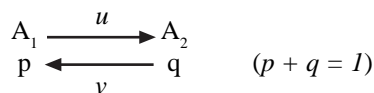
$$p_n = p_0 e^{-nu}$$

De aquí podemos deducir n :

$$L p_n = L p_0 + L e^{-nu} = L p_0 - nu;$$

$$n = (L p_0 - L p_n) / u = (1/u) L (p_0 / p_n)$$

Pero la situación más común es que la mutación recurrente sea reversible: el alelo A_2 también puede mutar a A_1 . Supongamos que A_1 muta a A_2 con una tasa u y que A_2 muta hacia A_1 con una tasa v por generación:



Si en un determinado momento las frecuencias de A_1 y A_2 son p y q , después de una generación la frecuencia de A_1 será:

$$p_1 = p - pu + qv,$$

y el cambio en la frecuencia de A_1 valdrá:

$$\Delta p = p_1 - p = qv - pu$$

Habrá equilibrio cuando $qv = pu$, es decir, cuando el número de alelos A_1 que cambian a A_2 sea el mismo que el número de alelos A_2 que cambian a A_1 .

Puesto que $p + q = 1$, $q = 1 - p$, por lo que

$$(1 - p) v = pu; \quad v - vp = pu; \quad v = pu + vp;$$

$$v = p (u + v); \quad p_{eq} = \frac{v}{u + v}$$

Igualmente, $p = 1 - q$

$$qv = (1 - q) u; \quad qv = u - uq;$$

$$u = qv + uq = q (u + v); \quad q_{eq} = \frac{u}{u + v}$$

Si la población no está en equilibrio, veamos cómo tiende hacia él:

$$\Delta p = qv - pu = (1-p) v - pu = v - pv - pu = \underline{v} - p (u + v);$$

Sustituimos \underline{v} por su valor en términos de p_{eq} :

$$\Delta p = p_{eq} (u + v) - p (u + v) = (p_{eq} - p) (u + v)$$

Es decir, Δp es proporcional a la distancia ($p_{eq} - p$) a que se encuentra la población del punto de equilibrio. El número de generaciones (n) para que la población alcance una determinada frecuencia (p_n) viene determinado por la siguiente expresión:

$$n = \frac{1}{u + v} L \frac{p_0 - p_{eq}}{p_n - p_{eq}} \quad (\text{Apéndice 6})$$

La selección

El papel de la selección radica en que los distintos individuos de una población difieren en viabilidad y/o fecundidad y, por lo tanto, contribuyen con número diferente de descendientes (y por tanto de genes) a la siguiente generación (ver Capítulo 7).

Se llama "eficacia biológica" (w) de un individuo a la proporción relativa de descendientes con que contribuye a la siguiente generación. La eficacia biológica es difícil de medir. En esencia, bastaría con contar el número de descendientes producidos por un individuo y compararlo con los producidos por el resto de los individuos de la población. Pero esto no es fácil de llevar a la práctica (ver Endler, 1986), por lo que se recurre al estudio de sus componentes, tales como la viabilidad, el éxito en el apareamiento, el número de hijos por camada, etc. Por ejemplo, en una población de insectos se contabilizaron 4050 larvas de genotipo $A_1 A_1$, 5860 $A_1 A_2$ y 2035 $A_2 A_2$. Varios meses después, en estado adulto, se encontraron 3550 $A_1 A_1$, 4305 $A_1 A_2$ y 1240 $A_2 A_2$. A partir de las tasas relativas de supervivencia de cada genotipo, se pueden calcular sus viabilidades relativas (Tabla 7).

Se llama "coeficiente de selección" (s) a la reducción proporcional en la eficacia biológica de cierto genotipo, en comparación con otro genotipo que se toma como patrón, y que suele ser el más favorecido por la selección. En el ejemplo anterior, $s=0.16$ para $A_1 A_2$ y $s= 0.3$ para $A_2 A_2$. La relación entre s y w es sencilla: $w= 1-s$, y $s= 1-w$.

Cambios en frecuencia producidos por la selección

En la mayoría de las plantas y animales superiores la selección tiene lugar primariamente en el estado diploide. En este caso existen tres genotipos posibles para una sola diferencia génica (A_1A_1 , A_1A_2 y A_2A_2), de modo que la eficacia de la selección depende, entre otras cosas, del grado de dominancia en eficacia biológica (h). Supongamos que la selección actúa contra el alelo A_2 . El grado de dominancia afectará solamente a la eficacia biológica del heterocigoto, por lo que los valores de eficacia biológica de los tres genotipos serán 1, $1-sh$ y $1-s$ para A_1A_1 , A_1A_2 y A_2A_2 , respectivamente (Tabla 8). Si h vale entre 0 y $1/2$, se trata de recesivos incompletos, y si h está comprendida entre $1/2$ y 1 serán dominantes incompletos.

La selección natural es la explicación más plausible para muchos caracteres adaptativos de los organismos. El principal supuesto inherente a la teoría de la selección natural es que algunos tipos hereditarios de una población tienen una cierta ventaja sobre otros para adaptarse, por su mayor supervivencia y/o capacidad reproductora, a los cambios ambientales. La consecuencia primaria de la selección es el cambio de las frecuencias génicas, pudiendo admitirse que en grandes poblaciones la selección es probablemente la fuerza más importante responsable del cambio de las frecuencias génicas.

El tratamiento matemático de la selección se debe principalmente a Haldane (1924-1932), Fisher (1930) y Wright (1931). El modelo general de la selección, teniendo en cuenta el grado de dominancia, se desarrolla partiendo de las frecuencias genotípicas iniciales que, al ser multiplicadas por sus correspondientes valores de eficacia biológica, darán las frecuencias resultantes de la acción de la selección. La suma de estas frecuencias constituye la eficacia biológica media de la población (\bar{w}), y a partir del valor relativo de la frecuencia de cada genotipo, con respecto a \bar{w} , podremos calcular las frecuencias alélicas de la siguiente generación (Tabla 9).

La eficacia biológica media de la población será: $\bar{w} = p^2 + 2pq - 2pqsh + q^2 - sq^2 = 1 - 2pqsh - sq^2$, y la nueva frecuencia de A_2 es:

$$q_1 = \frac{pq(1-sh) + q^2(1-s)}{1-2pqsh-sq^2} = \frac{pq-pqsh+q^2-sq^2}{1-2pqsh-sq^2} = \frac{q(p+q)-pqsh-sq^2}{1-2pqsh-sq^2} = \frac{q-pqsh-sq^2}{1-2pqsh-sq^2}$$

El cambio en la frecuencia de A_2 (Δq) es igual a q después de la selección (q_1) menos q antes de la selección (q):

Tabla 7

Cálculo de la eficacia biológica (medida por la viabilidad desde larva hasta adulto) y el coeficiente de selección en una población de insectos

	A_1A_1	A_1A_2	A_2A_2
Frecuencia en larvas	4050	5860	2035
Frecuencia en adultos	3550	4305	1240
Tasa de supervivencia (frecuencia en adultos /frecuencia en larvas)	0.88	0.73	0.61
Eficacia biológica relativa (w) (dividiendo por la mayor tasa de supervivencia)	1	0.84	0.70
Coeficiente de selección ($s = 1-w$)	0	0.16	0.30

Tabla 8

Selección en función de la dominancia

Grado de dominancia	Eficacia biológica			
	A_1A_1	A_1A_2	A_2A_2	
	1	$1-sh$	$1-s$	
$h=0$	1	1	$1-s$	Desventaja recesiva de A_2
$h=1$	1	$1-s$	$1-s$	Desventaja dominante de A_2
$h=1/2$	1	$1-1/2s$	$1-s$	Desventaja aditiva de A_2

Tabla 9

Efecto de una generación de selección sobre las frecuencias genotípicas.

$\bar{w} =$ Eficacia biológica media

Genotipos	A_1A_1	A_1A_2	A_2A_2	Total
Frecuencia inicial	p^2	$2pq$	q^2	1
Eficacia biológica	1	$1-sh$	$1-s$	
Frecuencia después de la selección	p^2	$2pq(1-sh)$	$q^2(1-s)$	$\bar{w} = 1 - 2pqsh - sq^2$

$$\begin{aligned} \Delta q = q_1 - q &= \frac{q - pqsh - sq^2 - q + 2pq^2sh + sq^3}{1 - 2pqsh - sq^2} = \\ &= \frac{-spqh - sq^2 + 2spq^2h + sq^3}{1 - 2pqsh - sq^2} = \frac{-spqh + 2spq^2h - spq^2}{1 - 2pqsh - sq^2} = \\ &= \frac{-spq(h - 2qh + q)}{1 - 2pqsh - sq^2} = \frac{-spq[q + h(1 - 2q)]}{1 - 2pqsh - sq^2} = \\ &= \frac{-spq[q + h(p - q)]}{1 - 2pqsh - sq^2} \end{aligned}$$

Para valores pequeños de s , el denominador es próximo a uno, por lo que el cambio en frecuencia es $\Delta q = -spq[q + h(p - q)]$. Se observa que i) la selección sólo actúa si existe variabilidad genética ($pq \neq 0$), ii) el cambio en frecuencia tiene signo negativo y conduce finalmente a la desaparición del alelo desfavorable, iii) el proceso está modulado por la frecuencia inicial (q en el corchete), por la magnitud del coeficiente de selección en contra de los homocigotos (s) y por la fracción de este coeficiente que se manifiesta en los heterocigotos (h , siempre que $p \neq q$).

Los efectos de la selección dependen del grado de dominancia

Si la desventaja de A_2 es recesiva ($h=0$), la eficacia biológica media de la población valdrá $\bar{w} = 1 - 2pqsh - sq^2 = 1 - sq^2$, la frecuencia de A_2 en la generación siguiente será

$$q_1 = \frac{q - pqsh - sq^2}{1 - 2pqsh - sq^2} = \frac{q - sq^2}{1 - sq^2} = \frac{q(1 - sq)}{1 - sq^2}$$
, y el cambio en frecuencia:

$$\Delta q = \frac{-spq[q + h(p - q)]}{1 - 2pqsh - sq^2} = \frac{-spq^2}{1 - sq^2}$$

De la misma forma, si la desventaja de A_2 es dominante ($h=1$):

$$\bar{w} = 1 - s + sp^2, \quad q_1 = \frac{q(1 - s)}{1 - s + sp^2} \quad \text{e} \quad \Delta q = \frac{-sp^2q}{1 - s + sp^2}$$

Si la desventaja de A_2 fuese aditiva ($h=1/2$):

$$\bar{w} = 1 - sq, \quad q_1 = \frac{q - 1/2sq(1 + q)}{1 - sq} \quad \text{e} \quad \Delta q = \frac{-1/2spq}{1 - sq}$$

Para cualquier tipo de acción génica, la selección es más eficaz, dado un valor de s , para valores intermedios de q , puesto que con valores de q próximos a 1 ó a 0, Δq es pequeño. La Figura 3 muestra la evolución de la frecuencia de un alelo (A_2), que confiere una desventaja selectiva ($s = 0.1$) y cuya frecuencia es inicialmente $q = 1$. Se observa que su frecuencia disminuye rápidamente si la desventaja de A_2 es recesiva o aditiva, pero tarda en disminuir si la desventaja de A_2 es dominante. Para comprender esto, pensemos en una población con elevada frecuencia de A_2 . El genotipo más frecuente será A_2A_2 y, en menor medida A_1A_2 , siendo los homocigotos A_1A_1 muy raros. Si la desventaja de A_2 es recesiva, los valores de eficacia biológica de los tres genotipos serán 1, 1 y $1-s$ para A_1A_1 , A_1A_2 y A_2A_2 , respectivamente. Como los dos genotipos predominantes en la población tienen diferente eficacia biológica, la selección actuará intensamente desde el primer momento, por lo que el cambio en la frecuencia de A_2 es grande desde el principio. En el caso aditivo, los tres genotipos difieren en eficacia biológica, por lo que el cambio en frecuencia siempre es rápido. Pero si la desventaja de A_2 es dominante, los valores de eficacia biológica serán 1, $1-s$ y $1-s$, de forma que los dos genotipos predominantes en la población, cuando q sea muy elevada (A_2A_2 y A_1A_2), tendrán la misma eficacia biológica y habrá poco sustrato para la selección hasta que aparezcan los homocigotos A_1A_1 que tienen la mayor eficacia biológica. En la Figura 3 se observa también que los máximos de-

crementos de q se dan pronto en los casos recesivo y aditivo, y tarde en el dominante. En cada caso, el mayor cambio se produce cuando la población alcanza la frecuencia máxima de heterocigotos.

Por un razonamiento similar, se deduce que la selección elimina más rápido los alelos deletéreos raros cuando son dominantes que cuando son recesivos (ver apartado siguiente), e incrementa más rápido la frecuencia de un nuevo alelo beneficioso cuando es dominante, como se observa frecuentemente con los genes de resistencia a insecticidas.

Tiempo necesario para los cambios en frecuencia producidos por la selección

Cuando s es pequeño, la eficacia biológica media valdrá prácticamente 1, por lo que las ecuaciones anteriores, que miden el cambio en frecuencia por selección, pueden resumirse al numerador simplemente:

$$\Delta q = -sq^2(1 - q), \quad \text{si la desventaja de } a \text{ es recesiva,}$$

$$\Delta q = -sq(1 - q)^2, \quad \text{si es dominante, y}$$

$$\Delta q = -\frac{1}{2}sq(1 - q), \quad \text{si la acción génica es aditiva.}$$

Además, el fenómeno se puede tratar como un proceso continuo, por lo que para averiguar el número de generaciones necesarias para que se produzca un determinado cambio en la frecuencia alélica, desde un valor inicial q hasta el valor q_n , podemos expresar las ecuaciones anteriores en forma diferencial. Si la desventaja es recesiva,

$$\frac{dq}{dt} = -sq^2(1 - q), \quad \text{por lo que} \quad \frac{1}{q^2(1 - q)}dq = -sdt.$$

Integrando esta ecuación diferencial entre las generaciones 0 y n (Apéndice 7a) obtenemos:

$$n = \frac{1}{s} \left[\frac{1}{q_n} - \frac{1}{q_0} + L \frac{q_0(1 - q_n)}{q_n(1 - q_0)} \right].$$
 Usando el mismo procedi-

$$\text{miento, obtenemos } n = \frac{1}{s} \left[\frac{1}{1 - q_0} - \frac{1}{1 - q_n} + L \frac{q_0(1 - q_n)}{q_n(1 - q_0)} \right], \text{ si}$$

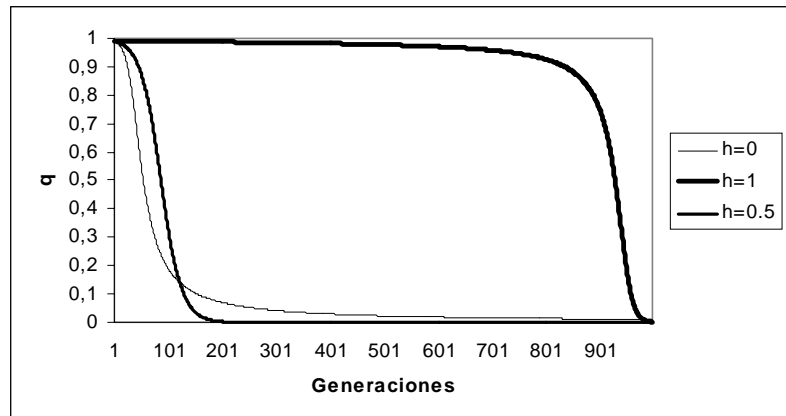


Figura 3. Evolución de la frecuencia de un alelo desfavorecido por la selección en función de que su acción génica sea recesiva ($h = 0$), dominante ($h = 1$) o aditiva ($h = 0.5$).

la desventaja es dominante (Apéndice 7b),

y $n = \frac{2}{s} L \frac{q_0(1-q_n)}{q_n(1-q_0)}$, si la acción génica es

aditiva (Apéndice 7c).

La Tabla 10 muestra un análisis comparativo del tiempo requerido para que la selección ($s=0.1$) efectúe un cambio en la frecuencia de A_2 desde $q_0=0.5$ hasta diferentes valores de q_n . La selección conduce al mismo resultado final en los tres casos (la eliminación del alelo desfavorable), pero el tiempo necesario es muy superior si éste es recesivo. Obsérvese que el cambio en la frecuencia del alelo desfavorable es $-spq^2$ en el caso recesivo y $-spq/2$ en el aditivo, y que basta que $q<0.5$ para que el segundo sea mayor que el primero.

Selección contra un letal recesivo

Un interesante caso particular es cuando el homocigoto recesivo (A_2A_2) es letal ($s=1$) (Tabla 11). La frecuencia de A_2 tras una generación de selección es:

$$q_1 = \frac{pq}{p^2 + 2pq} = \frac{pq}{p(p+2q)} = \frac{q}{p+q+q} = \frac{q}{1+q}$$

$$q_2 = \frac{q_1}{1+q_1} = \frac{\frac{q}{1+q}}{1+\frac{q}{1+q}} = \frac{q}{1+q+q} = \frac{q}{1+2q}$$

$$q_n = \frac{q}{1+nq}$$

De aquí podemos despejar n :

$$q_n = \frac{q}{1+nq}; \quad q_n(1+nq) = q; \quad q_n + nqq_n = q;$$

$$n = \frac{q - q_n}{qq_n} = \frac{1}{q_n} - \frac{1}{q}$$

El cambio en la frecuencia alélica en la primera generación será:

$$\Delta q = q_1 - q = \frac{q}{1+q} - q = \frac{q - q - q^2}{1+q} = \frac{-q^2}{1+q}$$

y el cambio en la generación n :

$$\Delta q_n = q_n - q_{n-1} = \frac{q}{1+nq} - \frac{q}{1+(n-1)q} =$$

$$= \frac{q[1+(n-1)q] - q(1+nq)}{(1+nq)[1+(n-1)q]} = \frac{q + (n-1)q^2 - q - nq^2}{(1+nq)[1+(n-1)q]} =$$

$$= \frac{nq^2 - q^2 - nq^2}{(1+nq)[1+(n-1)q]} = \frac{-q^2}{(1+nq)[1+(n-1)q]}$$

Tabla 10

Tiempo requerido para que la selección ($s=0.1$) efectúe un cambio en la frecuencia de A_2 desde $q_0=0.5$ hasta diferentes valores de q_n

Cambio en frecuencia		Nº de generaciones si la desventaja de A_2 es		
q_0	q_n	Recesiva	Dominante	Aditiva
0.5	0.1	102	31	44
0.5	0.01	1026	56	92
0.5	0.001	10049	79	138
0.5	0.0001	100072	102	184
0.5	0.00001	1000095	125	230

Tabla 11

Selección contra un gen letal recesivo

Genotipos	A_1A_1	A_1A_2	A_2A_2	Total
Frecuencia Inicial	p^2	$2pq$	q^2	1
Eficacia biológica	1	1	0	
Frecuencia después de la selección	p^2	$2pq$	0	$\bar{w} = p^2 + 2pq$

Tabla 12

Selección gamética

	Genotipos		
	A_1	A_2	Total
Frecuencia inicial	p	q	1
Eficacia biológica	1	$1-s$	
Frecuencia tras selección	p	$q(1-s)$	$\bar{w} = p+q-sq = 1-sq$

Por lo tanto, q_n disminuye tanto más deprisa cuanto mayor sea el valor inicial de q y menor el número de generaciones transcurridas. No obstante, como es de suponer que los letales recesivos muestren una frecuencia baja en las poblaciones naturales, su eliminación es un proceso lento.

Selección gamética

La selección puede tener lugar también en el estado haploide, como es el caso de algunos organismos que son haploides en todas o alguna de sus fases de vida, y en los gametos de los organismos diploides. Si la selección se produce en los gametos, la dominancia no afecta y todos los genotipos se expresan fenotípicamente. Por tanto, para un gen dialélico, puede considerarse que la población consta de genotipos A_1 con una frecuencia p y genotipos A_2 con una frecuencia q . Si se considera que la eficacia biológica de A_1 es 1 y la de A_2 se ve reducida por un coeficiente de selección s , una generación de selección causará los siguientes cambios mostrados en la Tabla 12. La frecuencia total de individuos que resisten a la selección ($1-sq$) constituye la eficacia biológica media de la pobla-

ción, y las frecuencias alélicas de la generación siguiente serán: $p_1 = p / (1-sq)$ y $q_1 = q (1-s) / (1-sq)$.

El cambio en la frecuencia de A_2 :

$$\begin{aligned} \Delta q &= q_1 - q = \frac{q(1-s)}{1-sq} - q = \frac{[q(1-s)] - [q(1-sq)]}{1-sq} = \\ &= \frac{q - sq - q + sq^2}{1-sq} = \frac{-sq + sq^2}{1-sq} = \frac{-sq(1-q)}{1-sq} \end{aligned}$$

Si la selección es débil, s es pequeño y sq del denominador puede ser eliminado, por lo que $\Delta q = -sq(1-q)$.

A partir de aquí se puede obtener el número de generaciones de selección que han de transcurrir para un determinado cambio de las frecuencias alélicas, siguiendo el mismo procedimiento matemático que en casos anteriores:

$$n = \frac{1}{s} \left(L \frac{q_0(1-q_n)}{q_n(1-q_0)} \right)$$

Obsérvese, comparando las ecuaciones resultantes, que este modelo es equivalente al caso aditivo de la selección cigótica, para valores pequeños de s .

Dos escalas para medir la selección

La selección natural puede formularse como una fuerza negativa, tal como hemos hecho hasta ahora, o bien como una fuerza positiva. Para un gen dialélico, cuando la selección actúa en perjuicio de uno de los alelos, a la vez, está actuando en beneficio del otro alelo. Cualquier disminución en la frecuencia del alelo A_2 irá acompañada del incremento en la frecuencia de A_1 . Puesto que las formulaciones positiva y negativa producen dos escalas diferentes, sería interesante conocer la equivalencia entre los coeficientes de selección en una y otra. Por ejemplo, ¿cuál sería la relación entre el coeficiente de selección responsable de la desventaja dominante de A_2 (s_1), que implica valores de eficacia biológica iguales a 1, $1-s_1$ y $1-s_1$ para los genotipos A_1A_1 , A_1A_2 y A_2A_2 , y el coeficiente de selección responsable de la ventaja recesiva de A_1 (s_2), que implica valores de eficacia biológica iguales a $1+s_2$, 1 y 1, respectivamente?. En la escala negativa, los valores de eficacia biológica se expresan en relación al genotipo con mayor eficacia biológica (A_1A_1). En la escala positiva, la eficacia biológica de A_1A_1 está expresada con respecto al valor 1 de los otros genotipos. Podemos hacer que la eficacia biológica de A_1A_2 y A_2A_2 sea 1 también en la escala negativa, dividiendo las eficacias biológicas de los tres genotipos (1, $1-s_1$ y $1-s_1$) entre $1-s_1$, y pasarían a ser: $1/(1-s_1)$, 1 y 1. Está claro que la eficacia biológica de A_1A_1 en la escala positiva ($1+s_2$) es equivalente a $1/(1-s_1)$ en la escala negativa. Partiendo de esa igualdad, podemos deducir la equivalencia entre los coeficientes de selección en las dos escalas:

$$\begin{aligned} 1+s_2 &= \frac{1}{1-s_1}; \text{ de forma que } 1-s_1 = \frac{1}{1+s_2}; \text{ y} \\ s_1 &= 1 - \frac{1}{1+s_2} = \frac{s_2}{1+s_2}. \end{aligned}$$

$$\text{De igual modo, } s_2 = \frac{1}{1-s_1} - 1 = \frac{s_1}{1-s_1}.$$

Por ejemplo, la desventaja dominante de A_2 determina, con $s_1 = 0.10$, los valores de eficacia biológica 1, 0.90 y 0.90 para los genotipos A_1A_1 , A_1A_2 y A_2A_2 , respectivamente. El cálculo de s_2 con la ecuación anterior arroja un valor de 0.11, que indica que los valores genotípicos de eficacia biológica equivalentes, en la otra escala, son 1.11, 1 y 1.

Para valores pequeños de s , las dos escalas son muy semejantes, pero la diferencia entre s_1 y s_2 crece exponencialmente cuando la selección va siendo más intensa.

Sobredominancia: ventaja del heterocigoto

Se produce sobredominancia cuando el genotipo más favorecido por la selección natural es el heterocigoto. En este caso desaparece el concepto de alelo favorable, que implica una mayor eficacia biológica para los individuos que lo portan en una o dos dosis, modulada de acuerdo con la acción génica concreta. Cuando hay sobredominancia, la mayor eficacia biológica deriva de la presencia de una dosis de cada alelo, mientras que la presencia de dos dosis de cualquiera de los alelos lleva a una menor eficacia biológica. Mientras que, en los casos vistos hasta ahora, la selección conduce a la fijación del alelo favorecido y a la eliminación del otro alelo, la sobredominancia conduce al mantenimiento del polimorfismo. La Tabla 13 muestra el planteamiento del modelo.

$$\begin{aligned} q_1 &= \frac{pq + q^2(1-t)}{1-sp^2-tq^2} = \frac{pq + q^2 - tq^2}{1-sp^2-tq^2} = \frac{q-tq^2}{1-sp^2-tq^2} = \frac{q(1-tq)}{1-sp^2-tq^2} \\ \Delta q &= q_1 - q = \frac{q(1-tq)}{1-sp^2-tq^2} - q = \frac{q-tq^2 - q + sp^2q + tq^3}{1-sp^2-tq^2} = \\ &= \frac{-tq^2 + sp^2q + tq^3}{1-sp^2-tq^2} = \frac{-tq^2 + sp^2q + tq^2(1-p)}{1-sp^2-tq^2} = \\ &= \frac{-tq^2 + sp^2q + tq^2 - tq^2p}{1-sp^2-tq^2} = \frac{sp^2q - tq^2p}{1-sp^2-tq^2} = \frac{pq(sp-tq)}{1-sp^2-tq^2} \end{aligned}$$

Cuando $sp=tq$, $\Delta q = 0$ y la población estará en equilibrio.

$$sp = t(1-p); sp = t - tp; t = sp + tp; t = p(s+t); p_{eq} = \frac{t}{s+t}$$

Igualmente:

$$s(1-q) = tq; s - sq = tq; s = tq + sq; s = q(s+t); q_{eq} = \frac{s}{s+t}$$

Estos valores (p_{eq} y q_{eq}) son independientes de las frecuencias iniciales en la población, y sólo dependen de los coeficientes de selección contra los homocigotos.

Como s y t son constantes, el equilibrio es "estable". Si, por cualquier causa, se fuerza el desequilibrio de las frecuencias alélicas, éstas vuelven otra vez a los valores de equilibrio.

La desviación del punto de equilibrio para un valor q dado puede relacionarse con Δq de la forma siguiente:

$$\Delta q = \frac{pq(sp - tq)}{1 - sp^2 - tq^2} = pq \frac{s(1 - q) - tq}{1 - sp^2 - tq^2} =$$

$$= pq \frac{s - sq - tq}{1 - sp^2 - tq^2} = pq \frac{s - q(s + t)}{1 - sp^2 - tq^2}$$

$$q_{eq} = -\frac{s}{s + t}; \quad s = q_{eq}(s + t)$$

$$\Delta q = pq \frac{q_{eq}(s + t) - q(s + t)}{1 - sp^2 - tq^2} =$$

$$= pq \frac{(s + t)(q_{eq} - q)}{1 - sp^2 - tq^2}$$

Según esta expresión, cuando q es mayor que q_{eq} , Δq es negativo y la frecuencia de A_2 disminuye tendiendo hacia q_{eq} . Cuando q es menor que q_{eq} , Δq es positivo y la frecuencia de A_2 aumenta tendiendo hacia q_{eq} . Por eso el equilibrio es estable.

La sobredominancia es un fenómeno muy raro que provoca que los alelos de un locus lleguen al equilibrio estable (Fig. 4). También es aplicable en el caso de genes letales. Si, por ejemplo, los genotipos A_1A_1 , A_1A_2 y A_2A_2 tienen eficacias biológicas $1-s$, 1 y 0 , respectivamente, por ser A_2 un alelo letal recesivo, pero el heterocigoto es superior al homocigoto normal, la frecuencia de

equilibrio sería: $q_{eq} = \frac{s}{s + 1}$, ya que $t = 1$.

El número de generaciones para un determinado cambio en frecuencia, en caso de sobredominancia, viene dado por la siguiente expresión:

$$n = \frac{1}{s + t} \left[\frac{1}{q_{eq}} L \frac{q_n}{q_0} + \frac{1}{q_{eq} - 1} L \frac{1 - q_0}{1 - q_n} + \frac{1}{q_{eq}(1 - q_{eq})} L \frac{q_{eq} - q_0}{q_{eq} - q_n} \right]$$

(Apéndice 7d)

Subdominancia: selección contra heterocigotos

El desarrollo del modelo es resumido en la Tabla 14, de donde se deduce que:

$$q_1 = \frac{pq(1 - s) + q^2}{1 - 2pqs} = \frac{pq - pqs + q^2}{1 - 2pqs} =$$

$$= \frac{pq + q^2 - pqs}{1 - 2pqs} = \frac{q - pqs}{1 - 2pqs} = \frac{q(1 - ps)}{1 - 2pqs}$$

Tabla 13

Selección a favor de heterocigotos (sobredominancia)

Genotipos	A_1A_1	A_1A_2	A_2A_2	Total
Frecuencia Inicial	p^2	$2pq$	q^2	1
Eficacia biológica	$1-s$	1	$1-t$	
Frecuencia después de la selección	$p^2(1-s)$	$2pq$	$q^2(1-t)$	$\bar{w} = 1 - sp^2 - tq^2$ *

* $\bar{w} = p^2(1-s) + 2pq + q^2(1-t) = p^2 - sp^2 + 2pq + q^2 - tq^2 = 1 - sp^2 - tq^2$

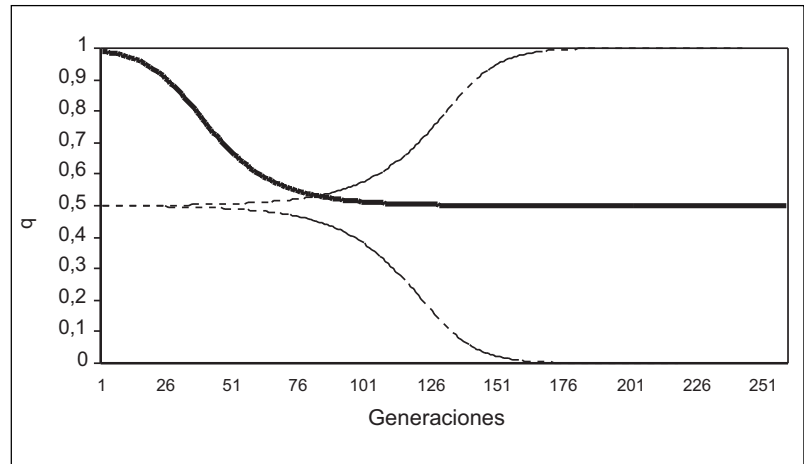


Figura 4. Equilibrio estable causado por selección a favor de heterocigotos (línea continua gruesa) y equilibrio inestable debido a selección contra heterocigotos (línea discontinua).

Tabla 14

Selección en contra de heterocigotos (subdominancia)

Genotipos	A_1A_1	A_1A_2	A_2A_2	Total
Frecuencia Inicial	p^2	$2pq$	q^2	1
Eficacia biológica	1	$1-s$	1	
Frecuencia después de la selección	p^2	$2pq(1-s)$	q^2	$\bar{w} = 1 - 2pqs$ *

* $\bar{w} = p^2 + 2pq(1-s) + q^2 = p^2 + 2pq - 2pqs + q^2 = 1 - 2pqs$

$$\Delta q = q_1 - q = \frac{q - pqs}{1 - 2pqs} - q = \frac{q - pqs - q + 2pq^2s}{1 - 2pqs} =$$

$$= \frac{-pqs + 2pq^2s}{1 - 2pqs} = \frac{pqs(2q - 1)}{1 - 2pqs}$$

De la ecuación anterior se deduce que el único valor de q en que habrá equilibrio ($\Delta q = 0$) es 0.5, pero este es un punto de equilibrio "inestable", porque cualquier valor de q diferente de 0.5 provocará que la frecuencia alélica se aleje del punto de equilibrio (Fig. 4). Así, cuando $q > 0.5$, Δq es positivo y q tiende a aumentar, alejándose aún más del punto de equilibrio. Cuando $q < 0.5$, Δq es negativo y q tiende a disminuir, alejándose igualmente del punto de equilibrio.

El número de generaciones para los cambios en frecuencia producidos en caso de subdominancia se estima mediante la siguiente expresión:

$$n = \frac{1}{s} \left[L \frac{q_0(1-q_0)(2q_n-1)^2}{q_n(1-q_n)(2q_0-1)^2} \right]$$

(Apéndice 7e)

Un ejemplo de selección contra los heterocigotos se presenta cuando una madre $rh^- (rr)$ espera un hijo $rh^+ (Rr)$ que muere como consecuencia de la eritroblastosis fetal. En este caso ocurre que el coeficiente de selección es directamente proporcional a la frecuencia de $r (q)$.

Selección y mutación

Hasta ahora hemos considerado los efectos que la mutación y la selección producen por separado sobre las frecuencias alélicas. Sin embargo, en la naturaleza la mutación y la selección se dan simultáneamente y los valores de las frecuencias génicas son influidos por ambos factores.

Es obvio que si la mutación y la selección ejercen sus presiones en la misma dirección los cambios en las frecuencias alélicas serán más rápidos que considerándolas por separado. Pero si ambos agentes oponen sus efectos entonces pueden llegar a anularse estos efectos mutuamente y la población a alcanzar una posición de “equilibrio estable”. Esta actuación simultánea de mutación y selección explica por qué genes asociados a una baja eficacia biológica (deletéreos o letales) se mantienen presentes en las poblaciones con una cierta frecuencia de equilibrio.

La mutación a un gen determinado es más efectiva en aumentar la frecuencia de ese gen cuanto menor es su frecuencia, puesto que es más probable que los genes sin mutar muten. En cambio, la eficacia de la selección es menor cuando el gen es raro.

Para calcular la frecuencia de equilibrio bastará igualar los valores que expresan los cambios en las frecuencias alélicas (Δq) en cada caso, cambiando de signo uno de dichos valores.

El cambio en frecuencia por mutación valdría:

$$\Delta q = up - vq = u(1-q) - vq$$

Si la desventaja de A_2 fuese recesiva, el cambio en frecuencia causado por la selección sería:

$$\Delta q = -\frac{sq^2(1-q)}{1-sq^2}$$

En el equilibrio: $u(1-q) - vq = \frac{sq^2(1-q)}{1-sq^2}$

Si s y v son pequeños:

$u(1-q) = sq^2(1-q)$; $u = sq^2$; por lo que :

$$q_{eq} = \sqrt{\frac{u}{s}}$$

Si A_2 es letal: $q_{eq} = \sqrt{u}$.

Si la desventaja de A_2 fuese dominante, llegaríamos a

la expresión $q_{eq} = \frac{u}{s}$, y si A_2 fuese letal $q_{eq} = u$. Finalmente, si la acción génica fuese aditiva obtendríamos

$q_{eq} = \frac{2u}{s}$, y si A_2 fuese letal $q_{eq} = 2u$.

Obsérvese que la frecuencia de equilibrio de un alelo dominante letal es simplemente la tasa de mutación. Esto es porque los individuos que llevan el alelo, ya sea en homocigosis o en heterocigosis, no pueden reproducirse, por lo que los únicos alelos de esta clase encontrados en una población serán los que hayan aparecido de nuevo por mutación en esa generación.

Suponiendo la misma tasa de mutación y los mismos coeficientes de selección, la frecuencia de equilibrio es mucho mayor para un letal recesivo que para un letal dominante (nótese que la raíz cuadrada de un número positivo menor que 1 es mayor que el número). Este resultado es de esperar porque los alelos recesivos se esconden de la selección en los heterocigotos.

Lastre genético

Muchos de los genes que se mantienen en las poblaciones naturales pueden ser desventajosos para sus portadores, ya sea en homocigosis, ya sea en determinadas combinaciones. Si consideramos la perfección genética como la ausencia de combinaciones génicas deletéreas, entonces la mayoría o todas las poblaciones son genéticamente imperfectas. El grado en que una población se aparta de la constitución genética perfecta se llama “lastre genético”, que va acompañado de la pérdida de una parte de los individuos por “muerte genética” (aquellos individuos que no contribuyen genéticamente a la siguiente generación). El lastre genético es, por tanto, la disminución de la eficacia biológica media de una población, asociada a los cambios en las frecuencias génicas producidas por cualquier agente. El lastre genético puede ser “manifiesto” (p. ej. genes deletéreos dominantes, genes deletéreos recesivos en homocigosis) o “velado” (genes deletéreos recesivos en heterocigosis).

Para calcular el lastre genético basta recordar que la proporción de individuos que sufren muerte genética, el lastre, es igual a uno menos la eficacia biológica media de la población ($L = 1 - \bar{w}$). El lastre por selección es, en general, $L = 2pqsh + sq^2$. Por ejemplo, si un gen es deletéreo en homocigosis ($h=0$), el lastre es: $L = sq^2$.

El lastre genético no es necesariamente una carga negativa que han de soportar las poblaciones, ya que su existencia aumenta la capacidad evolutiva de éstas ante cambios ambientales que conduzcan a cambios en la adaptabilidad de los distintos genotipos. En cierto sentido, el lastre genético es el precio que se paga por la evolución.

Dos son los componentes genéticos que contribuyen principalmente al lastre genético: la mutación y la segregación. El lastre mutacional es el grado en que una población se ve dañada por la mutación recurrente. Como la

mutación es la materia prima de la evolución, el lastre mutacional forma parte del lastre genético de todas las especies. Por ejemplo, en un locus el alelo A_2 es deletéreo en homocigosis, y surge por mutación con una tasa u . El lastre genético es $L=sq^2$. Cuando existe equilibrio mutación-selección,

$q_{eq} = \sqrt{\frac{u}{s}}$, y entonces:

$$L = sq_{eq}^2 = s \left(\sqrt{\frac{u}{s}} \right)^2 = s \frac{u}{s} = u ; L = u.$$

En los casos de desventaja aditiva y dominante de A_2 , el lastre ($L=2u$) es el doble que cuando la acción génica es recesiva.

El lastre segregacional se da en casos de polimorfismo equilibrado en los que los heterocigotos, favorecidos por la selección, segregan constantemente homocigotos desfavorables. El lastre segregacional se calcula a partir de la eficacia biológica media de una población donde está actuando la selección en favor del heterocigoto. En ese caso, la eficacia biológica media es $w = 1 - sp^2 - tq^2$, y el lastre vale: $L = sp^2 + tq^2$.

Si la población está en equilibrio, $sp = tq$ y, sustituyendo:

$$L = spp + tq^2 = tqp + tq^2 = t(pq + q^2) = tq = sp.$$

Por tanto, en el equilibrio $L = sp_{eq} = tq_{eq}$.

En el caso de letales equilibrados, $s=1$ y $t=1$, por lo que $L=p=q=0.5$.

Los cálculos anteriores son para un único locus. Los efectos acumulados del lastre segregacional de, por ejemplo, un centenar de loci sobredominantes serían, sin embargo, insostenibles para una población.

Evolución de la eficacia biológica media de la población

Hasta ahora hemos analizado los cambios en las frecuencias alélicas y genotípicas derivados de la actuación de la selección natural. Cabe preguntarse, a continuación, cuáles son los efectos de la selección sobre la eficacia biológica media de la población, ya que ésta puede ser un indicador del grado de adaptación de la población en el ambiente donde vive. Según el “teorema fundamental de Fisher”, el cambio en la eficacia biológica media de una población donde está actuando la selección natural, es proporcional a la varianza genética aditiva en la eficacia biológica media de la población. Como esta varianza no puede ser negativa, se deduce que la selección incrementa la eficacia biológica media de la población sobre la que actúa. La Figura 5 muestra el cambio en la eficacia biológica media

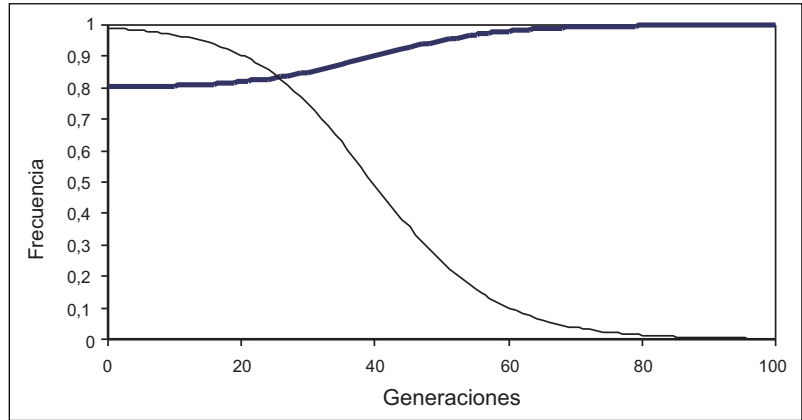


Figura 5. Incremento de la eficacia biológica media (línea gruesa) en una población donde está actuando la selección ($s = 0.1$) contra un alelo con acción génica aditiva, en paralelo con la disminución en frecuencia del alelo desfavorecido (línea fina).

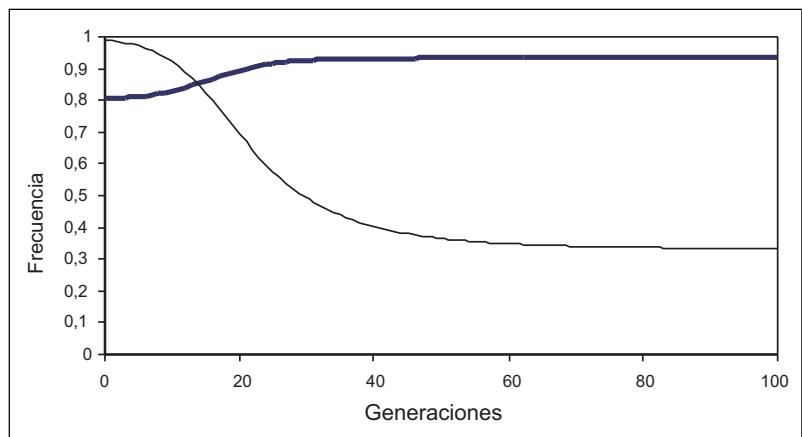


Figura 6. Incremento y estabilización de la eficacia biológica media (línea gruesa) en una población donde está actuando la selección a favor de heterocigotos ($s = 0.1$ y $t = 0.2$).

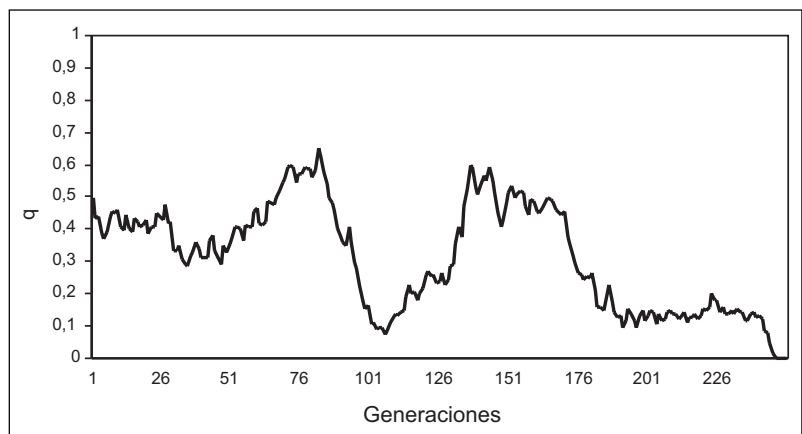


Figura 7. Desaparición de un alelo por deriva genética en una población con $N = 100$.

de una población donde está actuando la selección ($s = 0.1$) contra un alelo con acción génica aditiva. Se observa que la eficacia biológica media crece conforme decrece la frecuencia del alelo deletéreo. Esta es la forma en que la selección mejora la adaptación. Esto es cierto siempre que los valores de eficacia biológica asociados a cada genoti-

po se mantengan constantes. No lo es, por ejemplo, en el caso de selección dependiente de la frecuencia. En otros casos, existe un límite para el incremento en la eficacia biológica media. Cuando existe sobredominancia, la eficacia biológica media crece si las frecuencias alélicas no son las de equilibrio pero, cuando se alcanza éste, la varianza aditiva es nula y la eficacia biológica media se estabiliza en un valor inferior a 1 (Fig. 6). Esto es debido al lastre segregacional, es decir, a la generación de homocigotos con menor eficacia biológica en cada generación.

El valor de la eficacia biológica media (\bar{w}) de la población en el equilibrio puede calcularse sustituyendo los valores de p_{eq} y q_{eq} en la ecuación de la eficacia biológica media:

$$\begin{aligned}\bar{w} &= 1 - sp^2 - tq^2 = 1 - sp_e^2 - tq_e^2 = \\ &= 1 - \frac{st^2}{(s+t)^2} - \frac{ts^2}{(s+t)^2} = 1 - \frac{st^2 - ts^2}{(s+t)^2} = \\ &= 1 - \frac{st(t+s)}{(s+t)^2} = 1 - \frac{st}{s+t}\end{aligned}$$

En el caso de la figura anterior, $\bar{w} = 0.933$.

Deriva genética

Las poblaciones de organismos constan de un número limitado de individuos. Debido a que las poblaciones son finitas en número, las frecuencias génicas pueden cambiar por un puro proceso de azar conocido como “deriva genética”.

Supongamos que en cierta población existen dos alelos, A_1 y A_2 , con frecuencias 0.4 y 0.6. La frecuencia de A_1 en la siguiente generación puede ser menor (o mayor) de 0.4 debido simplemente a que, por azar, el alelo A_1 esté presente menos (o más) frecuentemente entre los gametos que forman los cigotos de esta generación.

La deriva genética es un caso particular de los errores de muestreo. La magnitud de los errores de muestreo es inversamente proporcional al tamaño de la muestra: cuanto menor es la muestra mayores son sus efectos. En el caso de los organismos, cuanto menor es el número de reproductores en una población, probablemente mayores serán los cambios en las frecuencias alélicas debidos a la deriva genética.

Es sencillo ver por qué hay una relación inversa entre el tamaño de la muestra y el error de muestreo: Si tiramos una moneda una vez, la probabilidad de obtener cara es 0.5, pero saldrá cara (en cuyo caso la frecuencia de caras es 1) o bien saldrá cruz (en cuyo caso la frecuencia de caras es 0). Si tiramos la moneda 10 veces, es probable que obtengamos varias caras y varias cruces; nos sorprenderíamos si obtuviésemos 10 caras, pero no si, por ejemplo, obtuviésemos 6 caras y 4 cruces. La frecuencia de caras sería 0.6 en vez de la esperada 0.5. La diferencia observada es sólo 0.1 en este caso. Si lanzamos la moneda 1000 veces, no esperaríamos obtener 1000 caras, ni

siquiera 600 caras y 400 cruces (frecuencia de cara= 0.6). No nos sorprendería, en cambio, obtener 504 caras y 496 cruces, en cuyo caso la frecuencia de caras sería 0.504 y la diferencia con el 0.5 esperado de tan sólo 0.004. Por tanto, cuanto mayor es la muestra más cerca está la frecuencia observada de la esperada (0.5).

En las poblaciones, cuanto mayor sea el número de individuos que dan origen a la siguiente generación más cerca estará la frecuencia observada (en la generación descendiente) de la frecuencia esperada (la de la generación parental). Por el contrario, si para comenzar una nueva generación se eligen tan sólo unos pocos progenitores, esa muestra tan pequeña puede no ser representativa de su generación y, por tanto, las frecuencias alélicas de la generación siguiente pueden desviarse mucho de las de la generación anterior.

Supongamos una población de N individuos diploides constituida por fecundación de $2N$ gametos tomados al azar de un acervo infinito de gametos. La probabilidad de que la muestra de $2N$ gametos contenga i del tipo A_1 viene dada por la expresión binomial:

$$\Pr(i) = \frac{(2N)!}{i!(2N-i)!} p^i q^{2N-i}, \text{ donde } p \text{ y } q \text{ son las frecuencias}$$

de los alelos A_1 y A_2 , respectivamente, en el acervo gamético completo (de forma que $p+q=1$), e i puede tomar cualquier valor entero entre 0 y $2N$. La nueva frecuencia alélica de A_1 en la población (p') es $i/2N$, es decir, el cociente entre el número de alelos A_1 (i) y el total de alelos ($2N$). En la siguiente generación, se repite el proceso de muestreo aleatorio, y la nueva probabilidad de que la población contenga un número dado de alelos A_1 (i) viene también expresada por la probabilidad binomial anteriormente mencionada, ahora para las frecuencias p' y q' . Por ello, la frecuencia alélica puede cambiar al azar de generación en generación (Fig. 7).

Modelo de Fisher-Wright

Aunque es virtualmente imposible predecir los cambios en frecuencia alélica, debidos a la deriva genética, en una sola población, el comportamiento promedio de las frecuencias alélicas en un número grande de poblaciones sí puede predecirse. Supongamos que una población inicial infinitamente grande se divide en un gran número de subpoblaciones de tamaño N y que se cumplen las siguientes condiciones: 1) Existe igual número de machos que de hembras, 2) el tamaño de cada subpoblación se mantiene constante, 3) el apareamiento es al azar dentro de cada subpoblación, 4) no existe migración entre subpoblaciones, 5) la mutación no afecta, y 6) la selección no actúa. El modelo se refiere a un conjunto de subpoblaciones, conceptualmente infinito, que se comportan todas ellas de la misma forma y de las cuales sólo una subpoblación, o varias, tienen existencia material. Fisher (1930) y Wright (1931) consideraron las consecuencias del muestreo binomial que ocurre en poblaciones pequeñas a lo largo de muchas generaciones. El modelo que desarrollaron, conocido como “modelo de Fisher-

Wright”, predice la distribución de frecuencias alélicas en un conjunto de poblaciones sometidas a la acción de la deriva genética.

Si una población tiene $2N$ alelos del mismo locus que pueden ser de dos tipos (A_1 y A_2), el estado de la población puede ser descrito por el número de alelos A_1 que contiene. Los estados posibles son tener 0, 1, 2, ..., $2N$ alelos A_1 . El estado no es más que una forma alternativa de representar la frecuencia alélica desde la ausencia de alelo

A_1 hasta su fijación $(0, \frac{1}{2N}, \frac{2}{2N}, \dots, 1)$.

Los estados 0 y $2N$ (frecuencias 0 y 1) se denominan “absorbentes” porque, en ausencia de mutación, una población que se encuentre en ellos no puede cambiar a otro estado. Si se encuentra en cualquier otro estado, sin embargo, puede cambiar a un estado diferente, es decir, a otra frecuencia alélica. Para cualquier cambio de estado, la probabilidad es mayor para los cambios pequeños (p. ej. pasar del estado 1 al 2) que para los grandes (del 1 al $2N$). La probabilidad de que la población derive desde el estado i al estado j se conoce como “probabilidad de transición” (T_{ij}), y se calcula, aplicando la ecuación de la probabilidad binomial, que vimos anteriormente, como:

$$T_{ij} = \frac{(2N)!}{j!(2N-j)!} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j}$$

Las probabilidades de transición pueden ponerse en una matriz \mathbf{T} cuyos elementos T_{ij} expresan la probabilidad de transición desde el estado i al estado j , pudiendo tomar i y j los valores 0, 1, 2, ..., $2N$. Esta matriz contiene todo lo necesario para predecir la distribución esperada de poblaciones en los distintos estados a lo largo de las generaciones. Este tipo de modelo matemático, expresado en términos de estados discretos con probabilidades fijas de ir de un estado a otro, se denomina “cadena de Markov”. Por ejemplo, para una población con $N=2$, la matriz de probabilidades de transición sería la mostrada en la Tabla 15, en la que cada elemento ha sido calculado con la ecuación anterior. Observe cómo la probabilidad de que la población cambie desde los dos estados absorbentes (0 y 4) a cualquier otro estado es cero.

La distribución de frecuencias de poblaciones en los diferentes estados (es decir, con las diferentes frecuencias alélicas) se obtiene a partir de la cadena de Markov. Sea X_i la frecuencia de poblaciones que tienen i alelos A_1 , es decir, que están en el estado i ($i=0, 1, 2, 3, \text{ ó } 4$), y sea \mathbf{X} el vector de las X_i 's. La distribución de poblaciones con las diferentes frecuencias alélicas (\mathbf{X}'), tras una generación de apareamiento aleatorio, es determinada por la matriz resultante de multiplicar este vector por la matriz de probabilidades de transición:

Tabla 15

Matriz de probabilidades de transición en una población con $N=2$.

Nº de alelos A_1 en la generación n	Nº de alelos A_1 en la generación $n+1$				
	0	1	2	3	4
0	1	0	0	0	0
1	0.316	0.422	0.211	0.047	0.004
2	0.062	0.250	0.375	0.250	0.062
3	0.004	0.047	0.211	0.422	0.316
4	0	0	0	0	1

Tabla 16

Evolución de la distribución de frecuencias de subpoblaciones con diferente número de alelos A_1 con frecuencia inicial 0.5 y $N=2$. H es la heterocigosis media de todas las subpoblaciones.

Generación	Número de alelos					H
	0	1	2	3	4	
0	0	0	1	0	0	0.5
1	0.0625	0.2500	0.375	0.2500	0.0625	0.375
2	0.1660	0.2109	0.2461	0.2109	0.1660	0.2812
3	0.2490	0.1604	0.1812	0.1604	0.2490	0.2109
4	0.3117	0.1205	0.1356	0.1205	0.3117	0.1582
5	0.3587	0.0904	0.1018	0.0904	0.3587	0.1187
n	0.5	0	0	0	0.5	0

$$\mathbf{X}' = \mathbf{X}\mathbf{T} = (X_0 \ X_1 \ X_2 \ \dots \ X_{2N}) \begin{pmatrix} T_{00} & T_{01} & T_{02} & \dots & T_{0,2N} \\ T_{10} & T_{11} & T_{12} & \dots & T_{1,2N} \\ T_{20} & T_{21} & T_{22} & \dots & T_{2,2N} \\ \dots & \dots & \dots & \dots & \dots \\ T_{2N,0} & T_{2N,1} & T_{2N,2} & \dots & T_{2N,2N} \end{pmatrix} = (X'_0 \ X'_1 \ X'_2 \ \dots \ X'_{2N})$$

donde $X'_j = \sum_i X_i T_{ij}$.

A lo largo del proceso dispersivo, en algún momento se alcanza una distribución rectangular (excluyendo los estados absorbentes), que se adelgaza con el paso de las generaciones hasta que llega un momento en que sólo permanecen los estados absorbentes. Por ejemplo, en un conjunto de subpoblaciones, con $N=2$, el fenómeno dispersivo de las frecuencias alélicas daría lugar a las distribuciones de poblaciones, con diferente número de alelos A_1 , mostradas en la Tabla 16. Obsérvese cómo las clases extremas (con 0 y 4 alelos) van aumentando en frecuencia en detrimento de las clases intermedias. Esas clases extremas (los estados absorbentes) están representadas por las poblaciones que se van fijando para uno u otro alelo. Al cabo de muchas generaciones, se habrán fijado todas las poblaciones.

Un resultado importante del análisis mediante la cadena de Markov es la “probabilidad de fijación” de un

alelo. Las cadenas de Markov convergen hacia una distribución estacionaria en que la probabilidad de estar en el estado i en la generación n permanece constante a lo largo de las generaciones. Este estado “estacionario” se caracteriza porque la distribución de probabilidades se mantiene uniforme, es decir, todas las poblaciones que mantienen variabilidad están en la misma frecuencia. En el estado estacionario se debe cumplir, por tanto, que $\pi_j = \sum_i \pi_i T_{ij}$. Supongamos que la frecuencia inicial del alelo A_j es π_0 . La única distribución de probabilidad estacionaria de la cadena de Markov que especifica el modelo de Fisher-Wright es: $\pi=(1-\pi_0, 0, 0, 0, \dots, \pi_0)$, donde el número de ceros es $2N-1$. Esta distribución estacionaria indica que, tras un tiempo suficientemente largo, la probabilidad de estar en el estado 0 (eliminación del alelo A_j) es $1-\pi_0$, y la probabilidad de estar en el estado $2N$ (fijación del alelo A_j) es π_0 . La consecuencia importante es que “la probabilidad de fijación de un alelo es simplemente su frecuencia inicial”.

A modo de ilustración con un ejemplo práctico, imaginemos el juego de ‘cara o cruz’ entre dos jugadores con las siguientes reglas: 1) Si sale cara, gana el jugador A y si sale cruz gana el jugador B; 2) si gana A recibe 1 euro de B, y si gana B recibe 1 euro de A; y 3) el juego termina cuando un jugador pierde todo su dinero. Según la regla 3, el juego se acaba cuando se llegue a uno de los dos estados de absorción posibles: A pierde todo o B pierde todo. ¿Quién ganará? Si la cantidad de dinero total es $nA + nB = N$, la probabilidad de que A gane todo es nA/N (probabilidad de que se fije el alelo A) y la de que B gane todo es nB/N (fijación del alelo B). Por lo tanto, aunque el juego parezca ‘equitativo’ porque cada jugador tiene una probabilidad 1/2 de ganar en cada lanzamiento, en realidad no lo es si nA es muy diferente de nB . Moraleja: si B es un casino cuya banca dispone de una cantidad ‘ilimitada’ de dinero, lo más probable es que acabe ganando siempre (sin necesidad de hacer trampas). Esto se conoce como ‘la ruina del jugador’.

En el ejemplo de la Tabla 16 se observa que, para una frecuencia inicial de 0.5 para el alelo A_j , la mitad de las poblaciones terminan fijándose para ese alelo. También se observa que la heterocigosis promedio de todas las subpoblaciones va disminuyendo, según la ecuación

$H_n = \left(1 - \frac{1}{2N}\right)^n H_0$, donde $\frac{1}{2N}$ es la velocidad de desintegración, es decir, la pérdida de heterocigosis por apareamiento aleatorio (ver apartado “Consanguinidad en poblaciones finitas”).

Fisher y Wright extendieron su modelo para incluir los efectos de la mutación y la selección. ¿Cómo afectan estas fuerzas a la matriz de transición? Con mutación recurrente reversible, los dos estados absorbentes desaparecen. La selección también afecta a los elementos de la matriz de probabilidad de transición, pero continúan siendo constantes y, por ello, la teoría de la cadena de Markov es aplicable también en este caso. No obstante, dada su complejidad, no consideraremos aquí estos modelos.

Fisher y Wright abordaron también la estimación del tiempo esperado para la fijación. Fisher (1922) apuntó que este problema podría resolverse aplicando la ecuación de difusión del calor a lo largo de una barra sólida, pero la solución tuvo que esperar a Kimura (1955). En esencia, Kimura estimó que el tiempo de fijación depende del tamaño de población y de la frecuencia alélica inicial. Por ejemplo, si ésta es 0.5, en $2N$ generaciones se alcanza el estado estacionario (quedando aún aproximadamente la mitad de las poblaciones sin fijar), y en $2.8N$ generaciones se habrán fijado todas las poblaciones. Si la frecuencia inicial de A_j es 0.1, sin embargo, se necesitan $4N$ generaciones para llegar a la fase estacionaria, con sólo un 10% de poblaciones sin fijar.

Censo efectivo de población

En poblaciones reales, donde el número de individuos que se reproducen (transmitiendo sus genes a la siguiente generación) es finito, debemos introducir el concepto de “censo efectivo” de la población (N_e), que representa “el tamaño de una población ideal que experimenta el mismo descenso de heterocigosis que la población real en estudio”. Por tanto, la magnitud del censo efectivo viene determinada por la intensidad de los efectos de la deriva genética. En una población de organismos con reproducción sexual, y sin posibilidades de autofecundación, compuesta de N_m machos y N_h hembras reproductores, el censo efectivo se calcula mediante la expresión:

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

Por ejemplo, una población con 3 machos, que se aparean con 300 hembras tendrá un censo efectivo, N_e , igual a 11. Vemos que las proporciones de sexos no equilibradas reducen considerablemente el censo efectivo poblacional, lo que incrementa la pérdida de heterocigosis por apareamiento aleatorio. La fórmula anterior siempre resulta en valores de N_e inferiores a la suma $N_m + N_h$, excepto cuando el número de machos es igual al de hembras, en cuyo caso no se utiliza la expresión anterior y N_e se calcula como $N_m + N_h + 1/2$.

Efectos de la deriva genética

El grado de desviación de las frecuencias alélicas (o la divergencia en frecuencia entre subpoblaciones) puede medirse matemáticamente mediante la desviación estándar de la distribución binomial de las frecuencias alélicas:

$\sigma = \sqrt{\frac{pq}{2N_e}}$. Por ejemplo, si comenzamos una población de organismos diploides con 2500 reproductores de cada sexo, y $p=q=0.5$, la desviación de las frecuencia alélicas,

en una generación, será $\sigma = \sqrt{\frac{0.5 \times 0.5}{10001}} = 0.005$. Esto significa que, en la generación siguiente, las frecuencias alélicas de esa población(es) fluctuarán alrededor de ciertos

valores que vendrán determinados por el intervalo de confianza $\pm 2\sigma$, lo que implica que, en el 95% de las poblaciones que se encuentren en esas mismas condiciones, las frecuencias alélicas fluctuarán entre $0.5 \pm 2 \times 0.005$, o sea, entre 0.49 y 0.51. Con sólo 2 reproductores, sin embargo, $\sigma = \sqrt{\frac{0.5 \times 0.5}{5}} = 0.224$, con lo que las frecuencias fluctuarán en el intervalo 0.50 ± 0.45 , lo que puede llevar cerca de la fijación a alguno de los alelos, en una sola generación.

Wright propuso que, en determinadas circunstancias, la deriva genética causada por pequeños censos efectivos de población puede dar lugar a cambios notables en las frecuencias alélicas. A menos que una población sea muy pequeña, los cambios en las frecuencias alélicas debidos a la deriva serán pequeños de una generación a otra, pero los efectos a lo largo de muchas generaciones pueden ser grandes. Si ningún otro proceso (mutación, migración, selección) afecta a las frecuencias alélicas en un locus, el resultado de la evolución será, en último término, la fijación de un alelo y la eliminación de todos los demás. Sin embargo, es improbable que sólo la deriva afecte a las frecuencias alélicas de cualquier locus durante largos períodos de tiempo, ya que probablemente la mutación, la migración y la selección actuarán en un momento u otro. Se puede aplicar una regla sencilla para determinar la importancia relativa de los efectos de la deriva respecto a los de los tres procesos direccionales mencionados. Si usamos x para representar a una de las fuerzas direccionales, ya sea la tasa de mutación (u), la tasa de migración (m) o el coeficiente de selección (s), se dice que los cambios en las frecuencias alélicas se deben a la deriva genética si, y sólo si, $N \cdot x \leq 1$ (Tabla 17).

Cuando se establece una población con muy pocos individuos pueden exagerarse las diferencias en la frecuencia de muchos genes con respecto a la población de origen. Mayr (1942) llamó a este fenómeno "efecto fundador". Las poblaciones de muchas especies que viven en las islas oceánicas, aunque ahora pueden constar de millones de individuos, son descendientes de uno o muy pocos colonizadores que llegaron hace tiempo por dispersión accidental. La situación es similar en lagos u otras extensiones de agua aisladas, en bosques y en otros sistemas ecológicos aislados. Debido a errores de muestreo las frecuencias

Tabla 17

Efecto relativo de varias fuerzas direccionales cuando actúan conjuntamente con la deriva genética

Proceso (x)	N	$N \cdot x \leq 1$	Resultado
$u = 10^{-5}$	100	$N \cdot u = 100 \cdot 10^{-5} = 10^{-3}$	deriva > mutación
	10^6	$N \cdot u = 10^6 \cdot 10^{-5} = 10$	mutación > deriva
$m = 0.01$	100	$N \cdot u = 100 \cdot 0.01 = 1$	deriva > migración
	1000	$N \cdot u = 1000 \cdot 0.01 = 10$	migración > deriva
$s = 0.1$	100	$N \cdot u = 100 \cdot 0.1 = 10$	selección > deriva
	10	$N \cdot u = 10 \cdot 0.1 = 1$	deriva > selección

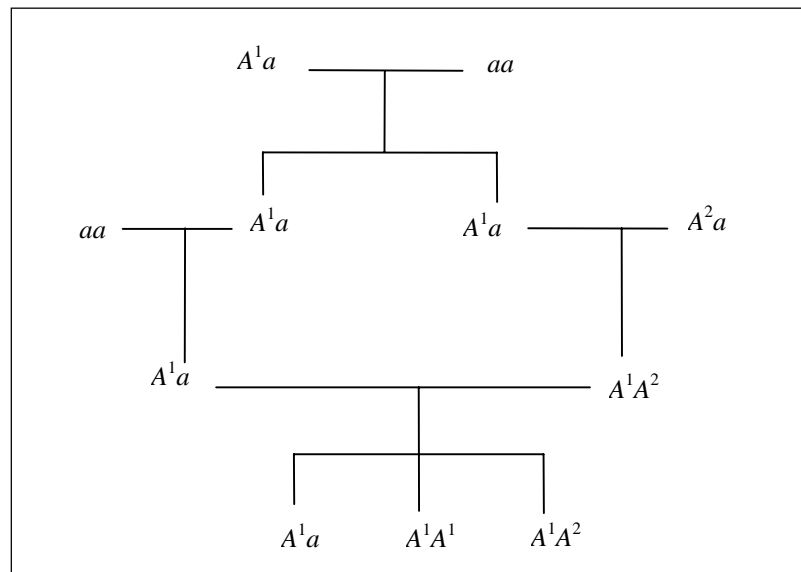


Figura 8. Fijándonos en los genotipos de los individuos de la última generación, los dos alelos A^1 del individuo del centro se consideran idénticos, porque son copia de un alelo que estaba presente en un antecesor común (primera generación). A^1 y a , en el individuo de la izquierda, son alelos diferentes porque tienen distinta expresión fenotípica. En el individuo de la derecha, A^1 y A^2 son alelos similares, porque tienen la misma expresión fenotípica pero distinta ascendencia.

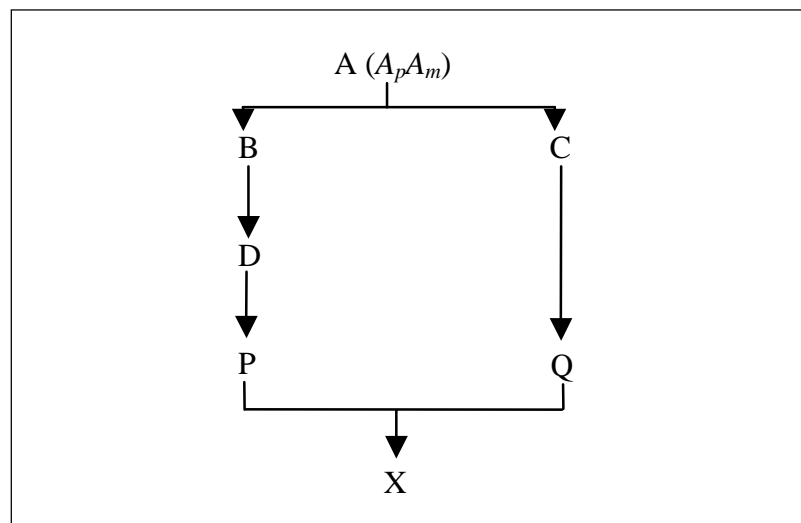


Figura 9. Genealogía para el cálculo del coeficiente de consanguinidad del individuo X.

alélicas de muchos loci, en los pocos colonizadores, pueden ser diferentes a las de la población de la que proceden, y ello puede tener efectos duraderos en la evolución de tales poblaciones aisladas. Por ejemplo, Carson (1983) estudió los patrones de bandas de los cromosomas politénicos de las glándulas salivares de las especies de *Drosophila* de la isla de Hawaii (la más joven del archipiélago hawaiano) y demostró que las especies de esta isla deben haberse originado a partir de un solo individuo fundador (o muy pocos) procedente de la vecina isla de Maui. Estos fundadores aislados portaban patrones de bandas únicos cuya pista puede seguirse en las especies descendientes actuales.

Una demostración del efecto fundador fue realizada por Dobzhansky y Pavlovsky (1957). Iniciaron poblaciones de laboratorio de *Drosophila pseudoobscura* a partir de una población en que cierta reordenación cromosómica del tercer cromosoma (denominada PP) tenía una frecuencia de 0.5. Había dos tipos de poblaciones, unas grandes iniciadas con 5000 individuos y otras pequeñas iniciadas con 20 individuos cada una. Después de año y medio (unas 18 generaciones) la frecuencia media de PP era aproximadamente 0.3 tanto en las poblaciones grandes como en las pequeñas, pero el rango de variación de las frecuencias era considerablemente mayor en las poblaciones pequeñas. Al iniciar las poblaciones pequeñas con pocos fundadores se originaba una considerable variación entre poblaciones en la frecuencia de PP.

Variaciones fortuitas en las frecuencias alélicas, similares a las del efecto fundador, ocurren cuando las poblaciones pasan por un “cuello de botella”. Cuando el clima y otras condiciones son desfavorables, las poblaciones pueden reducirse drásticamente en número y corren el riesgo de extinguirse. Tales poblaciones pueden recuperar su tamaño más tarde, pero la deriva puede alterar considerablemente sus frecuencias alélicas durante el cuello de botella. (Para más detalles sobre los efectos de la deriva genética, ver Hartl y Clark, 1989, y Fontdevila y Moya, 1999).

Consanguinidad

En la mayoría de las poblaciones la movilidad de los individuos, y en consecuencia la movilidad de sus gametos, se halla generalmente restringida. Es decir, los individuos tienden a aparearse con individuos situados en las proximidades. Ello implica que el apareamiento entre todos los miembros de una población no se realiza por lo general al azar, y el acervo génico de muchas poblaciones puede consistir en muchos subconjuntos pequeños (subpoblaciones), pudiendo desviarse algunos de ellos de las características generales del conjunto total.

La “consanguinidad” se produce por la “ocurrencia de apareamientos entre parientes a una frecuencia superior de la que cabría esperar si el apareamiento fuese aleatorio”. Puede haber consanguinidad en poblaciones grandes, como resultado de apareamiento no aleatorio, pero también puede aparecer en poblaciones pequeñas, aunque sean panmícticas (ver apartado 3.5.3), ya que la pro-

babilidad de que se den apareamientos al azar entre parientes es muy baja en una población grande, pero aumenta considerablemente en una población pequeña.

Como veremos, si estos apareamientos se producen con frecuencia, pueden tener poco efecto en el cambio de las frecuencias alélicas conjuntas, pero pueden tener un efecto importante en aumentar la frecuencia de homocigotos. Como regla general, la consanguinidad hará que los genes recesivos raros se presenten en homocigosis con una frecuencia superior que si ocurriese apareamiento aleatorio. Con esto la selección tendrá una mayor oportunidad de actuar sobre los genes deletéreos poco frecuentes.

Medida de la consanguinidad

Para medir la consanguinidad es necesario determinar el grado en que los dos alelos de un individuo cualquiera son “idénticos por descendencia”, es decir, son copias de un mismo alelo ancestral presente en un antepasado de ese individuo.

En la última generación del pedigrí mostrado en la Figura 8, el individuo A^1A^1 lleva dos alelos idénticos por descendencia, ya que son copia del alelo A^1 de su antecesor tres generaciones atrás. El individuo A^1A^2 también es homocigótico, pero sus alelos no son idénticos sino similares (con la misma expresión fenotípica pero procedentes de linajes diferentes), y el individuo A^1a es heterocigótico pues sus alelos son diferentes (con distinta expresión fenotípica).

La medida en que dos alelos de un individuo son idénticos viene expresada por el “coeficiente de consanguinidad”, F , que es “la probabilidad de que dos alelos de un locus determinado de un individuo sean idénticos por descendencia”.

Veamos cómo se calcula el coeficiente de consanguinidad en genealogías, es decir, cuando se conocen con exactitud los árboles genealógicos familiares. En este caso, puede calcularse con precisión el coeficiente de consanguinidad de los descendientes de parejas con antepasados comunes. Se basa en calcular la probabilidad de que se hereden alelos idénticos en cada etapa de la transmisión. En la genealogía de la Figura 9, para averiguar el coeficiente de consanguinidad de X (F_x) tendríamos que calcular cuál es la probabilidad de que X reciba alelos idénticos de su antecesor A a través de sus padres P y Q . Consideremos primero los ancestrales B y C . La probabilidad de que los dos reciban el mismo alelo de A (ya sea el de origen paterno A_p o el de origen materno A_m) es:

$$\text{prob. } A_p \text{ y } A_p + \text{prob. } A_m \text{ y } A_m = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

y la probabilidad de que reciban alelos diferentes es:

$$\text{prob. } A_p \text{ y } A_m + \text{prob. } A_m \text{ y } A_p = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

Pero, incluso si B y C reciben alelos diferentes de A , éstos podrían todavía ser idénticos si A tiene alguna consanguinidad previa, es decir, si existe la posibilidad de que los dos alelos de A sean idénticos, que viene dada por F_A . Por tanto, la probabilidad total de que B y C reciban alelos idénticos es la suma de la probabilidad de que reciban copia del mismo alelo de A (A_p y A_p o A_m y A_m) (que es

$\frac{1}{2}$) y la probabilidad de que aún recibiendo copia de alelos diferentes (A_p y A_m o A_m y A_p) éstos sean idénticos por la existencia de consanguinidad previa (que vale $\frac{1}{2} F_A$). Es decir: $\frac{1}{2} + \frac{1}{2} F_A = \frac{1}{2} (1 + F_A)$. Dicho de otra forma, $\frac{1}{2} (1 + F_A)$ es la probabilidad de que dos gametos de A tomados al azar lleven alelos idénticos. En cada paso posterior de la genealogía, $B \rightarrow D \rightarrow P \rightarrow X$ y $C \rightarrow Q \rightarrow X$ la probabilidad de que se transmita el mismo alelo es $\frac{1}{2}$. Por tanto, la probabilidad total de que X reciba alelos idénticos por descendencia será $F_x = \frac{1}{2} (1 + F_A) (\frac{1}{2})^3 (\frac{1}{2})^2 = (\frac{1}{2})^6 (1 + F_A)$.

En general, $F_x = (\frac{1}{2})^{n_1+n_2+1} (1 + F_A)$, donde n_1 y n_2 son el número de generaciones que separan los parentales (P y Q) de X del ancestral común A (3 y 2 en este caso). Si los progenitores del individuo problema tienen más de un antecesor común, entonces hay que sumar las probabilidades correspondientes a cada ancestral común:

$$F_x = \sum \left[\left(\frac{1}{2} \right)^{n_1+n_2+1} (1 + F_A) \right]$$

Al calcular el coeficiente de consanguinidad en genealogías hay que considerar todos los antecesores comunes a los progenitores del individuo problema, y considerar también todos los caminos que unen a los progenitores a través del antecesor común, siempre que sea en línea ascendente (desde un progenitor al ancestral) y descendente (desde el mismo ancestral al otro progenitor). Esto quiere decir que, en cada camino, no se puede pasar dos veces por el mismo individuo. El Apéndice 8 muestra algunos ejemplos.

El riesgo genético que entrañan los matrimonios consanguíneos radica en la mayor probabilidad de que, en la descendencia, se manifieste en homocigosis un gen deletéreo portado por uno de los antecesores. Si el gen está ligado al sexo, el método de cálculo de la consanguinidad en la genealogía sufre algunas variaciones (para una descripción detallada, consultar Fontdevila y Moya, 1999, p. 238).

Efecto de la consanguinidad sobre las frecuencias genotípicas

Puesto que la consanguinidad reduce la frecuencia de heterocigotos, veamos el cambio en las frecuencias genotípicas de la población. Si nos centramos en un locus con los alelos A_1 y A_2 , habrá una proporción F de homocigotos idénticos, de los cuales pF serán A_1A_1 y qF serán A_2A_2 . Por tanto, las frecuencias genotípicas de la población serán:

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

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

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

Vemos cómo el aumento en pqF de cada homocigoto procede de la reducción equivalente de $2pqF$ en la frecuencia de heterocigotos. Es de notar que la reducción de heterocigotos afecta por igual a p y q , por lo que la consanguinidad únicamente cambia las frecuencias genotípicas, no afecta a las alélicas.

Consanguinidad en poblaciones finitas

Como hemos comentado anteriormente, la probabilidad de que se den apareamientos entre parientes es muy baja en una población grande, pero puede ser grande en poblaciones pequeñas. Con apareamiento aleatorio, en una población infinita el coeficiente de consanguinidad es cero, pero en una población finita puede alcanzar un cierto valor. Por tanto, en poblaciones finitas puede estimarse F aun desconociendo las relaciones de parentesco, en base al tamaño de población.

Por ejemplo, supongamos 50 individuos diploides de los que ninguno lleva alelos idénticos a los de los demás. Es decir, hay 100 tipos de alelos no idénticos (es decir, similares o diferentes): el individuo 1 lleva los alelos A_1A_2 , el individuo 2 lleva los alelos A_3A_4 , el individuo 3 lleva los alelos A_5A_6 ... el individuo 50 lleva los alelos $A_{99}A_{100}$.

Supongamos que cada individuo produce tanto gametos masculinos como femeninos y que éstos se unen al azar, como ocurre, por ejemplo, en los animales hermafroditas marinos. Para el locus A habrá 100 tipos de gametos, cada uno con frecuencia $1/100$.

En la primera generación sólo es posible que surjan individuos homocigóticos para alelos idénticos mediante autofecundación. Como hay 50 individuos, la probabilidad de autofecundación es $1/50$. Puesto que los alelos de un mismo individuo, en esta primera generación no pueden ser idénticos, la probabilidad de que dos gametos de un mismo individuo lleven alelos idénticos es $1/2$ (las combinaciones posibles, para un individuo A_1A_2 , son: A_1A_1 , A_1A_2 , A_2A_1 y A_2A_2). Por tanto, la probabilidad de que un individuo de la primera generación lleve dos alelos idénticos por autofecundación es $(1/50) \cdot (1/2) = 1/100$.

En general, la probabilidad de que un individuo de la primera generación lleve alelos idénticos en una población de tamaño N es: $F_1 = \frac{1}{2N}$, que es la probabilidad de que se fecunden dos gametos con alelos idénticos, procedentes del mismo individuo.

En la segunda generación habrá de nuevo $2N$ tipos de gametos producidos por los nuevos progenitores, y la probabilidad de consanguinidad será otra vez $1/(2N)$ para las combinaciones de gametos procedentes del mismo individuo, pero ahora habrá que sumarle la probabilidad de que se fecunden gametos de individuos diferentes pero portadores de alelos idénticos debido a la consanguinidad acumulada en la generación anterior (p. ej. A'_1 podría ser idéntico a A'_8 , es decir, uno de los gametos producido por el individuo 1 de la primera generación podría llevar el alelo A'_1 que es idéntico al que lleva un gameto producido por el individuo 4, A'_8). Este último factor es igual al coeficiente de consanguinidad de la generación anterior (F_1) multiplicado por la probabilidad de que se unan gametos con alelos no idénticos [$1 - 1/(2N)$]. Por tanto:

$$F_2 = \frac{1}{2N} + \left(1 - \frac{1}{2N} \right) F_1 \quad \text{y} \quad F_3 = \frac{1}{2N} + \left(1 - \frac{1}{2N} \right) F_2,$$

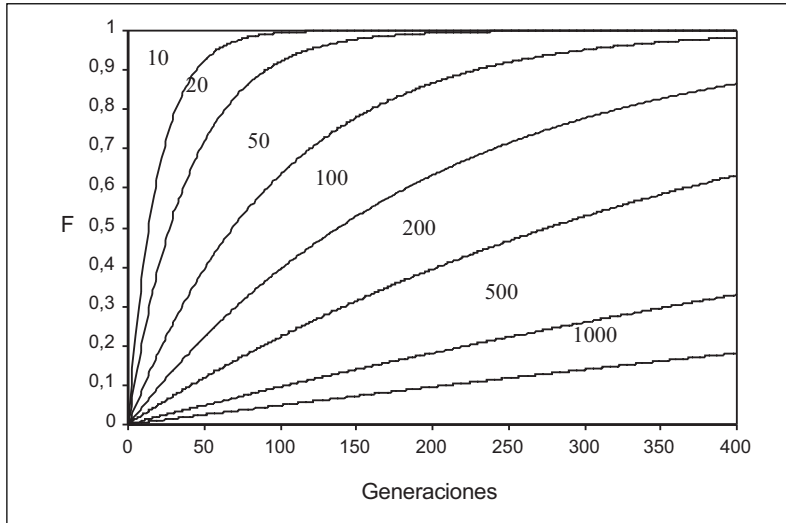


Figura 10. Consanguinidad (F) en poblaciones finitas.

Tabla 18
Efecto Wahlund

Muestra	Frecuencias alélicas		Frecuencias genotípicas		
	$p(A_1)$	$q(A_2)$	p^2	$2pq$	q^2
Subpoblación 1	0.1	0.9	0.01	0.18	0.81
Subpoblación 2	0.9	0.1	0.81	0.18	0.01
Población media	0.5	0.5	0.41	0.18	0.41
Población panmíctica	0.5	0.5	0.25	0.50	0.25

y en general: $F_n = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) F_{n-1}$

El coeficiente de consanguinidad (F) de una población nos permite estimar el grado en que el apareamiento de los gametos de una población de tamaño limitado se aparta del estado panmíctico ideal en el que ninguna pareja de alelos de un individuo se hallan relacionados.

A partir de F podemos definir el “índice de panmixia”, $P = 1 - F$, que es una medida de la proporción relativa de heterocigosis debida al apareamiento aleatorio, que inicialmente puede valer 1 y que se va perdiendo por consanguinidad. Sustituyendo en la ecuación anterior, obtenemos

$P_n = P_{n-1} \left(1 - \frac{1}{2N}\right)$ (Apéndice 9), donde $1/(2N)$ es

la “velocidad de desintegración”, es decir, la pérdida de heterocigosis por apareamiento aleatorio. En organismos dioicos, donde resulta imposible la autofecundación, la velocidad de desintegración es menor: $1/(2N+1)$, pero si el tamaño de población es bastante grande la velocidad de desintegración puede aproximarse a $1/(2N)$.

En poblaciones grandes, P_n puede expresarse en términos

de generaciones anteriores: $P_n = P_0 \left(1 - \frac{1}{2N}\right)^n$

(Apéndice 10). A partir de esta ecuación podemos deducir n. Cuando n es grande y N pequeño, esta ecuación

puede aproximarse a $P_n = P_0 e^{-\frac{n}{2N}}$, de donde podemos despejar n:

$$\frac{P_n}{P_0} = e^{-\frac{n}{2N}}; L \frac{P_n}{P_0} = -\frac{n}{2N}; n = -2NL \frac{P_n}{P_0}$$

y resolver cuestiones como, por ejemplo, ¿cuántas generaciones serían necesarias para reducir un cierto valor de heterocigosis a la mitad? Para ello, se ha de cumplir la condición $P_n = \frac{1}{2} P_0$. Sustituyendo,

$$n = -2NL \frac{\frac{1}{2} P_0}{P_0} = -2NL \frac{1}{2} = 1.39N$$

Por otra parte, podemos sustituir $P = 1 - F$ en la ecuación de P_n :

$$1 - F_n = (1 - F_0) \left(1 - \frac{1}{2N}\right)^n,$$

y si $F_0 = 0$, entonces $1 - F_n = \left(1 - \frac{1}{2N}\right)^n$, de

donde

$$F_n = 1 - \left(1 - \frac{1}{2N}\right)^n$$

Si comenzamos una población con $F=0$, los valores de F van aumentando progresivamente, tanto más pronunciadamente cuanto menor sea N (Fig. 10).

Subdivisión poblacional

Hemos visto que la consanguinidad genera un déficit de heterocigotos en las poblaciones. En poblaciones subdivididas, el efecto acumulado de la deriva genética sobre cada una de las subpoblaciones genera también un déficit de heterocigotos que se denomina “efecto Wahlund”. El déficit resulta de la predominancia (o fijación) de un alelo en unas poblaciones y del otro en otras. Por ejemplo, supongamos una población subdividida en dos subpoblaciones (Tabla 18). El déficit de heterocigotos es 0.32 (0.5-0.18) y equivale a $2pqF$. De aquí podemos calcular F, que en este caso vale 0.64 [0.32/(2×0.5×0.5)]. Este parámetro se denomina “índice de fijación” (F_{st}) (ver apartado siguiente) y oscila entre 0 y 1, según que las subpoblaciones tengan las mismas frecuencias alélicas o bien estén fijadas para alelos diferentes.

Estructura poblacional

En una población subdividida, podemos considerar tres niveles de diferenciación genética: dentro de cada subpoblación, entre subpoblaciones y en el conjunto de

subpoblaciones (es decir, en la metapoblación). Sewall Wright desarrolló un método de análisis de la variabilidad genética a estos tres (o más) niveles, mediante el estadístico F .

En el primer nivel, calculamos la heterocigosis media intra-subpoblacional (H_I):

$$H_I = \sum_i^k \frac{H_i}{k}$$

donde H_i es la heterocigosis observada en la subpoblación i y k el número de subpoblaciones.

Al segundo nivel, entre subpoblaciones, calculamos primero la heterocigosis esperada (H_S) para cada subpoblación (S), considerando los i alelos de cada locus:

$$H_S = 1 - \sum_i p_{i,s}^2$$

y después la H_S media de las k subpoblaciones:

$$\bar{H}_S = \sum_s^k \frac{H_S}{k}$$

Al nivel metapoblacional, calculamos la heterocigosis total (H_T) que se esperaría si todas las subpoblaciones constituyeran, en realidad, una única población panmíctica: donde \bar{p}_i es la frecuencia media del alelo i en la metapoblación.

En los mismos tres niveles, calculamos ahora los respectivos valores de F , que representa la disminución de la heterocigosis originada por apareamiento aleatorio.

En el nivel intra-subpoblacional, calculamos:

$$F_{IS} = \frac{\bar{H}_S - H_I}{\bar{H}_S}$$

al nivel inter-subpoblacional, calculamos:

$$F_{ST} = \frac{H_T - \bar{H}_S}{H_T}$$

y al nivel metapoblacional, calculamos:

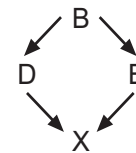
$$F_{IT} = \frac{H_T - H_I}{H_T}$$

Cuanto más difieran los tres índices de cero, más se alejará la metapoblación de una única población panmíctica. F_{IS} puede alcanzar valores positivos (déficit de heterocigotos), negativos (exceso de heterocigotos) o, simplemente cero cuando las subpoblaciones se encuentran en equilibrio Hardy-Weinberg. F_{ST} expresa el índice de fijación y varía entre cero (si todas las subpoblaciones tienen las mismas frecuencias alélicas) y uno (si las subpoblaciones están fijadas para alelos diferentes). Este índice es inversamente proporcional al número de emigrantes entre subpoblaciones por generación, ya que el flujo génico entre éstas se opone fuertemente al aumento de la divergencia genética entre ellas (para conocer más detalles del significado de estos índices, ver capítulo 7 en Fontdevila y Moya, 1999).

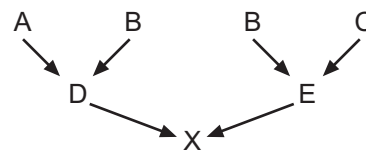
Consanguinidad sistemática

Para la obtención de cepas genéticamente uniformes, pueden construirse sistemas de consanguinidad tales que el procedimiento de apareamiento de cada generación de lugar a individuos, todos, con el mismo coeficiente de consanguinidad.

Para controlar el proceso, debemos predecir el cambio en consanguinidad que se observará a lo largo de las generaciones. Para deducirlo, usaremos el “coeficiente de parentesco” (f) entre dos individuos, que es la probabilidad de que, tomando un gameto de cada individuo, ambos gametos tengan copia de un mismo alelo ancestral. El coeficiente de parentesco entre dos individuos es igual al coeficiente de consanguinidad de un hijo de ambos. Para calcular el coeficiente de parentesco, es necesario representar las genealogías completas, de forma que los individuos de una generación sean el doble que los de la siguiente y la mitad que los de la anterior. Por ejemplo, la siguiente genealogía:



deberíamos representarla:



El coeficiente de parentesco entre los individuos D y E es:

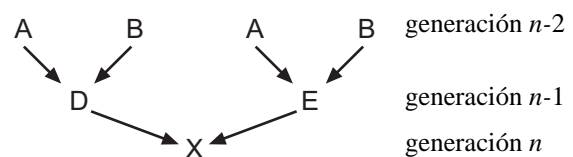
$$F_x = f_{DE} = 1/4 (f_{AB} + f_{AC} + f_{BB} + f_{BC})$$

Si desconocemos el parentesco entre los individuos de la generación más antigua, consideramos que f_{AB} , f_{AC} y f_{BC} valen cero, por lo que

$$f_{DE} = 1/4 f_{BB} = 1/4 \cdot 1/2 (1 + F_B);$$

y como $F_B = 0$, $f_{DE} = 1/4 f_{BB} = 1/4 \cdot 1/2 = 1/8$

El coeficiente de parentesco es muy útil para el control del nivel de consanguinidad cuando ésta se practica sistemáticamente para conseguir líneas lo más homocigóticas posible. En animales, el método más rápido es el cruzamiento sistemático entre hermanos:



Ahora A y B son hermanos. El coeficiente de parentesco entre D y E es:

$$f_{DE} = 1/4 (f_{AA} + f_{AB} + f_{BA} + f_{BB})$$

Tabla 19
Evolución de la consanguinidad en una genealogía con apareamiento entre hermanos

<i>n</i>	<i>F</i>	<i>n</i>	<i>F</i>	<i>n</i>	<i>F</i>	<i>n</i>	<i>F</i>	<i>n</i>	<i>F</i>
0	0	7	0.734	14	0.940	21	0.986	28	0.997
1	0	8	0.785	15	0.951	22	0.989	29	0.997
2	0.250	9	0.826	16	0.961	23	0.991	30	0.998
3	0.375	10	0.859	17	0.968	24	0.993	30	0.998
4	0.500	11	0.886	18	0.974	25	0.994	31	0.998
5	0.594	12	0.908	19	0.979	26	0.995	32	0.999
6	0.672	13	0.926	20	0.983	27	0.996	33	0.999

Tabla 20
Evolución de la consanguinidad con autofecundación

<i>n</i>	<i>F</i>	<i>n</i>	<i>F</i>
0	0	5	0.969
1	0.500	6	0.984
2	0.750	7	0.992
3	0.875	8	0.996
4	0.938	9	0.998

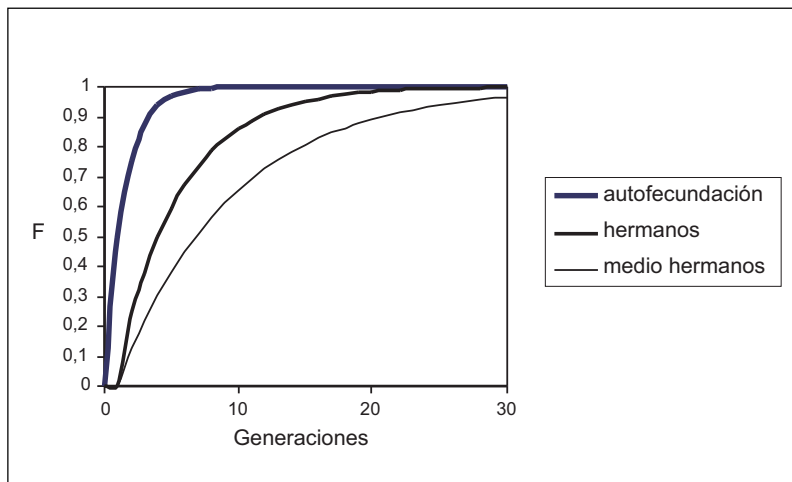


Figura 11. Evolución del coeficiente de consanguinidad (F) en tres clases de apareamientos sistemáticos: autofecundación, apareamiento entre hermanos y apareamientos entre medio hermanos.

Todos los individuos de la misma generación tienen el mismo coeficiente de consanguinidad, puesto que pertenecen a la misma genealogía. Por ello, $f_{AB} = f_{BA} = F_{n-1}$, es decir, el coeficiente de parentesco entre los individuos A y B es el coeficiente de consanguinidad de sus hijos de la generación *n*-1. Puesto que $F_A = F_B = F_{n-2}$, se cumple que $f_{AA} = \frac{1}{2}(1+F_A) = f_{BB} = \frac{1}{2}(1+F_B) = \frac{1}{2}(1+F_{n-2})$. El coeficiente de consanguinidad de la generación *n*, será:

$$F_x = f_{DE} = \frac{1}{4}(f_{AA} + f_{AB} + f_{BA} + f_{BB}) = \frac{1}{4}[\frac{1}{2}(1+F_{n-2}) + F_{n-1} + F_{n-1} + \frac{1}{2}(1+F_{n-2})];$$

$$F_x = \frac{1}{4}(1 + 2F_{n-1} + F_{n-2})$$

La Tabla 19 muestra la evolución de la consanguinidad en una genealogía donde los apareamientos se dan siempre entre hermanos. Se observa que en la generación 12 el coeficiente de consanguinidad supera el 90%, y en la 15 el 95%.

En organismos donde es posible la autofecundación, la homocigosis se consigue aun más rápido:



$$F_n = f_{AA} = \frac{1}{2}(1+F_A) = \frac{1}{2}(1+F_{n-1})$$

La Tabla 20 muestra la evolución de la consanguinidad, en este caso. Se observa que el coeficiente de consanguinidad ha superado ya el 90% en la cuarta generación, y el 99% en la séptima. La Figura 11 muestra la evolución comparada de la consanguinidad en tres sistemas de cruzamiento diferentes: autofecundación, apareamiento entre hermanos y apareamiento entre medio hermanos.

Homogamia y heterogamia

Los apareamientos consanguíneos son un tipo de apareamiento no panmíctico cuya consecuencia principal es el incremento de la homocigosis para todos los genes. Pero existen otros tipos de apareamientos no panmícticos realizados, no en base a lazos de parentesco, sino a la similitud fenotípica de los progenitores. Cuando el apareamiento se da preferentemente entre individuos más parecidos, hablamos de "homogamia". Si es entre individuos más diferentes, de "heterogamia". Por ejemplo, para cierto carácter monogénico, pueden darse apareamientos preferentemente entre fenotipos dominantes ($AA \times AA$, $AA \times Aa$, $Aa \times Aa$), por un lado, y entre recesivos ($aa \times aa$), por otro. Si estos cruzamientos homogámicos continúan durante mucho tiempo,

acabarán desapareciendo los heterocigotos y la población quedará dividida en dos razas, una será AA y la otra aa. A diferencia de los apareamientos consanguíneos, donde la homocigosis se alcanza para todos los loci, la homogamia no afecta a otros genes (Bb , Cc , ...), que no se harán necesariamente homocigóticos. La heterogamia, por el contrario, contribuye a mantener la heterocigosis. Para el gen *A/a*, los cruzamientos se darían preferentemente entre dominante por recesivo. Muchos organismos han desarrollado sistemas naturales heterogámicos que favorecen los cruzamientos entre fenotipos diferentes, evitando así la pérdida de heterocigosis. Un buen ejemplo lo

constituyen los mecanismos que impiden la autofecundación en plantas hermafroditas (alelos de autoesterilidad, posición relativa del estilo y las anteras, etc.).

Depresión consanguínea

Cuando se fuerza el modo natural de reproducción de una población alógama para inducir artificialmente la consanguinidad, se observa que, al aumentar la consanguinidad, hay una depresión de los caracteres relacionados con la eficacia biológica. La depresión consanguínea es la disminución de los valores fenotípicos medios de los caracteres relacionados con la capacidad reproductora o la eficiencia fisiológica.

En numerosas investigaciones se ha demostrado que la consanguinidad tiende a reducir la eficacia biológica. ¿Por qué? Puesto que la consanguinidad no cambia las frecuencias alélicas sino las genotípicas, aumentando la frecuencia de homocigotos en detrimento de los heterocigotos, la depresión consanguínea ha de estar ligada con los cambios en las frecuencias genotípicas. Por ejemplo, si una enfermedad causada por homocigosis para un alelo recesivo (*aa*) tiene una frecuencia q^2 en una población con apareamiento aleatorio, su frecuencia en una población consanguínea será $q^2 + pqF$. Podemos medir la depresión consanguínea por el incremento relativo en la frecuencia de homocigotos:

$$\frac{q^2 + pqF}{q^2} = \frac{q + pF}{q}$$

Cuanto menor sea la frecuencia del alelo recesivo deletéreo mayor será la depresión consanguínea. Aplicando la ecuación anterior para distintos valores de q y con $F=1/16$ (apareamiento entre primos hermanos) comprobamos la afirmación anterior (Tabla 21). Es probable que sean los genes deletéreos recesivos “descubiertos” por la consanguinidad los responsables de la depresión consanguínea

Tabla 21

Incremento de la depresión consanguínea con la disminución de la frecuencia alélica

q	$\frac{q + pF}{q}$
0.5	1.06
0.005	13.40
0.0005	126.00

Las líneas puras suelen mostrar variabilidad fenotípica debido a factores ambientales incontrolados, por lo que se dice que tienen una baja “homeostasis genética”. En contraste con la depresión consanguínea, el cruzamiento de dos líneas consanguíneas diferentes generalmente produce una descendencia híbrida genética y fenotípicamente uniforme, mostrando una mayor homeostasis genética y un mayor “vigor híbrido”, manifestado por un incremento de las características generales de la eficacia biológica.

La mayor homeostasis de los híbridos se explica porque las dos líneas consanguíneas probablemente son homocigóticas para distintos alelos, el híbrido expresa el genotipo dominante para muchos loci relacionados con la eficacia biológica. El vigor híbrido se ha utilizado mucho para incrementar la producción en muchos cultivos vegetales, tales como el maíz.

Conviene precisar que el verdadero vigor híbrido siempre resulta en un incremento en la capacidad reproductora, mientras que incrementos en caracteres que no mejoran directamente la eficacia biológica como, por ejemplo, el tamaño, se consideran más bien “exhuberancia”. Es el caso de la mula, el híbrido entre asno y yegua, que presenta mucha exhuberancia porque es superior a sus progenitores en muchos aspectos físicos importantes, pero no vigor híbrido, ya que siempre resulta estéril.

Variabilidad genética

La evolución desde los organismos primitivos hasta los más de dos millones de especies descritas en la actualidad, ha ocurrido gracias a numerosos cambios en la cantidad y calidad del material hereditario (ADN). Estos cambios son las “mutaciones”, debidas a errores en el almacenaje de la información genética y posterior propagación por replicación. Las mutaciones son la ‘fuente primaria de variabilidad genética’. La variabilidad se origina también por recombinación meiótica. Mientras que la mutación introduce nuevos alelos en la población, la recombinación genera nuevas combinaciones de genes que determinan la singularidad de cada individuo.

Podemos distinguir dos tipos principales de mutación: “génica” (o puntual), que afecta sólo a uno o a unos pocos nucleótidos de un gen, y “cromosómica”, que afecta al número de cromosomas, al número de genes de un cromosoma o a la ordenación de los genes dentro de un cromosoma.

Las mutaciones génicas pueden tener mayor o menor efecto sobre los organismos, dependiendo de cómo afectan a las proteínas correspondientes. Una sustitución de un par de bases que no origine cambios en la secuencia de aminoácidos de la proteína codificada, no tendrá efecto sobre la capacidad del organismo para sobrevivir y reproducirse. Las mutaciones que cambian uno o incluso varios aminoácidos también pueden tener efectos pequeños o nulos en el organismo, si no está afectada la función biológica esencial de la proteína codificada. Sin embargo, las consecuencias pueden ser severas si la sustitución afecta al lugar activo de una enzima o modifica de alguna manera la función esencial de la proteína.

Por otra parte, los efectos de las mutaciones dependen de las condiciones ambientales, es decir, puede ocurrir que una misma mutación resulte perjudicial o no a un individuo dependiendo del ambiente en que viva. Por ejemplo, en nuestra especie, la fenilcetonuria se debe a la homocigosis de un alelo recesivo e imposibilita a los individuos metabolizar la fenilalanina, lo que origina su retraso mental. Pero estos individuos pueden vivir normalmente con una dieta libre de fenilalanina.

Debido a la acción de la selección natural en el pasado, podemos admitir que las poblaciones se encuentran generalmente bien adaptadas en el ambiente donde viven. Por ello, las mutaciones nuevas suelen ser deletéreas. Sin embargo, puede ocurrir que una mutación de origen reciente resulte beneficiosa a sus portadores, si ocurren cambios ambientales o la población emigra a un nuevo hábitat. En estos casos, puede que la adaptación de la población en el nuevo ambiente le lleve hacia una nueva constitución genética óptima.

Tasas de mutación

La tasa de mutación de un gen indica su capacidad de mutación. Podemos definirla como la probabilidad de que ocurra una mutación por entidad biológica (virus, célula, individuo) y por generación.

Las tasas de mutación son, en general, menores en bacterias y otros microorganismos que en organismos pluricelulares. En *Escherichia coli* es del orden de 10^{-8} a 10^{-9} . En los seres humanos y otros organismos pluricelulares la tasa de mutación está generalmente comprendida entre 10^{-5} y 10^{-6} por locus, por generación y por gameto. Pero existe considerable variación de un gen a otro y de un organismo a otro. No se conocen las razones exactas de estas diferencias en las tasas de mutación, pero podrían ser un reflejo de la eficiencia relativa de los sistemas enzimáticos cuya función es reparar los errores generados durante la replicación.

Al nivel molecular, la tasa de mutación se expresa “por par de bases y ronda de replicación” (μ_b). Recientemente, Drake et al. (1998) han realizado un interesante análisis comparativo de la tasa de mutación espontánea en diversos organismos. En microorganismos con ADN como material genético, μ_b varía entre 7.2×10^{-7} en el bacteriófago M13 hasta 7.2×10^{-11} en el hongo *Neurospora crassa*, pasando por 5.4×10^{-10} en *Escherichia coli* y 2.2×10^{-10} en la levadura. Este rango tan amplio se reduce si consideramos el tamaño genómico de estos organismos, es decir, expresando la tasa de mutación “por genoma y ronda de replicación” (μ_g), siendo todos los valores próximos a 0.0034. En eucariotas superiores, μ_b también tiende a ser menor en los genomas pequeños, variando entre 2.3×10^{-10} en *Caenorhabditis elegans* y 5×10^{-11} en los seres humanos. La tasa por genoma (μ_g) es más variable que en el caso anterior (entre 0.018 y 0.49). Teniendo en cuenta que gran parte del genoma de los eucariotas superiores es susceptible de sufrir mutaciones neutras (debido a la existencia de los intrones y los espaciadores intergénicos, principalmente), y para una mejor comparación con los microorganismos, es mejor expresar la tasa de mutación por “genoma efectivo (la porción del genoma donde la mayoría de las mutaciones son deletéreas) y ronda de replicación” (μ_{eg}). En ese caso, los valores difieren mucho menos entre los eucariotas analizados (0.004-0.014) y, a su vez, difieren poco de los observados en microorganismos ($\mu_g = 0.0034$). Los organismos con ARN como material genético, sin embargo, son los que tienen las mayores tasas de mutación: $\mu_g \approx 1-2$ para los

virus líticos y $\mu_g \approx 0.1-0.2$ para los retrovirus. Para conocer las implicaciones evolutivas de las tasas de mutación espontánea, ver Drake et al. (1998) y el capítulo 25 de este libro.

Mutaciones inducidas y agentes mutagénicos

Las mutaciones que ocurren en la naturaleza, se denominan mutaciones espontáneas y, en general, son debidas a cambios aleatorios de la secuencia nucleotídica de los genes durante el proceso enzimático de la replicación. Pero las mutaciones también pueden ser el resultado de cualquier factor ambiental y, en este caso, se denominan mutaciones inducidas. La posibilidad de inducir las mutaciones, con el descubrimiento del poder mutagénico de las radiaciones y ciertos agentes químicos, supuso un enorme avance en la investigación genética. Con estas herramientas ya fue posible imitar a la naturaleza en la producción de variabilidad genética.

Entre los agentes mutagénicos, cabe destacar las radiaciones y los mutágenos químicos. La “radiación” es un proceso mediante el cual la energía se traslada en el espacio. Las radiaciones pueden ser “ionizantes” o “no-ionizantes”. Las ionizantes producen iones que pueden reaccionar física y químicamente al ponerse en contacto con moléculas biológicas. Las radiaciones ionizantes pueden ser ondulatorias o corpusculares, según estén compuestas por ondas electromagnéticas (rayos X y rayos gamma), o por partículas subatómicas que han adquirido una alta energía por su velocidad (rayos alfa y beta, protones acelerados, neutrones, etc.). Los pares iónicos inducidos por las radiaciones ionizantes dan lugar a radicales libres altamente reactivos que inducen transformaciones químicas en los genes (efectos genéticos), en los cromosomas (efectos citogenéticos), en los enzimas (efectos fisiológicos), etc. Los factores que determinan el efecto de la radiación ionizante son: el tipo de radiación, la dosis utilizada, el material irradiado y factores externos como la temperatura, el agua y el oxígeno.

La única radiación no-ionizante capaz de inducir mutaciones con frecuencia superior a la tasa espontánea es la luz ultravioleta. El efecto más importante de la luz ultravioleta se produce en las pirimidinas, formándose dímeros, especialmente, entre dos residuos de timina, aunque también se pueden formar, pero con menor frecuencia, dímeros citosina-citosina y timina-citosina. Se cree que los dímeros distorsionan la conformación del ADN e inhiben la replicación normal, siendo responsables, al menos en parte, de los efectos letales de la radiación ultravioleta en los microorganismos.

Nuestros genes se enfrentan continuamente a una radiación de fondo proveniente de elementos naturales, tales como el uranio y el torio del suelo, y el fósforo, potasio y calcio de nuestros huesos, y también de los rayos cósmicos.

Los “mutágenos químicos” tienen dos características generales importantes: su efecto retardado en la producción de mutaciones y su especificidad de acción. Algunos agentes químicos mutagénicos son, por ejemplo, i) los

agentes alquilantes como el gas mostaza, el etilmetano-sulfonato (EMS) y la nitrosoguanidina (NG), que ceden grupos alquilo (etilo o metilo) a grupos amino o ceto de los nucleótidos, alterando las afinidades de emparejamiento y provocando transiciones (sustitución de una base púrica por otra púrica, o bien una pirimidínica por otra pirimidínica); ii) los análogos de bases (halouracilos y derivados de la uridina, aminopurinas), que son moléculas que pueden sustituir a las purinas o a las pirimidinas durante la biosíntesis de los ácidos nucleicos; iii) los derivados del nitrógeno (hidrazina, hidroxilamina); iv) el peróxido de hidrógeno y peróxidos orgánicos; v) las sales metálicas y vi) los ésteres del ácido fosfórico.

Carácter preadaptativo de la mutación

Muchas veces se oye decir, en el lenguaje común, que cuando se hacen tratamientos continuados con insecticidas contra alguna plaga, al cabo de cierto tiempo resultan ineficaces, aceptándose la conclusión de que el propio insecticida había producido mutaciones en los individuos de la especie combatida, haciéndolos resistentes a la sustancia utilizada como insecticida. Si eso fuera cierto, el propio ambiente sería quien induciría la aparición de mutaciones beneficiosas, lo que implicaría un concepto lamarckiano: el carácter postadaptativo de la mutación, ya que supone que la mutación surge como una consecuencia “dirigida” por el propio ambiente.

Realmente el mecanismo adaptativo es al contrario: en la población existe variabilidad genética sobre la que actúa el ambiente favoreciendo o seleccionando a los que mejor se adaptan a él: es el carácter “preadaptativo” de la mutación. En el caso de los insecticidas, en la población ya existía la mutación que proporcionaba la resistencia a la droga antes de aplicar los tratamientos. Al aplicar éstos, el ambiente favoreció a los individuos que ya eran resistentes y, consecuentemente, éstos aumentaron en frecuencia. Podemos decir que la mutación es un proceso que ocurre al azar en dos sentidos: aunque podemos predecir la probabilidad con que aparecerá una mutación, no podemos predecir en qué copia ocurrirá. El proceso es estocástico, no determinista, y en segundo lugar, la mutación ocurre independientemente de que al organismo le resulte beneficiosa o perjudicial en ese ambiente. Es la selección natural quien, posteriormente, hará que la frecuencia de ese nuevo alelo mutante aumente o no, según resulte beneficioso o perjudicial, en función de las condiciones ambientales en que viva esa población.

El carácter preadaptativo de la mutación quedó claramente demostrado por varios experimentos clásicos, como la prueba de la fluctuación (Luria y Delbrück, 1943), la prueba de la réplica en placa (Lederberg y Lederberg, 1952) y la selección de moscas resistentes al DDT sin haber entrado en contacto con el insecticida (Bennet, 1960).

Resultados recientes han sugerido la posibilidad de que algunos organismos, tales como las bacterias, puedan “seleccionar” un conjunto específico de mutaciones producidas como resultado de presiones ambientales (Cairns

et al. 1988, Hall 1988, 1990). Cairns y sus colaboradores diseñaron un protocolo para detectar “mutaciones no aleatorias” que surgiesen en respuesta a factores del ambiente en el que se cultivaban las bacterias (*E. coli*). Los resultados de estos experimentos sugirieron que algunas mutaciones ventajosas podrían ser dirigidas por el ambiente en condiciones de presión nutricional (ayuno). En este caso, las bacterias podrían activar mecanismos que crean un estado hipermutable en genes que incrementan la supervivencia. El experimento de Hall demostró una respuesta adaptativa semejante. Estos fenómenos no se pueden explicar aún de forma adecuada ya que se ignoran los mecanismos fundamentales y las tasas de mutación en las células que no están creciendo. Sus interpretaciones han sido criticadas por otros investigadores y actualmente son tema de debate entre los genéticos.

Tipos de mutaciones génicas y sistemas de reparación

Puesto que la lectura de la secuencia génica se realiza en tripletes y cada secuencia de tres nucleótidos específica un único aminoácido en el polipéptido, cualquier cambio que modifique o interrumpa esta secuencia proporciona una base suficiente para la mutación. El cambio más simple es la sustitución de un solo nucleótido, pero puede haber también inserciones y deleciones de uno o más nucleótidos. La Tabla 22 muestra una clasificación de las mutaciones puntuales indicando los principales tipos de cambio en el ADN y sus efectos funcionales en las proteínas.

Los organismos disponen de diversos sistemas enzimáticos para reparar los daños en el ADN de varias formas diferentes. Estos sistemas de reparación son esenciales para la supervivencia de los organismos, y su fallo conduce a un aumento en la tasa de mutación. La reparación puede llevarse a cabo mediante “sistemas que neutralizan la acción de mutágenos potenciales”. Por ejemplo, las enzimas superóxido dismutasa y catalasa catalizan la conversión del superóxido de hidrógeno en peróxido de hidrógeno y de éste en agua, respectivamente. También puede realizarse por “reversión directa de la lesión”. Este sistema consiste en eliminar directamente el error y repararlo a continuación. Algunas lesiones pueden ser reparadas de esta forma, como, por ejemplo, los fotodímeros mutagénicos producidos por la luz ultravioleta. En este proceso de reparación está implicada una enzima, la fotoliasa, que necesita un fotón de luz para revertir total o parcialmente la lesión (Fig. 12). Otras enzimas que eliminan errores de forma directa son las transferasas del grupo alquilo, que retiran los grupos alquilo que se hayan añadido a las posiciones 0-6 de la guanina por acción de mutágenos como la NG y el EMS.

La reparación “por escisión” comienza con una endonucleasa que rompe un enlace fosfodiéster a cada lado de la lesión, quedando un “hueco” que se rellena por síntesis de reparación y es sellado por una ligasa. Esta vía de reparación es importante en humanos. Existe una enfermedad, xeroderma pigmentosum, originada por la carencia de una de las enzimas de la reparación por escisión.

Tabla 22
Tipos de mutaciones génicas

Tipo	Consiste en
Transición	Sustitución de una purina por otra purina, o de una pirimidina por otra pirimidina (Ejemplos: A sustituido por G, G por A, C por T y T por C).
Transversión	Sustitución de una purina por una pirimidina, o viceversa (Ejemplos: A sustituido por C, A por T, G por T, G por C, T por G, T por A, C por A y C por G).
Cambio de fase	Inserción o delección de uno o varios nucleótidos que, cuando no es múltiplo de tres, cambia la fase de lectura del ARNm y, a partir de ahí, se altera la composición aminoacídica del polipéptido o se interrumpe su síntesis.
No sinónima (o de reemplazamiento)	Cambio a otro codón que determina un aminoácido diferente.
Sinónima	Cambio a otro codón que determina el mismo aminoácido.
Sin sentido	Cambio a un codón de terminación.

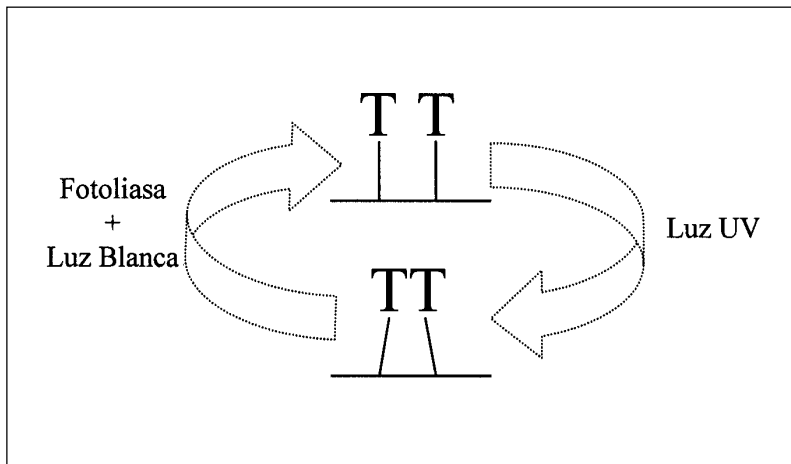


Figura 12. La luz ultravioleta produce dímeros de timina que son reparados por la enzima fotoliasa. La enzima reconoce el fotodímero, se une a él y en presencia de luz blanca, rompe el dímero en sus monómeros originales.

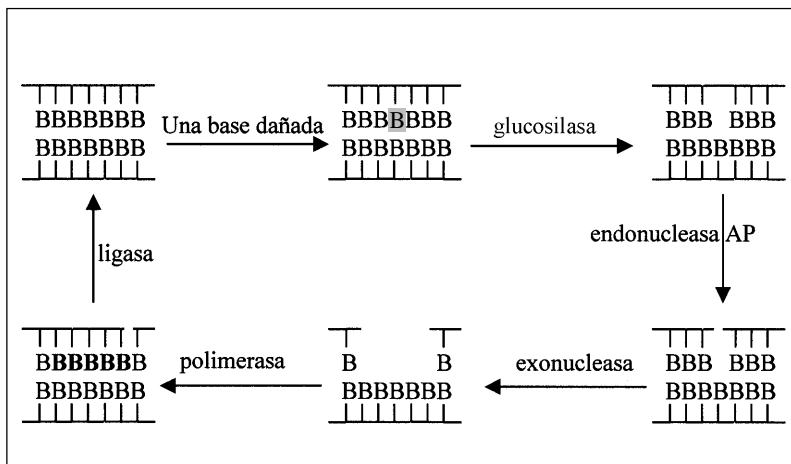


Figura 13. Reparación del ADN. La enzima glucosilasa elimina la base alterada y deja un sitio AP. Las endonucleasas AP reconocen los sitios AP y cortan el enlace fosfodiéster. Una exonucleasa elimina el fragmento de ADN, dejando un hueco que es relleno por la ADN polimerasa I. Posteriormente una ligasa sella las muescas.

Los pacientes son muy sensibles a la luz y suelen morir de cáncer de piel antes de los treinta años de edad. Los sistemas de escisión específicos, que reparan las lesiones que escapan al mecanismo general de escisión, son llevados a cabo mediante la intervención de dos enzimas: las glucosilasas del ADN, que eliminan la base alterada, y las endonucleasas AP que reconocen sitios apurínicos o apirimidínicos e introducen un corte en la cadena afectada para que luego actúen la exonucleasa (que escinde un fragmento de ADN), la polimerasa (que sintetiza nuevo ADN) y la ligasa (que sella la muesca) (Fig. 13).

Finalmente, la “reparación postreplicativa” se realiza mediante sistemas que reconocen las lesiones incluso después de la replicación. Uno de ellos es el sistema de “reparación de emparejamientos erróneos”, que detecta las bases mal emparejadas, determina cuál de las dos bases es la incorrecta (ello es posible porque puede distinguir la cadena molde, que está metilada en las secuencias GATC, de la cadena recién sintetizada, que tarda varios minutos en metilarse). Por último, este sistema escinde la base incorrecta y rellena el hueco por síntesis reparadora. Otra vía de reparación postreplicativa es la “reparación por recombinación”: el sistema de replicación, en ocasiones, ante una lesión, deja un hueco y continúa de nuevo la replicación. El hueco se rellena posteriormente con ADN procedente de la molécula hermana. Si la replicación se interrumpe totalmente, el cortocircuito a este bloqueo requiere la actuación del “sistema SOS”, en el que se activan genes que permiten una replicación menos exacta y posibilitan la supervivencia de la célula a cambio de cierto grado de mutagénesis.

Mutaciones cromosómicas

Algunas mutaciones cromosómicas producen cambios en la localización de los genes sobre los cromosomas (p. ej., las inversiones y las translocaciones), otras cambian el número de genes en los cromosomas (delecciones y duplicaciones) y otras el número de cromosomas (fusión y fisión céntricas, aneuploidía y poliploidía).

Inversiones y translocaciones

La “inversión” es una alteración estructural mediante la cual un segmento cromosómico cambia de sentido dentro del pro-

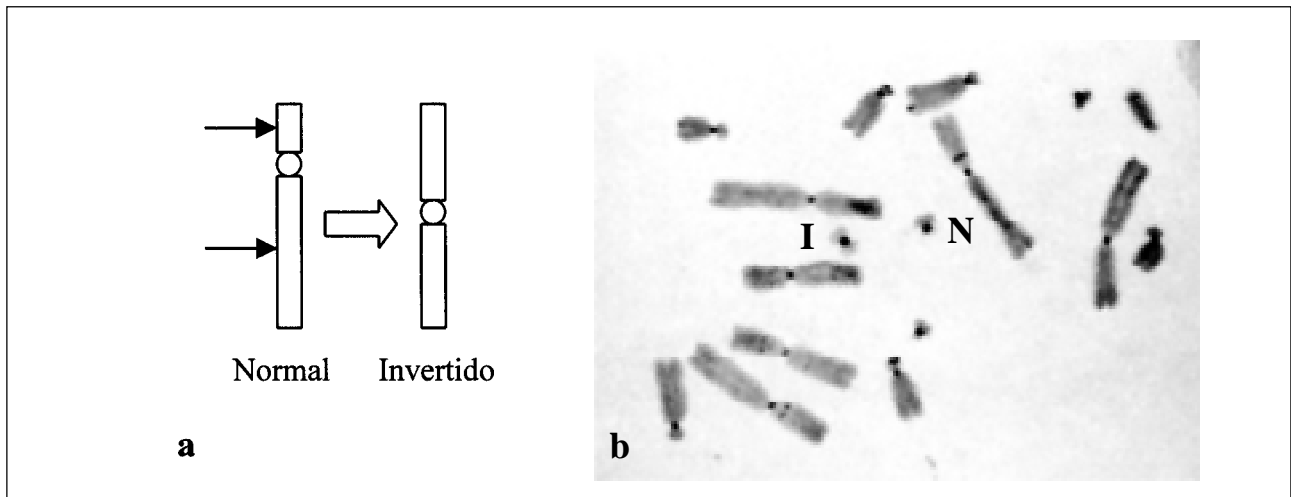


Figura 14. La inversión puede producir cambios en la relación del tamaño de los brazos cromosómicos. a) Debido a una inversión pericéntrica, un cromosoma acrocéntrico pasa a ser metacéntrico. b) Metafase mitótica de un embrión del saltamontes *Chorthippus vagans*, heterocigoto para una inversión pericéntrica en el cromosoma M_7 (N, cromosoma normal, I, cromosoma invertido).

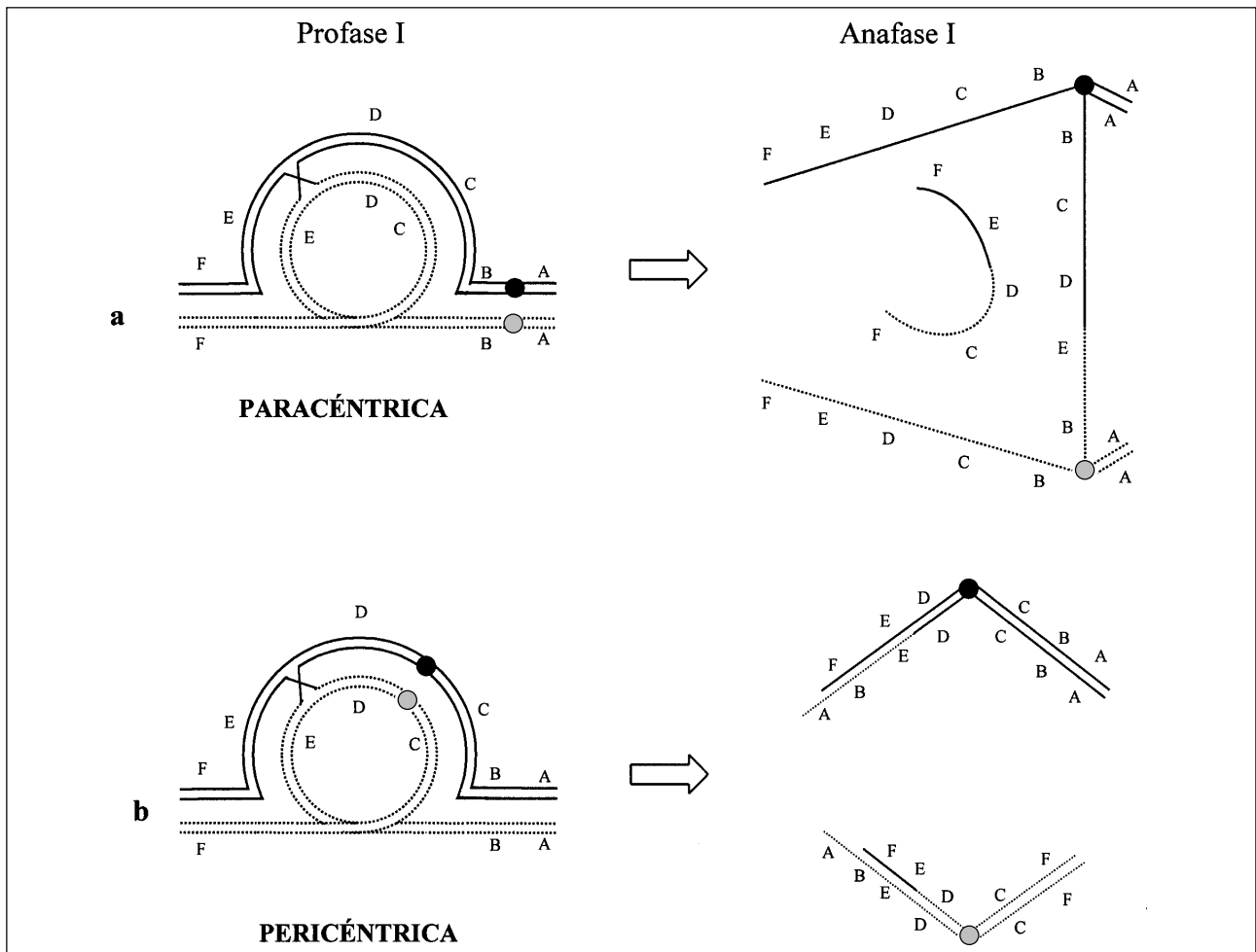


Figura 15. Consecuencias genéticas de la sinapsis y el sobrecruzamiento meióticos en individuos heterocigotos para inversiones cromosómicas. a) Inversiones paracéntricas: Formación de un bucle de apareamiento en profase I (cigotene-paquitene) y un puente dicéntrico más un fragmento acéntrico en anafase I. El fragmento acéntrico acabará perdiéndose, por no tener centrómero, y los gametos que reciban las cromátidas participantes en el puente (una vez que éste se rompa) serán inviables por faltarles algunos marcadores genéticos. b) Inversiones pericéntricas: El centrómero está incluido dentro del bucle de apareamiento y, en anafase I y metafase II, pueden observarse cromosomas con cromátidas desiguales (si el segmento invertido no es simétrico respecto al centrómero) debido a duplicaciones y deleciones en las cromátidas que participaron en el sobrecruzamiento. Los gametos que reciban estas cromátidas serán, por tanto, inviables.

pio cromosoma, modificando la ordenación de los loci contenidos en él. Una inversión es siempre intersticial, es decir, supone la rotura del cromosoma en dos puntos. Las inversiones pueden ser simples o complejas, dependiendo de si sólo cambia de sentido un segmento cromosómico o varios. Además, pueden ser “pericéntricas” (si el segmento invertido contiene el centrómero) o “paracéntricas” (si no lo contiene). La inversión pericéntrica puede cambiar el tamaño relativo de los brazos cromosómicos y producir un cromosoma de aspecto muy distinto del original. Por ejemplo, un cromosoma acrocéntrico puede transformarse en uno metacéntrico mediante una inversión que incluya longitudes desiguales del cromosoma situadas a derecha e izquierda del centrómero (Figura 14). Gran parte de la evolución cariotípica ha sido, presumiblemente, producida por estas alteraciones cromosómicas.

Las consecuencias genéticas de las inversiones dependen del comportamiento citológico durante la meiosis en los individuos heterocigotos. Para que el cromosoma normal y el invertido puedan llevar a cabo el apareamiento homólogo completo, es necesario que se forme un bucle en cigotene. La formación de quiasmas en la región invertida genera un puente dicéntrico y un fragmento acéntrico en anafase I, en el caso de las inversiones paracéntricas, y cromosomas con cromátidas desiguales, en el caso de las pericéntricas (Fig. 15). El resultado es que, aproximadamente, la mitad de los gametos son recombinantes e inviables. La otra mitad, los gametos parentales, son viables, y la mitad de ellos contienen el cromosoma invertido con la secuencia alélica intacta. Podemos decir, por tanto, que las inversiones son supresoras de la recombinación, ya que los genes contenidos en el segmento invertido mantienen una combinación alélica constante y tienden a transmitirse como una unidad, constituyendo lo que Darlington y Mather (1949) definieron como un “supergen”.

A pesar de la teórica semiesterilidad de los individuos heterocigóticos, las inversiones son muy frecuentes en algunos grupos, constituyendo polimorfismos estables. ¿Cómo pueden, entonces, llegar a ser frecuentes las nuevas inversiones? En *Drosophila*, donde son muy frecuentes los polimorfismos para inversiones paracéntricas, existen mecanismos que impiden el sobrecruzamiento entre la secuencia normal y la invertida ya que los machos tienen meiosis aquíasmática (con lo que todos sus gametos son viables), y en las hembras se produce la migración preferencial de las cromátidas parentales (normal e invertida) intactas, hacia la ovótida, eliminándose los puentes y fragmentos en los corpúsculos polares y evitando así la infertilidad gamética. En los saltamontes, las secuencias normal e invertida sufren apareamiento recto no homólogo (Del Cerro y Santos, 1995), o bien, asinapsis, con lo que los heterocigotos tampoco ven reducida su fertilidad (Fig. 16). Estos comportamientos evitan la subdominancia que, teóricamente, sería de esperar para estos polimorfismos. El mantenimiento de supergenes en las regiones invertidas, por la ausencia de recombinación, podría representar, en algunas circunstancias, una cierta ventaja que ayuda a explicar la frecuencia de estos polimorfismos.

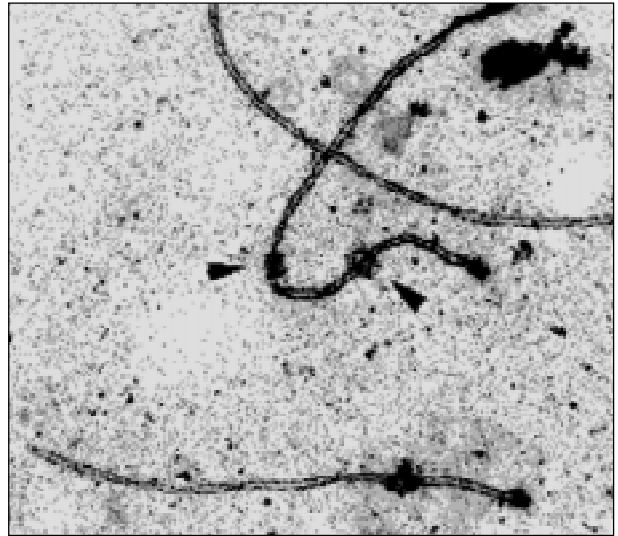


Figura 16. Una forma de evitar la inviabilidad gamética en heterocigotos para inversiones es el apareamiento recto no homólogo. La figura muestra los dos elementos laterales del complejo sinaptonémico en un individuo del saltamontes *Chorthippus vagans* heterocigótico para la inversión mostrada en la Figura 3b. Obsérvese que los centrómeros (flechas) de los dos cromosomas están situados en diferente posición. (Fotografía cortesía de Juan Luis Santos).

Una “translocación” es un cambio estructural en el que algún segmento cromosómico cambia de posición relativa dentro del complemento cromosómico, modificando, por tanto, los grupos de ligamiento. Un segmento cromosómico puede cambiar de posición dentro del mismo cromosoma (translocación intracromosómica), o bien cambiar de cromosoma (translocación intercromosómica). Dentro de estas últimas, las translocaciones recíprocas implican un intercambio de segmentos entre dos cromosomas, pudiendo afectar a brazos cromosómicos completos, o casi completos, cuando las roturas se producen en lugares próximos a los centrómeros (translocaciones de brazo completo). Los tipos más frecuentes en el reino animal son la fusión y la disociación céntrica, mediante las cuales, respectivamente, dos cromosomas telocéntricos se fusionan por sus centrómeros para dar lugar a un solo cromosoma con dos brazos o, por el contrario, un cromosoma con dos brazos se escinde transversalmente por el centrómero originando dos cromosomas telocéntricos (Fig. 17).

Cuando ocurre una translocación recíproca, el heterocigoto posee cuatro cromosomas con homología parcial. Ello determina que el apareamiento homólogo en meiosis requiera la formación de una cruz paquiténica con dos brazos constituidos por los segmentos cromosómicos intercambiados y los otros dos por el resto de las regiones cromosómicas (Fig. 18). Estos últimos brazos contienen los segmentos intersticiales, las regiones comprendidas entre los centrómeros y los puntos de rotura, que son de especial significación, como veremos después.

La configuración que se observará en metafase I dependerá del lugar donde haya quiasmas y del tipo de orien-

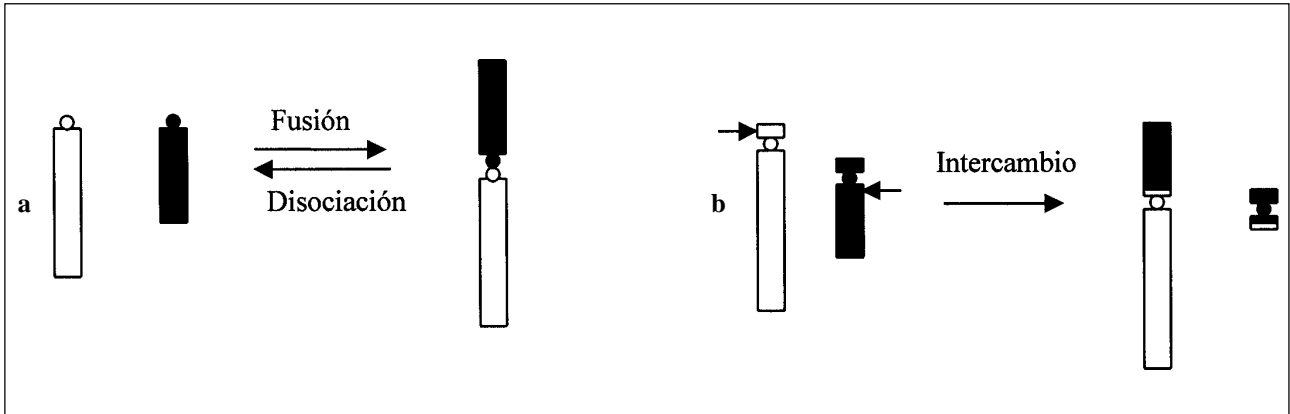


Figura 17. Cambios en el número cromosómico originados por fusiones y disociaciones céntricas. a) La fusión céntrica, en sentido estricto, se produce mediante la fusión de dos cromosomas telocéntricos a nivel de sus centrómeros, dando lugar a un único cromosoma con dos brazos, que será más o menos dicéntrico en función de a qué nivel se produzcan las roturas centroméricas. El proceso es reversible y, mediante la disociación céntrica, un cromosoma con dos brazos se divide transversalmente por su centrómero originando dos cromosomas telocéntricos. b) Un resultado equivalente al de la fusión céntrica se produce mediante una translocación recíproca entre dos cromosomas acrocéntricos, con puntos de rotura próximos a los centrómeros, pero en este caso se produce un minicromosoma que generalmente se pierde.

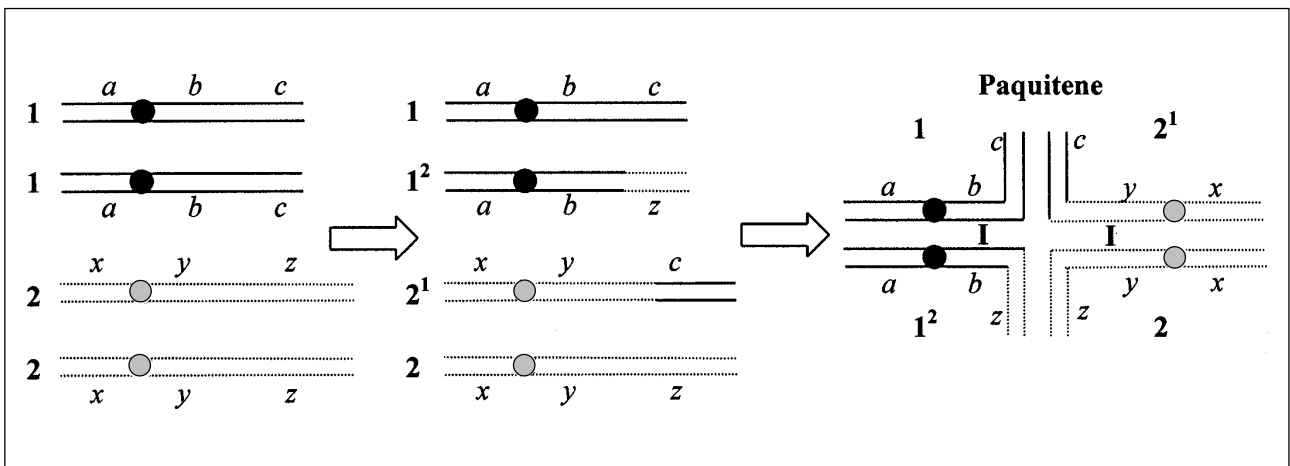


Figura 18. Translocación recíproca y formación de la cruz paquiténica durante la profase I de la meiosis, por apareamiento de los cuatro cromosomas que muestran homología parcial: 1, 1², 2¹ y 2. I= segmentos intersticiales.

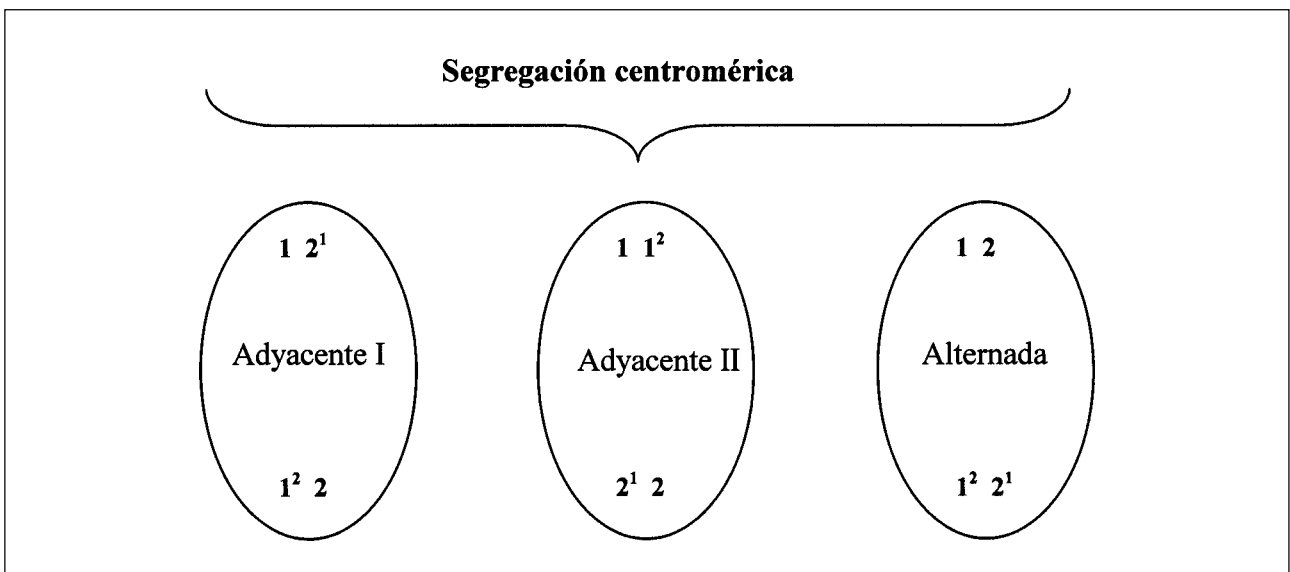


Figura 19. Tipos de segregación centromérica en anafase I de la meiosis.

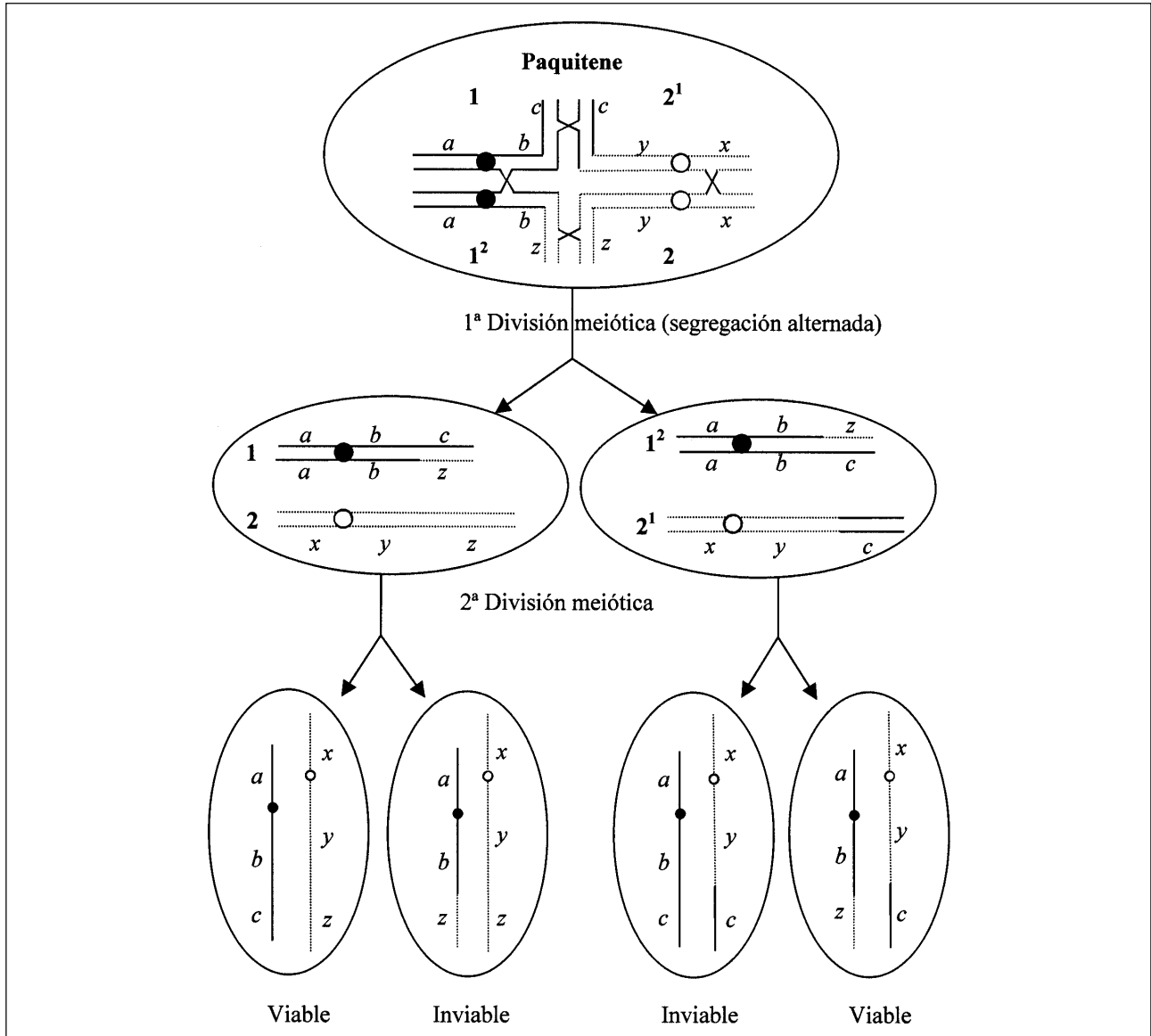


Figura 20. Inviabilidad gamética causada por la formación de quiasmas en los segmentos intersticiales.

tación centromérica. Si los cuatro cromosomas permanecen asociados por la existencia de quiasmas en los cuatro brazos, formarán un anillo que puede mostrar tres tipos de segregación centromérica (Fig. 19): “adyacente I” (con los centrómeros homólogos segregando a polos opuestos) que produce el 100% de gametos inviables, “adyacente II” (con los centrómeros homólogos segregando al mismo polo) que también produce el 100% de gametos inviables, y “alternada” (una configuración en zig-zag, en forma de ocho, con los dos cromosomas translocados yendo al mismo polo y los normales al otro) cuyo resultado es el 100% de gametos viables. Pero esto sólo es cierto si no hubo quiasmas en los segmentos intersticiales. Si los hubiera, se esperaría un 50% de gametos genéticamente desequilibrados (Fig. 20).

Para que un individuo heterocigótico para una translocación recíproca no vea reducida su fertilidad gamética, deben darse tres condiciones simultáneamente: 1) que se apareen los cuatro cromosomas implicados, 2) que no se formen quiasmas en los segmentos intersticiales,

y 3) que los centrómeros orienten alternadamente. En realidad, es bastante improbable que se den las tres condiciones, por lo que las translocaciones recíprocas producen severas reducciones de fertilidad en los heterocigotos. Estas reducciones son menores cuando se intercambian brazos cromosómicos completos, como en el caso de la fusión céntrica o translocación robertsoniana, porque, en ese caso, no existen segmentos intersticiales.

Con la reducción que las translocaciones recíprocas producen en la eficacia biológica de los heterocigotos, es difícil explicar cómo pueden aumentar en frecuencia y llegar a fijarse en una población, es decir, que todos los individuos sean homocigotos para la translocación. Existen varios mecanismos que pueden ayudar a una translocación a aumentar en frecuencia a pesar de su subdominancia, entre los que destacan los procesos reproductivos alternativos a la reproducción sexual (que evitan la meiosis), tales como diferentes modos de reproducción vegetativa. De esta forma, la translocación puede multiplicarse a partir del primer individuo heterocigótico y aumentar

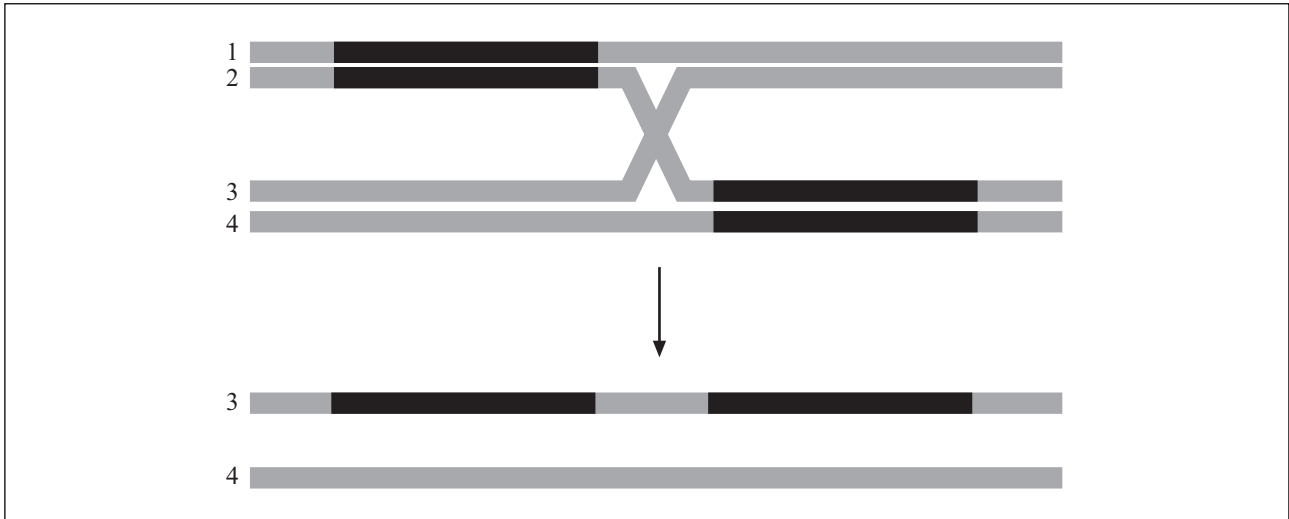


Figura 21. Duplicación (cromátida número 2) y deleción (cromátida número 3) del marcador representado en negro, mediante sobrecruzamiento desigual.

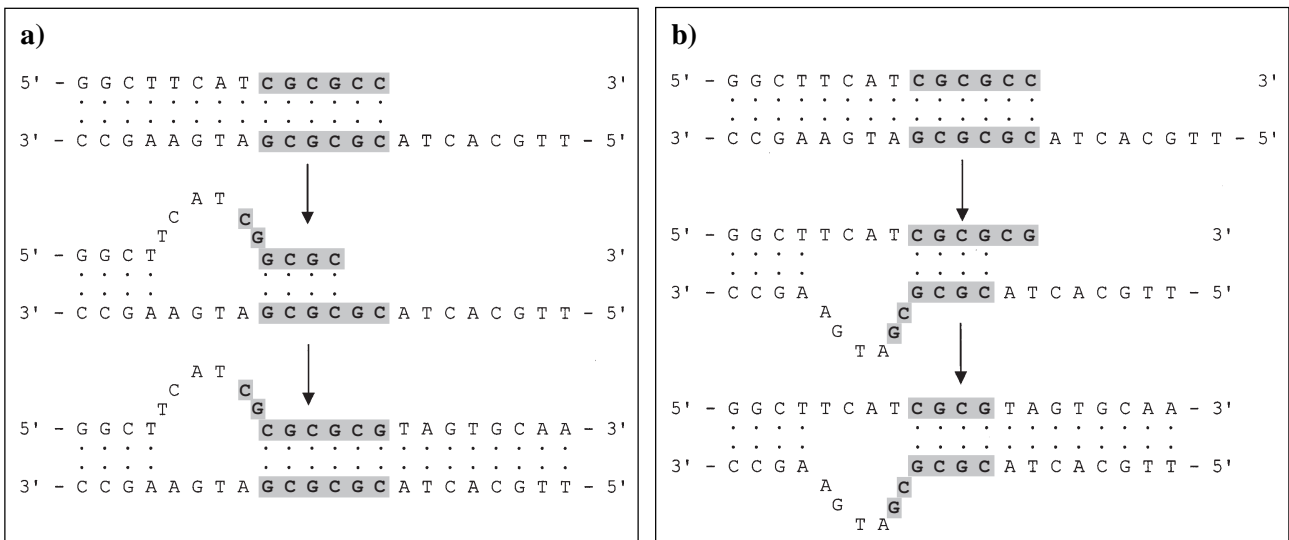


Figura 22. Origen de duplicaciones y deleciones por deslizamiento de hebra. a) El deslizamiento de dos bases de la secuencia GC repetida, en la cadena copia (deslizamiento de hebra en dirección 3'→5'), da lugar a la duplicación de una unidad GC. b) Si el deslizamiento ocurre en la cadena molde, el resultado es la deleción de una unidad GC.

en frecuencia en la población. Por ello, los pocos polimorfismos para translocaciones recíprocas que se han encontrado en poblaciones naturales, afectaban a especies vegetales, donde existen muchos tipos de reproducción vegetativa. Otros procesos, como la deriva genética y el impulso meiótico, pueden contribuir también a explicar el incremento en frecuencia de las translocaciones subdominantes.

En algunas especies vegetales, como *Oenothera lamarckiana*, las translocaciones forman parte importante de su sistema genético, ya que las plantas de esta especie son todas heterocigóticas para translocaciones múltiples que afectan a 12 de sus 14 cromosomas. En los eucariotas, tanto la fusión como la disociación céntrica han jugado un importante papel en la evolución cariotípica de muchos grupos (King 1995).

Deleciones y duplicaciones

Una “deleción” supone la pérdida de una parte de un brazo cromosómico. Una deficiencia pequeña dentro de un gen (deleción intragénica), inactiva a dicho gen y tiene el mismo efecto que otras mutaciones nulas de ese gen. Las deleciones intragénicas se distinguen de los cambios de un solo nucleótido porque no son reversibles. Las deleciones multigénicas (que afectan a varios genes) tienen consecuencias más graves. Generalmente resultan letales en homocigosis, e incluso un organismo heterocigótico para una deleción puede no sobrevivir debido a la alteración del equilibrio génico normal. Citológicamente, una deleción se puede detectar en los individuos heterocigóticos debido a la aparición de bucles cuando se produce el apareamiento homólogo.

La aparente dominancia de alelos recesivos (pseudodominancia) es también indicativo de la existencia de deleciones ya que la falta del segmento correspondiente, “desenmascara” alelos recesivos presentes en el homólogo normal que, inesperadamente, se ponen de manifiesto. Debido a su carácter deletéreo, las deleciones han podido jugar un papel menor en la evolución, pues sólo son toleradas si el segmento delecionado tiene escaso contenido genético como, por ejemplo, cuando ocurren fusiones o disociaciones y se pierde algún centrómero con la heterocromatina que lo rodea.

La existencia de copias extra de algunas regiones cromosómicas se conoce como “duplicación”. Las regiones duplicadas pueden estar situadas una junto a otra (duplicaciones en tándem), o bien pueden tener diferentes localizaciones, pudiendo incluso estar en cromosomas diferentes (duplicación insercional). En meiosis, los heterocigotos para una duplicación en tándem forman un bucle que representa la región extra no emparejada. Tanto las duplicaciones como las deleciones, pueden originarse por sobrecruzamiento desigual cuando los dos cromosomas homólogos aparean erróneamente (Fig. 21), o bien por deslizamiento de hebra durante la replicación (Fig. 22).

En ocasiones se producen duplicaciones o ampliaciones masivas de segmentos cromosómicos, que posteriormente pueden ser heterocromatinizados. Se trata de segmentos cromosómicos supernumerarios, que constituyen un polimorfismo muy frecuente en algunos organismos como, por ejemplo, los saltamontes (Fig. 23).



Figura 23. Doble segmento supernumerario (dos flechas) en el cromosoma M_6 del saltamontes *Chorthippus jacobsi*. Obsérvese su ausencia en el M_6 estándar (una flecha).

Los segmentos duplicados pueden mantenerse intactos, o bien divergir evolutivamente. Así, podemos considerar cinco tipos de secuencias de ADN reconocibles como duplicadas:

- Duplicaciones de un único gen ancestral, seguidas de la evolución divergente de los genes duplicados. Es el caso de los genes de la mioglobina y los de las distintas cadenas de las hemoglobinas (Fig. 24).
- Duplicaciones que permanecen idénticas en secuencia y función. De esa forma se consigue la producción

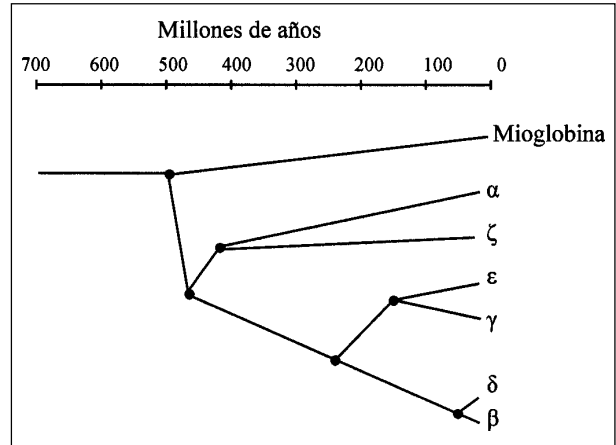


Figura 24. Evolución de la familia génica de las globinas mediante duplicaciones acontecidas en los últimos 500 millones de años.

masiva de la sustancia codificada. Es el caso de los genes que codifican para ARN transferente, ARN ribosómico, ARN para histonas, etc.

- Secuencias cortas de ADN que se encuentran repetidas hasta más de un millón de veces, aunque no todas las copias permanezcan idénticas. Constituyen el ADN altamente repetitivo, que suele formar parte de los centrómeros, los telómeros y otras regiones cromosómicas, y cuya función aun no está totalmente comprendida.
- Duplicaciones dentro de los propios genes, como se deduce de la presencia de regiones repetidas de aminoácidos (dominios polipeptídicos) dentro de la cadena pesada de clase gamma de la inmunoglobulina G, en la haptoglobina a-2 humana, en la ferredoxina y en otras proteínas.
- Pseudogenes: segmentos de ADN no funcional que poseen una alta homología de secuencia con respecto a genes funcionales conocidos.

Podemos deducir, por tanto, que las duplicaciones han jugado un papel importante en el origen y evolución de muchos genes y, en general, en la evolución del genoma.

Aneuploidía y poliploidía

En un individuo **aneuploide**, la dotación cromosómica aumenta o disminuye en uno o más cromosomas. Un individuo diploide normal, es disómico para cada cromosoma del cariotipo. Si falta un cromosoma de un par, el individuo se denomina monosómico y si faltan los dos cromosomas de un par, nulisómico. Los polisómicos tienen un cromosoma representado por más de dos homólogos: tres veces en los trisómicos ($2n+1$), cuatro veces en los tetrasómicos ($2n+2$), y así sucesivamente. Casi un tercio de los abortos humanos espontáneos corresponde a fetos con un número anormal de cromosomas, siendo la mayoría trisómicos o monosómicos.

La aneuploidía se origina generalmente por no-disyunción cromosómica. La trisomía del par 21 de humanos

(Síndrome de Down), por ejemplo, se origina generalmente por no disyunción meiótica de este cromosoma y, en algunos casos (aproximadamente un 10%), por translocación entre el cromosoma 21 y un cromosoma mediano.

A veces los cromosomas adicionales son tan diferentes de los cromosomas normales (o cromosomas A), en naturaleza y comportamiento meiótico, que se les considera otra clase distinta de elemento genético. Son los cromosomas supernumerarios, o cromosomas B, que se comportan como parásitos genómicos y se mantienen en las poblaciones naturales de muchas especies merced a sus mecanismos de acumulación que les confieren ventaja en la transmisión, aún ejerciendo efectos deletéreos sobre los individuos portadores (ver Capítulo 27). Muchas veces, los cromosomas B podrían haber derivado de cromosomas A extra resultantes de aneuploidía.

Si los individuos poseen más de dos dotaciones cromosómicas completas, se denominan “poliploides”. Pueden ser triploides ($3n$), tetraploides ($4n$), pentaploides ($5n$) y así sucesivamente. En los grupos taxonómicos que normalmente son diploides, un individuo que posea una sola dotación cromosómica (n) se denomina monoploide (por ejemplo, los machos de las abejas, avispas y hormigas) para distinguirlo de aquellas especies cuyos individuos tienen normalmente un solo juego de cromosomas (haploides).

En los animales, la poliploidía es un fenómeno raro y sólo se conoce en especies hermafroditas, como las lombrices de tierra, o en especies con hembras partenogenéticas, como algunas polillas, escarabajos y salamandras. La poliploidía es rara en animales porque altera el equilibrio entre autosomas y cromosomas sexuales necesario para la determinación del sexo y, en segundo lugar, porque los animales suelen presentar fecundación cruzada y, por tanto, un poliploide no puede cruzarse consigo mismo. En plantas, sin embargo, la poliploidía puede ser frecuente en poblaciones naturales presentando, los distintos grupos, varios niveles de ploidía. Las plantas poliploides suelen ser de mayor tamaño y, por ello, son de interés económico en agricultura. Las plantas tetraploides y, en general, con un número par de dotaciones cromosómicas, no experimentan muchas dificultades durante la meiosis; pero las plantas con número impar de dotaciones cromosómicas ($3n$, $5n$...) generalmente son estériles ya que durante la meiosis se originan univalentes que pueden dar lugar a gametos desequilibrados. Esta esterilidad resulta ventajosa, desde el punto de vista comercial, en el caso de los árboles frutales, porque evita la formación de semillas no comestibles y la planta puede reproducirse vegetativamente por esquejes.

Se pueden distinguir dos tipos de poliploidía dependiendo de si los genomas extras son de la misma especie, es decir, homólogos (autoploidía), o bien de especies diferentes, homeólogos (aloploidía). La “autoploidía” puede originarse por fallos en la división reductora durante la gametogénesis, originando gametos diploides. Esto también puede ser provocado por sustancias antimitóticas, como la colchicina. Otros posibles orígenes son la activación y fusión al óvulo de los corpúsculos polares, la

polispermia (fecundación de un óvulo por más de un espermatozoide), o bien, puede ser provocada por choques térmicos, método usualmente utilizado para la obtención, por ejemplo, de truchas triploides. La “aloploidía” se origina cuando se fecundan gametos de dos especies diferentes. El híbrido resultante generalmente es estéril debido a dificultades en el apareamiento de los cromosomas homeólogos durante la meiosis. No obstante, si se logra la tetraploidía, cada cromosoma poseerá su homólogo y podrán formarse gametos diploides viables. En la naturaleza, la aloploidía parece haber sido un motor importante en la formación de nuevas especies vegetales. Un alopoloide natural interesante es el trigo de panadería *Triticum aestivum* ($6n=42$), que es un alohexaploide formado por dos dotaciones de cada una de tres especies ancestrales.

Cuantificación de la variabilidad genética

La genética de poblaciones se ocupa necesariamente de la variación genotípica, pero sólo la variación fenotípica puede ser observada directamente. La relación entre fenotipo y genotipo varía en simplicidad de carácter a carácter. La mayoría de los estudios experimentales de genética de poblaciones se han centrado en caracteres con relaciones sencillas entre fenotipo y genotipo y no en caracteres cuantitativos en los que no siempre es posible emitir afirmaciones precisas sobre su variación genotípica. Para poder realmente cuantificar la variabilidad de una población es necesario disponer de una muestra al azar de loci, y considerar tanto los invariables (monomórficos) como los variables (polimórficos). La electroforesis en gel ha sido la técnica más utilizada, desde principios de los años 60, para obtener estimas de la variación genética en poblaciones naturales de muchos organismos. En la actualidad, su condición de marcadores genéticos neutros hace a las isoenzimas, y otros marcadores moleculares, muy útiles para la construcción de filogenias.

Las técnicas electroforéticas proporcionan los genotipos de los individuos de una muestra: cuántos son homocigóticos, cuántos son heterocigóticos y para qué alelos. Con el fin de obtener una estima de la cantidad de variación en una población, se estudian 20 o más loci genéticos. Es deseable resumir la información obtenida para todos los loci de forma que exprese el grado de variabilidad de la población y permita compararla con otras poblaciones. Para ello, se utilizan dos índices de variación genética, el grado de polimorfismo y la heterocigosis. El primero viene expresado por la “proporción de loci polimórficos” (P), en una población. Por ejemplo, si utilizando técnicas electroforéticas examinamos 30 loci genéticos en cierto organismo, y no encontramos variación en 9 loci, pero existe más de un alelo diferente en los otros 21 loci. Podemos decir que $21/30 = 0.70$ de los loci son polimórficos en esa población, o que el grado de polimorfismo en la población es 0.70. Para evitar que el tamaño de muestra pueda afectar al valor de P , se suele adoptar un “criterio de polimorfismo”. Un criterio usado a menudo es considerar un locus como polimórfico úni-

camente cuando el alelo más frecuente tiene una frecuencia no superior a 0.95. Entonces, a medida que se examinan más individuos, se pueden encontrar variantes adicionales, pero, por término medio, la fracción de loci polimórficos no cambiará.

Una medida mejor de la variabilidad genética es la frecuencia promedio de individuos heterocigóticos por locus, o simplemente la “heterocigosis” (H) de la población. Esta es una buena medida de la variación porque estima la probabilidad de que dos alelos tomados al azar de la población sean diferentes. La heterocigosis observada es un buen indicador de la variación genética en poblaciones de individuos diploides con apareamiento al azar, pero no refleja bien la cantidad de variación genética en poblaciones de organismos haploides o en las poblaciones en las que hay autofecundación, como en algunas plantas, o en organismos en los que los apareamientos entre parientes son comunes. En una población que siempre se reproduce por autofecundación, la mayoría de los individuos serán homocigóticos, aunque diferentes individuos pueden llevar diferentes alelos si el locus es variable en la población. Habrá también más homocigotos en una población en la que los cruzamientos entre parientes sean comunes que en una población donde no ocurran, aun cuando las frecuencias alélicas sean idénticas en ambas poblaciones.

Esta dificultad puede ser superada calculando la “diversidad génica” para el locus i (D_i) a partir de las frecuencias alélicas y asumiendo equilibrio Hardy-Weinberg; la frecuencia de heterocigotos para un locus es:

$$h = 1 - \sum_{i=1}^m p_i^2$$

donde p_i es la frecuencia de los i alelos de ese locus, m es el número de alelos observado en ese locus y p_i^2 es la frecuencia de homocigotos. En una población diploide con apareamiento aleatorio, h es la “heterocigosis esperada” y coincide con la diversidad génica (D_i).

Estimas electroforéticas de la variación

En 1966, con la publicación de un estudio en seres humanos y dos en moscas *Drosophila*, se aplicaron por primera vez las técnicas electroforéticas para estimar la variación genética en poblaciones naturales. Desde entonces se han estudiado numerosas poblaciones de muchos organismos. Se ha analizado por este método un gran número de especies, incluyendo bacterias, hongos, plantas superiores, vertebrados e invertebrados. Los resultados son notablemente homogéneos entre las distintas especies, e indican que aproximadamente un tercio de los loci estructurales son polimórficos, y la heterocigosis media de una población, teniendo en cuenta todos los loci estudiados, es alrededor del 10%. Esto significa que, rastreando en el genoma de un individuo, deberíamos encontrar que cerca de 1 de cada 10 loci está en condición heterocigótica, y que alrededor de 1 de cada 3 loci tiene 2 ó más alelos segregando en la población. Ello representa un inmenso potencial de variación para la evolución.

Para tener una idea del significado de las grandes cantidades de variación genética encontradas en las poblaciones naturales, consideremos nuestra especie, donde se ha detectado un 6.7% de heterocigosis por electroforesis. Si suponemos que tenemos 30 000 loci génicos estructurales, cualquier persona podría ser heterocigótica en 2010 de sus loci ($30000 \times 0,067$) y podría producir, teóricamente, 2^{2010} (10^{605}) clases de gametos diferentes. (Un individuo heterocigótico en un locus puede producir dos clases diferentes de gametos, uno con cada alelo; un individuo heterocigótico en n loci génicos puede producir 2^n gametos diferentes). Estos datos indican que casi todos los individuos son únicos desde el punto de vista genético y que, por tanto, las poblaciones humanas son muy diversas. No obstante, tras cuantificar las frecuencias alélicas de 7 sistemas de grupos sanguíneos y 7 enzimas del suero sanguíneo en 7 razas: caucásicos, negros africanos, mongoloides, aborígenes del sur de Asia, amerindios, oceánicos y aborígenes australianos, Lewontin (1972) observó que la diversidad entre individuos de la misma población (85.4%) es mucho mayor que la observada entre poblaciones de la misma raza (8.3%), y que la observada entre razas (6.3%). Por tanto, la diferenciación genética entre razas humanas es mucho más pequeña que la que existe dentro de ellas. De ahí que “el conocimiento racial de un individuo proporciona poca información acerca de su estructura genética. Cada persona tiene un genotipo único, que es diferente del de cualquier otra persona, sea o no de la misma raza”.

La desventaja de la electroforesis es que sólo detecta variaciones de los genes estructurales que codifican para proteínas solubles y no para otros genes como los reguladores o los que codifican para proteínas insolubles. Además, las técnicas electroforéticas no detectan aquellas sustituciones de nucleótidos en el ADN que no cambian las propiedades electroforéticas de las proteínas.

Variación al nivel del ADN

Durante muchos años, los genéticos evolutivos esperaron el momento en que fue posible medir la variación al nivel último, es decir, en la secuencia de nucleótidos de los genes. Ese momento llegó al principio de los 80, cuando los adelantos técnicos hicieron la secuenciación del ADN relativamente fácil. Otra técnica útil es la reacción en cadena de la polimerasa (PCR), que se usa para amplificar pequeños segmentos de ADN a partir de cantidades pequeñas de tejido, a fin de obtener copias suficientes para la secuenciación. Para analizar muestras grandes de genes, tal como se requiere para estimar la variación genética de una población, es útil otra técnica (RFLPs), en la que se usan enzimas de restricción para revelar polimorfismos para la longitud de los fragmentos de restricción generados. Este método no proporciona una secuencia de nucleótidos completa, pero en cambio revela si los genes difieren en la presencia o ausencia de las dianas de restricción específicas de las enzimas utilizadas. Permite detectar variación para diferentes marcadores genéticos como las repeticiones en tandem en número variable

(VNTRs). Los loci VNTRs humanos, por ejemplo, están formados por una unidad central de unos 15-100 nucleótidos, repetida un número variable de veces (10-1000) y que, a su vez, puede presentarse una, varias o muchas veces en el genoma, por lo que son muy variables de unos individuos a otros. Se habla de “minisatélites” si se trata de secuencias de unos 50pb repetidas hasta 10kb-20kb. Puesto que cada individuo posee un patrón característico para estas secuencias, se considera cada patrón como una identificación genética individual (se le denomina huella genética, del inglés “fingerprint”). Este polimorfismo tiene aplicación directa en estudios de medicina legal.

Otro tipo de secuencias variables, de distribución más general y más cortas que los minisatélites, son los “microsatélites”, constituidos por repeticiones en tandem de 1-5pb que se repiten hasta tramos de 200pb y localizadas cada 10kb en eucariotas. Están flanqueadas por secuencias conservadas que confieren especificidad a cada locus microsatélite. También se pueden analizar otro tipo de secuencias polimórficas, los RAPDs (del inglés Random Amplified Polymorphic DNA). Los RAPDs se generan amplificando ADN genómico con cebadores de oligonucleótidos arbitrarios. Cada banda del gel indica un locus en el que se pueden detectar dos alelos; presencia o ausencia del producto de amplificación. Los métodos de RFLPs y RAPDs originan fragmentos de diferente tamaño debido a la sustitución de bases en las dianas de restricción o en los cebadores, respectivamente, pero el método más preciso para medir variabilidad es la secuenciación del ADN. Este método ha revelado que existe una gran variabilidad molecular en el genoma (aproximadamente un 1% de sustituciones nucleotídicas por genoma).

El primer estudio de la variación mediante secuenciación de ADN fue llevado a cabo por Martin Kreitman (1983), quién secuenció una región de 2721 pares de bases en *Drosophila melanogaster*, que incluye el sitio que codifica para la enzima alcohol deshidrogenasa (ADH). Las poblaciones de esta especie son polimórficas para dos alelos electroforéticos: “rápido” (*AdhF*) y “lento” (*AdhS*). Kreitman secuenció 11 copias del gen, procedentes de tres regiones del Viejo Mundo y tres regiones americanas. Seis de las copias secuenciadas codificaban para el alelo *AdhS*, y las otras cinco para el alelo *AdhF*. La región secuenciada incluye los cuatro exones y los tres intrones del gen para la ADH, así como las regiones no codificadoras flanqueantes. La mayor parte de la variabilidad era causada por sustituciones, siendo mayor en los intrones (2.4% de sitios nucleotídicos variables) que en los exones (1.8%) y en las regiones flanqueantes (0.7%). Además, seis sitios en las regiones no codificadoras variaban para la presencia o ausencia de segmentos cortos de pares de bases (polimorfismos para inserción/delección). El descubrimiento más llamativo fue que 13 de las 14 variaciones de la región codificadora eran sustituciones sinónimas, que no alteran la secuencia de aminoácidos de la enzima. Una sola sustitución de reemplazamiento, treonina por lisina, diferencia los alelos *AdhS* y *AdhF*. Por tanto, sólo uno de

los 2721 sitios nucleotídicos, que constituían la región secuenciada, variaba como para alterar el producto del gen y potencialmente afectar la bioquímica y fisiología del organismo.

Para cuantificar la variabilidad al nivel del ADN, se utiliza el parámetro de “diversidad nucleotídica” (p), que indica el número medio de diferencias nucleotídicas, por sitio, entre dos secuencias de la población elegidas al azar. Es una medida de la divergencia media entre los alelos (haplotipos) de la población. Cuando se comparan k secuencias de una población, la diversidad nucleotídica vale:

$$\pi = \frac{k}{k-1} \sum_{i,j} f_i f_j \pi_{ij}$$

donde f_i y f_j son las frecuencias de los haplotipos i y j en esa población y π_{ij} es la divergencia nucleotídica entre los haplotipos, que viene dada por:

$$\pi_{ij} = \frac{2n_{ij}}{n_i + n_j}$$

donde n_i y n_j son el número de nucleótidos de cada haplotipo, y n_{ij} es el número de nucleótidos diferentes.

Cuando se dispone de los patrones de restricción en un gel, el estudio de las bandas nos permite calcular I , el “índice de semejanza de restricción”, que nos indica la proporción de bandas comunes entre dos secuencias:

$$I = \frac{2r_{xy}}{r_x + r_y}$$

donde r_x es el número de bandas obtenidas para la secuencia x , r_y el número obtenido para la secuencia y , y r_{xy} el número de bandas comunes a ambas secuencias. A partir de I , se puede calcular el índice de divergencia (D):

$$D = 1 - \left[\frac{-I + \sqrt{I^2 + 8I}}{2} \right]^{\frac{1}{f}}$$

donde f es el número de nucleótidos de que consta la diana de restricción.

Este índice también es posible calcularlo a partir de otro parámetro, el “índice de semejanza nucleotídica” (S), que se puede obtener comparando los patrones de restricción (número de dianas comunes) entre dos mapas:

$$S = \frac{2n_{xy}}{n_x + n_y}$$

donde n_x y n_y son el número de dianas en las secuencias x e y , respectivamente, y n_{xy} es el número de dianas comunes a ambas secuencias. Podemos expresar D en función de S :

$$D = \frac{-\ln S}{f}$$

donde f es el número de bases reconocido por la restrictasa utilizada. Si consideramos una muestra de n individuos, el número de medidas de la divergencia entre secuencias es $n(n-1)/2$ y, promediando, se obtiene la diversidad nucleotídica media.

Es posible calcular el polimorfismo (P_{nuc}) y la heterocigosis (H_{nuc}) nucleotídicos cuando se tienen observaciones de RFLPs. Para una muestra de n individuos, en los que se estudia una secuencia determinada mediante restrictasas, si m es el número total de dianas diferentes, k es el número de ellas que son polimórficas, y si el número de secuencias cortadas en una diana i (c_i) varía desde 1 hasta n , el número total de lugares de corte es $c = \sum c_i \cdot S c_i$. Si f es el número de nucleótidos reconocidos y suponemos que el polimorfismo de restricción se genera por sustitución de una sola base, entonces:

$$P_{nuc} = \frac{c - n(m - k)}{fc}, \text{ y } H_{nuc} = \frac{nc - \sum c_i^2}{fc(n - 1)}.$$

La unidad de diversidad nucleotídica entre dos secuencias es el nucleótido y no el locus o el haplotipo. La diver-

sidad en el caso de loci enzimáticos (el polimorfismo enzimático se refiere a genes completos), no es comparable con la diversidad nucleotídica ya que hay sustituciones que no afectan a la movilidad de la proteína y, por tanto, no se pueden detectar por electroforesis. Actualmente se conocen datos de los polimorfismos para microsatélites que dan valores de heterocigosis muy grandes. En humanos, por ejemplo, hay aproximadamente unos 50000 a 100000 loci de microsatélites, que pueden mostrar hasta más de 10 alelos, con una heterocigosis del 80%.

Agradecimientos

Nuestro más sincero agradecimiento a Carlos López-Fanjul y Mauro Santos por sus valiosos consejos para mejorar el texto. ■

.....

Apéndices

Apéndice 1

En el saltamontes *Caledia captiva*, la síntesis de la enzima málica se encuentra controlada por un locus autosómico (*Me*) con dos alelos codominantes (*c* y *d*). Al analizar una población natural de dicha especie se encontraron 31 individuos de genotipo *cc*, 62 de genotipo *cd* y 7 *dd*. Si consideramos que n_{cc} es el número de individuos *cc*, n_{cd} el de individuos *cd*, n_{dd} el de individuos *dd* y N el total de individuos de la población, podemos calcular las frecuencias genotípicas (X_{cc} , Y_{cd} y Z_{dd}) del siguiente modo:

$$X_{cc} = \frac{n_{cc}}{N} = \frac{31}{100} = 0.31 \quad (31\%)$$

$$Y_{cd} = \frac{n_{cd}}{N} = \frac{62}{100} = 0.62 \quad (62\%)$$

$$Z_{dd} = \frac{n_{dd}}{N} = \frac{7}{100} = 0.07 \quad (7\%)$$

Y las frecuencias alélicas se pueden calcular a partir del número de individuos observados:

$$p_c = \frac{n_{cc} + \frac{1}{2}n_{cd}}{N} = \frac{31 + \frac{1}{2}62}{100} = 0.62,$$

o bien de las frecuencias genotípicas:

$$p_c = X_{cc} + \frac{1}{2}Y_{cd} = 0.31 + \frac{0.62}{2} = 0.62, \text{ siendo la frecuencia del otro alelo: } q_d = 1 - p = 0.38.$$

Apéndice 2

Supongamos que analizamos un locus con dos alelos codominantes (A_1 y A_2) en una población que se inicia con las frecuencias 0.4 A_1A_1 , 0.4 A_1A_2 y 0.2 A_2A_2 . Las frecuencias alélicas serán: $p(A_1) = 0.4 + 1/2(0.4) = 0.6$, y $q(A_2) = 0.2 + 1/2(0.4) = 0.4$.

Si el apareamiento es al azar, ¿cuáles serían las frecuencias alélicas en la generación siguiente? Para calcularlo, hemos de tener en cuenta que la frecuencia del apareamiento al azar entre dos genotipos cualesquiera puede predecirse en base a las frecuencias de dichos genotipos en la población:

	Machos		
Hembras	A_1A_1 (0.4)	A_1A_2 (0.4)	A_2A_2 (0.2)
A_1A_1 (0.4)	0.16	0.16	0.08
A_1A_2 (0.4)	0.16	0.16	0.08
A_2A_2 (0.2)	0.08	0.08	0.04

Las frecuencias genotípicas en la descendencia serán 0.36 A_1A_1 , 0.48 A_1A_2 y 0.16 A_2A_2 :

Tipo de cruce	Nº de cruces		Proporción de descendientes		
	por sexo	Frecuencia	A_1A_1	A_1A_2	A_2A_2
$A_1A_1 \times A_1A_1$	1	0.16	0.16		
$A_1A_1 \times A_1A_2$	2	0.32	0.16	0.16	
$A_1A_1 \times A_2A_2$	2	0.16		0.16	
$A_1A_2 \times A_1A_2$	1	0.16	0.04	0.08	0.04
$A_1A_2 \times A_2A_2$	2	0.16		0.08	0.08
$A_2A_2 \times A_2A_2$	1	0.04			0.04
TOTAL	9	1.00	0.36	0.48	0.16

Aunque las frecuencias genotípicas de la población se han visto alteradas por el apareamiento aleatorio, las frecuencias alélicas no han cambiado:

$$p(A_1) = 0.36 + 1/2(0.48) = 0.6 \text{ y } q(A_2) = 0.16 + 1/2(0.48) = 0.4$$

Si se dan las condiciones de Hardy-Weinberg (esencialmente: panmixia, tamaño poblacional grande y ausencia de diferencias en eficacia biológica entre los individuos), las frecuencias alélicas de la siguiente generación serán las mismas que las de la generación parental. Por ejemplo, imaginemos que la población inicial hubiese tenido las mismas frecuencias alélicas pero diferentes frecuencias genotípicas (0.25 A_1A_1 , 0.70 A_1A_2 y 0.05 A_2A_2). En la siguiente generación, las frecuencias genotípicas serían 0.36 A_1A_1 , 0.48 A_1A_2 y 0.16 A_2A_2 :

Tipo de cruce	Nº de cruces		Proporción de descendientes		
	por sexo	Frecuencia relativa	A_1A_1	A_1A_2	A_2A_2
$A_1A_1 \times A_1A_1$	1	$0.25 \times 0.25 = 0.0625$	0.0625		
$A_1A_1 \times A_1A_2$	2	$2 \times 0.25 \times 0.70 = 0.35$	0.1750	0.1750	
$A_1A_1 \times A_2A_2$	2	$2 \times 0.25 \times 0.05 = 0.025$		0.0250	
$A_1A_2 \times A_1A_2$	1	$0.70 \times 0.70 = 0.49$	0.1225	0.2450	0.1225
$A_1A_2 \times A_2A_2$	2	$2 \times 0.70 \times 0.05 = 0.07$		0.0350	0.0350
$A_2A_2 \times A_2A_2$	1	$0.05 \times 0.05 = 0.0025$			0.0025
TOTAL	9	1.00	0.36	0.48	0.16

Obsérvese que en los dos supuestos hemos alcanzado las mismas frecuencias genotípicas en la descendencia (0.36, 0.48 y 0.16) aún partiendo de diferentes frecuencias genotípicas en los progenitores. Pero nótese también que en ambos casos las frecuencias alélicas iniciales eran las mismas (0.6 y 0.4). Por tanto, “en condiciones de apareamiento al azar (panmixia) en una población de gran tamaño en la que todos los genotipos son igualmente viables y fecundos, las frecuencias genotípicas de una generación particular dependen de las frecuencias alélicas de la generación anterior y no de las frecuencias genotípicas».

Si, en cualquiera de los dos casos anteriores, obtenemos una generación más, partiendo de las frecuencias genotípicas obtenidas en la primera generación (0.36, 0.48 y 0.16), observamos que las frecuencias genotípicas también permanecen estables:

Tipo de cruce	Nº de cruces		Proporción de descendientes		
	por sexo	Frecuencia relativa	A_1A_1	A_1A_2	A_2A_2
$A_1A_1 \times A_1A_1$	1	0.1296	0.1296		
$A_1A_1 \times A_1A_2$	2	0.3456	0.1728	0.1728	
$A_1A_1 \times A_2A_2$	2	0.1152		0.1152	
$A_1A_2 \times A_1A_2$	1	0.2304	0.0576	0.1152	0.0576
$A_1A_2 \times A_2A_2$	2	0.1536		0.0768	0.0768
$A_2A_2 \times A_2A_2$	1	0.0256			0.0256
TOTAL	9	1.00	0.36	0.48	0.16

Se dice que la población había alcanzado el equilibrio en la primera generación. Este equilibrio genotípico, basado en las frecuencias alélicas estables y en el apareamiento aleatorio, es el principio de Hardy-Weinberg, que constituye el punto de partida de la genética de poblaciones.

Apéndice 3

En el ejemplo mostrado en el texto, la diferencia en frecuencia alélica entre machos y hembras ($dif = p_m - p_h$) se reduce a la mitad en cada generación:

$$\begin{aligned} dif_0 &= 0.2 - 0.5 = -0.3 \\ dif_i &= 0.5 - 0.35 = 0.15 \\ dif_{ii} &= 0.35 - 0.425 = -0.075 \\ dif_{iii} &= 0.425 - 0.3875 = 0.0375 \\ dif_{iv} &= 0.3875 - 0.40625 = -0.01875 \\ dif_v &= 0.40625 - 0.396875 = 0.009375 \\ dif_{vi} &= 0.396875 - 0.4015625 = -0.0046875 \end{aligned}$$

El cambio de signo, de generación en generación, denota los cambios en frecuencia mostrados en la Figura 2. Es fácil demostrar que dif se reduce a la mitad en cada generación. Ya sabemos que $p_m = p_h$ y que $p_h = (p_m + p_h)/2$. Denominemos dif_0 a la diferencia entre sexos para la frecuencia alélica en la generación inicial ($dif_0 = p_m - p_h$) y dif_1 a esa diferencia en la primera generación ($dif_1 = p_m - p_h$). Por tanto, $dif_1 = p_m - p_h = p_h - (p_m + p_h)/2 = p_h - (1/2 p_m + 1/2 p_h) = -1/2 p_m - 1/2 p_h = -(p_m - p_h)/2$. Como $p_m - p_h = dif_0$, $p_h - p_m = -dif_0$, y $dif_1 = -1/2 dif_0$.

Es posible calcular el número de generaciones necesarias para que la diferencia en frecuencia alélica entre sexos se reduzca hasta un cierto valor. Como hemos visto,

$$\begin{aligned} dif_i &= -1/2 dif_0 \\ dif_{ii} &= -1/2 dif_i = (-1/2) (-1/2 dif_0) = (1/2)^2 dif_0 \\ dif_{iii} &= -1/2 dif_{ii} = (-1/2) (1/2)^2 dif_0 = -(1/2)^3 dif_0 \\ dif_{iv} &= -1/2 dif_{iii} = (-1/2) [-(1/2)^3 dif_0] = (1/2)^4 dif_0 \\ \dots\dots\dots \\ dif_n &= -1/2 dif_{n-1} = (-1/2) [(1/2)^{n-1} dif_0] = -(1/2)^n dif_0 \end{aligned}$$

Puesto que el signo depende de si han transcurrido número par o impar de generaciones, prescindimos de él para lo que sigue:

$$\begin{aligned} dif_n &= (1/2)^n dif_0 ; \\ (1/2)^n &= dif_n / dif_0 \\ \log (1/2)^n &= \log (dif_n / dif_0), \\ n \log (1/2) &= \log (dif_n / dif_0); \\ n &= \log (dif_n / dif_0) / \log (1/2); \end{aligned}$$

(el cociente de logaritmos es igual al cociente del logaritmo de las inversas): $n = \log (dif_0 / dif_n) / \log 2$.

Puesto que dif_n tiende a cero, n tiende a infinito. No obstante, con la ecuación anterior podemos calcular el número de generaciones necesarias para que dif_n alcance un determinado valor. Por ejemplo, ¿cuántas generaciones se necesitarían en el ejemplo anterior para que la diferencia en frecuencia alélica entre sexos (dif_n) sea menor del 5%? La ecuación anterior indica que serían tres.

Apéndice 4

Estima de las frecuencias de equilibrio en poblaciones naturales

a) Loci autosómicos

Cuando se trata de un carácter codominante, pueden distinguirse fenotípicamente todos los genotipos. Ello permite aplicar la prueba de χ^2 de bondad de ajuste para probar el equilibrio H-W. Por ejemplo, en la población del saltamontes *Caledia captiva*, a la que nos hemos referido anteriormente, las frecuencias genotípicas observadas eran 31 (*cc*), 62 (*cd*) y 7 (*dd*), lo que implicaba una frecuencia $p=0.62$ para c y $q=0.38$ para d . El número de individuos de cada genotipo, esperados según el principio de Hardy-Weinberg, son $p^2N = 0.62^2 \cdot 100 = 38.4$, $2pqN = 2 \cdot 0.62 \cdot 0.38 \cdot 100 = 47.1$, y $q^2N = 0.38^2 \cdot 100 = 14.4$, para los genotipos *BB*, *BS* y *SS*, respectivamente. Aplicamos ahora la prueba de chi-cuadrado de bondad de ajuste, con la hipótesis nula (H_0) de que se dan las condiciones del equilibrio Hardy-Weinberg, mediante la ecuación:

$$\chi^2 = \sum \frac{(o - e)^2}{e}$$

donde o y e son el número de individuos observado y esperado, respectivamente, de un genotipo determinado. En el ejemplo que nos ocupa, el valor de la χ^2 experimental, que resulta de la aplicación de la ecuación anterior (9.97), es mayor que el valor teórico para una χ^2 con un grado de libertad y nivel de significación del 5%, que es 3.84, por lo que rechazamos H_0 y concluimos que la población no está en equilibrio H-W para ese locus.

Si el carácter es dominante, sin embargo, no puede distinguirse fenotípicamente el genotipo heterocigótico (*Aa*) del homocigoto dominante (*AA*). Por ello no pueden obtenerse directamente las frecuencias alélicas, como en el caso de codominancia, ya que se desconocen dos de las frecuencias genotípicas.

Como primera aproximación, podríamos calcular la frecuencia q como la raíz cuadrada de q^2 , la frecuencia del homocigoto recesivo. Pero este método sólo es válido cuando la población está en equilibrio H-W, porque si no lo está la mayor o menor proporción de heterocigotos (apartándose de su frecuencia genotípica de equilibrio) podría hacer que nuestra estima de las frecuencias alélicas fuese errónea. Por ejemplo, imaginemos dos poblaciones donde el fenotipo recesivo constituye el 25% de los individuos, pero en una de ellas, que está en equilibrio, la frecuencia de heterocigotos es el 50% y en la otra, que no lo está, es el 25%. Calculando q como la raíz cuadrada de la frecuencia de homocigotos recesivos, obtendríamos una estima de $q=0.5$ en ambas poblaciones. Esta estima sería correcta en la población que está en equilibrio, pero no en la segunda, donde $q = 0.25 + 0.25/2 = 0.375$.

Entonces, ¿cómo se investiga el equilibrio en estos casos? Snyder (1932) ideó un método que sirve para averiguar si una población está en equilibrio para un carácter dominante, cuando se tienen datos de los tipos de apareamientos. La prueba se basa en la frecuencia esperada de recesivos producidos por las combinaciones pater- nas de apareamientos. Si existiese equilibrio, podríamos calcular la proporción esperada de descendientes recesivos en los apareamientos dominante por dominante y en los apareamientos dominante por recesivo:

La proporción de recesivos esperada en cruces dominante x dominante es:

Genotipos	Fenotipos	Frecuencia de cruces	Frecuencia esperada de recesivos
<i>AA</i> × <i>AA</i>	Dominante × dominante	$p^4 + 4p^3q + 4p^2q^2$	p^2q^2
(2) <i>AA</i> × <i>Aa</i>			
<i>Aa</i> × <i>Aa</i>			
(2) <i>Aa</i> × <i>aa</i>	Dominante × recesivo	$4pq^3 + 2p^2q^2$	$2pq^3$
(2) <i>AA</i> × <i>aa</i>			

La proporción de recesivos esperada en cruces dominante x dominante es:

$$\frac{p^2q^2}{p^4 + 4p^3q + 4p^2q^2} = \frac{p^2q^2}{p^2(p^2 + 4pq + 4q^2)} = \frac{q^2}{p^2 + 4pq + 4q^2} = \frac{q^2}{(p + 2q)^2} = \frac{q^2}{(p + q + q)^2} = \frac{q^2}{(1 + q)^2}$$

Y en cruces dominante × recesivo:

$$\frac{2pq^3}{4pq^3 + 2p^2q^2} = \frac{pq^3}{2pq^3 + p^2q^2} = \frac{pq^3}{pq^2(2q + p)} = \frac{q}{p + 2q} = \frac{q}{1 + q}$$

Así disponemos de dos proporciones de la población cuyos valores teóricos esperados pueden compararse con las frecuencias observadas. Basándose en esto, Snyder llevó a cabo una prueba respecto al equilibrio de los alelos para la capacidad de gustar la feniltiocarbamida:

Apareamiento	Progenitores	N	Frecuencia de no gustadores en la descendencia				
			Descendencia		en la descendencia		
			gustadores	no gustadores	total	observada	esperada
gust x gust		425	929	130	1059	0.123	0.122
gust x no gust		289	483	278	761	0.365	0.349
no gust x no gust		86	(5)	218	223		
Total		800	1417	626	2043		

Los cinco individuos entre paréntesis no serían de esperar del cruce de dos progenitores recesivos.

La hipótesis nula (H_0) es que la población está en equilibrio, por lo que calculamos q como la raíz cuadrada de la frecuencia de homocigotos recesivos entre los progenitores. De los 1600 progenitores, 461 eran no gustadores, luego $q^2 = 461/1600 = 0.288$, y $q = 0.288^{0.5} = 0.537$. Por tanto, la frecuencia esperada de descendientes recesivos en los cruces gustador x gustador sería: $q^2 / (1 + q)^2 = 0.537^2 / 1.537^2 = 0.122$, y la de dominantes (gustadores) sería $1 - 0.122 = 0.878$. En total, se esperarían $0.122 \times 1059 = 129.2$

descendientes no gustadores y $0.878 \times 1059 = 929.8$ gustadores. Una chi cuadrado, con un grado de libertad, no rechazaría la hipótesis nula:

$$\chi^2 = \frac{(130 - 129.2)^2}{129.2} + \frac{(929 - 929.8)^2}{929.8} = 0.0056$$

Igualmente, la frecuencia esperada de descendientes recesivos en cruces gustador \times no gustador sería:

$$q / (1 + q) = 0.537 / 1.537 = 0.349,$$

y la de dominantes sería $1 - 0.349 = 0.651$. En total, se esperarían $0.349 \times 761 = 265.6$ individuos no gustadores y $0.651 \times 761 = 495.4$ gustadores. La chi cuadrado correspondiente, con un grado de libertad, valdría:

$$\chi^2 = \frac{(278 - 265.6)^2}{265.6} + \frac{(483 - 495.4)^2}{495.4} = 0.89$$

Por tanto, no rechazamos la hipótesis de que la población está en equilibrio H-W para este carácter.

b) Loci ligados al sexo

Cuando se trata de genes con codominancia, primero comparamos la frecuencia alélica entre sexos, mediante un test de chi cuadrado de contingencia, y luego probamos el equilibrio genotípico en las hembras mediante el test chi cuadrado de bondad de ajuste, como en el caso de un locus autosómico (subapartado a). Cuando hay dominancia, sin embargo, no pueden identificarse dos de los genotipos de hembras (AA y Aa), pero los dos genotipos en machos (A y a) son claramente diferenciables. Entonces pueden utilizarse las estimas de las frecuencias alélicas en machos para calcular las frecuencias genotípicas esperadas en hembras. Por ejemplo, sea A/a un gen ligado al sexo, siendo A dominante sobre a , y con las siguientes frecuencias fenotípicas observadas en machos: 99 A y 28 a , y en hembras: 150 A_- y 4 aa . ¿Cómo podemos probar el equilibrio?

Las frecuencias alélicas en machos ($q = 28 / (99 + 28) = 0.22$, y $p = 0.78$) pueden utilizarse para calcular las frecuencias genotípicas esperadas en el equilibrio para las hembras:

$$p^2 AA = 0.78^2 = 0.61$$

$$2pq Aa = 2 \times 0.78 \times 0.22 = 0.34$$

$$q^2 aa = 0.22^2 = 0.05$$

Como los genotipos AA y Aa son fenotípicamente indistinguibles, sólo podemos probar el equilibrio fenotípico. Las frecuencias fenotípicas esperadas serían: $0.61 + 0.34 = 0.95 A_-$ y $0.05 aa$, que en 154 individuos supone 146.3 y 7.7, respectivamente. Un test chi cuadrado, con un grado de libertad, sugiere que existe equilibrio en las hembras:

$$\chi^2 = \frac{(150 - 146.3)^2}{146.3} + \frac{(4 - 7.7)^2}{7.7} = 1.87$$

c) Loci con alelos múltiples

Cuando en un locus hay más de dos alelos, el equilibrio genotípico viene determinado por el desarrollo multinomial ($p + q + r + \dots$), siendo p, q, r, \dots las frecuencias de los alelos A_1, A_2, A_3, \dots

Por ejemplo, con tres alelos, A_1, A_2 y A_3 , con frecuencias respectivas p, q y r , en el equilibrio H-W esperaríamos 6 genotipos diferentes en la proporción: $p^2 (A_1A_1)$, $2pq (A_1A_2)$, $2pr (A_1A_3)$, $q^2 (A_2A_2)$, $2qr (A_2A_3)$ y $r^2 (A_3A_3)$. Si los tres alelos fuesen codominantes, cada genotipo tendría su propio fenotipo diferenciable y las frecuencias genotípicas pueden deducirse directamente de las frecuencias fenotípicas. Las frecuencias alélicas se calculan entonces a partir de las frecuencias genotípicas:

$$p = \frac{n_{A_1A_1} + \frac{1}{2}n_{A_1A_2} + \frac{1}{2}n_{A_1A_3}}{N}$$

$$q = \frac{n_{A_2A_2} + \frac{1}{2}n_{A_1A_2} + \frac{1}{2}n_{A_2A_3}}{N}$$

$$r = \frac{n_{A_3A_3} + \frac{1}{2}n_{A_1A_3} + \frac{1}{2}n_{A_2A_3}}{N}$$

donde $n_{A_1A_1}, n_{A_1A_2}, \dots$, son los números de individuos de cada genotipo y N es el número total de individuos estudiados. El equilibrio se probaría mediante un test chi cuadrado.

Cuando hay dominancia, sin embargo, la situación se complica. Veamos el ejemplo de los grupos sanguíneos ABO. Existen tres alelos, siendo dos de ellos codominantes (I^A e I^B) y ambos dominantes sobre el tercero (i) ($I^A = I^B > i$).

Hay 4 fenotipos posibles, determinados por 6 genotipos diferentes:

Fenotipo	Genotipos
A	$I^A I^A, I^A i$
B	$I^B I^B, I^B i$
AB	$I^A I^B$
0	ii

Las frecuencias genotípicas en el equilibrio son: $p^2 (I^A I^A)$, $2pr (I^A i)$, $q^2 (I^B I^B)$, $2qr (I^B i)$, $2pq (I^A I^B)$ y $r^2 (ii)$. El cálculo de las frecuencias alélicas se realiza por estimación a partir de las frecuencias fenotípicas:

1) La suma de individuos de los grupos B y 0 es $q^2 + 2qr + r^2 = (q + r)^2$. Por tanto, podemos calcular q como 1 menos la raíz cuadrada de la frecuencia relativa de individuos de los grupos B + 0.

2) La suma de individuos de los grupos A y 0 es $p^2 + 2pr + r^2 = (p + r)^2$. Por tanto, podemos calcular p como 1 menos la raíz cuadrada de la frecuencia relativa de individuos de los grupos A + 0.

3) Los individuos de grupo 0 son homocigóticos recesivos y, por tanto, su frecuencia es r^2 . De aquí podemos deducir que r es la raíz cuadrada de la frecuencia relativa de individuos de grupo 0.

La siguiente Tabla resume un ejemplo en 192 individuos de Gales:

a)	A	B	AB	0	Total	
número de individuos	63	31	6	92	192	
frecuencia relativa	0.3281	0.1615	0.0312	0.4792		
b)						
Fenotipos	N	frec.	(frec.) ^{0.5}	estima	corrección	frecuencia alélica
B+0	123	0.6407	0.8004	p=0.1996	0.1996/0.9933	p' = 0.2010
A+0	155	0.8073	0.8985	q=0.1015	0.1015/0.9933	q' = 0.1022
0	92	0.4792	0.6922	r=0.6922	0.6922/0.9933	r' = 0.6968
Total	192			0.9933		1.0000

Utilizando las frecuencias alélicas corregidas pueden calcularse las frecuencias fenotípicas esperadas en el equilibrio y compararse con las observadas mediante un test de chi cuadrado, que arroja el valor 0.55 y concluimos que existe equilibrio fenotípico en esa población, para este carácter.

Apéndice 5

En la generación siguiente se verificaría que $r^n = r' \cdot cd_1$, donde

$$d_1 = r'u' - s't' = (r - cd_0)(u - cd_0) - (s + cd_0)(t + cd_0) =$$

$$= ru - ucd_0 - rcd_0 + c^2d_0^2 - st - ctd_0 - csd_0 - c^2d_0^2 =$$

$$= ru - st - cd_0(u + r + s + t) = ru - st - cd_0 = d_0 - cd_0,$$

de donde $d_1 = d_0(1 - c)$ y, por tanto, en la generación enésima $d_n = d_0(1 - c)^n$. Al ser $0 < c < 1/2$, $1 - c$ es siempre menor que 1 y, por tanto, $(1 - c)^n$ tiende a cero al crecer n .

Volviendo a las frecuencias gaméticas, tenemos que,

$$r' = r - cd_0$$

$$r'' = r' - cd_1 = r - cd_0 - cd_0(1 - c)$$

$$r''' = r'' - cd_2 = r - cd_0 - cd_0(1 - c) - cd_0(1 - c)^2$$

.....

$$r^n = r - cd_0[1 + (1 - c) + \dots + (1 - c)^{n-1}] = r - cd_0 \frac{1 - (1 - c)^n}{1 - (1 - c)},$$

porque la suma de una progresión de razón $(1 - c)$ es:

$$s = \frac{a_1 - a_n r}{1 - r} = \frac{1 - [(1 - c)^{n-1}(1 - c)]}{1 - (1 - c)}$$

Por tanto, $r^n = r - d_0 [1 - (1 - c)^n]$ y, por el mismo razonamiento:

$$s^n = s + d_0 [1 - (1 - c)^n]$$

$$t^n = t + d_0 [1 - (1 - c)^n]$$

$$u^n = u - d_0 [1 - (1 - c)^n]$$

al crecer n indefinidamente obtenemos los siguientes valores de equilibrio:

$$r_\infty = r - d_0$$

$$s_\infty = s + d_0$$

$$t_\infty = t + d_0$$

$$u_\infty = u - d_0$$

que son independientes de c .

Apéndice 6

El número de generaciones necesarias para que se efectúe un determinado cambio en la frecuencia alélica, por mutación recurrente reversible, puede calcularse admitiendo que, como Δp es muy pequeño, podemos considerar la igualdad $\Delta p = (p_{eq} - p)(u + v)$ como una ecuación diferencial sustituyendo Δp por dp/dt , donde t es el tiempo en unidades de generación:

$$\frac{dp}{dt} = (u + v)(p_{eq} - p); \quad \frac{dp}{(p_{eq} - p)} = (u + v)dt$$

Integrando entre los momentos inicial y final de un período determinado:

$$\int_0^n \frac{1}{p_{eq} - p} dp = \int_0^n (u + v) dt; \quad - \int_0^n \frac{1}{p - p_{eq}} dp = \int_0^n (u + v) dt;$$

$$[-L(p - p_{eq})]_0^n = (u + v)[t]_0^n; \quad -L \frac{p^n - p_{eq}}{p_0 - p_{eq}} = (u + v)n;$$

$$n = \frac{1}{u + v} L \frac{p_0 - p_{eq}}{p_n - p_{eq}}$$

Por ejemplo, si $u = 6 \times 10^{-5}$, $v = 4 \times 10^{-5}$ y $p_0 = 0.9$, ¿cuántas generaciones deberán transcurrir hasta que p_n valga 0.8?

$$p_{eq} = \frac{v}{u + v} = \frac{4 \times 10^{-5}}{10 \times 10^{-5}} = 0.4, \text{ y}$$

$$n = \frac{1}{10 \times 10^{-5}} L \frac{0.9 - 0.4}{0.8 - 0.4} = 2232 \text{ generaciones.}$$

De lo anteriormente expuesto se deduce que, con tasas normales de mutación (del orden de 10^{-4} a 10^{-8}), los cambios en frecuencias génicas debidos a mutación son tan pequeños que sólo son importantes a una escala evolutiva muy larga, y difícilmente serían detectables en experimentos a menos que se trate de microorganismos.

Apéndice 7

Cálculo del número de generaciones necesarias para los cambios por selección

a) Desventaja recesiva:

$$\Delta q = \frac{-sq^2(1-q)}{1-sq^2}; \text{ si } s \text{ es pequeño, } \Delta q = -sq^2(1-q).$$

Transformándolo en una ecuación diferencial:

$$\frac{dq}{dt} = -sq^2(1-q); \quad \frac{1}{q^2(1-q)} dq = -sdt$$

Primero descomponemos la fracción:

$$\frac{1}{q^2(1-q)} = \frac{A}{q^2} + \frac{B}{q} + \frac{C}{(1-q)}$$

$$\forall q \Rightarrow 1 = A(1-q) + Bq(1-q) + Cq^2$$

$$q=1 \rightarrow C=1$$

$$q=0 \rightarrow A=1$$

$$q=0.5 \rightarrow 1 = 0.5A + 0.25B + 0.25C = \\ = 0.5 + 0.25B + 0.25 \rightarrow B=1$$

Por tanto, la integral que debemos resolver es:

$$\int_0^n \frac{1}{q^2} dq + \int_0^n \frac{1}{q} dq + \int_0^n \frac{1}{1-q} dq = -s \int_0^n dt$$

$$\left[-\frac{1}{q} + Lq - L(1-q) \right]_0^n = -sn$$

$$-\frac{1}{q_n} + Lq_n - L(1-q_n) + \frac{1}{q_0} - Lq_0 + L(1-q_0) = -sn$$

$$\frac{1}{q_n} - Lq_n + L(1-q_n) - \frac{1}{q_0} + Lq_0 - L(1-q_0) = sn$$

$$n = \frac{1}{s} \left[\frac{1}{q_n} - \frac{1}{q_0} + L \frac{q_0(1-q_n)}{q_n(1-q_0)} \right]$$

b) Desventaja dominante:

$$\frac{dq}{dt} = -sq(1-q)^2; \quad \frac{1}{q(1-q)^2} dq = -sdt$$

Descomponemos la fracción:

$$\frac{1}{q(1-q)^2} = \frac{A}{q} + \frac{B}{(1-q)^2} + \frac{C}{(1-q)}$$

$$\forall q \Rightarrow 1 = A(1-q)^2 + Bq + Cq(1-q)$$

$$q=1 \rightarrow B=1$$

$$q=0 \rightarrow A=1$$

$$q=0.5 \rightarrow 1 = 0.25A + 0.5B + 0.25C = \\ = 0.25 + 0.5 + 0.25C \rightarrow C=1$$

Por tanto, la integral que debemos resolver es:

$$\int_0^n \frac{1}{q} dq + \int_0^n \frac{1}{(1-q)^2} dq + \int_0^n \frac{1}{1-q} dq = -s \int_0^n dt$$

$$\left[Lq + \frac{1}{1-q} - L(1-q) \right]_0^n = -sn$$

$$Lq_n + \frac{1}{1-q_n} - L(1-q_n) - Lq_0 - \frac{1}{1-q_0} + L(1-q_0) = -sn$$

$$-Lq_n - \frac{1}{1-q_n} + L(1-q_n) + Lq_0 + \frac{1}{1-q_0} - L(1-q_0) = sn$$

$$n = \frac{1}{s} \left[\frac{1}{1-q_0} - \frac{1}{1-q_n} + L \frac{q_0(1-q_n)}{q_n(1-q_0)} \right]$$

c) Acción génica aditiva:

$$\frac{dq}{dt} = -\frac{1}{2} sq(1-q), \text{ de donde } \frac{2}{q(1-q)} dq = -sdt.$$

Siguiendo el mismo procedimiento que en los casos anteriores, obtenemos la ecuación:

$$n = \frac{2}{s} L \frac{q_0(1-q_n)}{q_n(1-q_0)}$$

d) Sobredominancia:

$$\Delta q = pq \frac{(s+t)(q_{eq}-q)}{1-sp^2-tq^2}, \text{ y si } s \text{ es pequeño:}$$

$$\Delta q = pq(s+t)(q_{eq}-q)$$

$$\frac{1}{q(1-q)(q_{eq}-q)} dq = (s+t)dt$$

$$\frac{A}{q} + \frac{B}{1-q} + \frac{C}{(q_{eq}-q)} = 1;$$

$$\forall q \rightarrow 1 = A(1-q)(q_{eq}-q) + Bq(q_{eq}-q) + Cq(1-q)$$

$$q=1 \rightarrow B = \frac{1}{q_{eq}-1}$$

$$q=0 \rightarrow A = \frac{1}{q_{eq}}$$

$$q=q_{eq} \rightarrow C = \frac{1}{q_{eq}(1-q_{eq})}$$

$$\frac{1}{q_{eq}} + \frac{1}{q_{eq}-1} + \frac{1}{q_{eq}(1-q_{eq})} =$$

$$= \frac{1}{q_{eq}} \cdot \frac{1}{q} + \frac{1}{q_{eq}-1} \cdot \frac{1}{1-q} + \frac{1}{q_{eq}(1-q_{eq})} \cdot \frac{1}{q_{eq}-q}$$

Por tanto,

$$\frac{1}{q_{eq}} \int_0^n \frac{1}{q} dq + \frac{1}{q_{eq}-1} \int_0^n \frac{1}{1-q} dq + \frac{1}{q_{eq}(1-q_{eq})} \int_0^n \frac{1}{q_{eq}-q} dq = \int_0^n (s+t) dt$$

$$\left[\frac{1}{q_{eq}} Lq - \frac{1}{q_{eq}-1} L(1-q) - \frac{1}{q_{eq}(1-q_{eq})} L(q_{eq}-q) \right]_0^n = (s+t)n$$

$$n = \frac{1}{s+t} \left[\frac{1}{q_{eq}} Lq_n - \frac{1}{q_{eq}-1} L(1-q_n) - \frac{1}{q_{eq}(1-q_{eq})} L(q_{eq}-q_n) - \left(\frac{1}{q_{eq}} Lq_0 - \frac{1}{q_{eq}-1} L(1-q_0) - \frac{1}{q_{eq}(1-q_{eq})} L(q_{eq}-q_0) \right) \right]$$

$$n = \frac{1}{s+t} \left[\frac{1}{q_{eq}} L \frac{q_n}{q_0} + \frac{1}{q_{eq}-1} L \frac{1-q_0}{1-q_n} + \frac{1}{q_{eq}(1-q_{eq})} L \frac{q_{eq}-q_0}{q_{eq}-q_n} \right]$$

e) Subdominancia:

Con valores pequeños de s :

$$\Delta q = spq(2q-1); \text{ y por tanto}$$

$$\frac{dq}{dt} = spq(2q-1) = s(1-q)q(2q-1)$$

$$\frac{1}{q(1-q)(2q-1)} dq = s dt;$$

$$\frac{1}{q(1-q)(2q-1)} = \frac{A}{q} + \frac{B}{1-q} + \frac{C}{2q-1}$$

$$\forall q \rightarrow 1 = A(1-q)(2q-1) + Bq(2q-1) + Cq(1-q)$$

$$q=1 \rightarrow B=1$$

$$q=0 \rightarrow A=-1$$

$$q=0.5 \rightarrow C=4$$

Por tanto,

$$\left(-\frac{1}{q} + \frac{1}{1-q} + \frac{4}{2q-1} \right) dq = s dt ;$$

$$\int_0^n -\frac{1}{q} dq + \int_0^n \frac{1}{1-q} dq + 4 \int_0^n \frac{1}{2q-1} dq = \int_0^n s dt ;$$

$$\left[-Lq - L(1-q) + 2L(2q-1) \right]_0^n = sn ;$$

$$\left[L(2q-1)^2 - (Lq + L(1-q)) \right]_0^n = sn ;$$

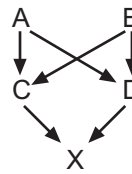
$$\left[L \frac{(2q-1)^2}{q(1-q)} \right]_0^n = sn ;$$

$$n = \frac{1}{s} \left[L \frac{(2q_n-1)^2}{q_n(1-q_n)} - L \frac{(2q_0-1)^2}{q_0(1-q_0)} \right];$$

$$n = \frac{1}{s} \left[L \frac{q_0(1-q_0)(2q_n-1)^2}{q_n(1-q_n)(2q_0-1)^2} \right]$$

Apéndice 8

Hijos de hermanos:

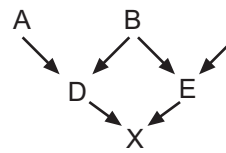


Ancestrales comunes: A y B
Vías: C→A→D
C→B→D

$$F_x = (\frac{1}{2})^3 \cdot (1+F_A) + (\frac{1}{2})^3 \cdot (1+F_B)$$

$$F_A = 0, F_B = 0; F_x = (\frac{1}{2})^3 + (\frac{1}{2})^3 = 1/4$$

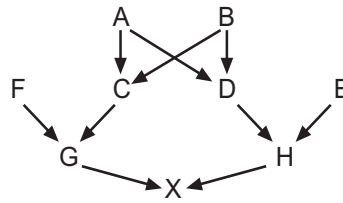
Hijos de hermanastros:



El único ancestres común es B
Vía: D→B→E

$$F_x = (\frac{1}{2})^3 \cdot (1+F_B) = 1/8$$

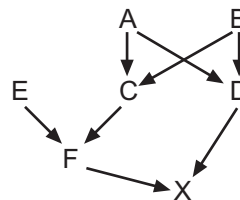
Hijos de primos hermanos:



Ancestrales comunes: A y B
Vías: G→C→A→D→H
G→C→B→D→H

$$F_x = (\frac{1}{2})^5 + (\frac{1}{2})^5 = 1/16$$

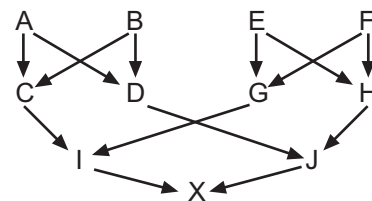
Hijos de tío(a) sobrina(o):



Ancestrales comunes: A y B
Vías: F→C→A→D
F→C→B→D

$$F_x = (\frac{1}{2})^4 + (\frac{1}{2})^4 = 1/8$$

Hijos de dobles primos hermanos:



Antecedentes comunes: A, B, E y F

Vías: I→V→A→D→J
I→C→B→D→J
I→G→E→H→J
I→G→F→H→J

$$F_x = (\frac{1}{2})^5 \cdot 4 = 1/8$$

Apéndice 9

$$1 - P_n = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) (1 - P_{n-1})$$

$$-P_n = -1 + \frac{1}{2N} + 1 - P_{n-1} + \frac{1}{2N} P_{n-1} - \frac{1}{2N}$$

$$-P_n = -P_{n-1} + \frac{1}{2N} P_{n-1}$$

$$P_n = P_{n-1} - \frac{1}{2N} P_{n-1}$$

$$P_n = P_{n-1} \left(1 - \frac{1}{2N}\right)$$

Apéndice 10

$$P_n = P_{n-1} \left(1 - \frac{1}{2N}\right) = \left[P_{n-2} \left(1 - \frac{1}{2N}\right) \right] \left(1 - \frac{1}{2N}\right) =$$

$$= P_{n-2} \left(1 - \frac{1}{2N}\right)^2; P_n = P_0 \left(1 - \frac{1}{2N}\right)^n$$

Bibliografía

- BENNET, J. 1960. A comparison of selective methods and a test of the pre-adaptation hypothesis. *Heredity* 15: 65-77.
- CAIRNS, J., OVERBAUGH, J. y MILLER, S. 1988. The origin of mutants. *Nature* 335: 142-145.
- CARSON, H.L. 1983. Chromosomal sequences and interisland colonizations in Hawaiian *Drosophila*. *Genetics* 103: 465-482.
- DARLINGTON, C.D. y MATHER, K. 1949. *The elements of Genetics*. Allen and Unwin, London.
- DEL CERRO A.L. y SANTOS J.L. 1995. Synapsis in grasshopper bivalents heterozygous for centric shifts. *Genome* 38: 616-622.
- DOBZHANSKY, T. y PAVLOSKY, O. 1957. An experimental study of interaction between genetics drift and natural selection. *Evolution* 11: 311-319.
- DRAKE, J.W., CHARLESWORTH, B., CHARLESWORTH, D. y CROW, J.F. 1998. Rates of spontaneous mutation. *Genetics* 148: 1667-1686.
- ENDLER, J.A. 1986. *Natural selection in the wild*. Princeton University Press, Princeton, N.J.
- FISHER, R.A. 1922. On the dominance ratio. *Proc. Roy. Soc. Edin.* 42: 321-431.
- FISHER, R.A. 1930. *The genetical theory of natural selection*. 1ª ed. Clarendon, Oxford.
- FONTDEVILA, A. y MOYA, A. 1999. *Introducción a la genética de poblaciones*. Editorial Síntesis, Madrid.
- HALDANE, J.B.S. 1924-1932. A mathematical theory of natural and artificial selection. *Proc. Camb. Phil. Soc.* (Colección de artículos).
- HALL, B.G. 1988. Adaptive evolution that requires multiple spontaneous mutations. I. Mutations involving an insertion sequence. *Genetics* 120: 887-897.
- HALL, B.G. 1990. Spontaneous point mutations that occur more often when advantageous than when neutral. *Genetics* 126: 5-16.
- HARDY, G.H. 1908. Mendelian proportions in a mixed population. *Science* 28: 49-50.
- HARTL, D.L. y CLARK, A.G. 1989. *Principles of population genetics*. 2nd ed. Sinauer, Sunderland.
- KIMURA, M. 1955. Solution of a process of random genetic drift with a continuous model. *Proc. Natl. Acad. Sci. USA* 41: 144-150.
- KING, M. 1995. *Species evolution. The role of chromosome change*. Cambridge University Press, Cambridge.
- KREITMAN, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* 304: 412-417.
- LEDERBERG, J. y LEDERBERG, E.M. 1952. Replica plating and indirect selection of bacterial mutants. *J. Bact.* 63: 393-406.
- LURIA, S.E. y DELBRÜCK, M. 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28: 491-511.
- MAYR, E. 1942. *Systematics and the Origin of Species*. Columbia Univ. Press, New York.
- SNYDER, L.H. 1932. *Studies in human inheritance. IX. The inheritance of taste deficiency in man*. *Ohio J. Sci.* 32: 436-440.
- WEINBERG, W. 1908. Über den Nachweis der Vererbung beim Menschen. *Jahresh. Verein f. vaterl. Naturk. in Württemberg* 64: 368-382.
- WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97-159.

Lecturas recomendadas

- (1) FONTDEVILA, A. y MOYA, A. 1999. *Introducción a la genética de poblaciones*. Editorial Síntesis, Madrid. El más reciente y completo tratado en castellano sobre genética de poblaciones.
- (2) HARTL, D.L. y CLARK, A.G. 1989. *Principles of population genetics*. 2nd ed. Sinauer Ass. Uno de los mejores libros de texto sobre el tema.
- (3) LI, W.-H. 1997. *Molecular evolution*. Sinauer, Sunderland. Excelente para profundizar sobre la variabilidad al nivel del ADN y la evolución molecular, con una puesta al día muy reciente.