

24
38



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

EL EQUILIBRIO DE HARDY - WEINBERG PARA
DOS LOCI CON VARIOS ALELOS.

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
M A T E M Á T I C O
P R E S E N T A ;

ALBERTO MANUEL ALDAMA GARISOAIN



TESIS CON
FALLA DE ORIGEN



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ÍNDICE

INTRODUCCIÓN .-	2
CAPÍTULO I .- EL MODELO H-W ORIGINAL Y SU EXTENSIÓN INMEDIATA	5
CAPÍTULO II .- EL CASO DE GENES LIGADOS AL SEXO	31
CAPÍTULO III .- EL MODELO MATRICIAL DE LI : DOS LOCI, DOS ALELOS	43
CAPÍTULO IV .- DOS LOCI CON UN NÚMERO ARBITRARIO DE ALELOS	59
CAPÍTULO V .- CONCLUSIONES Y PERSPECTIVAS	95
APÉNDICE A .- LA CONDICIÓN DE EQUILIBRIO H-W EN LA GENÉTICA DE POBLACIONES : CONSIDERACIONES HISTÓRICAS	109
APÉNDICE B .- BASES FENOMENOLÓGICAS DE LA GENÉTICA DE POBLACIONES	112
APÉNDICE C .- CUESTIONES TÉCNICAS	127
BIBLIOGRAFÍA .-	149

INTRODUCCIÓN

La finalidad de este trabajo es hacer una presentación más general del resultado clásico de la Genética de poblaciones conocido como "Ley de Hardy-Weinberg". A partir de trabajos posteriores al descubrimiento independiente por G. H. Hardy y W. Weinberg en 1908 - trabajos en los que se proponen extensiones al original en el sentido de considerar más de un solo locus - , se observa que en estos casos el equilibrio en las frecuencias genotípicas no se alcanza en la primera generación, sino que debe pasar un tiempo infinito para ello; sin embargo, el comportamiento de dichas frecuencias es tal que su diferencia con respecto al estado de equilibrio, en la generación t , puede medirse como α^t , siendo α un número con valor absoluto estrictamente menor que uno; asimismo, se ha mostrado que el estado de equilibrio corresponde al producto de las frecuencias marginales génicas haploides iniciales.

La generalidad de este trabajo consiste en considerar un número arbitrario de alelos en cada locus, aunque se mantiene la restricción de estudiar solamente al caso de dos loci. Por supuesto, las condiciones bajo las cuales se procede son esencialmente las mismas que en los trabajos arriba mencionados: una población infinita,

independencia de los loci (o localización en pares cromosómicos distintos) y apareamientos al azar. También se supone que las generaciones están separadas y que la población no se halla sujeta a selección ni a presiones de mutación o migración.

En el primer capítulo hago una descripción detallada - que pretende ser una presentación pedagógica - del resultado original para un locus y dos alelos, extendiéndolo inmediatamente al caso de un número arbitrario de alelos. El segundo capítulo considera también el caso de un locus pero con la particularidad de referirse al par de cromosomas sexuales; esto es, se trata el caso de genes ligados al sexo. En el tercer capítulo presento la extensión al caso de dos loci hecha por C.C. Li en 1955, en la que se consideran dos alelos en cada locus; este capítulo, así como el anterior, ilustran el hecho de que -salvo para valores iniciales particulares de las probabilidades iniciales- el equilibrio no se alcanza en la primera generación, pero que existe un comportamiento asintótico de tendencia a dicho equilibrio y que la forma que tiene la distribución en este estado es la de productos de las distribuciones marginales. Es precisamente este último resultado el que tiene por objetivo esta tesis: en el cuarto capítulo se parte del método seguido por Li y se generaliza a considerar un número arbitrario de alelos conservando la particularidad del tratamiento al caso de dos loci.

Por último, en el quinto capítulo se esboza la posibilidad de continuar este trabajo comenzando con el próximo paso en la generalización,

que consistiría en la consideración de un número arbitrario de loci con un número arbitrario de alelos en cada locus.

Uno de los apéndices contiene las demostraciones de carácter técnico que no se dan en el cuerpo principal del texto por no desviar la atención hacia razonamientos que no por ser muy sencillos dejan de ser un poco largos y tediosos. Los otros dos apéndices contienen, uno una breve reseña histórica del uso del concepto de equilibrio Hardy-Weinberg, y el otro las bases fenomenológicas - breve y esquemáticamente presentadas para el lector no familiarizado - sobre las que se asienta el tratamiento matemático de este modelo en particular y de la Genética de Poblaciones en general.

CAPÍTULO I

EL MODELO H-W ORIGINAL Y SU EXTENSIÓN INMEDIATA

EL MODELO ORIGINAL : UN LOCUS , DOS ALELOS

(UNA PRESENTACIÓN PEDAGÓGICA)

Considérese una población mendeliana , esto es, una población cuyo acervo génico es constante y en la que las características hereditarias obedecen las leyes de Mendel . Considérese asimismo una característica hereditaria determinada por un solo gen autosómico y supóngase que éste puede presentarse en una de dos formas o alelos , digamos A_1 y A_2 . En cuanto a esta característica entonces , la población puede clasificarse de acuerdo a tres diferentes genotipos :

$$A_1 A_1 , \quad A_1 A_2 \quad \text{y} \quad A_2 A_2 ,$$

a los que denominaremos , por el momento , homocigoto dominante , heterocigoto y homocigoto recesivo , respectivamente .

Supóngase que dicha población inicialmente se halla distribuida de tal modo que las probabilidades de encontrar - eligiendo al azar - un individuo con genotipo $A_1 A_1$, uno con genotipo $A_1 A_2$ y uno con genotipo $A_2 A_2$ son , respectivamente , d , h y r , teniéndose , por supuesto :

$$d + h + r = 1 \quad \dots \quad (I.1)$$

Para esta población, que se reproduce sexualmente, podemos suponer que la distribución dada es la misma tanto para machos como para hembras; atendiendo entonces a las diferencias tanto genotípicas como sexuales, se tienen 9 posibles maneras diferentes de formar pareja, a los que listamos en la tabla I.

Si añadimos la hipótesis de que los apareamientos ocurren al azar, esto es, que el proceso de formación de pareja es aleatorio, obten que el hecho de que determinado macho encuentre a determinada hembra es azaroso, entonces esto equivale a suponer que el evento consistente en el encuentro de ese determinado macho con esa determinada hembra es equivalente a la ocurrencia simultánea de dos sucesos independientes: la aparición de ese macho y esa hembra.

Como hemos supuesto que las probabilidades genotípicas dadas arriba son iguales para machos y para hembras, las probabilidades asociadas a la formación de las distintas parejas listadas en la tabla I pueden calcularse del siguiente modo:

$$\begin{aligned}
 \Pr(\text{formación de la pareja } \sigma^x \times \phi^y) &= \\
 &= \Pr(\text{genotipo } \sigma^x \text{ es } X \text{ y genotipo } \phi \text{ es } Y) \\
 &= \Pr(\text{genotipo } \sigma^x \text{ es } X) \times \Pr(\text{genotipo } \phi \text{ es } Y)
 \end{aligned}$$

$$= P_r(X) \cdot P_r(Y)$$

debido a la independencia de los eventos, como hemos dicho arriba. (X , Y representan a cualquiera de los genotipos considerados)

Nótese que

$$d^2 + h^2 + r^2 + 2(dh + dr + hr) = (d + h + r)^2 = 1$$

lo que era de esperarse desde el punto de vista de la teoría clásica de probabilidades, ya que las nueve parejas constituyen un "conjunto exhaustivo de eventos", esto es, sus elementos son mutuamente excluyentes y complementarios.

TABLA I

# Pareja	Pareja $\sigma \times \varphi$	Probabilidad de formación	PROCESOS FECUNDATIVOS (Probabilidades Condicionales)			
			$A_1 \rightarrow A_1$	$A_1 \rightarrow A_2$	$A_2 \rightarrow A_1$	$A_2 \rightarrow A_2$
1	$A_1 A_1 \times A_1 A_1$	d^2	1	0	0	0
2	$A_1 A_1 \times A_1 A_2$	$d h$	$\frac{1}{2}$	$\frac{1}{2}$	0	0
3	$A_1 A_1 \times A_2 A_2$	$d r$	0	1	0	0
4	$A_1 A_2 \times A_1 A_1$	$h d$	$\frac{1}{2}$	0	$\frac{1}{2}$	0
5	$A_1 A_2 \times A_1 A_2$	h^2	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
6	$A_1 A_2 \times A_2 A_2$	$h r$	0	$\frac{1}{2}$	0	$\frac{1}{2}$
7	$A_2 A_2 \times A_1 A_1$	$r d$	0	0	1	0
8	$A_2 A_2 \times A_1 A_2$	$r h$	0	0	$\frac{1}{2}$	$\frac{1}{2}$
9	$A_2 A_2 \times A_2 A_1$	r^2	0	0	0	1

Ahora consideremos a los procesos de fecundación que darán origen a los nuevos individuos, a quienes veremos como integrantes de la generación 1.

Por lo que sabemos acerca de los procesos de gametogénesis, si el genotipo de un individuo es, digamos XX' , entonces sus gametos serán la mitad X y la otra mitad X' , esto es, cada gameto posee información genética igual a la mitad de la información genotípica que tienen las otras células (somáticas) del individuo dado. Suponer cierto esto (la primera ley de Mendel) equivale a decir que la probabilidad de que un determinado gameto posea una de las dos formas alélicas depende del genotipo del individuo que produce dicho gameto. Entonces, mediante el uso de la definición de probabilidad condicional, se tiene que, si X es un alelo distinto de X' :

$$\Pr(\text{gameto sea } X \mid \text{genotipo es } XX) = 1$$

$$\Pr(\text{gameto sea } X \mid \text{genotipo es } XX') = 1/2$$

$$\Pr(\text{gameto sea } X \mid \text{genotipo es } X'X') = 0$$

Ahora, el proceso fecundativo consistente en la fecundación de un huevo X' por un espermia X (que simbolizamos escribiendo $X \rightarrow X'$) puede ser considerado como el encuentro azaroso de éstos, ya que la forma alélica de uno no condiciona la del otro; esto es, este evento constituye la ocurrencia simultánea de dos eventos

independientes. Por eso :

$$P_r (X \rightarrow X') = P_r (X) P_r (X')$$

Para una pareja dada, por ejemplo el caso de la pareja B, se tiene:

$$\begin{aligned} P_r (A_1 \rightarrow A_1) &= P_r (\text{gameto masculino } A_1 \text{ y gameto femenino } A_1) \\ &= P_r (\text{gameto masculino } A_1) \cdot P_r (\text{gameto femenino } A_1) \\ &= P_r (\text{gameto } A_1 | \text{genotipo } A_1 A_1) \cdot P_r (\text{gameto } A_1 | \text{genotipo } A_1 A_1) \\ &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \end{aligned}$$

En esta forma se calcularon las probabilidades asignadas a cada uno de los cuatro procesos fecundativos que son posibles, condicionados a realizarse por cada pareja dada y que se listan en el renglón correspondiente a dicho pareja, en la columna respectiva a cada proceso fecundativo.

Una vez calculadas estas probabilidades condicionales procederemos a su vez a calcular, con base en éstas, las probabilidades asociadas a los distintos genotipos para la siguiente generación, esto es, para la generación 1. Osea: puesto que los individuos de la generación 1 provienen de procesos fecundativos llevados a cabo por los gametos de los diferentes integrantes de las parejas formadas en la generación inicial - a la cual llamaremos generación 0 (cero)- y puesto que las probabilidades asociadas tanto a la formación de la

pareja como a los posibles procesos fecundativos subsecuentes han sido ya calculadas, será posible obtener de aquí las probabilidades genotípicas de la nueva generación. Sólo falta aún decir que también suponemos, como hipótesis adicional, que consideramos a las generaciones separadas, esto es, la población de la generación siguiente consta exclusivamente de los hijos de individuos de la generación, los cuales se considera que han desaparecido.

Obtendremos entonces las probabilidades genotípicas de la nueva generación paso a paso, para los dos casos A_1A_1 y A_1A_2 pensando que el tercero, A_2A_2 , es análogo al primero pues los papeles de los símbolos A_1 y A_2 son simétricos.

Pues bien, la probabilidad de que nazca un individuo con genotipo A_1A_1 debe ser igual a la probabilidad de que alguno de los 9 tipos de parejas haya llevado a cabo el proceso fecundativo $A_1 \rightarrow A_1$ que es el único de los 4 posibles que puede darle origen.

Entonces, para la primera generación (generación 1) :

$$\Pr(A_1A_1) = \Pr(A_1 \rightarrow A_1) \quad ;$$

como dicho proceso fecundativo fue llevado a cabo por alguna de las nueve parejas posibles, podemos escribir :

$$\begin{aligned}
Pr(A_1 \rightarrow A_1) &= Pr(A_1 \rightarrow A_1, \text{ y el proceso lo realice alguna de las 9 parejas}) \\
&= Pr(A_1 \rightarrow A_1, \text{ y (el proceso lo realice la pareja 1 o} \\
&\quad \text{" " " " " " 2 o} \\
&\quad \text{etc.} \\
&\quad \text{" " " " " " 9)}) \\
&= Pr((A_1 \rightarrow A_1, \text{ y el proceso lo realice la pareja 1}) \\
&\quad \cup (A_1 \rightarrow A_1, \text{ " " " " " " 2}) \\
&\quad \cup \text{ etc.} \\
&\quad \cup (A_1 \rightarrow A_1, \text{ " " " " " " 9})) \\
&= Pr(A_1 \rightarrow A_1, \text{ y el proceso lo realice la pareja 1}) \\
&+ Pr(A_1 \rightarrow A_1, \text{ " " " " " " 2}) \\
&+ \text{ etc.} \\
&+ Pr(A_1 \rightarrow A_1, \text{ " " " " " " 9}) \\
&= Pr(A_1 \rightarrow A_1 | \text{ pareja 1}) \cdot Pr(\text{pareja 1}) \\
&+ Pr(A_1 \rightarrow A_1 | \text{ pareja 2}) \cdot Pr(\text{pareja 2}) \\
&+ \text{ etc.} \\
&+ Pr(A_1 \rightarrow A_1 | \text{ pareja 9}) \cdot Pr(\text{pareja 9})
\end{aligned}$$

Explicamos: el evento cuya probabilidad nos interesa es " $A_1 \rightarrow A_1$ ", así, de modo absoluto; sin embargo, de este modo no tenemos manera de calcular su probabilidad asociada. Por otro lado, existe un evento seguro: "la fecundación sea realizada por alguna de las nueve parejas". Sabemos que si " S " es un evento seguro, entonces $Pr(S) = 1$, por lo que, de

acuerdo con la fórmula para probabilidades condicionales, si "E" es cualquier evento :

$$Pr(E \mid S) = Pr(E \mid S) Pr(S) = Pr(E \mid S),$$

pero siendo "S" seguro, es claro que $Pr(E \mid S) = Pr(E)$ pues efectivamente S ha de suceder. En nuestro caso: la probabilidad de que "el proceso fecundativo realizado sea $A_1 \rightarrow A_1$ " (en el papel de "E") debe ser igual a la probabilidad de que "el proceso fecundativo realizado sea $A_1 \rightarrow A_1$ " y "este proceso sea realizado por alguna de las 9 parejas", lo cual constituye, en escritura común, lo establecido en la primera de la serie de igualdades de atrás.

Ahora, el evento "la fecundación sea realizada por alguna de las nueve parejas" puede enunciarse como: "la fecundación sea realizada por la pareja 1 o por la pareja 2 o... etc, o por la pareja 9" que es a su vez lo mismo que "la fecundación sea realizada por la pareja 1" o "la fecundación sea realizada por la pareja 2" o "etc. ..." o "la fecundación sea realizada por la pareja 9", lo que se traduce en la segunda y tercera igualdades por la distributividad de los conectivos lógicos. Aquí, las nueve últimas disyunciones constituyen nueve eventos mutuamente excluyentes, por lo que la probabilidad de que suceda alguno es igual a la suma de las correspondientes probabilidades individuales, dando así origen a la cuarta igualdad. La última igualdad se debe al hecho de que, en cada caso, la

probabilidad del proceso $A_1 \rightarrow A_1$ depende de la pareja que habrá de realizarlo. Entonces la probabilidad de que suceda simultáneamente el proceso $A_1 \rightarrow A_1$ y que este proceso sea realizado por una pareja determinada, será igual - usando el teorema conocido para probabilidades condicionales:

$Pr(X \text{ y } Y) = Pr(X|Y) Pr(Y)$ - el producto de la probabilidad de que suceda $A_1 \rightarrow A_1$ bajo la hipótesis de que la pareja que ha de realizarlo sea esa determinada, por la probabilidad de que dicha pareja se encuentre, ya que suponemos que una vez que se ha formado la pareja, ésta habrá de realizar algún proceso fecundativo de los cuatro posibles, con las respectivas probabilidades condicionales calculadas.

En la forma en como hemos construido la Tabla I, podemos ver que el resultado buscado - la probabilidad absoluta del proceso $A_1 \rightarrow A_1$ - se encuentra multiplicando los números correspondientes a las probabilidades condicionales asociadas al proceso $A_1 \rightarrow A_1$ con los correspondientes a las probabilidades de formación de pareja y sumando dichos productos; esto, como sabemos, es el producto interno de los vectores columna "probabilidades condicionales para el proceso $A_1 \rightarrow A_1$ " con "probabilidades de formación de pareja", o sea:

$$Pr(A_1 \rightarrow A_1) = \begin{bmatrix} d^2 \\ dh \\ dr \\ hd \\ h^2 \\ hr \\ rd \\ rh \\ r^2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1/2 \\ 0 \\ 1/2 \\ 1/4 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} &= d^2 + 1/2 dh + 1/2 hd + 1/4 h^2 \\ &= d^2 + dh + 1/4 h^2 \\ &= (d + 1/2 h)^2 \end{aligned}$$

Ahora, el genotipo $A_1 A_2$ puede formarse a partir de dos procesos fecundativos distintos: $A_1 \rightarrow A_2$ y $A_2 \rightarrow A_1$, por lo que

$$Pr(\text{genotipo } A_1 A_2) = Pr(\text{"proceso fecundativo } A_1 \rightarrow A_2" \text{ o "proceso fecundativo } A_2 \rightarrow A_1")$$

Como estos dos procesos constituyen eventos mutuamente excluyentes, se tendrá que la probabilidad anterior es igual a la suma:

$$Pr(\text{"proceso fecundativo } A_1 \rightarrow A_2") + Pr(\text{"proceso fecundativo } A_2 \rightarrow A_1");$$

cada uno de estos términos se calcula ahora mediante un procedimiento totalmente análogo al del caso $A_1 \rightarrow A_1$; se tiene, portanto:

$$\begin{aligned} \Pr(A_1 \rightarrow A_2) &= \Pr(A_1 \rightarrow A_2 \text{ y sea realizado por alguna de las 9 parejas}) \\ &= \Pr(\text{"}A_1 \rightarrow A_2\text{" y sea realizado por la pareja 1 o} \\ &\quad \text{sea realizado por la pareja 2 o} \\ &\quad \text{etc... sea realizado por la pareja 9"} \end{aligned}$$

por la distributividad de los conectivos Lógicos

$$\begin{aligned} &= \Pr(\text{"}A_1 \rightarrow A_2 \text{ y sea realizado por la pareja 1"} \\ &\quad \text{"}A_1 \rightarrow A_2 \text{ " " " " " " " 2"} \\ &\quad \text{etc.} \end{aligned}$$

por la mutua exclusividad de los eventos

$$\begin{aligned} &= \Pr(A_1 \rightarrow A_2 \text{ y sea realizado por la pareja 1}) \\ &+ \Pr(A_1 \rightarrow A_2 \text{ " " " " " " " 2}) \\ &+ \text{etc.} \\ &+ \Pr(A_1 \rightarrow A_2 \text{ " " " " " " " 9}) \end{aligned}$$

por la consistencia de los eventos

$$\begin{aligned} &= \Pr(A_1 \rightarrow A_2 \mid \text{pareja 1}) \Pr(\text{pareja 1}) \\ &+ \Pr(A_1 \rightarrow A_2 \mid \text{pareja 2}) \Pr(\text{pareja 2}) \\ &+ \text{etc.} \\ &+ \Pr(A_1 \rightarrow A_2 \mid \text{pareja 9}) \Pr(\text{pareja 9}) \end{aligned}$$

de la Tabla I

$$= \begin{bmatrix} d^2 \\ dh \\ dr \\ hd \\ h^2 \\ hr \\ rd \\ rh \\ r^2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1/2 \\ 1 \\ 0 \\ 1/4 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$= 1/2 dh + dr + 1/4 h^2 + 1/2 hr$$

y de manera similar: $Pr(A_2 \rightarrow A_1) = 1/2 hd + 1/4 h^2 + rd + 1/2 rh$
 por lo que

$$\begin{aligned} Pr(\text{genotipo } A_1 A_2) &= Pr(A_1 \rightarrow A_2) + Pr(A_2 \rightarrow A_1) \\ &= (1/2 dh + 1/4 h^2 + 1/2 hr) + (1/2 hd + 1/4 h^2 + rd + 1/2 rh) \\ &= dh + hr + 2dr + 1/2 h^2 \\ &= 2(d + 1/2 h)(r + 1/2 h) \end{aligned}$$

Por último, la probabilidad $Pr(A_1 A_2)$ puede calcularse de manera completamente análoga a la probabilidad $Pr(A_1 A_1)$ ya sea siguiendo paso a paso el procedimiento hasta aquí descrito, o bien aprovechando la simetría del problema y notando que si cambiásemos el símbolo A_1 por el símbolo A_2 , con el correspondiente

cambio de d por r , el cálculo sería exactamente el mismo y tendríamos:

$$\Pr(A_2 A_2) = (r + \frac{1}{2}k)^2$$

Resumiendo, tenemos que para la generación 1 las probabilidades asignadas a los distintos genotipos son:

$$\begin{aligned} d_1 &\equiv \Pr(A_1 A_1) = (d + \frac{1}{2}k)^2 \\ h_1 &\equiv \Pr(A_1 A_2) = 2(d + \frac{1}{2}k)(r + \frac{1}{2}k) \\ r_1 &\equiv \Pr(A_2 A_2) = (r + \frac{1}{2}k)^2 \quad \dots \text{(I.2)} \end{aligned}$$

que denotamos como d_1 , h_1 y r_1 , respectivamente, para recuperar la notación empleada inicialmente y donde usamos el subíndice 1 para indicar la generación a que se refieren.

Si comenzamos de nuevo - manteniendo la hipótesis de que los apareamientos ocurren al azar - a partir de las nuevas probabilidades d_1 , h_1 y r_1 , tendremos exactamente la misma situación que en un principio; esto es, formaremos una tabla análoga a la Tabla I conteniendo las probabilidades calculadas de la misma forma para, después de proceder de igual manera a como hasta aquí lo hemos hecho, llegar a la conclusión de que, para la generación 2, ha de tenerse:

$$d_2 \equiv Pr(A_1 A_1) = (d_1 + \frac{1}{2} h_1)^2$$

$$h_2 \equiv Pr(A_1 A_2) = 2(d_1 + \frac{1}{2} h_1)(r_1 + \frac{1}{2} h_1)$$

$$r_2 \equiv Pr(A_2 A_2) = (r_1 + \frac{1}{2} h_1)^2$$

y por tanto, repitiendo este proceso t veces tendremos, para la generación t :

$$d_t \equiv Pr(A_1 A_1) = (d_{t-1} + \frac{1}{2} h_{t-1})^2$$

$$h_t \equiv Pr(A_1 A_2) = 2(d_{t-1} + \frac{1}{2} h_{t-1})(r_{t-1} + \frac{1}{2} h_{t-1})$$

$$r_t \equiv Pr(A_2 A_2) = (r_{t-1} + \frac{1}{2} h_{t-1})^2$$

... (I.3)

siempre y cuando la hipótesis de que los apareamientos ocurren al azar se mantenga, pues dicha hipótesis ha sido fundamental para el cálculo efectuado al buscar las probabilidades asociadas a la formación de parejas, pudiendo así considerar a tales eventos como ocurrencias simultáneas de sucesos independientes y dar por ello una distribución de probabilidad correspondiente a los nueve términos del desarrollo de $(d_t + h_t + r_t)^2$.

Antes de proseguir con las implicaciones que tienen estos resultados, notemos que si hacemos

$$p \equiv d + \frac{1}{2} h \quad , \quad q \equiv r + \frac{1}{2} h \quad \dots \text{(I.4)}$$

se tiene:

$$p+q=1, \quad d_1=p^2, \quad h_1=2pq, \quad r_1=q^2 \quad \dots (I.5)$$

directamente de (I.1) y (I.2), o sea que las probabilidades asignadas a los diversos genotipos, en la primera generación, corresponden a los términos del desarrollo del binomio $(p+q)^2$, cosa que no necesariamente tenía por qué haber sucedido para los números - arbitrarios, aunque sujetos a la condición (I.1) - correspondientes a las probabilidades asociadas a tales genotipos en la población inicial (generación 0): d, h y r , respectivamente, como puede comprobarse fácilmente.

Con esto último en mente observemos que

$$\begin{aligned} d_2 &= (d_1 + \frac{1}{2} h_1)^2 && \text{de (I.3)} \\ &= (p^2 + pq)^2 && \text{de (I.2 y 4)} \\ &= (p(p+q))^2 \\ &= p^2 \\ &= (d + \frac{1}{2} h)^2 && \text{de (I.4)} \\ &= d, && \text{de (I.2)} \end{aligned}$$

Análogamente - recordando que los papeles jugados por d y r son simétricos - puede verse fácilmente que $r_2 = r_1$, y por tanto que también $h_2 = h_1$. Es inmediato ver aquí que también podemos seguir los mismos pasos, ahora para d_3 ,

h_3 y r_3 , teniéndose que $d_3 = d_2$, $h_3 = h_2$ y $r_3 = r_2$, por lo que procediendo de la misma forma tendremos:

$$d_{t+1} = d_t = d_1, \quad h_{t+1} = h_t = h_1, \quad r_{t+1} = r_t = r_1,$$

lo que significa que, a partir de la primera generación, las proporciones en que se hallan presentes los genotipos A_1A_1 , A_1A_2 y A_2A_2 en la población considerada, y bajo la hipótesis fundamental de que los apareamientos ocurren al azar, no cambian con el tiempo y son iguales a los correspondientes términos del desarrollo del binomio $(p+q)^2$ con p y q dados por (I.4).

Una última consideración, que será de gran utilidad después, es la siguiente: podemos pensar en los distintos tipos de gametos que produce la población, que en nuestro caso son dos: A_1 y A_2 , pues no dependen de la forma que los hace actuantes en el proceso fecundativo, es decir, no importa si son espermas u óvulos puesto que el gene que estamos estudiando es autosómico - no pertenece a cromosomas sexuales o determinantes del sexo -; podemos preguntarnos acerca de la probabilidad con la que están distribuidos tales gametos en la población, obten por la distribución de probabilidad de los tipos alélicos, pues nos hallamos en el caso de un solo locus.

Pensemos entonces que, por lo ya referido sobre gametogénesis,

cada individuo "porta" dos "etiquetas", cada una con la mitad de la información genética correspondiente a su genotipo; así entonces, los homocigos A_1A_1 portan dos "etiquetas" A_1 , mientras que los heterocigos A_1A_2 tienen una "etiqueta" A_1 y otra A_2 . Si dichas "etiquetas" fuesen "depositadas" en una urna y de allí se extrajese una al azar: ¿cuál es la probabilidad de que ésta sea, digamos, A_1 ?

Puede verse fácilmente que, en este caso, se debe tener, en la generación inicial:

$$Pr(A_1) = d + \frac{1}{2}h = p$$

análogamente:

$$Pr(A_2) = r + \frac{1}{2}h = q$$

en la generación 1: $Pr(A_1) = d_1 + \frac{1}{2}h_1 = p^2 + pq = p(p+q) = p$

$$Pr(A_2) = r_1 + \frac{1}{2}h_1 = q^2 + pq = q(p+q) = q$$

y por lo que ya sabemos sobre el comportamiento de las proporciones genotípicas de las generaciones subsecuentes, podemos decir que para toda generación:

$$Pr(A_1) = p, \quad Pr(A_2) = q$$

Si pensamos ahora que el engendrar a un nuevo individuo es equivalente a juntar dos "etiquetas" extraídas al azar, tendremos que pensar en dos urnas, una masculina y otra femenina, ambas con la misma distribución de probabilidad para sus "etiquetas".

tas", y la probabilidad de juntar una etiqueta específica de la urna masculina con una específica de la urna femenina es un evento que consiste en la ocurrencia simultánea de dos eventos independientes, entonces:

$$\begin{aligned}
 \Pr(A_1 A_1) &= \Pr(A_1 \text{ de urna masc. y } A_1 \text{ de urna fem.}) \\
 &= \Pr(A_1 \sigma \text{ y } A_1 Q) \\
 &= \Pr(A_1 \sigma) \Pr(A_1 Q) \\
 &= \Pr(A_1) \Pr(A_1) \\
 &= p^2
 \end{aligned}$$

$$\begin{aligned}
 \Pr(A_1 A_2) &= \Pr((A_1 \sigma \text{ y } A_2 Q) \cup (A_2 \sigma \text{ y } A_1 Q)) \\
 &= \Pr(A_1 \sigma \text{ y } A_2 Q) + \Pr(A_2 \sigma \text{ y } A_1 Q) \\
 &= \Pr(A_1 \sigma) \Pr(A_2 Q) + \Pr(A_2 \sigma) \Pr(A_1 Q) \\
 &= \Pr(A_1) \Pr(A_2) + \Pr(A_2) \Pr(A_1) \\
 &= 2pq
 \end{aligned}$$

y también: $\Pr(A_2 A_2) = q^2$

es decir, la misma distribución de probabilidad que habíamos deducido antes estudiando a los posibles parejas de individuos y sus uniones al azar.

Tenemos entonces que es lo mismo hablar de unión al azar de los individuos - formación de parejas al azar - que de la unión al azar de los gametos, pensando a éstos como antes abstractos o "etiquetas"

depositadas en una urna o acervo génico. Esto será de gran utilidad más adelante pues nos ahorraríamos el tener que considerar a todas las posibles parejas para estudiar la dinámica de la herencia bajo la hipótesis de apareamientos al azar.

En conclusión, para el caso de considerar un locus con dos alelos, y bajo la hipótesis de que a partir de cierto momento los apareamientos ocurren al azar, se tiene que, desde la primera generación, la distribución de probabilidad para los diversos genotipos es la que corresponde a los términos del desarrollo del binomio $(p+q)^2$, siendo p y q las probabilidades de los gametos, pensando a éstos como distribuidos en un acervo génico de la población, siendo dicha distribución genotípica la misma que si se supone una unión aleatoria de los gametos, y se tiene que tal distribución se mantendrá constante a lo largo de las generaciones ulteriores si la hipótesis de apareamientos al azar se sostiene en todo momento.

De una población con estas características se dice que está en equilibrio génico (de Hardy - Weinberg); es decir, la distribución de probabilidad para los genotipos no cambia con el paso de las generaciones. En este caso, el equilibrio debido a la aleatoriedad de los apareamientos se denomina equilibrio de Hardy - Weinberg o Ley de Hardy - Weinberg al resultado obtenido, pues estos dos científicos, matemático británico el primero, médico alemán el segundo, llegaron ambos por primera vez e independientemente a estas conclusiones en el año de 1908.

LA EXTENSIÓN INMEDIATA : UN LOCUS , VARIOS ALELOS .

Supongamos ahora una situación similar a la analizada anteriormente ; esto es, veremos el caso de un solo locus autosómico para una población mendeliana cuya reproducción pueda suponerse que se realiza formando parejas al azar.

Como ya sabemos de los resultados anteriores, esto es lo mismo que suponer el encuentro azaroso de los gametos, suponiendo a éstos como entes abstractos que son extraídos aleatoriamente de un conjunto infinito (acervo génico) en el que se hallan presentes con una determinada distribución inicial de probabilidad dada por la distribución original de los genotipos diploides que les dan origen mediante reglas sencillas de gametogénesis (primera ley de Mendel).

Suponemos entonces que en el locus autosómico α se tienen n alelos posibles, digamos A_1, A_2, \dots, A_n . Se tendrán entonces $\frac{n(n+1)}{2}$ posibles genotipos diploides :

$$\begin{array}{l}
 A_1 A_1, A_1 A_2, A_1 A_3, \dots, A_1 A_n, \\
 A_2 A_2, A_2 A_3, \dots, A_2 A_n, \\
 \vdots \\
 A_{n-1} A_{n-1}, A_{n-1} A_n, \\
 A_n A_n.
 \end{array}$$

$$\text{Hagamos } G \equiv \{A_i \mid i \in [1, n]\}$$

$$Z \equiv \{A_i A_j \mid i \in [1, n], j \in [i, n]\}$$

(usamos la notación $[a, b]$ de intervalo cerrado común para subconjuntos de \mathbb{R} , pero restringiendo su uso a los números enteros; esto es, $[a, b]$ es el conjunto de enteros x tales que $a \leq x \leq b$; asimismo: (a, b) es el conjunto de enteros tales que $a < x < b$, etc. Esto facilitará la notación y lo usaremos así de aquí en adelante en todo el trabajo. En caso de requerir usar a todo el intervalo real usaremos la siguiente notación alternativa: $[,]_{\mathbb{R}}$ o $(,)_{\mathbb{R}}$, etc.)

G y Z son el conjunto de gametos y el conjunto de zigotos, respectivamente.

Supongamos una distribución inicial de probabilidad en los zigotos o genotipos diploides dada por:

$$Pr(A_i A_j) = P_{ij} \quad i \in [1, n], j \in [i, n]$$

$$\text{con } \sum_{i=1}^n \sum_{j=i}^n P_{ij} = 1 \quad \dots \quad (\text{I.5})$$

Si concebimos a los gametos como los entes abstractos pertenecientes al arbo geneico infinito ya mencionado y pensamos en su distribución de probabilidad, podemos escribir, para la misma generación inicial, o generación cero:

$$\Pr(A_k) = P_k \quad \text{con} \quad \sum_{k=1}^n P_k = 1 \quad \dots \text{(I.6)}$$

$k \in [1, n]$

Los procesos gametogénicos, por hipótesis, deben dar lugar a las siguientes relaciones gametogénicas:

$$P_k = P_{kk} + \frac{1}{2} \sum_{i=1}^{k-1} P_{ik} + \frac{1}{2} \sum_{j=k+1}^n P_{kj} \quad \dots \text{(I.7)}$$

$$k \in [1, n]$$

pues los gametos A_k provienen de los diversos genotipos diploides de acuerdo a lo siguiente:

todos los gametos de los individuos	$A_k A_k$	
la mitad de los gametos de los individuos	$A_i A_k$	$i < k$
la mitad de los gametos de los individuos	$A_k A_j$	$j > k$

Ahora, suponiendo que las generaciones están separadas, tendremos que para la siguiente generación la distribución de probabilidad para los genotipos diploides o zigóticos de dicha generación (generación 1) dependerá exclusivamente de los procesos de apareamiento y fecundación.

Si suponemos que los apareamientos ocurren al azar y de aquí infer-

rimos que esto es lo mismo que suponer el apareamiento al azar de los gametos, tendremos que habrá n^2 procesos fecundativos diferentes

$$A_i \rightarrow A_j \quad i, j \in [1, n]$$

Las siguientes relaciones de fecundación se cumplen entonces:

$$P_{i,j}(1) = \begin{cases} P_i^2 & i=j \\ 2P_i P_j & i \neq j \end{cases} \quad \dots \quad (I.8)$$

Para esta nueva generación deben tenerse, al momento de la gametogénesis, las mismas relaciones gametogénicas (I.7), sólo que ahora con el correspondiente cambio en la asignación de tiempos generacionales; esto es:

$$P_h(1) = P_{hh}(1) + \frac{1}{2} \sum_{i=1}^{h-1} P_{ih}(1) + \frac{1}{2} \sum_{j=h+1}^n P_{hj}(1)$$

por lo que, de las relaciones de fecundación (I.8), tenemos lo que podemos llamar relaciones hereditarias haploides:

$$\begin{aligned} P_h(1) &= P_h^2 + \frac{1}{2} \sum_{i=1}^{h-1} (2P_i P_h) + \frac{1}{2} \sum_{j=h+1}^n (2P_h P_j) \\ &= P_h \left(P_h + \sum_{i=1}^{h-1} P_i + \sum_{j=h+1}^n P_j \right) \end{aligned}$$

$$= P_k \left(\sum_{i=1}^u P_i \right) = P_k \quad \dots \quad (\text{I.9})$$

$\forall k \in [1, u]$

En fin, si podemos mantener las relaciones gametogénicas (primera ley de Mendel) y las relaciones de fecundación (apareamiento al azar de los gametos) durante toda generación, podremos decir:

$$P_k(t) = P_{kk}(t) + \frac{1}{2} \sum_{i=1}^{k-1} P_{ik}(t) + \frac{1}{2} \sum_{j=k+1}^u P_{kj}(t)$$

... relaciones gametogénicas

$$P_{ij}(t+1) = \begin{cases} P_i^2(t) & i=j \\ 2P_i(t)P_j(t) & i \neq j \end{cases} \quad t \in \mathbb{N}$$

... relaciones de fecundación

donde la variable t indica, en cada caso, la generación de que se trata.

Se tiene entonces, siguiendo a (I.9), por inducción:

$$P_k(t) = P_k(1) = P_k \quad k \in [1, u] \quad , \quad t \in \mathbb{N}$$

por lo que

$$P_{ij}(t) = P_{ij}(1) = \begin{cases} P_i^2 & i=j \\ 2P_i P_j & i \neq j \end{cases} \quad \begin{matrix} i \in [1, u] \\ j \in [1, u] \\ t \in \mathbb{N} \end{matrix}$$

que es la condición de equilibrio H-W buscada.

CAPÍTULO II

EL CASO DE GENES LIGADOS AL SEXO

Consideraremos ahora el caso de un carácter hereditario determinado por un solo gen que tiene dos formas alélicas, digamos (1) y (2), y cuyo locus está situado en el cromosoma sexual X (*). Se tienen tres genotipos femeninos: X_1X_1 , X_1X_2 , X_2X_2 , y dos genotipos masculinos: X_1Y , X_2Y . Supongamos una población mendeliana con esta característica y para la cual inicialmente se tenga:

$$P_r(X_1X_1) = d, \quad P_r(X_1X_2) = h, \quad P_r(X_2X_2) = r$$

con $d + h + r = 1$... (II.1)

en la subpoblación femenina, y

$$P_r(X_1Y) = d', \quad P_r(X_2Y) = r' \quad \text{con } d' + r' = 1$$

... (II.2)

en la subpoblación masculina.

Notemos que al asignar estos valores a las proporciones en que se hallan inicialmente los diversos genotipos diploides, hemos implícitamente separado a la población en dos subpoblaciones, la femenina y la masculina, y dentro de cada una de estas han sido asignadas distribuciones de probabilidad diferentes; o sea que la probabilidad asociada

(*). Pedimos al lector no familiarizado con la división de los cromosomas sexuales consultar el Apéndice B, donde se discute brevemente la razón del tratamiento especial que aquí hacemos.

a un determinado genotipo femenino se ha tomado como la proporción de individuos con dicho genotipo con respecto al total de la subpoblación femenina. De manera completamente análoga en el caso masculino. Esto, en otras palabras, equivale a tomar las probabilidades condicionales asignadas a un determinado genotipo bajo la hipótesis de que pertenece a uno u otro sexo, según el caso.

Elaboraremos una tabla (Tabla II) semejante a la que construimos en el caso de cromosomas autosómicos, es decir, no sexuales, y deberemos poner especial cuidado en mantener a ambas subpoblaciones separadas, sobre todo en el momento de calcular las probabilidades condicionales asociadas a los diversos procesos fecundativos; por tanto, habrá procesos fecundativos que originen machos y procesos fecundativos que originen hembras, y tales procesos serán considerados distintos y por separado, lo que quiere decir que calcularemos las probabilidades condicionales de que se efectúen los procesos fecundativos determinados no sólo bajo la hipótesis de que dichos procesos son realizados por una determinada pareja, sino también bajo la hipótesis de que dicho proceso originará a un individuo con un sexo determinado.

De nuevo, por supuesto, supondremos también, con carácter central, que los apareamientos ocurren al azar.

TABLA II

PAREJA	Probabilidad de formación	PROCESOS FECUNDATIVOS (Probabilidades condicionales)					
		MACHOS		HEMBRAS			
$\sigma \times \varphi$		$Y \rightarrow X_1$	$Y \rightarrow X_2$	$X_1 \rightarrow X_1$	$X_1 \rightarrow X_2$	$X_2 \rightarrow X_1$	$X_2 \rightarrow X_2$
$X_1 Y \times X_1 X_1$	d'd	1	0	1	0	0	0
$X_1 Y \times X_1 X_2$	d'h	1/2	1/2	1/2	1/2	0	0
$X_1 Y \times X_2 X_2$	d'r	0	1	0	1	0	0
$X_2 Y \times X_1 X_1$	r'd	1	0	0	0	1	0
$X_2 Y \times X_1 X_2$	r'h	1/2	1/2	0	0	1/2	1/2
$X_2 Y \times X_2 X_2$	r'r	0	1	0	0	0	1

Siguiendo un método similar al del caso autosómico, con la novedad de que ahora tratamos con una población dividida en dos subpoblaciones distintas - y así lo seguiremos considerando para todas las generaciones subsiguientes - tendremos, para la generación 1:

$$\begin{aligned}
 d'_1 \equiv Pr(X_1, Y) &= d'd + \frac{1}{2}d'h + r'd + \frac{1}{2}r'h \\
 &= d'(d + \frac{1}{2}h) + r'(d + \frac{1}{2}h) \\
 &= (d' + r')(d + \frac{1}{2}h) \\
 &= d + \frac{1}{2}h
 \end{aligned}$$

análogamente: $r'_1 \equiv Pr(X_2, Y) = r + \frac{1}{2}h$

$$d_1 \equiv Pr(X_1, X_1) = d'(d + \frac{1}{2}h)$$

$$h_1 \equiv Pr(X_1, X_2) = d'(r + \frac{1}{2}h) + r'(d + \frac{1}{2}h)$$

$$r_1 \equiv Pr(X_2, X_2) = r'(r + \frac{1}{2}h)$$

... (II.3)

Es inmediato ver que, comenzando ahora con estas nuevas cantidades y postulando las mismas hipótesis mencionadas arriba, se tendrá, pensando inductivamente:

$$d'_{t+1} \equiv Pr(X_t, Y) = d_t + \frac{1}{2}h_t$$

$$d_{t+1} \equiv Pr(X_t, X_t) = d'_t(d_t + \frac{1}{2}h_t)$$

$$h_{t+1} \equiv Pr(X_t, X_{t+1}) = d'_t(r_t + \frac{1}{2}h_t) + r'_t(d_t + \frac{1}{2}h_t)$$

$$r'_{t+1} \equiv Pr(X_{t+1}, Y) = r_t + \frac{1}{2}h_t$$

$$r_{t+1} \equiv Pr(X_{t+1}, X_{t+1}) = r'_t(r_t + \frac{1}{2}h_t)$$

... (II.4)

Las fórmulas (II.4) relacionan a las probabilidades genotípicos diploides en la generación $t+1$ con aquellos de la generación t .

Si hacemos ahora a P'_t la probabilidad génica haploide (proporción de gametos) asociada al alelo (1) en la población masculina al tiempo t (en generaciones), y q'_t a la correspondiente al alelo (2) en la misma población, tendremos:

$$P'_t = d'_t \quad , \quad q'_t = r'_t$$

y de manera semejante, llamando P_t y q_t a las probabilidades génicas haploides respectivos en la población femenina - en la generación t - tendremos que las proporciones génicas haploides de la generación $t+1$:

$$P'_{t+1} \quad , \quad q'_{t+1} \quad ; \quad P_{t+1} \quad , \quad q_{t+1}$$

en términos de las proporciones genotípicas diploides de la misma generación ($t+1$), son :

$$P'_{t+1} = d'_{t+1} \quad , \quad q'_{t+1} = r'_{t+1} \quad \text{para los machos}$$

$$P_{t+1} = d_{t+1} + \frac{1}{2} h_{t+1} \quad , \quad q_{t+1} = r_{t+1} + \frac{1}{2} h_{t+1}$$

para las hembras.

Usando (II.4) para expresar las proporciones génicas de la generación $t+1$ en términos de las proporciones genotípicas de la generación anterior (t) obtenemos, para los machos:

$$P'_{t+1} = d_t + \frac{1}{2} h_t, \quad q'_{t+1} = r_t + \frac{1}{2} h_t,$$

lo cual muestra el hecho, inmerso en el modelo, de que la característica en cuestión, en los machos, se debe sólo a la composición genotípica de las hembras de la generación anterior.

Para las hembras, en cambio, la composición génica está influida tanto por la composición genotípica de las hembras como por la de los machos de la generación anterior, según se muestra a continuación:

$$\begin{aligned} P'_{t+1} &= d'_t (d_t + \frac{1}{2} h_t) + \frac{1}{2} d'_t (r_t + \frac{1}{2} h_t) + \frac{1}{2} r'_t (d_t + \frac{1}{2} h_t) \\ &= d'_t d_t + \frac{1}{2} d'_t h_t + \frac{1}{2} d'_t r_t + \frac{1}{4} d'_t h_t + \frac{1}{2} r'_t (d_t + \frac{1}{2} h_t) \\ &= \frac{1}{2} d'_t d_t + \frac{1}{2} d'_t d_t + \frac{1}{2} d'_t h_t + \frac{1}{2} d'_t r_t + \frac{1}{4} d'_t h_t + \frac{1}{2} r'_t (d_t + \frac{1}{2} h_t) \\ &= \frac{1}{2} d'_t d_t + \frac{1}{2} d'_t h_t + \frac{1}{2} d'_t r_t + \frac{1}{2} d'_t d_t + \frac{1}{4} d'_t h_t + \frac{1}{2} r'_t (d_t + \frac{1}{2} h_t) \\ &= \frac{1}{2} d'_t (d_t + h_t + r_t) + \frac{1}{2} d'_t (d_t + \frac{1}{2} h_t) + \frac{1}{2} r'_t (d_t + \frac{1}{2} h_t) \end{aligned}$$

$$= \frac{1}{2} d'_t (d_t + h_t + r_t) + \frac{1}{2} (d_t + \frac{1}{2} h_t) (d'_t + r'_t)$$

y como $d_t + h_t + r_t = 1$ y también $d'_t + r'_t = 1$, se obtiene, finalmente:

$$P_{t+1} = \frac{1}{2} d'_t + \frac{1}{2} (d_t + \frac{1}{2} h_t) ;$$

una deducción completamente análoga nos lleva a establecer:

$$q_{t+1} = \frac{1}{2} r'_t + \frac{1}{2} (r_t + \frac{1}{2} h_t)$$

Estas relaciones muestran claramente cómo contribuyen las composiciones genotípicas de machos y hembras de la generación t en la composición génica de las hembras de la generación $t+1$. (Porque, puesto que en d'_t y r'_t (la composición genotípica de los machos) influye sólo la composición genotípica de las hembras de la generación previa ($t-1$), esto nos lleva a conjeturar que las diferencias en las proporciones génicas de machos y hembras deben tender a desaparecer. Para ver que el modelo muestra esto, fijémonos en la diferencia $P'_{t+1} - P_{t+1}$ como sigue:

$$\begin{aligned} P'_{t+1} - P_{t+1} &= (d_t + \frac{1}{2} h_t) - \left[\frac{1}{2} d'_t + \frac{1}{2} (d_t + \frac{1}{2} h_t) \right] \\ &= -\frac{1}{2} \left[d'_t - (d_t + \frac{1}{2} h_t) \right] \end{aligned}$$

$$= -\frac{1}{2} [P'_t - P_t]$$

de donde se deduce inmediatamente que :

$$P'_t - P_t = (-\frac{1}{2})^t (P' - P) \quad \dots \text{ (II.5)}$$

siendo P' y P las proporciones génicas iniciales

$$P' = d' \quad \text{y} \quad P = d + \frac{1}{2} h$$

Vemos así que : $\lim_{t \rightarrow \infty} (P'_t - P_t) = 0 \quad \dots \text{ (II.6)}$

Podemos decir que , de una generación a otra , la diferencia entre las probabilidades génicas haploides para la subpoblación masculina y la femenina , respectivamente , se reduce a la mitad , siendo esta diferencia unas veces positiva y otras negativa , en forma alterna , con cada paso de generación en generación . Además , del resultado (II.5) , podemos poner también lo siguiente :

$$P_t = (-\frac{1}{2})^t (P - P') + P'_t$$

y como $P'_t = P'_{t-1}$, como puede verse fácilmente , se tiene :

$$\begin{aligned}
 P_t &= (-1/2)^t (P - P') + P_{t-1} \\
 &= (-1/2)^t (P - P') + (-1/2)^{t-1} (P - P') + P_{t-2} \\
 &= \dots \text{etc.} \quad (\text{recursivamente}) \\
 &= (P - P') \sum_{x=1}^t (-1/2)^x + P \\
 &= (P - P') \frac{(-1/2) (1 - (-1/2)^t)}{(1 - (-1/2))} + P \\
 &= 1/3 (P - P') ((-1/2)^t - 1) + P
 \end{aligned}$$

por lo que, en el límite:

$$P_\infty = 1/3 (P - P') + P = 2/3 P + 1/3 P' \quad \dots \quad (\text{II.7})$$

lo que implica que $q_\infty = 2/3 q + 1/3 q'$, habiendo definido a q , q' y q_∞ de modo análogo para el alelo (2), y que las distribuciones en equilibrio sean:

$$P_r(X_1 X_1) = P_\infty^2, \quad P_r(X_1 X_2) = 2 P_\infty q_\infty, \quad P_r(X_2 X_2) = q_\infty^2$$

$$P_r(X_1 Y) = P_\infty q_\infty, \quad P_r(X_2 Y) = q_\infty P_\infty \quad \dots \quad (\text{II.8})$$

Mostremos esto: en la generación t

$$P_r(X_1 X_1) = d_1 (d_1 + 1/2 h_t) = P_t' P_t = P_{t-1} P_t \quad \text{de (II.4)}$$

por lo que $\lim_{t \rightarrow \infty} Pr(X_1 X_2 | \text{gen } t) = \lim_{t \rightarrow \infty} P_{t-1} P_t = \lim_{t \rightarrow \infty} P_t^2 = P_{\infty}^2$

análogamente, en la generación t :

$$\begin{aligned} Pr(X_1 X_2) &= d_t'(r_t + \frac{1}{2} h_t) + r_t'(d_t + \frac{1}{2} h_t) \\ &= P_t' q_t + q_t' P_t = P_{t-1} q_t + q_{t-1} P_t \end{aligned}$$

por lo que $\lim_{t \rightarrow \infty} Pr(X_1 X_2 | \text{gen } t) = \lim_{t \rightarrow \infty} (P_{t-1} q_t + q_{t-1} P_t) = 2P_{\infty} q_{\infty}$

etc.

Entonces, en el caso de genes ligados al sexo, la condición de equilibrio H-W se expresa de modo similar al caso autosómico; esto es, como la distribución cuyos términos corresponden a los correspondientes del desarrollo del binomio $(P_{\infty} + q_{\infty})^2$, pero esto restringiendo el análisis a la parte femenina de la población; al incluir a la subpoblación masculina, la condición de equilibrio exige que los genotipos masculinos tengan una probabilidad igual a la de los genotipos haploides de la población femenina; esto es, P_{∞} y q_{∞} , respectivamente.

Para una población que no se halla inicialmente en equilibrio, la tendencia a éste es muy rápida, pues como puede verse de (II.5), podemos poner: $P_t - P_{t-1} = (-\frac{1}{2})^t (P - P')$,

esto es, la diferencia entre las proporciones génicas haploides (para la población femenina o la masculina, indistintamente) disminuye como $1/2 |t$ siendo t el número de generaciones transcurrido.

CAPÍTULO III

EL MODELO MATRICIAL DE LI: ^(*) DOS LOCI, DOS
ALELOS POR LOCUS.

(*) Véase en el Apéndice A un esbozo biográfico del genetista C.C. Li.

Pensemos ahora en dos loci independientes, esto es, que se localizan en pares cromosómicos (autosómicos) diferentes. Al pensar en los gametos o formas haploides, tenemos entonces dos cromosomas diferentes que nos interesan; llamémosles cromosoma (o locus) α y cromosoma (o locus) β . Si suponemos que en el locus α hay dos formas alélicas posibles - digamos A_1 y A_2 - y que en el locus β también hay dos posibles alelos - digamos B_1 y B_2 -, tendremos 9 (3') posibles genotipos diploides:

$A_1 A_1 B_1 B_1$	$A_1 A_1 B_1 B_2$	$A_1 A_1 B_2 B_2$
$A_1 A_2 B_1 B_1$	$A_1 A_2 B_1 B_2$	$A_1 A_2 B_2 B_2$
$A_2 A_2 B_1 B_1$	$A_2 A_2 B_1 B_2$	$A_2 A_2 B_2 B_2$

Supongamos una población infinita en la cual, inicialmente, tales genotipos se hallen distribuidos de acuerdo a la siguiente distribución de probabilidad (esto es, eligiendo al azar un individuo, éste tendrá el genotipo dado con la probabilidad mencionada):

$$\begin{array}{lll}
 \Pr(A_1 A_1 B_1 B_1) \equiv P_{11} & \Pr(A_1 A_1 B_1 B_2) \equiv P_{12} & \Pr(A_1 A_1 B_2 B_2) \equiv P_{13} \\
 \Pr(A_1 A_2 B_1 B_1) \equiv P_{21} & \Pr(A_1 A_2 B_1 B_2) \equiv P_{22} & \Pr(A_1 A_2 B_2 B_2) \equiv P_{23} \\
 \Pr(A_2 A_2 B_1 B_1) \equiv P_{31} & \Pr(A_2 A_2 B_1 B_2) \equiv P_{32} & \Pr(A_2 A_2 B_2 B_2) \equiv P_{33}
 \end{array}$$

y podemos considerar que, para toda generación $t \in \mathbb{N} \cup \{0\}$, se tiene la matriz

$$P(t) \equiv \begin{bmatrix} P_{11}(t) & P_{12}(t) & P_{13}(t) \\ P_{21}(t) & P_{22}(t) & P_{23}(t) \\ P_{31}(t) & P_{32}(t) & P_{33}(t) \end{bmatrix}$$

$$\text{con } \sum_{i=1}^3 \sum_{j=1}^3 P_{ij}(t) = 1 \quad \dots \text{ (III.1)}$$

siendo $P_{ij}(t)$ la correspondiente probabilidad del genotipo diploide (i, j) - según el esquema dado arriba - en la generación t .

Podemos definir a las distribuciones marginales para el locus α :

$$\begin{aligned} P_r(A_1 A_1 | \text{gen } t) &= P_r(A_1 A_1 B_1 B_1 \circ A_1 A_1 B_1 B_2 \circ A_1 A_1 B_2 B_2 | \text{gen } t) \\ &= P_r(A_1 A_1 B_1 B_1 | \text{gen } t) + P_r(A_1 A_1 B_1 B_2 | \text{gen } t) + P_r(A_1 A_1 B_2 B_2 | \text{gen } t) \\ &= P_{11}(t) + P_{12}(t) + P_{13}(t) \equiv d_\alpha(t), \end{aligned}$$

$$\begin{aligned} P_r(A_1 A_2 | \text{gen } t) &= P_r(A_1 A_2 B_1 B_1 \circ A_1 A_2 B_1 B_2 \circ A_1 A_2 B_2 B_2 | \text{gen } t) \\ &= P_r(A_1 A_2 B_1 B_1 | \text{gen } t) + P_r(A_1 A_2 B_1 B_2 | \text{gen } t) + P_r(A_1 A_2 B_2 B_2 | \text{gen } t) \end{aligned}$$

$$= P_{21}(t) + P_{22}(t) + P_{23}(t) \equiv h_{\alpha}(t)$$

$$\begin{aligned} P_r(A_2 A_2 | g_{\alpha t}) &= P_r(A_2 A_2 B_1 B_1 \cdot A_2 A_2 B_1 B_1 \cdot A_2 A_2 B_2 B_2 | g_{\alpha t}) \\ &= P_r(A_2 A_2 B_1 B_1 | g_{\alpha t}) + P_r(A_1 A_1 B_1 B_2 | g_{\alpha t}) + P_r(A_2 A_2 B_2 B_2 | g_{\alpha t}) \\ &= P_{31}(t) + P_{32}(t) + P_{33}(t) \equiv r_{\alpha}(t) \end{aligned}$$

esto es:

$$d_{\alpha}(t) \equiv P_r(A_1 A_1 | g_{\alpha t}) = \sum_{j=1}^3 P_{1j}(t)$$

$$h_{\alpha}(t) \equiv P_r(A_1 A_2 | g_{\alpha t}) = \sum_{j=1}^3 P_{2j}(t)$$

$$r_{\alpha}(t) \equiv P_r(A_2 A_2 | g_{\alpha t}) = \sum_{j=2}^3 P_{3j}(t)$$

con $d_{\alpha}(t) + h_{\alpha}(t) + r_{\alpha}(t) = 1$... (III.2)

análogamente, para el locus β :

$$d_{\beta}(t) \equiv P_r(B_1 B_1 | g_{\alpha t}) = \sum_{i=1}^3 P_{i1}(t)$$

$$h_{\beta}(t) \equiv P_r(B_1 B_2 | g_{\alpha t}) = \sum_{i=1}^3 P_{i2}(t)$$

$$r_{\beta}(t) \equiv P_r(B_2 B_2 | g_{\alpha t}) = \sum_{i=2}^3 P_{i3}(t)$$

... (III.2')

$$\text{con } d_p(t) + h_p(t) + r_p(t) = 1 \quad (t \in \mathbb{N} \cup \{0\})$$

Ahora bien, para esta población se tiene que los individuos que la integran pueden producir 4 (2^2) tipos de gametos:

$$\begin{array}{cc} A_1 B_1 & A_1 B_2 \\ A_2 B_1 & A_2 B_2 \end{array}$$

los cuales, de nuevo como elementos abstractos del acervo genético, podemos suponer distribuidos, en la generación t , de acuerdo a:

$$P_r(A_1 B_1 | \text{gen}t) \equiv P_{11}(t) \quad P_r(A_1 B_2 | \text{gen}t) \equiv P_{12}(t)$$

$$P_r(A_2 B_1 | \text{gen}t) \equiv P_{21}(t) \quad P_r(A_2 B_2 | \text{gen}t) \equiv P_{22}(t)$$

y considerar a la matriz

$$P(t) \equiv \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix}$$

$$\text{con } \sum_{i=1}^2 \sum_{j=1}^2 P_{ij}(t) = 1 \quad (t \in \mathbb{N} \cup \{0\})$$

... (III.3)

También tenemos las distribuciones marginales haploides para el locus α :

$$P_{\alpha}(t) \equiv P_r(A_i | \text{gamet}) = P_{i1}(t) + P_{i2}(t)$$

$$q_{\alpha}(t) \equiv P_r(A_2 | \text{gamet}) = P_{21}(t) + P_{22}(t)$$

$$\text{con } P_{\alpha}(t) + q_{\alpha}(t) = 1 \quad \dots (\text{III} \cdot 4)$$

y para el locus β :

$$P_{\beta}(t) \equiv P_r(B_i | \text{gamet}) = P_{i1}(t) + P_{i2}(t)$$

$$q_{\beta}(t) \equiv P_r(B_2 | \text{gamet}) = P_{21}(t) + P_{22}(t)$$

$$\text{con } P_{\beta}(t) + q_{\beta}(t) = 1 \quad \dots (\text{III} \cdot 4')$$

Ahora, estos gametos serán producidos de acuerdo con las leyes de la gametogénesis mendeliana y en particular de acuerdo con la segunda ley de Mendel acerca de la segregación independiente de los caracteres pues, por hipótesis, hemos supuesto que los dos cromosomas en cuestión son diferentes, lo que es lo mismo que decir que los loci son independientes. Entonces cada genotipo diploide producirá determinado tipo de gametos y lo hará en una determinada proporción. Por ejemplo, el genotipo $A_1A_1B_1B_1$ produce todos sus gametos del tipo A_1B_1 , el genotipo $A_1A_2B_1B_1$ produce la mitad de sus gametos del tipo A_1B_1 y la otra mitad del tipo A_2B_1 , mientras que el genotipo $A_1A_2B_1B_2$ produce los cuatro tipos de gametos (A_1B_1 , A_1B_2 , A_2B_1 y A_2B_2),

todos en la misma proporción, esto es, en la proporción $1/4$.

Todo esto nos lleva a las siguientes relaciones genéticas:

$$P_{11}(t) = P_{11}(t-1) + \frac{1}{2} P_{12}(t-1) + \frac{1}{2} P_{21}(t-1) + \frac{1}{4} P_{22}(t-1)$$

$$P_{12}(t) = P_{12}(t-1) + \frac{1}{2} P_{11}(t-1) + \frac{1}{2} P_{22}(t-1) + \frac{1}{4} P_{21}(t-1)$$

$$P_{21}(t) = P_{21}(t-1) + \frac{1}{2} P_{11}(t-1) + \frac{1}{2} P_{22}(t-1) + \frac{1}{4} P_{12}(t-1)$$

$$P_{22}(t) = P_{22}(t-1) + \frac{1}{2} P_{12}(t-1) + \frac{1}{2} P_{21}(t-1) + \frac{1}{4} P_{11}(t-1)$$

$$t \in \mathbb{N} \cup \{0\}$$

... (III.5)

Por otro lado, bajo el supuesto de que la próxima generación nacera de los procesos fecundativos realizados por las parejas formadas al azar de entre los individuos de la población, y teniendo en cuenta que esto es lo mismo que suponer el apareamiento al azar de los gametos, y suponiendo por último que las generaciones están separadas, esto es, que los individuos de la siguiente generación (generación $t+1$) sólo son los individuos engendrados en la forma ya descrita por sus padres de la generación anterior (generación t), tendremos las siguientes relaciones de fecundación:

$$P_{11}(t+1) = P_{11}^2(t)$$

$$P_{12}(t+1) = 2 P_{11}(t) P_{12}(t)$$

$$P_{13}(t+1) = P_{12}^2(t)$$

$$P_{21}(t+1) = 2 P_{11}(t) P_{21}(t)$$

$$P_{22}(t+1) = 2(P_{11}(t)P_{21}(t) + P_{11}(t)P_{22}(t))$$

$$P_{23}(t+1) = 2 P_{12}(t) P_{21}(t)$$

$$P_{31}(t+1) = P_{21}^2(t)$$

$$P_{32}(t+1) = 2 P_{21}(t) P_{22}(t)$$

$$P_{33}(t+1) = P_{22}^2(t)$$

... (III. 6)

Esto es así ya que, por ejemplo, el genotipo $A_1 A_1 B_1 B_1$ sólo puede engendrarse mediante el proceso $A_1 B_1 \rightarrow A_1 B_1$; el genotipo $A_1 A_2 B_1 B_1$ puede engendrarse de dos maneras: $A_1 B_1 \rightarrow A_2 B_1$ y $A_2 B_1 \rightarrow A_1 B_1$ (ambas con la misma probabilidad asociada); mientras que el genotipo $A_1 A_2 B_1 B_2$ puede ser engendrado de cuatro maneras diferentes, a saber: $A_1 B_1 \rightarrow A_2 B_2$, $A_2 B_2 \rightarrow A_1 B_1$, $A_1 B_2 \rightarrow A_2 B_1$ y $A_2 B_1 \rightarrow A_1 B_2$ (las dos primeras con igual probabilidad y las dos últimas asimismo).

Bien, para la nueva generación (generación $t+1$) se repiten los mismos procesos gametogénicos cuya probabilidad se describe en (III.5), por lo que:

$$P_{11}(t+1) = P_{11}^2(t+1) + \frac{1}{2} P_{12}(t+1) + \frac{1}{2} P_{21}(t+1) + \frac{1}{4} P_{22}(t+1)$$

$$= P_{11}^*(t) + P_{11}(t)P_{11}(t) + P_{11}(t)P_{11}(t) + \frac{1}{2} (P_{11}(t)P_{22}(t) + P_{22}(t)P_{11}(t))$$

(usando (III.6))

$$= P_{11}(t) (P_{11}(t) + P_{11}(t) + P_{11}(t) + P_{22}(t))$$

$$- \frac{1}{2} (P_{11}(t)P_{22}(t) - P_{22}(t)P_{11}(t))$$

$$= P_{11}(t) - \frac{1}{2} d(t)$$

siendo
$$d(t) \equiv \begin{vmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{vmatrix} = \det P(t)$$

... (III.7)

Procediendo en forma análoga con los otros términos, puede mostrarse que:

$$P_{11}(t+1) = P_{11}(t) - \frac{1}{2} d(t)$$

$$P_{12}(t+1) = P_{12}(t) + \frac{1}{2} d(t)$$

$$P_{21}(t+1) = P_{21}(t) + \frac{1}{2} d(t)$$

$$P_{22}(t+1) = P_{22}(t) - \frac{1}{2} d(t)$$

... (III.8)

Hay que dejar claras dos cosas: primero, para la generación inicial - en cuyas probabilidades no ponemos un signo particular para indicarnos - se tiene:

$$P_{11} + P_{12} + P_{21} + P_{22} = 1$$

por lo que

$$P_{11}(1) = P_{11} (P_{11} + P_{12} + P_{21} + P_{22}) - \frac{1}{2} d$$

$$= P_{11} - \frac{1}{2} d \quad \text{con} \quad d \equiv \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}$$

asimismo:

$$P_{12}(1) = P_{12} + \frac{1}{2} d$$

$$P_{21}(1) = P_{21} + \frac{1}{2} d$$

$$P_{22}(1) = P_{22} - \frac{1}{2} d$$

por lo que

$$P_{11}(1) + P_{12}(1) + P_{21}(1) + P_{22}(1) = 1$$

justificando así la cuarta igualdad en la deducción de $P_{11}(t+1)$ en el caso $t=1$; por tanto, en segundo lugar, y mediante un sencillo razonamiento inductivo usando (III.8) en el caso $t=1$, tendremos $P_{11}(2) + P_{12}(2) + P_{21}(2) + P_{22}(2) = 1$, por lo que inductivamente podemos dar por ciertas las relaciones (III.8) en el caso general.

Observemos ahora que

$$d(1) = \begin{pmatrix} P_{11}(1) & P_{12}(1) \\ P_{21}(1) & P_{22}(1) \end{pmatrix}$$

$$= \begin{vmatrix} P_{11} - \frac{1}{2}d & P_{12} + \frac{1}{2}d \\ P_{21} + \frac{1}{2}d & P_{22} - \frac{1}{2}d \end{vmatrix}$$

de (III.8)

$$= \begin{vmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{vmatrix} + \begin{vmatrix} P_{11} & \frac{1}{2}d \\ P_{12} & -\frac{1}{2}d \end{vmatrix} + \begin{vmatrix} -\frac{1}{2}d & P_{12} \\ \frac{1}{2}d & P_{22} \end{vmatrix} + \begin{vmatrix} -\frac{1}{2}d & \frac{1}{2}d \\ \frac{1}{2}d & -\frac{1}{2}d \end{vmatrix}$$

$$= d - \frac{1}{2}d (P_{11} + P_{12} + P_{21} + P_{22}) = d - \frac{1}{2}d = \frac{1}{2}d$$

por lo que, en general: $d(t+1) = \frac{1}{2}d(t)$, lo que recursivamente nos lleva a

$$d(t) = \frac{1}{2^t} d \quad t \in \mathbb{N}$$

... (III.9)

entonces, en (III.8) podemos poner:

$$P_{11}(t+1) = P_{11}(t) - \frac{1}{2}d(t) = P_{11}(t) - \frac{1}{2}2^{-t}d$$

$$= P_{11}(t-1) - \frac{1}{2}2^{-t}d - \frac{1}{2}2^{-t}d$$

etc.

$$= P_{11} - d \sum_{x=1}^{t+1} \left(\frac{1}{2}2^x\right) = P_{11} - d(1 - \frac{1}{2}2^{t+1})$$

y análogamente para las otras probabilidades; esto es:

$$P_{11}(t) = P_{11} - d(1 - \frac{1}{2}t)$$

$$P_{12}(t) = P_{12} + d(1 - \frac{1}{2}t)$$

$$P_{21}(t) = P_{21} + d(1 - \frac{1}{2}t)$$

$$P_{22}(t) = P_{22} - d(1 - \frac{1}{2}t)$$

... (III.10)

por lo que en el límite, cuando $t \rightarrow \infty$:

$$P_{11}^{\infty} = P_{11} - d$$

$$P_{12}^{\infty} = P_{12} + d$$

$$P_{21}^{\infty} = P_{21} + d$$

$$P_{22}^{\infty} = P_{22} - d$$

... (III.11)

teniéndose claramente que $d_{\infty} = 0$

Además - y esto será muy importante - :

$$P_{11}^{\infty} = P_{11} - d = P_{11} - (P_{11}P_{22} - P_{12}P_{21})$$

$$= P_{11}(1 - P_{22}) + P_{12}P_{21}$$

$$= P_{11}(P_{11} + P_{12} + P_{21}) + P_{12}P_{21}$$

$$= P_{11}P_{11} + P_{11}P_{12} + P_{11}P_{21} + P_{12}P_{21}$$

$$= (P_{11} + P_{12})(P_{11} + P_{21})$$

$$= P_{11}P_{11}$$

de (III.4, 4')

estos, de manera análoga para las otras probabilidades, se tiene:

$$P_{11}^{\infty} = P_{\alpha} P_{\beta}$$

$$P_{12}^{\infty} = P_{\alpha} q_{\beta}$$

$$P_{21}^{\infty} = q_{\alpha} P_{\beta}$$

$$P_{22}^{\infty} = q_{\alpha} q_{\beta}$$

... (III.12)

Así, las proporciones de equilibrio son los productos de las distribuciones marginales iniciales, las cuales permanecen constantes, como veremos adelante. Pero antes, de (III.12) y las relaciones de fecundación (III.6), escribamos la matriz zigótica P^{∞} en equilibrio:

$$P^{\infty} = \begin{bmatrix} P_{11}^{\infty} & P_{12}^{\infty} & P_{13}^{\infty} \\ P_{21}^{\infty} & P_{22}^{\infty} & P_{23}^{\infty} \\ P_{31}^{\infty} & P_{32}^{\infty} & P_{33}^{\infty} \end{bmatrix}$$

$$= \begin{bmatrix} P_{\alpha}^2 P_{\beta}^2 & 2 P_{\alpha}^2 P_{\beta} q_{\beta} & P_{\alpha}^2 q_{\beta}^2 \\ 2 P_{\alpha} q_{\alpha} P_{\beta}^2 & 4 P_{\alpha} q_{\alpha} P_{\beta} q_{\beta} & 2 P_{\alpha} q_{\alpha} q_{\beta}^2 \\ q_{\alpha}^2 P_{\beta}^2 & 2 q_{\alpha}^2 P_{\beta} q_{\beta} & q_{\alpha}^2 q_{\beta}^2 \end{bmatrix}$$

... (III.13)

donde podemos observar lo siguiente: dado cualquier $t \in \mathbb{N}$, la matriz de distribución de probabilidad genotípica diploide, de (III.4 y 4') se escribe así:

$$P(t+1) = \begin{matrix} P_{11}(t+1) & P_{12}(t+1) & P_{13}(t+1) \\ P_{21}(t+1) & P_{22}(t+1) & P_{23}(t+1) \\ P_{31}(t+1) & P_{32}(t+1) & P_{33}(t+1) \end{matrix}$$

$$= \begin{matrix} P_{11}^2(t) & 2 P_{11}(t) P_{12}(t) & P_{12}^2(t) \\ 2 P_{11}(t) P_{21}(t) & 2 (P_{11}(t) P_{22}(t) + P_{12}(t) P_{21}(t)) & 2 P_{12}(t) P_{23}(t) \\ P_{21}^2(t) & 2 P_{21}(t) P_{22}(t) & P_{22}^2(t) \end{matrix}$$

donde las distribuciones marginales para la generación $t+1$ resultan:

$$\begin{aligned} d_{\alpha}(t+1) &= P_{11}(t+1) + P_{12}(t+1) + P_{13}(t+1) \\ &= P_{11}^2(t) + 2 P_{11}(t) P_{12}(t) + P_{12}^2(t) \\ &= (P_{11}(t) + P_{12}(t))^2 = P_{\alpha}^2(t) && \text{de (III.4)} \\ &= (P_{11} - d(1-1/2^t) + P_{12} + d(1-1/2^t))^2 && \text{de (III.10)} \\ &= (P_{11} + P_{12})^2 = P_{\alpha}^2 && \text{de (III.4)} \\ &= P_{11}^2 + 2 P_{11} P_{12} + P_{12}^2 \\ &= P_{11}(1) + P_{12}(1) + P_{13}(1) && \text{de (II.6)} \\ &= d_{\alpha}(1) \end{aligned}$$

y análogamente se demuestra que $h_{\alpha}(t+1) = h_{\alpha}(1) = 2 P_{\alpha} q_{\alpha}$

$$r_{\alpha}(t+1) = r_{\alpha}(1) = q_{\alpha}^2$$

estates, en el locus α , y atendiendo sólo al locus α , la población se halla en equilibrio desde la generación 1. Lo mismo sucede para el locus β , pues se demuestra de manera análoga que

$$\begin{aligned}d_{\beta}(t+1) &= d_{\beta}(1) = p_{\beta}^2 \\h_{\beta}(t+1) &= h_{\beta}(1) = 2p_{\beta}q_{\beta} \\r_{\beta}(t+1) &= r_{\beta}(1) = q_{\beta}^2\end{aligned}$$

Todo esto significa que, si atendemos a uno solo de los loci, α o β , la población se halla en equilibrio. Hay equilibrio H-W en cada locus individual, cosa que era de esperarse pues ya habíamos demostrado esto en el capítulo I. Sin embargo, atendiendo a ambos loci simultáneamente, es decir, si vemos a los genotipos individuales en cuanto a la estructura fina de ambos loci, puede verse que éstos tienen diferentes probabilidades entre una generación y otra, a menos que la cantidad d , definida por (III.7), sea cero. Esto puede verse de nuevo a partir de la matriz de distribución de probabilidad genotípica diploide.

Por ejemplo, para el elemento correspondiente a $Pr(A_1A_1B_1B_1)$ (el elemento del renglón 1, columna 1) se tiene, para la generación t :

$$\begin{aligned}P_{11}(t) &= P_{11}^2(t-1) = \left(P_{11} - d\left(1 - \frac{1}{2^{t-1}}\right)\right)^2 \\&= P_{11}^2 - 2P_{11}d\left(1 - \frac{1}{2^{t-1}}\right) + d^2\left(1 - \frac{1}{2^{t-1}}\right)^2\end{aligned}$$

mientras que para la generación $t+1$:

$$P_{11}(t+1) = P_{11}^2 - 2P_{11}d(1 - \frac{1}{2}t) + d^2(1 - \frac{1}{2}t)^2$$

es lo es: $P_{11}(t+1) \neq P_{11}(t)$ omeuz, que $d = 0$.

Por tanto, podemos decir que el determinante definido por (III.7) es una medida de la desviación con respecto al equilibrio en ambos loci en que se halla una población en un momento dado. Además, para una población que esté inicialmente fuera de equilibrio ($d \neq 0$), pero para la cual puede suponerse que las condiciones H-W se cumplen, se tendrá que la aproximación al equilibrio, con el paso de las generaciones - esto es, la tendencia al equilibrio - será como $\frac{1}{2}t$.

Dadas las condiciones expuestas en todas las generaciones, se tiene que las probabilidades genotípicas diploides se acercan a las de equilibrio tan rápido como la sucesión $\frac{1}{2}t$ converge a cero; dichas probabilidades de equilibrio son iguales a los productos de las correspondientes probabilidades marginales para cada locus, las cuales, desde la primera generación, se conservan constantes, resultado que está ya contemplado en el planteamiento original para un locus. Esto implica, por último, que las proporciones genotípicas diploides en equilibrio son los productos de los correspondientes términos de los binomios $(P_a + q_a)^2$ y $(P_b + q_b)^2$ formados por las distribuciones marginales haploides iniciales.

CAPÍTULO IV

DOS LOCI CON UN NÚMERO ARBITRARIO DE ALELOS

DOS LOCI, TRES ALELOS.

Comenzaremos tratando primero el caso particular de tres alelos por locus. Llamemos α y β , como antes, a los dos loci y llamemos A_1, A_2, A_3 y B_1, B_2, B_3 a los respectivos tres alelos diferentes para cada locus. Se tienen los siguientes 9 ($= 3^2$) gametos posibles:

$$\begin{array}{lll} A_1 B_1 & A_1 B_2 & A_1 B_3 \\ A_2 B_1 & A_2 B_2 & A_2 B_3 \\ A_3 B_1 & A_3 B_2 & A_3 B_3 \end{array}$$

con la siguiente distribución de probabilidad para cada generación $t \in \mathbb{N} \cup \{0\}$ (a la distribución inicial le pondremos $t=0$ o bien omitiremos incluso esta y cualquier otra indicación), a la que llamamos matriz de distribución genética (o haploide) de probabilidad:

$$P(t) \equiv \begin{bmatrix} P_{11}(t) & P_{12}(t) & P_{13}(t) \\ P_{21}(t) & P_{22}(t) & P_{23}(t) \\ P_{31}(t) & P_{32}(t) & P_{33}(t) \end{bmatrix}$$

$$\text{con } \sum_{i=1}^3 \sum_{j=1}^3 P_{ij}(t) = 1$$

... (IX.1)

$$\text{siendo } P_{ij}(t) \equiv P_r(A_i B_j | \text{gen } t)$$

$$i, j \in \{1, 2, 3\}$$

La condición de normalidad $\sum_{i=1}^3 \sum_{j=1}^3 P_{ij}(t) = 1$ está sujeta a ser demostrada su compatibilidad con las ecuaciones de gametogénesis y de fecundación que presentaremos adelante. Solo podemos decir, por el momento, que tal condición se supone para la generación inicial ($t=0$). El mismo comentario vale para la distribución genotípica diploide que inmediatamente presentamos.

Así entonces, para cada locus se tienen 6 ($= \frac{3(3+1)}{2}$) genotipos diploides posibles (por ejemplo, en α : $A_1A_1, A_1A_2, A_1A_3, A_2A_1, A_2A_3, A_3A_3$), así que se tienen 36 genotipos diploides en cuanto a los dos loci se refiere:

$A_1A_1B_1B_1$	$A_1A_1B_1B_2$	$A_1A_1B_1B_3$	$A_1A_1B_2B_1$	$A_1A_1B_2B_2$	$A_1A_1B_2B_3$	$A_1A_1B_3B_3$
$A_1A_2B_1B_1$	$A_1A_2B_1B_2$	$A_1A_2B_1B_3$	$A_1A_2B_2B_1$	$A_1A_2B_2B_2$	$A_1A_2B_2B_3$	$A_1A_2B_3B_3$
$A_1A_3B_1B_1$	$A_1A_3B_1B_2$	$A_1A_3B_1B_3$	$A_1A_3B_2B_1$	$A_1A_3B_2B_2$	$A_1A_3B_2B_3$	$A_1A_3B_3B_3$
$A_2A_2B_1B_1$	$A_2A_2B_1B_2$	$A_2A_2B_1B_3$	$A_2A_2B_2B_1$	$A_2A_2B_2B_2$	$A_2A_2B_2B_3$	$A_2A_2B_3B_3$
$A_2A_3B_1B_1$	$A_2A_3B_1B_2$	$A_2A_3B_1B_3$	$A_2A_3B_2B_1$	$A_2A_3B_2B_2$	$A_2A_3B_2B_3$	$A_2A_3B_3B_3$
$A_3A_3B_1B_1$	$A_3A_3B_1B_2$	$A_3A_3B_1B_3$	$A_3A_3B_2B_1$	$A_3A_3B_2B_2$	$A_3A_3B_2B_3$	$A_3A_3B_3B_3$

cuyas probabilidades - para todo t - pueden arreglarse de acuerdo al mismo patrón en la matriz de distribución genotípica diploide:

$$P(t) \equiv \begin{bmatrix} P_{11}(t) & P_{12}(t) & P_{13}(t) & P_{14}(t) & P_{15}(t) & P_{16}(t) \\ P_{21}(t) & P_{22}(t) & P_{23}(t) & P_{24}(t) & P_{25}(t) & P_{26}(t) \\ P_{31}(t) & P_{32}(t) & P_{33}(t) & P_{34}(t) & P_{35}(t) & P_{36}(t) \\ P_{41}(t) & P_{42}(t) & P_{43}(t) & P_{44}(t) & P_{45}(t) & P_{46}(t) \\ P_{51}(t) & P_{52}(t) & P_{53}(t) & P_{54}(t) & P_{55}(t) & P_{56}(t) \\ P_{61}(t) & P_{62}(t) & P_{63}(t) & P_{64}(t) & P_{65}(t) & P_{66}(t) \end{bmatrix}$$

$$\text{con } \sum_{i=1}^6 \sum_{j=1}^6 P_{ij}(t) = 1 \quad \dots \text{ (IV. 2)}$$

La hipótesis de independencia de los loci y las leyes de Mendel nos llevan a las siguientes relaciones gametogénicas:

$$P_{11}(t) = P_{11}(t) + \frac{1}{2} (P_{12}(t) + P_{13}(t) + P_{14}(t) + P_{15}(t)) + \frac{1}{4} (P_{16}(t) + P_{21}(t) + P_{22}(t) + P_{23}(t))$$

$$P_{12}(t) = P_{14}(t) + \frac{1}{2} (P_{12}(t) + P_{15}(t) + P_{14}(t) + P_{34}(t)) + \frac{1}{4} (P_{11}(t) + P_{25}(t) + P_{31}(t) + P_{35}(t))$$

$$P_{13}(t) = P_{16}(t) + \frac{1}{2} (P_{13}(t) + P_{15}(t) + P_{26}(t) + P_{36}(t)) + \frac{1}{4} (P_{13}(t) + P_{25}(t) + P_{33}(t) + P_{35}(t))$$

$$P_{21}(t) = P_{21}(t) + \frac{1}{2} (P_{22}(t) + P_{27}(t) + P_{43}(t) + P_{51}(t)) + \frac{1}{4} (P_{11}(t) + P_{25}(t) + P_{31}(t) + P_{35}(t))$$

$$P_{22}(t) = P_{44}(t) + \frac{1}{2} (P_{14}(t) + P_{41}(t) + P_{45}(t) + P_{54}(t)) + \frac{1}{4} (P_{11}(t) + P_{25}(t) + P_{31}(t) + P_{35}(t))$$

$$P_{23}(t) = P_{46}(t) + \frac{1}{2} (P_{22}(t) + P_{27}(t) + P_{45}(t) + P_{56}(t)) + \frac{1}{4} (P_{13}(t) + P_{25}(t) + P_{33}(t) + P_{35}(t))$$

$$P_{31}(t) = P_{21}(t) + \frac{1}{2} (P_{21}(t) + P_{31}(t) + P_{62}(t) + P_{63}(t)) + \frac{1}{4} (P_{21}(t) + P_{33}(t) + P_{37}(t) + P_{51}(t))$$

$$P_{32}(t) = P_{64}(t) + \frac{1}{2} (P_{34}(t) + P_{54}(t) + P_{64}(t) + P_{67}(t)) + \frac{1}{4} (P_{22}(t) + P_{35}(t) + P_{51}(t) + P_{55}(t))$$

$$P_{33}(t) = P_{66}(t) + \frac{1}{2} (P_{36}(t) + P_{66}(t) + P_{63}(t) + P_{65}(t)) + \frac{1}{4} (P_{33}(t) + P_{35}(t) + P_{53}(t) + P_{55}(t))$$

... (IV. 3)

La hipótesis H-W de apareamientos al azar la traducimos en el encuentro azaroso de los gametos, lo que nos lleva a las siguientes relaciones de fecundación:

$$P_{11}(t+1) = P_{11}^2(t)$$

$$P_{21}(t+1) = 2 P_{11}(t) P_{21}(t)$$

$$P_{31}(t+1) = 2 P_{11}(t) P_{31}(t)$$

$$P_{41}(t+1) = P_{21}^2(t)$$

$$P_{51}(t+1) = 2 P_{21}(t) P_{31}(t)$$

$$P_{61}(t+1) = P_{31}^2(t)$$

$$P_{12}(t+1) = 2 P_{11}(t) P_{12}(t)$$

$$P_{22}(t+1) = 2 (P_{11}(t) P_{22}(t) + P_{12}(t) P_{21}(t))$$

$$P_{32}(t+1) = 2 (P_{11}(t) P_{32}(t) + P_{12}(t) P_{31}(t))$$

$$P_{42}(t+1) = 2 P_{21}(t) P_{12}(t)$$

$$P_{52}(t+1) = 2 (P_{11}(t) P_{32}(t) + P_{22}(t) P_{31}(t))$$

$$P_{62}(t+1) = 2 P_{31}(t) P_{32}(t)$$

$$P_{13}(t+1) = 2 P_{11}(t) P_{13}(t)$$

$$P_{23}(t+1) = 2 (P_{11}(t) P_{23}(t) + P_{12}(t) P_{21}(t))$$

$$P_{33}(t+1) = 2 (P_{33}(t) P_{11}(t) + P_{13}(t) P_{31}(t))$$

$$P_{43}(t+1) = 2 P_{21}(t) P_{23}(t)$$

$$P_{53}(t+1) = 2 (P_{21}(t) P_{33}(t) + P_{23}(t) P_{31}(t))$$

$$P_{63}(t+1) = 2 P_{31}(t) P_{33}(t)$$

$$P_{14}(t+1) = P_{12}^2(t)$$

$$P_{24}(t+1) = 2 P_{12}(t) P_{23}(t)$$

$$P_{34}(t+1) = 2 P_{12}(t) P_{32}(t)$$

$$P_{44}(t+1) = P_{22}^2(t)$$

$$P_{54}(t+1) = 2 P_{22}(t) P_{32}(t)$$

$$P_{64}(t+1) = P_{32}^2(t)$$

... (IX. 4)

$$P_{15}(t+1) = 2 P_{12}(t) P_{13}(t)$$

$$P_{25}(t+1) = 2 (P_{12}(t) P_{23}(t) + P_{13}(t) P_{21}(t))$$

$$P_{35}(t+1) = 2 (P_{13}(t) P_{32}(t) + P_{12}(t) P_{33}(t))$$

$$P_{45}(t+1) = 2 P_{22}(t) P_{23}(t)$$

$$P_{55}(t+1) = 2 (P_{22}(t) P_{33}(t) + P_{32}(t) P_{23}(t))$$

$$P_{65}(t+1) = 2 P_{32}(t) P_{33}(t)$$

$$P_{16}(t+1) = P_{13}^2(t)$$

$$P_{26}(t+1) = 2 P_{13}(t) P_{23}(t)$$

$$P_{36}(t+1) = 2 P_{13}(t) P_{33}(t)$$

$$P_{46}(t+1) = P_{23}^2(t)$$

$$P_{56}(t+1) = 2 P_{23}(t) P_{33}(t)$$

$$P_{66}(t+1) = P_{33}^2(t)$$

Escribiendo ahora las relaciones (IV.3) correspondientes a la generación $t=1$ y las proporciones $P_{rs}(1)$ ($r, s \in [1,3]$) de acuerdo con (IV.4), puede demostrarse (ver Apéndice C) que

$$\begin{aligned} P_{11}(1) &= P_{11} - \frac{1}{2} \left(\begin{vmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{13} \\ P_{21} & P_{23} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{12} \\ P_{31} & P_{32} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{13} \\ P_{31} & P_{33} \end{vmatrix} \right) \\ &= P_{11} - \frac{1}{2} \left(\begin{vmatrix} P_{11} & P_{11} \\ P_{11} & P_{11} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{12} \\ P_{11} & P_{12} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{13} \\ P_{11} & P_{13} \end{vmatrix} + \right. \\ &\quad + \begin{vmatrix} P_{11} & P_{11} \\ P_{21} & P_{21} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{12} \\ P_{21} & P_{21} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{13} \\ P_{21} & P_{23} \end{vmatrix} + \\ &\quad \left. + \begin{vmatrix} P_{11} & P_{11} \\ P_{31} & P_{31} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{12} \\ P_{31} & P_{32} \end{vmatrix} + \begin{vmatrix} P_{11} & P_{13} \\ P_{31} & P_{33} \end{vmatrix} \right) \end{aligned}$$

y análogamente para las otras proporciones; esto es, que

$$P_{ik}(1) = P_{ik} - \frac{1}{2} \sum_{j=1}^3 \sum_{l=1}^3 \begin{vmatrix} P_{ik} & P_{il} \\ P_{jk} & P_{jl} \end{vmatrix} \quad \dots \text{(IV.5)}$$

$i, k \in [1,3]$

en general, para la generación $t+1$:

$$P_{ik}(t+1) = P_{ik}(t) - \frac{1}{2} \sum_{j=1}^3 \sum_{l=1}^3 \begin{vmatrix} P_{ik}(t) & P_{il}(t) \\ P_{jk}(t) & P_{jl}(t) \end{vmatrix} \quad \dots \text{(IV.5')}$$

$i, k \in [1,3]$

Por lo que, si la población se hallase en un principio en equilibrio, debería tenerse que para $i, k \in [1,3]$, las sumas de los determinantes

$$\sum_{j=1}^3 \sum_{k=1}^3 \begin{vmatrix} P_{jk} & P_{kj} \\ P_{jk} & P_{jk} \end{vmatrix}$$

deberían ser iguales a cero.

Puede mostrarse que, de modo análogo al caso de dos alelos, dichos determinantes de 2×2 - y por tanto las sumas - tienden a cero conforme crece el número de generaciones, t , aumenta. Sin embargo, haremos la deducción de las proporciones en equilibrio de otro modo - por ser un camino más corto - que también nos dará las proporciones genotípicas haploides para toda generación como función de las distribuciones iniciales, el tiempo t en generaciones y las distribuciones marginales iniciales.

Así pues, definimos a tales distribuciones marginales, para el locus α , como:

$$P_{\alpha i}(t) \equiv P_r(A_i | \text{gen } t) = \sum_{j=1}^3 P_{ij}(t) \quad \dots \text{(IV.6)}$$

$i \in \{1, 2, 3\}$

y para el locus β :

$$P_{\beta k}(t) \equiv P_r(B_k | \text{gen } t) = \sum_{l=1}^3 P_{lk}(t)$$

$k \in \{1, 2, 3\}$

teniéndose que, para la generación inicial:

$$\sum_{i=1}^3 P_{\alpha i} = 1 \quad , \quad \sum_{k=1}^3 P_{\beta k} = 1 \quad \text{como consecuencia de}$$

la condición (III.1) en $t = 0$.

PRIMER RESULTADO BÁSICO : LAS FRECUENCIAS EN LA GENERACIÓN 1

Para la generación 1 se tiene, de (III.5) :

$$\begin{aligned}
 P_{ikh}^{(1)} &= P_{ikh} - \frac{1}{2} \left(\sum_{j=1}^3 \sum_{l=1}^3 (P_{ikh} P_{jkl} - P_{jkh} P_{ilk}) \right) \\
 &= P_{ikh} - \frac{1}{2} \left(\sum_{j=1}^3 \sum_{l=1}^3 P_{ikh} P_{jkl} - \sum_{j=1}^3 \sum_{l=1}^3 P_{jkh} P_{ilk} \right) \\
 &= P_{ikh} - \frac{1}{2} P_{ikh} \sum_{j=1}^3 \sum_{l=1}^3 P_{jkl} + \frac{1}{2} \sum_{j=1}^3 P_{jkh} \sum_{l=1}^3 P_{ilk} \\
 &= P_{ikh} - \frac{1}{2} P_{ikh} + \frac{1}{2} \sum_{j=1}^3 P_{jkh} P_{kil} \quad \text{de (III.1 y 6)} \\
 &= \frac{1}{2} P_{ikh} + \frac{1}{2} P_{jkh} P_{kil} \quad \text{de nuevo por (IV.6)} \\
 &\dots \text{ (IV.7)}
 \end{aligned}$$

y esto $\forall i, k \in [1, 3]$. Esto es, la probabilidad de cada genotipo haploide en la primera generación está compuesta por dos términos: la mitad de la probabilidad correspondiente en la generación inicial - el valor que tenía antes - más la mitad de la probabilidad que tendría - y que tendrá en equilibrio - si hubiera independencia estocástica en la distribución estudiada. Este es el resultado crucial prácticamente de todo el trabajo. Lo que sigue en esta sección será estudiar el comportamiento general entre las generaciones t y $t+1$ para poder conocer el comportamiento de las frecuencias sólo como función del tiempo y las frecuencias iniciales. Sin em-

bergo, antes tenemos que analizar dos cosas: la compatibilidad de la condición de normalidad - y su verificación - para todo t , y el comportamiento en el tiempo de las distribuciones marginales.

Pues bien, usando (IV.7) tenemos, para la generación 1:

$$\begin{aligned}
 \sum_{i=1}^3 \sum_{k=1}^3 P_{ik}(1) &= \sum_{i=1}^3 \sum_{k=1}^3 \left(\frac{1}{2} P_{ik} + \frac{1}{2} P_{\beta k} P_{\alpha i} \right) \\
 &= \frac{1}{2} \sum_{i=1}^3 \sum_{k=1}^3 P_{ik} + \frac{1}{2} \sum_{i=1}^3 \sum_{k=1}^3 P_{\beta k} P_{\alpha i} \\
 &= \frac{1}{2} + \frac{1}{2} \sum_{i=1}^3 P_{\alpha i} \sum_{k=1}^3 P_{\beta k} \\
 &= \frac{1}{2} + \frac{1}{2} \left(\sum_{i=1}^3 P_{\alpha i} \right) \left(\sum_{k=1}^3 P_{\beta k} \right) \\
 &= \frac{1}{2} + \frac{1}{2} = 1 \quad \dots \text{(IV.8)}
 \end{aligned}$$

Donde hemos usado la condición de normalidad en la generación inicial y el correspondiente modo de expresarla mediante los marginales.

Tenemos entonces mostrado que la condición de normalidad se verifica en la generación 1 de manera compatible con la dinámica implícita en las relaciones (IV.3 y 4).

Vemos entonces asegurada otro resultado parcial, para la generación 1, que nos será muy fácil extrapolar a todo t y que será de tremenda importancia.

INVARIANCIA DE LOS MARGINALES

La distribución marginal para el locus α en la generación t , para $i \in \{1, 3\}$, podemos escribirla como:

$$\begin{aligned} P_{\alpha i}(t) &= \sum_{k=1}^3 P_{i k}(t) = \sum_{k=1}^3 \left(\frac{1}{2} P_{i k} + \frac{1}{2} P_{\alpha i} P_{\beta k} \right) \\ &= \frac{1}{2} \sum_{k=1}^3 P_{i k} + \frac{1}{2} P_{\alpha i} \sum_{k=1}^3 P_{\beta k} \\ &= \frac{1}{2} P_{\alpha i} + \frac{1}{2} P_{\alpha i} = P_{\alpha i} \end{aligned}$$

... (IV.9)

análogamente para el locus β : $P_{\beta h}(t) = P_{\beta h}$

esto es, las distribuciones marginales se mantienen, en la generación t , iguales a las correspondientes de la generación inicial.

EL RESULTADO BÁSICO GENERAL

Tenemos ahora los elementos para establecer la siguiente

Proposición IV. 10. - Para todo $t \in \mathbb{N}$ se cumple:

$$a) \sum_{i=1}^3 \sum_{k=1}^3 P_{i k}(t) = 1, \quad b) P_{\alpha i}(t) = P_{\alpha i}, \quad P_{\beta k}(t) = P_{\beta k}$$

$i, k \in \{1, 3\}$

$$c) P_{i k}(t+1) = \frac{1}{2} P_{i k}(t) + \frac{1}{2} P_{\alpha i} P_{\beta k}$$

La proposición (IV.10) se demuestra por inducción, en una versión más general, en el Apéndice C, siguiendo esencialmente los mismos pasos que en (IV.5, 8 y 9). Lo importante, por supuesto, es el resultado c) de (IV.10):

$$P_{ik}(t+1) = \frac{1}{2} P_{ik}(t) + \frac{1}{2} P_{\alpha i} P_{\beta k} \quad i, k \in \{1, 2\} \\ \forall t \in \mathbb{N}$$

ecuación que se resuelve en:

$$P_{ik}(t) = \frac{1}{2^t} P_{ik} + \frac{2^t - 1}{2^t} P_{\alpha i} P_{\beta k} \quad i, k \in \{1, 2\}$$

por lo que en el límite $t \rightarrow \infty$:

$$P_{ik}^{\infty} = P_{\alpha i} P_{\beta k} \quad i, k \in \{1, 2\} \quad \dots \text{(IV.11)}$$

que corresponde a la distribución genotípica haploide en equilibrio.

Esto es, las proporciones haploides en equilibrio corresponden a los productos de las distribuciones marginales iniciales para cada locus.

La matriz de la distribución haploide en equilibrio queda:

$$P^{\infty} = \begin{bmatrix} P_{\alpha 1} P_{\beta 1} & P_{\alpha 1} P_{\beta 2} & P_{\alpha 1} P_{\beta 3} \\ P_{\alpha 2} P_{\beta 1} & P_{\alpha 2} P_{\beta 2} & P_{\alpha 2} P_{\beta 3} \\ P_{\alpha 3} P_{\beta 1} & P_{\alpha 3} P_{\beta 2} & P_{\alpha 3} P_{\beta 3} \end{bmatrix}$$

y la matriz de distribución diploide en equilibrio :

$$P^0 = \begin{bmatrix} P_{A_1}^1 P_{A_1}^1 & 2 P_{A_1}^1 P_{A_1}^1 P_{A_2} & 2 P_{A_1}^1 P_{A_1}^1 P_{A_3} & P_{A_1}^1 P_{A_2}^1 & 2 P_{A_1}^1 P_{A_2}^1 P_{A_3} & P_{A_1}^1 P_{A_3}^1 \\ 2 P_{A_1}^1 P_{A_2}^1 P_{A_1}^1 & 4 P_{A_1}^1 P_{A_2}^1 P_{A_1}^1 P_{A_2} & 4 P_{A_1}^1 P_{A_2}^1 P_{A_1}^1 P_{A_3} & 2 P_{A_1}^1 P_{A_2}^1 P_{A_2}^1 & 4 P_{A_1}^1 P_{A_2}^1 P_{A_2}^1 P_{A_3} & 2 P_{A_1}^1 P_{A_2}^1 P_{A_3}^1 \\ 2 P_{A_1}^1 P_{A_3}^1 P_{A_1}^1 & 4 P_{A_1}^1 P_{A_3}^1 P_{A_1}^1 P_{A_2} & 4 P_{A_1}^1 P_{A_3}^1 P_{A_1}^1 P_{A_3} & 2 P_{A_1}^1 P_{A_3}^1 P_{A_2}^1 & 4 P_{A_1}^1 P_{A_3}^1 P_{A_2}^1 P_{A_3} & 2 P_{A_1}^1 P_{A_3}^1 P_{A_3}^1 \\ P_{A_2}^1 P_{A_1}^1 & 2 P_{A_2}^1 P_{A_1}^1 P_{A_2} & 2 P_{A_2}^1 P_{A_1}^1 P_{A_3} & P_{A_2}^1 P_{A_2}^1 & 2 P_{A_2}^1 P_{A_2}^1 P_{A_3} & P_{A_2}^1 P_{A_3}^1 \\ 2 P_{A_2}^1 P_{A_3}^1 P_{A_1}^1 & 4 P_{A_2}^1 P_{A_3}^1 P_{A_1}^1 P_{A_2} & 4 P_{A_2}^1 P_{A_3}^1 P_{A_1}^1 P_{A_3} & 2 P_{A_2}^1 P_{A_3}^1 P_{A_2}^1 & 4 P_{A_2}^1 P_{A_3}^1 P_{A_2}^1 P_{A_3} & 2 P_{A_2}^1 P_{A_3}^1 P_{A_3}^1 \\ P_{A_3}^1 P_{A_1}^1 & 2 P_{A_3}^1 P_{A_1}^1 P_{A_2} & 2 P_{A_3}^1 P_{A_1}^1 P_{A_3} & P_{A_3}^1 P_{A_2}^1 & 2 P_{A_3}^1 P_{A_2}^1 P_{A_3} & P_{A_3}^1 P_{A_3}^1 \end{bmatrix}$$

Queda por resolver la cuestión de cómo identificar, en el caso general, la forma específica que debe tomar cada elemento de la matriz P como uno de los términos del desarrollo de

$$(P_{A_1} + P_{A_2} + P_{A_3})^2 (P_{A_1} + P_{A_2} + P_{A_3})^2 ;$$

por tanto, haremos ahora una descripción del método empleado para identificar, dado un genotipo diploide, a los pares de gametos que

puedan intervenir en su génesis; así como, dado un gameto específico, identificar a los genotipos diploides que pueden producirlo en la gametogénesis. Consecuentemente, haremos ver dicho método de identificación en el caso general de varios alelos - el mismo número - en cada locus.

VARIOS ALELOS : EL MÉTODO

Tenemos de nuevo dos loci: α y β , y tenemos ahora un número entero positivo mayor que 1, digamos n , alelos en cada locus; podemos considerar al conjunto de gametos G como el producto cartesiano $G_\alpha \times G_\beta$, siendo

$$G_\alpha \equiv \{A_i \mid i \in [1, n]\} \quad \text{el conj. de alelos en el locus } \alpha$$

$$\text{y } G_\beta \equiv \{B_k \mid k \in [1, n]\} \quad \text{el conj. de alelos en el locus } \beta$$

y ordenarlo - análogamente a como ya se ha ilustrado - como un arreglo matricial

$$\begin{array}{cccc} A_1 B_1 & A_1 B_2 & \dots & A_1 B_n \\ A_2 B_1 & A_2 B_2 & \dots & A_2 B_n \\ \vdots & & & \\ A_n B_1 & A_n B_2 & \dots & A_n B_n \end{array}$$

y al cual asociamos - en la generación t - la distribución de probabilidad dada por la matriz

$$P(t) \equiv \begin{bmatrix} P_{11}(t) & P_{12}(t) & \dots & P_{1n}(t) \\ P_{21}(t) & P_{22}(t) & \dots & P_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1}(t) & P_{n2}(t) & \dots & P_{nn}(t) \end{bmatrix} \dots (\text{IV.12})$$

siendo $P_{ik}(t) \equiv \text{Pr}(A_i B_k | \text{gen } t)$ y teniéndose, por supuesto:

$\sum_{i=1}^n \sum_{k=1}^n P_{ik}(t) = 1$, condición que habremos de mostrar compatible con las relaciones de gametogénesis y fecundación y que podemos pedir para la generación inicial, omitiendo indicar en ese caso el valor $t=0$.

Pensemos ahora por el momento en un solo locus, digamos el locus α , y pensemos en los posibles procesos fecundativos; puesto que el proceso $A_i \rightarrow A_j$ es distinto de $A_j \rightarrow A_i$ a menos que, y solo en ese caso: $i=j$, en tal caso el conjunto de procesos fecundativos - en cuanto al locus α se refiere - puede expresarse mediante el producto cartesiano $G_{\alpha}^2 = G_{\alpha} \times G_{\alpha}$. Sin embargo, cuando pensamos en los posibles genotipos diploides que pueden ser engendrados por estos procesos fecundativos, aquéllos son tales

que se tiene $A_i A_j = A_j A_i$, esto es, los pares $A_i A_j$ y $A_j A_i$ son indistinguibles, por lo que la manera de construir el conjunto de genotipos diploides puede ser ésta:

$$Z_n \equiv \left\{ A_i A_j \mid i \in [1, n], j \in [i, n] \right\}, \text{ con el posible orde-}$$

namiento esencialmente igual al ya visto en el Capítulo I :

$$\begin{array}{ccccccc} A_1 A_1 & A_1 A_2 & A_1 A_3 & \dots & A_1 A_n & & \\ & A_2 A_2 & A_2 A_3 & \dots & A_2 A_n & & \\ & & A_3 A_3 & \dots & A_3 A_n & & \\ & & & & \vdots & & \\ & & & & A_{n-1} A_{n-1} & A_{n-1} A_n & \\ & & & & & A_n A_n & \end{array}$$

que, como puede verse, consta de $\frac{n(n+1)}{2}$ elementos. Para el locus β todo lo hasta aquí dicho es análogo.

Ahora, como queremos tener en cuenta a los dos loci simultáneamente, tenemos que considerar a todas las posibles parejas de la forma

$A_i A_j B_k B_l$. Es claro que dicha pareja debe considerarse igual a la pareja $B_k B_l A_i A_j$, pues el orden de la escritura es intrascendente, pero hay un orden implícito al poner letras diferentes para los diferentes loci considerados, los cuales no poseen otro orden de modo natural. Entonces la pareja $A_i A_j B_k B_l$ sí es diferente a la pareja $A_k A_l B_i B_j$.

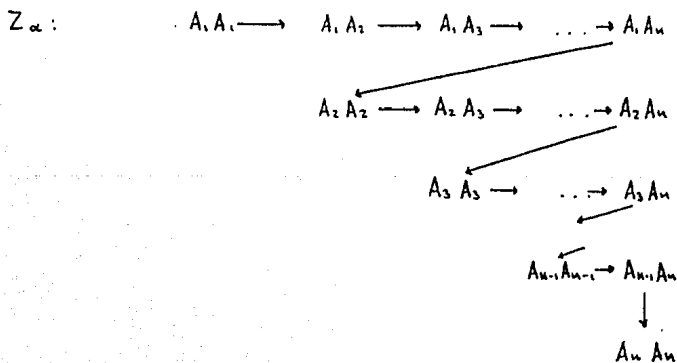
a menos que $k=i$ y $l=j$; esto implica que el conjunto de posibles genotipos diploides, atendiendo a ambos loci, deberá ser equivalente al producto cartesiano $Z \equiv Z_\alpha \times Z_\beta$, es decir, al conjunto

$$Z \equiv \left\{ (A_i A_j, B_k B_l) \mid A_i A_j \in Z_\alpha, B_k B_l \in Z_\beta \right\}$$

que es igual a $\left\{ A_i A_j B_k B_l \mid i, j \in [1, n]; k, l \in [1, n] \right\}$

pues en ambos loci la situación es completamente análoga. Esto implica que Z - el conjunto de zigotos - tiene $\left(\frac{n(n+1)}{2} \right)^2$ elementos.

Para dar una ordenación matricial - lo que ya hemos hecho en el caso ilustrado de $n=3$ - de Z , requerimos de ordenar linealmente a cada uno de los conjuntos Z_α y Z_β . Esto puede hacerse del siguiente modo - ya ilustrado también - :



Esto es, asociando biunívocamente a los conjuntos

$$\left[1, \frac{n(n+1)}{2}\right] = \left\{ r \in \mathbb{N} \mid 1 \leq r \leq \frac{n(n+1)}{2} \right\} \quad \text{con} \quad \left\{ (i, j) \mid i \in [1, n], j \in [i, n] \right\}$$

mediante la siguiente regla:

$$\begin{aligned} \sigma(r) &= (1, r) && \text{si } r \in [1, n] \\ &= (2, 2+r-n-1) && \text{si } r \in (n, n+(n-1)] \\ &\vdots \\ &= (3, 3+r-(n+(n-1))-1) && \text{si } r \in (n+(n-1), n+(n-1)+(n-2)] \\ &\vdots \\ &\text{etc.} \end{aligned}$$

que haremos explícita mediante las siguientes consideraciones:

considerando a la función
$$p(i) \equiv \sum_{x=0}^i (n-x) = \frac{(i+1)(2n-i)}{2}$$

para $i \in [0, n]$ y con $p(-1) \equiv 0 \quad \dots \text{(IV.13)}$

(Véase Apéndice C) tenemos que los intervalos siguientes son
ajenos por construcción y su unión es el intervalo $I \equiv (0, \frac{n(n+1)}{2}]$:

$$\begin{array}{ll}
 I_0 \equiv (0, n] & \text{con } \# I_0 = n \\
 I_1 \equiv (n, n+(n-1)] & \text{con } \# I_1 = n-1 \\
 I_2 \equiv (n+(n-1), n+(n-1)+(n-2)] & \text{con } \# I_2 = n-2 \\
 \vdots & \\
 I_i \equiv (\rho(i-1), \rho(i)] & \text{con } \# I_i = n-i \\
 \vdots & i \in [0, n] \\
 I_{n-1} \equiv (\rho(n-2), \rho(n-1)] = \left\{ \frac{n(n+1)}{2} \right\} & \text{con } \# I_{n-1} = 1 \\
 I_n \equiv (\rho(n-1), \rho(n)] = \emptyset & \text{con } \# I_n = 0
 \end{array}$$

Entonces, puesto que dado $r \in [1, \frac{n(n+1)}{2}]$ existe un único $i \in [1, n]$ tal que $r \in I_{i-1}$, podemos hacer

$$\sigma : [1, \frac{n(n+1)}{2}] \rightarrow \{(i, j) \mid i \in [1, n], j \in [i, n]\}$$

como $\sigma_1(r) = i$ si y sólo si $r \in I_{i-1}$

$$y \quad \sigma_2(r) = r - \frac{(i-1)(2n-i)}{2} = j$$

o sea: $\sigma(r) = \left(i, r - \frac{(i-1)(2n-i)}{2} \right)$ si $r \in I_{i-1}$
 ... (IV.15)

Se demuestra en el Apéndice C que σ tiene las propiedades de "ordenamiento" para las parejas (i, j) consideradas, así como de biyectividad; se tiene que

dato $(k, \ell) \in \{(i, j) \mid i \in [1, n], j \in [i, n]\}$:

$$\sigma^{-1}(k, \ell) = \ell + \frac{(k-1)(2n-k)}{2} \quad \dots \text{(IV.16)}$$

Todo esto nos permite hacer el siguiente arreglo matricial para el conjunto de genotipos diploides Z :

$A_1 A_1 B_1 B_1$	$A_1 A_1 B_1 B_2$	$A_1 A_1 B_1 B_3$	\dots	$A_1 A_1 B_n B_n$
$A_1 A_2 B_1 B_1$	$A_1 A_2 B_1 B_2$	$A_1 A_2 B_1 B_3$	\dots	$A_1 A_2 B_n B_n$
$A_1 A_3 B_1 B_1$	$A_1 A_3 B_1 B_2$	$A_1 A_3 B_1 B_3$	\dots	$A_1 A_3 B_n B_n$
\vdots				
$A_n A_n B_1 B_1$	$A_n A_n B_1 B_2$	$A_n A_n B_1 B_3$	\dots	$A_n A_n B_n B_n$

y considerar a la matriz $P(t)$ de orden $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$
 dado $t \in \mathbb{N} \cup \{0\}$ con

$$P_{rs}(t) \equiv \Pr(A_{\sigma_1(t)} A_{\sigma_2(t)} B_{\sigma_3(t)} B_{\sigma_4(t)} \mid \text{gen } t)$$

$$\text{para } r, s \in [1, \frac{n(n+1)}{2}] \quad \dots \text{(IV.17)}$$

Ahora bien, dado el gameto $A_i B_k$ (con $i, k \in [1, n]$) procederemos a encontrar al conjunto de zigotos que pueden dar génesis a dicho gameto (dicho conjunto es igual al conjunto de zigotos en cuya fecundación puede intervenir $A_i B_k$), así como a la probabili-

dad con que se da dicha génesis (gametogénesis).

Recordando a la función $\rho: [-1, n] \rightarrow [0, \frac{n(n+1)}{2}]$

y a los conjuntos I_i , $i \in [0, n]$ definidos por (IV.13) y (IV.14), respectivamente, puede demostrarse (ver Apéndice C) que:

- A_i está en forma doble (homocigosis en el locus α) en el i -ésimo renglón cuyo número es $\rho(i-2)+1$, así que hacemos a la función $\delta: [1, n] \rightarrow [1, \frac{n(n+1)}{2}]$ como $\delta(i) \equiv \rho(i-2)+1 \dots$ (IV.18)

esto es, tendremos homocigosis para el alelo A_i en el locus α en el i -ésimo elemento del conjunto $D_i \equiv \{ \delta(i) \}$ $i \in [1, n] \dots$ (IV.19)

$$\left(D_1 = \{1\}, D_2 = \{n+1\}, \dots, D_n = \left\{ \frac{n(n+1)}{2} \right\} \right)$$

- A_i está en forma sencilla (heterocigosis en α) a la izquierda en todos los renglones correspondientes al conjunto

$$L_i \equiv I_{i-1} \setminus D_i \quad i \in [1, n] \quad \dots \text{ (IV.20)}$$

$$\left(\text{teniéndose } \#L_i = n-i \quad \forall i \in [1, n] \right)$$

- A_i estará en forma sencillo a la derecha en el conjunto

$$R_i \equiv \left\{ \rho(x) + i - (x+1) \mid x \in [-1, i-3] \right\} \quad i \in \{1, n\} \dots$$

$$\dots \text{(III. 21)}$$

(teniéndose $R_1 = \emptyset$, $R_2 = \{2\}$, \dots , $R_n = \{n, n(n-1), \dots, \frac{n(n+1)}{2} - 1\}$)
 con $\# R_i = i-1$)

Podemos resumir diciendo que A_i está doble en D_i , sencillo a la izquierda en L_i y sencillo a la derecha en R_i ; análogamente, B_k es doble en D_k , sencillo a la izquierda en L_k y sencillo a la derecha en R_k . Por tanto, para el arreglo matricial diploide se tiene:

- $A_i B_k$ está en forma homociga en ambos loci en $D_i \times D_k$
- $A_i B_k$ está en forma homociga en α , heterociga en β , en $D_i \times (L_k \cup R_k)$
- $A_i B_k$ está en forma heterociga en α , homociga en β , en $(L_i \cup R_i) \times D_k$
- $A_i B_k$ está en forma heterociga en ambos loci en $(L_i \cup R_i) \times (L_k \cup R_k)$

Por todo esto, podemos afirmar que, en este caso, las hipótesis de la herencia mendeliana nos llevan a las siguientes relaciones genotípicas:

$$P_{ik}(t) = P_{D_i D_k}(t) + \frac{1}{2} \left(\sum_{\alpha \in L_i \cup R_i} P_{D_i \alpha}(t) + \sum_{\beta \in L_i \cup R_i} P_{\beta D_k}(t) \right)$$

$$+ \frac{1}{4} \left(\sum_{\alpha \in L_i \cup R_i} \sum_{\beta \in L_k \cup R_k} P_{\alpha \beta}(t) \right)$$

$$\dots \text{(IV. 22)}$$

Consideremos ahora al elemento $(r, s) \in [1, \frac{n(n+1)}{2}]^2$:

- en el locus α hay homocigosis ssi $\sigma_1(r) = \sigma_2(r)$, y se puede ver con facilidad que en ese caso $r = \delta(\sigma_1(r))$, o sea que $r \in D_{\sigma_1(r)}$ - y lo análogo para el locus β -
- en el locus α hay heterocigosis ssi $\sigma_1(r) < \sigma_2(r)$, y en este caso $\delta(\sigma_1(r)) < r$, $r \in L_{\sigma_1(r)}$ y $r \in R_{\sigma_1(r)}$ - y lo análogo para el locus β -

Por tanto, para el zigoto que ocupa el lugar (r, s) pueden tenerse las posibles situaciones siguientes:

- en (r, s) hay homocigosis en ambos loci: entonces $\sigma_1(r) = \sigma_2(r)$, $\sigma_1(s) = \sigma_2(s)$, $r \in D_{\sigma_1(r)}$, $s \in D_{\sigma_1(s)} \Rightarrow (r, s) \in D_{\sigma_1(r)} \times D_{\sigma_1(s)}$ y por ello, ante la hipótesis H-W deberá tenerse: $P_{r,s}(t+1) = P_{\sigma_1(r)\sigma_1(s)}^2(t)$ (1)
 - en (r, s) hay homocigosis en α y heterocigosis en β : entonces $\sigma_1(r) = \sigma_2(r)$ y $\sigma_1(s) < \sigma_2(s)$; o sea: $(r, s) \in D_{\sigma_1(r)} \times (L_{\sigma_1(s)} \cup R_{\sigma_1(s)})$
 - si -análogamente- (r, s) tiene heterocigosis en α y homocigosis en β , $\sigma_1(r) < \sigma_2(r)$ y $\sigma_1(s) = \sigma_2(s)$; por tanto: $(r, s) \in (L_{\sigma_1(r)} \cup R_{\sigma_1(r)}) \times D_{\sigma_1(s)}$
- en el caso anterior: $P_{r,s}(t+1) = 2 P_{\sigma_1(r)\sigma_1(s)}^{(1)} P_{\sigma_1(r)\sigma_1(s)}^{(1)}$
- en el caso presente: $P_{r,s}(t+1) = 2 P_{\sigma_1(r)\sigma_1(s)}^{(1)} P_{\sigma_2(r)\sigma_1(s)}^{(1)}$
- si - por último - en (r, s) se tiene heterocigosis en ambos loci:

$\sigma_1(r) < \sigma_2(r)$, $\sigma_1(s) < \sigma_2(s)$ $\Rightarrow (r,s) \in (L_{\sigma_1(1)} \cup R_{\sigma_1(1)}) \times (L_{\sigma_1(2)} \cup R_{\sigma_1(2)})$, por lo cual

$$P_{rs}(t+1) = 2 \left(P_{\sigma_1(1)\sigma_1(2)}^{(t)} P_{\sigma_1(1)\sigma_1(2)}^{(t+1)} + P_{\sigma_1(1)\sigma_2(2)}^{(t)} P_{\sigma_2(1)\sigma_1(2)}^{(t+1)} \right)$$

o sea, tenemos las siguientes relaciones de fecundación :

$$P_{rs}(t+1) = \begin{cases} P_{\sigma_1(1)\sigma_1(2)}^{(t)} & \text{si } \sigma_1(1) = \sigma_2(1) \text{ y } \sigma_1(2) = \sigma_2(2) \\ 2 P_{\sigma_1(1)\sigma_1(2)}^{(t)} P_{\sigma_1(1)\sigma_1(2)}^{(t+1)} & \text{si } \sigma_1(1) = \sigma_2(1) \text{ y } \sigma_1(2) < \sigma_2(2) \\ 2 P_{\sigma_1(1)\sigma_1(2)}^{(t)} P_{\sigma_2(1)\sigma_1(2)}^{(t+1)} & \text{si } \sigma_1(1) < \sigma_2(1) \text{ y } \sigma_1(2) = \sigma_2(2) \\ 2 \left(P_{\sigma_1(1)\sigma_1(2)}^{(t)} P_{\sigma_2(1)\sigma_2(2)}^{(t+1)} + P_{\sigma_1(1)\sigma_1(2)}^{(t)} P_{\sigma_2(1)\sigma_1(2)}^{(t+1)} \right) & \text{si } \sigma_1(1) < \sigma_2(1) \text{ y } \sigma_1(2) < \sigma_2(2) \end{cases}$$

... (IX.23)

Tomemos ahora las relaciones gametogénicas (IX.22) para $t=1$ y sustituycmos los términos correspondientes a las probabilidades genotípicas diploides de acuerdo con las relaciones de fecundación (IX.23) entre la generación 0 - omitiendo la indicación en este caso - y la generación 1 :

Dado entonces cualquier $(i,h) \in [1,4]^2$, tenemos, de (IX.22) para la generación 1 :

$$P_{L_h}(1) = P_{\delta(1)\delta(h)}(1) + \frac{1}{2} \left(\sum_{s \in L_h} P_{\delta(1)s}(1) + \sum_{r \in L_{UR_i}} P_{r\delta(h)}(1) \right) + \frac{1}{4} \left(\sum_{r \in L_{UR_i}} \sum_{s \in L_{UR_h}} P_{rs}(1) \right)$$

Ahora, puesto que $\forall x \in [1,4]$ se tiene $L_x \cap R_x = \emptyset$ (ver Apéndice C), podemos escribir la expresión anterior como:

$$\begin{aligned} &= P_{\delta(1)\delta(h)}(1) + \frac{1}{2} \left(\sum_{s \in L_h} P_{\delta(1)s}(1) + \sum_{s \in R_h} P_{\delta(1)s}(1) \right) \\ &+ \sum_{r \in L_i} P_{r\delta(h)}(1) + \sum_{r \in R_i} P_{r\delta(h)}(1) \\ &+ \frac{1}{4} \left(\sum_{r \in L_i} \sum_{s \in L_h} P_{rs}(1) + \sum_{r \in L_i} \sum_{s \in R_h} P_{rs}(1) + \sum_{r \in R_i} \sum_{s \in L_h} P_{rs}(1) + \sum_{r \in R_i} \sum_{s \in R_h} P_{rs}(1) \right) \\ &= P_{\sigma_1(\delta(1))\sigma_1(\delta(h))}^2 + \frac{1}{2} \left(\sum_{s \in L_h} z(P_{\sigma_1(\delta(1))\sigma_1(s)} P_{\sigma_1(\delta(1))\sigma_2(s)}) \right) \\ &+ \sum_{s \in R_h} z(P_{\sigma_1(\delta(1))\sigma_1(s)} P_{\sigma_1(\delta(1))\sigma_2(s)}) \\ &+ \sum_{r \in L_i} z(P_{\sigma_1(r)\sigma_1(\delta(h))} P_{\sigma_1(r)\sigma_1(\delta(h))}) + \sum_{r \in R_i} z(P_{\sigma_1(r)\sigma_1(\delta(h))} P_{\sigma_1(r)\sigma_2(\delta(h))}) \\ &+ \frac{1}{4} \left(\sum_{r \in L_i} \sum_{s \in L_h} z(P_{\sigma_1(r)\sigma_1(s)} P_{\sigma_1(r)\sigma_2(s)} + P_{\sigma_1(r)\sigma_2(s)} P_{\sigma_2(r)\sigma_1(s)}) \right) \end{aligned}$$

$$\begin{aligned}
 & + \sum_{r \in R_i} \sum_{s \in R_h} 2 (P_{\sigma_i(t), \sigma_i(s)} P_{\sigma_i(t), \sigma_i(s)} + P_{\sigma_i(t), \sigma_i(s)} P_{\sigma_i(t), \sigma_i(s)}) \\
 & + \sum_{r \in R_i} \sum_{s \in L_h} 2 (P_{\sigma_i(t), \sigma_i(s)} P_{\sigma_i(t), \sigma_i(s)} + P_{\sigma_i(t), \sigma_i(s)} P_{\sigma_i(t), \sigma_i(s)}) \\
 & + \sum_{r \in R_i} \sum_{s \in R_h} 2 (P_{\sigma_i(t), \sigma_i(s)} P_{\sigma_i(t), \sigma_i(s)} + P_{\sigma_i(t), \sigma_i(s)} P_{\sigma_i(t), \sigma_i(s)}) \quad)
 \end{aligned}$$

habiendo usado (IV. 23) para $t = 0$.

$$\begin{aligned}
 & = P_{i,h}^2 + \sum_{s \in L_h} P_{i,h} P_{i,\sigma_i(s)} + \sum_{s \in R_h} P_{i,\sigma_i(s)} P_{i,h} \\
 & \quad + \sum_{r \in L_i} P_{i,h} P_{\sigma_i(r),h} + \sum_{r \in R_i} P_{\sigma_i(r),h} P_{i,h} \\
 & + \frac{1}{2} \left(\sum_{r \in L_i} \sum_{s \in L_h} (P_{i,h} P_{\sigma_i(r),\sigma_i(s)} + P_{i,\sigma_i(r)} P_{\sigma_i(r),h}) \right. \\
 & \quad + \sum_{r \in L_i} \sum_{s \in R_h} (P_{i,\sigma_i(s)} P_{\sigma_i(r),h} + P_{i,h} P_{\sigma_i(r),\sigma_i(s)}) \\
 & \quad + \sum_{r \in R_i} \sum_{s \in L_h} (P_{\sigma_i(r),h} P_{i,\sigma_i(s)} + P_{\sigma_i(r),\sigma_i(s)} P_{i,h}) \\
 & \quad \left. + \sum_{r \in R_i} \sum_{s \in R_h} (P_{\sigma_i(r),\sigma_i(s)} P_{i,h} + P_{\sigma_i(r),h} P_{i,\sigma_i(s)}) \right)
 \end{aligned}$$

habiendo hecho uso de las propiedades de σ : $\sigma_i(\sigma(x)) = x$,

$$\sigma_i(y) = x \text{ si } y \in L_x, \quad \sigma_i(y) = x \text{ si } y \in R_x$$

(Apéndice C)

por último, manipulando ciertos términos de la expresión anterior, puede escribirse:

$$\begin{aligned}
 P_{\lambda k}^{(1)} = & P_{\lambda k} \left[P_{\lambda k} + \sum_{s \in L_k} P_{\lambda r_s(s)} + \sum_{s \in R_k} P_{\lambda r_s(s)} \right. \\
 & + \sum_{r \in L_i} P_{r_2(s)k} + \sum_{r \in R_i} P_{r_1(s)k} \\
 & + \sum_{r \in L_i} \sum_{s \in L_k} P_{r_2(s) r_1(s)} + \sum_{r \in L_i} \sum_{s \in R_k} P_{r_2(s) r_1(s)} \\
 & \left. + \sum_{r \in R_i} \sum_{s \in L_k} P_{r_1(s) r_2(s)} + \sum_{r \in R_i} \sum_{s \in R_k} P_{r_1(s) r_2(s)} \right] \\
 & + \frac{1}{2} \left[\sum_{r \in L_i} \sum_{s \in L_k} (P_{\lambda r_s(s)} P_{r_1(s)k} - P_{\lambda k} P_{r_1(s) r_2(s)}) \right. \\
 & + \sum_{r \in L_i} \sum_{s \in R_k} (P_{\lambda r_s(s)} P_{r_2(s)k} - P_{\lambda k} P_{r_2(s) r_1(s)}) \\
 & + \sum_{r \in R_i} \sum_{s \in L_k} (P_{r_1(s)k} P_{\lambda r_s(s)} - P_{r_1(s) r_2(s)} P_{\lambda k}) \\
 & \left. + \sum_{r \in R_i} \sum_{s \in R_k} (P_{r_1(s)k} P_{\lambda r_s(s)} - P_{r_1(s) r_2(s)} P_{\lambda k}) \right]
 \end{aligned}$$

$\lambda, k \in \{1, n\}$

... (IV.24)

Dejemos aquí por un momento esta expresión para $P^{(1)}$ en términos de $P^{(0)}$ y devolvamos la atención hacia la definición (IV.6) y generalicémosla para el caso de n alelos:

Para todo $t \in \mathbb{N} \cup \{0\}$:

$$P_{\alpha i}(t) \equiv P_r(A_i | \text{gen } t) = \sum_{j=1}^n P_{ij}(t) \quad , \quad i \in \{1, \dots, n\}$$

$$P_{\beta k}(t) \equiv P_r(B_k | \text{gen } t) = \sum_{j=1}^n P_{jk}(t) \quad , \quad k \in \{1, \dots, n\}$$

... (IV.25)

las probabilidades marginales para el locus α y para el locus β , respectivamente.

Por otro lado, observemos que si $r \in D_i \Rightarrow \sigma_1(r) = \sigma_2(r) = i$;
 si $r \in L_i \Rightarrow \sigma_1(r) = i < \sigma_2(r) = r - \frac{(i-1)(2n-i)}{2}$
 y $r \in L_i \Rightarrow r \in (\rho(i-2)+1, \rho(i-1)]$ (por (IV.15, 19, 20))
 lo cual asu vez implica, por la forma de σ_2 , que

$$\begin{aligned} \sigma_2(r) &\in \left(\rho(i-2)+1 - \frac{(i-1)(2n-i)}{2}, \rho(i-1) - \frac{(i-1)(2n-i)}{2} \right) \\ &= (i, n] \quad (\text{Ver Apéndice C}) \end{aligned}$$

esto es: $r \in L_i \Rightarrow \sigma_1(r) = i, \quad i < \sigma_2(r) \leq n$

Por último, si $r \in R_i \Rightarrow \sigma_2(r) = i$, por lo que

$$i = r - \frac{(\sigma_1(r)-1)(2n-\sigma_1(r))}{2}, \quad \text{pero también}$$

$\exists x \in [-1, i-3]$ tal que $r = \rho(x) + (i - (x-1))$ por (IV.21)

entonces: $f(i-1) + (i - (i-1) + 1) \leq r \leq f(i-3) + (i - (i-3) + 1)$
 pues f es creciente; pero esta desigualdad es la misma que

$$i \leq r \leq \frac{(i-2)(2n - (i-3))}{2} + 2 = f(i-3) + 2$$

$\Rightarrow \sigma_1(i) \leq \sigma_1(r) \leq \sigma_1(f(i-3) + 2)$ pues σ_1 creciente

$\Rightarrow 1 \leq \sigma_1(r) \leq i-1$ por def. de σ_1

En resumen: $r \in D_i \Rightarrow \sigma_1(r) = \sigma_2(r) = i$

$r \in L_i \Rightarrow \sigma_1(r) = i < \sigma_2(r) \leq n$

$r \in R_i \Rightarrow 1 \leq \sigma_1(r) < i = \sigma_2(r)$

para cualquier $i \in [1, n]$; por tanto se tiene que

$$\{\sigma_1(r) \mid r \in D_i\} \cup \{\sigma_2(r) \mid r \in L_i\} \cup \{\sigma_1(r) \mid r \in R_i\} = [1, n] \quad \dots (IV.26)$$

para cualquier $i \in [1, n]$, dada la cardinalidad de cada uno de los conjuntos involucrados y las desigualdades arriba establecidas.

Por todo lo anterior, se tiene:

$$P_{2h}(t) + \sum_{s \in L_h} P_{i \sigma_2(s)}(t) + \sum_{s \in R_h} P_{i \sigma_1(s)}(t) = \sum_{j=1}^n P_{ij}(t) = P_{oi}(t) \quad \dots (IV.27)$$

también:

$$\sum_{r \in L_i} P_{\sigma_2(r)h}(t) + \sum_{r \in R_i} P_{\sigma_1(r)h}(t) = \sum_{k=1}^n P_{ih}(t) - P_{ih}(t) = P_{ph}(t) - P_{ih}(t) \quad \dots (IV.27')$$

g, enfin :

$$\begin{aligned}
 & \sum_{r \in L_i} \sum_{s \in L_h} P_{\sigma_r(i) \sigma_s(i)}(t) + \sum_{r \in L_i} \sum_{s \in R_h} P_{\sigma_r(i) \sigma_s(i)}(t) \\
 + & \sum_{r \in R_i} \sum_{s \in L_h} P_{\sigma_r(i) \sigma_s(i)}(t) + \sum_{r \in R_i} \sum_{s \in R_h} P_{\sigma_r(i) \sigma_s(i)}(t) = \\
 & = \sum_{r \in L_i} \left(\sum_{s \in L_h} P_{\sigma_r(i) \sigma_s(i)}(t) + \sum_{s \in R_h} P_{\sigma_r(i) \sigma_s(i)}(t) \right) \\
 & + \sum_{r \in R_i} \left(\sum_{s \in L_h} P_{\sigma_r(i) \sigma_s(i)}(t) + \sum_{s \in R_h} P_{\sigma_r(i) \sigma_s(i)}(t) \right) \\
 & = \sum_{r \in L_i} \left(\sum_{l=1}^n P_{\sigma_r(i) l}(t) - P_{\sigma_r(i) h}(t) \right) \\
 & + \sum_{r \in R_i} \left(\sum_{l=1}^n P_{\sigma_r(i) l}(t) - P_{\sigma_r(i) h}(t) \right) \quad \text{dc (IV.26)} \\
 & = \sum_{r \in L_i} \left(P_{\alpha \sigma_r(i)}(t) - P_{\sigma_r(i) h}(t) \right) + \sum_{r \in R_i} \left(P_{\alpha \sigma_r(i)}(t) - P_{\sigma_r(i) h}(t) \right) \\
 & = \sum_{j=1}^n \left(P_{\alpha j}(t) - P_{j h}(t) \right) - \left(P_{\alpha i}(t) - P_{i h}(t) \right)
 \end{aligned}$$

... (IV.27")

Por último, hagamos

$$\begin{aligned}
 -d_{ik}(t) &= \sum_{r \in I_i} \sum_{s \in R_k} (P_{i\sigma_r(s)}(t) P_{\sigma_r(s)k}(t) - P_{ik}(t) P_{\sigma_r(s)\sigma_r(s)}(t)) \\
 &+ \sum_{r \in I_i} \sum_{s \in R_k} (P_{i\sigma_r(s)}(t) P_{\sigma_r(s)k}(t) - P_{ik}(t) P_{\sigma_r(s)\sigma_r(s)}(t)) \\
 &+ \sum_{r \in R_i} \sum_{s \in R_k} (P_{\sigma_r(s)k}(t) P_{i\sigma_r(s)}(t) - P_{\sigma_r(s)\sigma_r(s)}(t) P_{ik}(t)) \\
 &+ \sum_{r \in R_i} \sum_{s \in R_k} (P_{\sigma_r(s)k}(t) P_{i\sigma_r(s)}(t) - P_{\sigma_r(s)\sigma_r(s)}(t) P_{ik}(t)) \\
 &= \sum_{r \in I_i} \left(\sum_{s \in R_k} (P_{i\sigma_r(s)}(t) P_{\sigma_r(s)k}(t) - P_{ik}(t) P_{\sigma_r(s)\sigma_r(s)}(t)) \right. \\
 &\quad \left. + \sum_{s \in R_k} (P_{i\sigma_r(s)}(t) P_{\sigma_r(s)k}(t) - P_{ik}(t) P_{\sigma_r(s)\sigma_r(s)}(t)) \right) \\
 &+ \sum_{r \in R_i} \left(\sum_{s \in R_k} (P_{\sigma_r(s)k}(t) P_{i\sigma_r(s)}(t) - P_{\sigma_r(s)\sigma_r(s)}(t) P_{ik}(t)) \right. \\
 &\quad \left. + \sum_{s \in R_k} (P_{\sigma_r(s)k}(t) P_{i\sigma_r(s)}(t) - P_{\sigma_r(s)\sigma_r(s)}(t) P_{ik}(t)) \right) \\
 &= \sum_{r \in I_i} \left(\sum_{k=1}^n (P_{ik}(t) P_{\sigma_r(s)k}(t) - P_{ik}(t) P_{\sigma_r(s)k}(t)) \right. \\
 &\quad \left. - (P_{ik}(t) P_{\sigma_r(s)k}(t) - P_{ik}(t) P_{\sigma_r(s)k}(t)) \right) \\
 &+ \sum_{r \in R_i} \left(\sum_{k=1}^n (P_{\sigma_r(s)k}(t) P_{ik}(t) - P_{\sigma_r(s)k}(t) P_{ik}(t)) \right. \\
 &\quad \left. - (P_{\sigma_r(s)k}(t) P_{ik}(t) - P_{\sigma_r(s)k}(t) P_{ik}(t)) \right)
 \end{aligned}$$

se restan determinantes iguales a cero

$$\begin{aligned}
&= \sum_{r \in R_i} \left(\sum_{k=1}^n (P_{i, r}(t) P_{\sigma_r(i), k}(t) - P_{i, k}(t) P_{\sigma_r(i), r}(t)) \right) \\
&+ \sum_{r \in R_i} \left(\sum_{k=1}^n (P_{\sigma_r(i), k}(t) P_{i, r}(t) - P_{\sigma_r(i), r}(t) P_{i, k}(t)) \right) \\
&= \sum_{j=1}^n \left(\sum_{k=1}^n (P_{i, j}(t) P_{j, k}(t) - P_{i, k}(t) P_{j, j}(t)) \right. \\
&\quad \left. - (P_{i, j}(t) P_{i, k}(t) - P_{i, k}(t) P_{i, j}(t)) \right)
\end{aligned}$$

se restan determinantes cero; por lo cual:

$$-d_{ik}(t) = \sum_{j=1}^n \sum_{l=1}^n (P_{i, j}(t) P_{j, k}(t) - P_{i, k}(t) P_{j, j}(t))$$

obten:

$$\begin{aligned}
d_{ik}(t) &= \sum_{j=1}^n \sum_{l=1}^n (P_{i, k}(t) P_{j, l}(t) - P_{i, l}(t) P_{j, k}(t)) \\
&= \sum_{j=1}^n \sum_{l=1}^n \begin{vmatrix} P_{i, k}(t) & P_{i, l}(t) \\ P_{j, k}(t) & P_{j, l}(t) \end{vmatrix}
\end{aligned}$$

... (IV. 27''')

pero $-d_{ik}(t)$ también puede escribirse así:

$$\begin{aligned}
-d_{ik}(t) &= \sum_{j=1}^n \sum_{l=1}^n P_{i, l}(t) P_{j, k}(t) - \sum_{j=1}^n \sum_{l=1}^n P_{i, k}(t) P_{j, l}(t) \\
&= \sum_{j=1}^n P_{j, k}(t) \sum_{l=1}^n P_{i, l}(t) - P_{i, k}(t) \sum_{j=1}^n \sum_{l=1}^n P_{j, l}(t) \\
&= \left(\sum_{j=1}^n P_{j, k}(t) \right) \left(\sum_{l=1}^n P_{i, l}(t) \right) - P_{i, k}(t) \left(\sum_{j=1}^n \sum_{l=1}^n P_{j, l}(t) \right) \\
&= P_{\beta, k}(t) P_{\alpha, i}(t) - P_{i, k}(t) \left(\sum_{j=1}^n \sum_{l=1}^n P_{j, l}(t) \right)
\end{aligned}$$

... (IV. 27''''')

Colectando ahora los resultados parciales (IV. 27, 27', 27'', 27''', 27'''')) y llevándolos a la relación (IX. 24) entre las proporciones genotípicas haploides de la generación 0 a la generación 1, tenemos:

$$\begin{aligned}
 P_{ik}(1) &= P_{ih} \left[P_{ai} + P_{bh} - P_{ih} + \sum_{j \neq i} (P_{aj} - P_{jh}) - (P_{ai} - P_{ih}) \right] \\
 &\quad + \frac{1}{2} \left[P_{bh} P_{ai} - P_{ih} \sum_{j \neq i} \sum_{k \neq i} P_{jk} \right] \\
 &= P_{ih} \left[P_{ai} + P_{bh} - P_{ih} + \sum_{j \neq i} P_{aj} - \sum_{j \neq i} P_{jh} - P_{ai} + P_{ih} \right] \\
 &\quad + \frac{1}{2} \left[P_{bh} P_{ai} - P_{ih} \right] \quad \begin{array}{l} \text{por la condición de normalidad} \\ \text{inicial:} \end{array} \\
 &= P_{ih} \left[P_{ai} + P_{bh} - P_{ih} + 1 - P_{bh} - P_{ai} + P_{ih} \right] \\
 &\quad + \frac{1}{2} \left[P_{bh} P_{ai} - P_{ih} \right] \\
 &= P_{ih} - \frac{1}{2} P_{ih} + \frac{1}{2} P_{ai} P_{bh} \\
 &= \frac{1}{2} P_{ih} + \frac{1}{2} P_{ai} P_{bh} \quad \dots \text{ (IV. 28) }
 \end{aligned}$$

$$i, h \in \{1, n\}$$

resultado que es esencialmente el mismo que (IV. 7) en el caso $n=3$ alelos.

$$\begin{aligned}
 \text{entonces: } \sum_{i=1}^n \sum_{k=1}^n P_{ik}(1) &= \sum_{i=1}^n \sum_{k=1}^n \left(\frac{1}{2} P_{\alpha i} P_{\beta k} + \frac{1}{2} P_{ik} \right) \\
 &= \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n P_{\alpha i} P_{\beta k} + \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n P_{ik} \\
 &= \frac{1}{2} + \frac{1}{2} = 1 \quad \dots \text{(IV.28')}
 \end{aligned}$$

esto es, la condición de normalidad para la generación t es compatible con la ecuación de transición (III.28). También se tiene:

$$\begin{aligned}
 P_{\alpha i}(1) &= \sum_{\beta=1}^n P_{i\beta}(1) = \sum_{\beta=1}^n \left(\frac{1}{2} P_{\alpha i} P_{\beta\beta} + \frac{1}{2} P_{i\beta} \right) \\
 &= \frac{1}{2} P_{\alpha i} \sum_{\beta=1}^n P_{\beta\beta} + \frac{1}{2} \sum_{\beta=1}^n P_{i\beta} \\
 &= \frac{1}{2} P_{\alpha i} + \frac{1}{2} P_{\alpha i} = P_{\alpha i}
 \end{aligned}$$

análogamente en el torus β , por lo que

$$P_{\alpha i}(1) = P_{\alpha i}, \quad P_{\beta k}(1) = P_{\beta k}, \quad i, k \in \{1, n\} \dots \text{(IV.28'')}$$

Con estos resultados (III.28, 28', 28''), y puesto que la relación (IV.24) se ha propuesto -implícitamente y gracias a las relaciones (III.22 y 23) - para la transición entre las generaciones t y $t+1$, proponemos que los resultados mencionados se cumplen para todo valor del tiempo generacional t . Esto es, tenemos la siguiente proposición (que demostramos por inducción en el Apéndice C):

Proposición (IV.29) :- Para todo $t \in \mathbb{N}$ se cumple:

$$a) \sum_{i=1}^n \sum_{k=1}^n P_{ik}(t) = 1 \quad \begin{array}{l} \text{compatibilidad de la condición} \\ \text{de normalidad} \end{array}$$

$$b) P_{\alpha i}(t) = P_{\alpha i}, \quad P_{\beta k}(t) = P_{\beta k} \quad \text{invariancia de los marginales}$$

$$c) P_{ik}(t+1) = \frac{1}{2} P_{ik}(t) + \frac{1}{2} P_{\alpha i} P_{\beta k} \quad \text{la relación fundamental}$$

Es la ecuación (IV.29 c) , por supuesto, la ecuación importante, y se resuelve inmediatamente como

$$P_{ik}(t) = \frac{1}{2^t} P_{ik} + \frac{2^t - 1}{2^t} P_{\alpha i} P_{\beta k} \quad \dots \text{(IV.30)} \\ i, k \in [1, n], \quad t \in \mathbb{N}$$

por lo que , en el límite cuando $t \rightarrow \infty$:

$$P_{ik}^{\infty} = P_{\alpha i} P_{\beta k} \quad i, k \in [1, n] \quad \dots \text{(IV.31)}$$

que es la forma de la distribución haploide en equilibrio H-W.

De la distribución haploide en equilibrio (IV.31) y las relaciones de fecundación (IV.23) se tiene la distribución diploide en equilibrio:

$$P_{rs}^m = \begin{cases} P_{\sigma_1(r)}^2 P_{\beta\sigma_1(s)}^2 & \text{si } \sigma_1(r) = \sigma_2(r), \sigma_1(s) = \sigma_2(s) \\ 2 P_{\sigma_1(r)}^2 P_{\beta\sigma_1(s)} P_{\beta\sigma_2(s)} & \text{si } \sigma_1(r) = \sigma_2(r), \sigma_1(s) < \sigma_2(s) \\ 2 P_{\sigma_1(r)} P_{\sigma_2(r)} P_{\beta\sigma_1(s)}^2 & \text{si } \sigma_1(r) < \sigma_2(r), \sigma_1(s) = \sigma_2(s) \\ 4 P_{\sigma_1(r)} P_{\sigma_2(r)} P_{\beta\sigma_1(s)} P_{\beta\sigma_2(s)} & \text{si } \sigma_1(r) < \sigma_2(r), \sigma_1(s) < \sigma_2(s) \end{cases}$$

$$r, s \in [1, \frac{n(n+1)}{2}] \quad \dots \text{ (IV.32)}$$

Sólo resta decir aquí que existen dos maneras de medir la diferencia que hay, en un momento dado, entre la distribución genotípica en ese momento y la eventual distribución de equilibrio H-W (IV.31):

primera: de (IV.30), las diferencias

$$P_{ik}(t) - P_{ik}$$

segunda: de (IV.28''') las sumas de los determinantes 2x2

$$d_{ik}(t) = \sum_{j=1}^n \sum_{l=1}^n \begin{vmatrix} P_{jk}(t) & P_{il}(t) \\ P_{jl}(t) & P_{ik}(t) \end{vmatrix}$$

si estos últimos valen cero, la población se hallará en equilibrio H-W para ambos loci considerados simultáneamente; en caso de ser dis-

tiutos de cero, y si se dan las condiciones de independencia mendeliana y a apareamientos al azar, dicha diferencia se reducirá a la mitad en la generación siguiente y esto se repetirá en las generaciones subsiguientes, por lo que, en términos prácticos, podrá suponerse que se ha alcanzado el equilibrio después de un número suficientemente grande de generaciones.

CAPÍTULO V

CONCLUSIONES Y PERSPECTIVAS

Como hemos visto, se considere a G - el conjunto de gametos - como el producto cartesiano $G_\alpha \times G_\beta$, siendo G_α y G_β los conjuntos de alelos en el locus α y en el locus β , respectivamente, a los cuales hemos identificado con $[1, u]$; por eso, G puede verse como $[1, u]^2$; ordenando el conjunto de pares $\{(i, j) \mid 1 \leq i \leq j \leq u\}$ con el conjunto $[1, \binom{u}{2}]$ mediante la función σ identificamos a Z_α (y a Z_β) con $[1, \binom{u}{2}]$ y al conjunto de zigotos, Z , con $[1, \binom{u}{2}]^2$.

Llevamos por ello a ambos conjuntos - y a sus distribuciones de probabilidad - al arreglo matricial que permite una fácil identificación de los gametos producidos por un determinado zigoto, así como de la proporción en que es producido por dicho zigoto, para construir las relaciones de gametogénesis mendeliana; igualmente sencilla resulta la identificación de todos los posibles zigotos que pueden ser producidos mediante fecundación por un gametocitado de otros posibles gametos, así como la identificación de estos últimos, pudiendo llevar esto a la forma de las relaciones de fecundación. Lo que resta es una manipulación tediosa pero sencilla para llegar a la conclusión de que el estado de equilibrio H-W tiene la distribución de probabilidad consistente en los términos del producto $(P_{\alpha 1} + P_{\alpha 2} + \dots + P_{\alpha u})^2 (P_{\beta 1} + P_{\beta 2} + \dots + P_{\beta u})^2$ siendo $P_{\alpha i}$, $P_{\beta k}$ las probabilidades marginales iniciales para los loci α , β respectivamente, asociadas a los alelos i, k también respectivamente. La propia manipulación muestra que la tendencia al equilibrio es como la convergencia de la sucesión $(\frac{1}{2})^t$.

El esquema más general que se permite bajo las condiciones H-W de este trabajo - población infinita, loci independientes, aleatoriedad de los procesos de fecundación, etc. - es el de considerar al conjunto G de gametos como el producto cartesiano

$$G = \prod_{\alpha \in [1, m]} G_{\alpha}$$

siendo G_{α} el conjunto de alelos en el locus α ($\alpha \in [1, m]$), que puede tener cardinalidad n_{α} .

Z , el conjunto de zigotos, también puede expresarse así:

$$Z = \prod_{\alpha \in [1, m]} Z'_{\alpha}$$

con Z'_{α} definido como el conjunto de pares (i, j) con $1 \leq i \leq j \leq n_{\alpha}$ ordenado por

$$\sigma : Z_{\alpha} = [1, \binom{n_{\alpha}}{2}] \rightarrow Z'_{\alpha}$$

Las distribuciones de probabilidad, haploides y diploides, se definen como antes:

$$P(t) : G \rightarrow [0, 1]_{\mathbb{R}} \quad , \quad \sum_{g \in G} P_g(t) = 1$$

$$P(\cdot) : Z \rightarrow [0, 1]_{\mathbb{R}} \quad \sum_{z \in Z} P_z(\cdot) = 1$$

Para establecer las relaciones gametogénicas H-W y las relaciones de fecundación adecuadas, debemos establecer algunas definiciones y resultados que menciono a continuación:

- Se definen, como antes:

$$D_g^\alpha = \{ \delta(g_{\alpha i}) \} \quad , \quad L_g^\alpha = I_{g_{\alpha i-1}} \setminus D_g^\alpha \quad ,$$

$$R_g^\alpha = \{ \rho(x) + g_{\alpha i} - (x_{i+1}) \mid x \in [-1, g_{\alpha i-1}] \} \quad ; \quad g_{\alpha i} \in [1, u_{\alpha}]$$

para cada locus $\alpha \in [1, m]$, con ρ , σ , δ adecuadas

- Se construye el conjunto

$$K_g^\mu = \bigcup_{\psi \in \Psi_\mu} \left(\prod_{\alpha \in [1, m]} S_g^\alpha(\psi) \right)$$

$$\text{con } S_g^\alpha(\psi) = \begin{cases} D_g^\alpha & \alpha \in [1, m] \setminus \psi([1, \mu]) \\ L_g^\alpha \cup R_g^\alpha & \alpha \in \psi([1, \mu]) \end{cases}$$

siendo $\Psi_\mu = \{ \psi : [1, \mu] \rightarrow [1, m] \mid \psi \text{ creciente} \}$

con $\mu \in [0, m]$

Se demuestra que K_g^m es el conjunto de zigotos heterocigos en m loci, uno de cuyos gametos producidos es g (o en cuya concepción g puede intervenir).

- Las relaciones de gametogénesis serían de la forma

$$P_g(t) = \sum_{\mu=0}^m \frac{1}{2} \mu \left(\sum_{\tau \in \mathcal{P}_\mu} \left(\sum_{z \in \mathbb{T}S_g^+(t)} P_z(t) \right) \right)$$

Ahora, hacemos
$$h_z = \left\{ \alpha \in [1, m] \mid \sigma_1(z(\alpha)) < \sigma_2(z(\alpha)) \right\}$$

y
$$H_z = \left\{ \alpha \in [1, m] \mid \sigma_1(z(\alpha)) = \sigma_2(z(\alpha)) \right\}$$

$$z \in Z$$

los conjuntos de loci donde el zigoto z es heterocigo y homocigo, respectivamente.

Hacemos
$$\eta_\mu = \left\{ z \in Z \mid \#h_z = \mu \right\}$$

el conjunto de zigotos heterocigos en m loci.

y
$$T_m = \left\{ \tau: [1, m] \rightarrow \{1, 2\} \right\} = \{1, 2\}^m$$

el conjunto de todas las funciones de $[1, m]$ a $\{1, 2\}$.

- se demuestra que $\forall z \in Z \quad \exists! \mu \in [0, m], \exists! \gamma \in \mathcal{T}_\mu$
 tal que $\mathcal{T}([1, \mu]) = h_z$ y podemos construir a
 $g(z, \gamma) \in G$ para todo $\gamma \in \mathcal{T}_\mu$, como

$$[g(z, \gamma)](\omega) = \begin{cases} \sigma_1(z(\omega)) = \sigma_2(z(\omega)) & \omega \in [1, \mu] \setminus \mathcal{T}([1, \mu]) \\ \sigma_1(z(\omega)) \quad (\neq \sigma_2(z(\omega))) & \omega \in \mathcal{T}(\gamma^{-1}([1, 1])) \\ \sigma_2(z(\omega)) \quad (\neq \sigma_1(z(\omega))) & \omega \in \mathcal{T}(\gamma^{-1}([1, 2])) \end{cases}$$

esto es, para cada $\gamma \in \mathcal{T}_\mu$ $g(z, \gamma)$ es un gameto
 que puede producir a z mediante la fecundación con otro
 gameto.

- dado entonces $\mu \in [0, m]$, $z \in \mathcal{M}_\mu$, hacemos

$$W_z^\mu = \{ g(z, \gamma) \in G \mid \gamma \in \mathcal{T}_\mu \}$$

y a $\Theta : W_z^\mu \rightarrow W_z^\mu$ como $\Theta(g(z, \gamma)) = \Theta g(z, \delta)$

$$\text{con} \quad [\Theta g(z, \delta)](\omega) = \begin{cases} \sigma_1(z(\omega)) = \sigma_2(z(\omega)) & \omega \in [1, \mu] \setminus \mathcal{T}([1, \mu]) \\ \sigma_2(z(\omega)) \quad (\neq \sigma_1(z(\omega))) & \omega \in \mathcal{T}(\delta^{-1}([1, 1])) \\ \sigma_1(z(\omega)) \quad (\neq \sigma_2(z(\omega))) & \omega \in \mathcal{T}(\delta^{-1}([1, 2])) \end{cases}$$

esto es, $\Theta g(z, \gamma)$ es aquel otro gameto; esto es,

el proceso fecundativo $g(z, r) \rightarrow og(z, r)$ produce al cigoto z .

- Se proponen entonces las siguientes relaciones de fecundación:

$$P_z(t+1) = \sum_{r \in T_u} P_{g(z, r)}(t) P_{og(z, r)}(t) \quad \text{ssi } z \in \eta_u$$

por lo que debería tenerse que la relación entre las distribuciones haploides de la generación t a la generación $t+1$ fuera:

$$P_g(t+1) = \sum_{u=0}^m \frac{1}{2^u} \left(\sum_{r \in \mathcal{L}_u} \left(\sum_{z \in \Pi S_g^r(t)} \left(\sum_{r \in T_u} P_{g(z, r)}(t) P_{og(z, r)}(t) \right) \right) \right)$$

relación suficientemente complicada que no está resuelta aún, pero que debería llevar a la situación de equilibrio H-W:

$$P_g^\infty = \prod_{\alpha < 1} P_{\alpha g(\alpha)}$$

siendo $P_{\alpha g(\alpha)}$ la probabilidad marginal para el locus α en el alelo $g(\alpha)$.

Los pasos a seguir para resolver esta conjetura serían los siguientes:

primero, resolver el problema para pocos alelos en cada locus, i.e. $n_a = 2$ para todo $a \in [1, m]$,

segundo, tratar el problema de varios alelos pero conservando el caso del mismo número por locus, i.e. $n_a = n \quad \forall a \in [1, m]$, y entonces tratar de ver el problema aquí planteado.

Un problema que surge después es tratar el caso en el que los loci no son necesariamente independientes, sino que pueden estar ligados. Esto quiere decir que dos loci o más pueden hallarse en el mismo par cromosómico. En este caso, para cada locus, en la situación de heterosis, la probabilidad de segregación de un determinado alelo no es $1/2$ como en el caso de independencia, sino r_a , y la del otro alelo $(1-r_a)$, siendo $r_a \in (0, 1)$ la probabilidad de entrecruzamiento entre los dos cromosomas del par homólogo, exactamente en el locus. (Ver, por ejemplo, Ayala [3]).

Si la población es infinita y el sistema de apareamiento es completamente al azar (las principales hipótesis H-W), conjeturo que debe también tener lugar un equilibrio de la forma H-W en este caso de loci ligados, aunque quizás dicha distribución de equilibrio tome una forma diferente a la de los productos de expresiones del tipo $(P_{a_1} + P_{a_2} + \dots + P_{a_{n_a}})^2$.

Por último, hay que explorar la posibilidad de introducir en

este esquema a fenómenos como la selección, la mutación, la migración, etc. que por supuesto anularían el estado de equilibrio H-W, pero pueden dar lugar a otros estados de equilibrio. (Ver, por ejemplo, Akin [2] , Wright [25])

Sin embargo, en todos estos casos el postular una población suficientemente grande - o infinita - , siendo una gran ventaja desde el punto de vista del tratamiento matemático, puede ser un factor importante para impedir la modelación de dinámicas genéticas más realistas y quizá muy importantes para la teoría de la evolución.

APÉNDICE A

LA CONDICIÓN DE EQUILIBRIO H-W EN LA GENÉTICA DE POBLACIONES . CONSIDERACIONES HISTÓRICAS

En "El Origen de las Especies" Darwin muestra el hecho de la gran variabilidad presente en todas las poblaciones de organismos, la tendencia a la multiplicación del número de individuos de cada población y la consiguiente supervivencia diferencial de unos organismos con respecto a otros, proceso que denomina selección natural.

Al especular respecto a la fuente y el mantenimiento de dicha variabilidad, Darwin debe forzosamente enfrentarse al aparente dilema de la herencia: si "lo igual engendra lo igual", entonces ¿cómo aparece la variabilidad? Sin evidencias experimentales, Darwin especula sobre diversos mecanismos de herencia y, entre otras cosas, propone que una posible fuente de variación es una cierta forma de herencia de caracteres adquiridos (herencia blanda) y también que los caracteres heredados por los padres en cierta forma se mezclan (herencia mezcladora); este último mecanismo "unifor mizaria", a la larga, a la población, disminuyendo la variación, mientras que el primer mecanismo tendería a suministrar variación de modo continuo, manteniéndose así el equilibrio dinámico que se observa en la Naturaleza.

Posteriormente, Galton y Pearson, trabajando con caracteres cuantitativos continuos (estatura, peso, etc.) en diversas poblaciones - sobre todo humanas - desarrollan las bases de la estadística moderna, entre ellos el concepto de correlación mediante cuyo uso en la comparación de medidas antropométricas entre padres e hijos o en general entre parientes de diversos grados, Galton llega en 1897 a la conclusión de que en efecto el mecanismo de la herencia mezcladora explica perfectamente los valores de los coeficientes de correlación observados.

Si se considerara una posible fuente de variación, las diferencias en los caracteres hereditarios deberían tender a desaparecer en la razón de $\frac{1}{2}$ por generación. Sin embargo esto no es así para todos los caracteres: los descendientes de una pareja formada por un individuo de ojos café y un individuo de ojos azules tienen ojos café todos; los caracteres no se mezclan.

Cuando los experimentos de Mendel son repetidos - en forma independiente - por De Vries, Correns y Tschermak y se redescubren las leyes de Mendel, se entabla toda una batalla entre biometristas (la escuela biométrica inglesa creada por Galton y Pearson) y mendelianos.

Surgen dos problemas típicos: ¿cómo responde la genética mendeliana a los resultados de los biometristas? y ¿cómo es que puede resolver la cuestión del mantenimiento de la variabilidad?

Antes de proseguir la narración acerca de cómo se resolvieron estas cuestiones, describiré brevemente los resultados de los experimentos de Mendel y su interpretación en términos de la simbología y los conceptos que utilizó en esta tesis.

El primer experimento consiste en la cruce de dos estirpes de chícharos (Pisum sativum): amarillos (A_m , que siempre dan semillas amarillas) y verdes (V_e , que siempre dan semillas verdes). La primera cruce es: $A_m \times V_e$, cuyos descendientes todos son amarillos (no hay mezcla); llamamos a estos descendientes integrantes de la primera generación filial o F_1 . Cuando se cruzan a dos individuos F_1 se tiene que $\frac{3}{4}$ del total de descendientes F_2 son amarillos y $\frac{1}{4}$ verdes; el carácter verde "se espera" a manifestarse hasta la segunda generación y por ello se llama "recesivo"; el amarillo "domina" en la primera generación y por eso se le denomina "dominante". La explicación de este hecho se hace minuciosamente en el Apéndice B y puede ilustrarse mediante el uso del simbolismo usado en el Capítulo I:

$A_1 A_1$ es amarillo, $A_2 A_2$ es verde

La primera cruce $A_1 A_1 \times A_2 A_2$ da por resultado

$A_1 A_2$ (100%) en F_1 (amarillos)

La segunda cruce $A_1 A_2 \times A_1 A_2$ da por resultado

$A_1 A_1$ (25%), $A_1 A_2$ (50%), $A_2 A_2$ (25%) en F_2 ,

esto es, 75 % amarillos (A_1A_1 puros, A_1A_2 híbridos)
y 25 % verdes.

El segundo experimento consiste en la combinación de dos tipos de caracteres: el exterior, color de la semilla (con dos formas: amarilla y verde), y otro, la textura de la semilla (con dos formas también: lisa, dominante, y rugosa, recesiva). La interpretación de los resultados se da también en el Apéndice B y se ilustra aquí usando la simbología del Capítulo III:

$A_1A_1B_1B_1$ (amarillo liso) \times $A_2A_2B_2B_2$ (verde rugoso)

que da por resultado: $A_1A_2B_1B_2$ (amarillo liso) 100% en F_1

$A_1A_2B_1B_2 \times A_1A_2B_1B_2$

que da por resultado:

$A_1A_1B_1B_1$ ($1/16$), $A_1A_2B_1B_1$ ($2/16$), $A_1A_1B_1B_2$ ($2/16$), $A_1A_2B_1B_2$ ($4/16$)
amarillos lisos: $9/16$

$A_1A_1B_2B_2$ ($1/16$), $A_1A_2B_2B_2$ ($2/16$)
amarillos rugosos: $3/16$

en F_2

$A_2A_2B_1B_1$ ($1/16$), $A_2A_2B_1B_2$ ($2/16$)
verdes lisos: $3/16$

y $A_2A_2B_2B_2$ ($1/16$) verdes rugosos: $1/16$

Este comportamiento evidentemente discreto (o factorial, como le dice Mendel) contradice la hipótesis de la herencia mezcladora.

Es en medio de la polémica desatada por el redescubrimiento de estas leyes que entre los años 1902 y 1908 varios científicos - muchos de ellos matemáticos - como G. U. Yule, Galton, Bateson, Castle, R. C. Punnett, defienden unos la hipótesis de la herencia mezcladora y otros la refutan mediante argumentos más o menos correctos pero con cálculos erróneos, apoyándose en las hipótesis mendelianas. Es entonces cuando Punnett discute el problema con G. H. Hardy quien, una vez enterado de los postulados y las hipótesis de Mendel, resuelve el problema a favor del modelo mendeliano de la herencia, al menos en cuanto al mantenimiento de la variabilidad pues, como hizo ver Hardy - y como presenta esta tesis -, en el caso de apareamientos al azar las frecuencias genotípicas no cambian (no tenderá a desaparecer ninguna de las formas variantes) y se distribuirán como los términos del binomio $(p + q)^2$; esto es, seguirá habiendo variabilidad en el seno de la población. Al mismo tiempo (1908) e independientemente el médico alemán W. Weinberg presenta las mismas conclusiones en la Sociedad de Historia Natural de Stuttgart.

La cuestión acerca de los resultados de los biometristas la resuelve unos años después R. A. Fisher haciendo ver que los caracteres involucrados que se ven continuos, lo son por la interacción

de dos factores: primero, porque hay varios genes (varios loci) involucrados en la determinación de dichos caracteres, tendiéndose entonces muchas combinaciones posibles que establecerían una distribución binomial de probabilidad parecida a la distribución normal que observan los biometristas, como por ejemplo en el caso de la estatura, para los diferentes valores de tales rasgos; y en segundo lugar por la intervención del medio ambiente que, actuando de modo diverso sobre los individuos de la población, tiende a "suavizar" dicha distribución binomial haciéndola en efecto normal.

Otros trabajos que incluyen varios loci o varios alelos en situaciones particulares son efectuados por R. B. Robbins en 1918, pero no es sino hasta 1955 que C. C. Li [19] introduce el modelo matricial para el caso de dos loci con dos alelos, el cual presenta en el capítulo III. En una parte del citado trabajo, al presentar el modelo simple de un locus con dos alelos, Li hace ver que antes de Hardy y Weinberg, Castle, en 1903 ya había demostrado el resultado que conocemos como equilibrio de Hardy - Weinberg (H-W) y propone que se le llame de Castle - Hardy - Weinberg (C-H-W), pero la fuerza de la costumbre ha impedido que este uso se generalice.

La condición de equilibrio H-W se ha usado como punto de

partida en todos los trabajos clásicos de los fundadores de la Genética de Poblaciones: R.A. Fisher, J.B.S. Haldane y S. Wright en el desarrollo de los modelos teóricos propuestos para estudiar la dinámica de las distribuciones de probabilidad de las diferentes familias de genes en el seno de las poblaciones, en función de los conocimientos que brindan la Genética Clásica (citológica) y la Genética Molecular.

Actualmente se habla de Genética Matemática de Poblaciones y existen múltiples enfoques teóricos para abordar el estudio de los modelos: geométrico-diferenciales, estocásticos (cadenas de Markov), algebraicos, en teoría de grafos, teoría de difusión, etc. y la mayoría de ellos parten de la condición H-W para observar el comportamiento de poblaciones menos ideales quitando una o varias de las hipótesis inherentes al modelo H-W original. En particular, cuando S. Wright se da cuenta de que el tamaño finito de las poblaciones puede ser un factor de gran importancia en la evolución de las frecuencias génicas debido al "efecto de muestreo", se inicia toda una nueva era en la investigación matemática de la genética de poblaciones que ha desembocado en una teoría alternativa y complementaria de la llamada Síntesis Moderna (Ver Huxley [14]): la teoría neutral de la evolución molecular de Kimura [15].

Por supuesto, este abandono del supuesto de poblaciones muy grandes (o infinitas) implica directamente el abandonar la

posibilidad de equilibrio en el sentido de Hardy-Weinberg.

Cabe sólo añadir que una población en equilibrio H-W no evoluciona pues, por definición, la evolución implica el cambio en el tiempo de las distribuciones de frecuencias génicas.

APÉNDICE B

BASES FENOMENOLÓGICAS DE LA GENÉTICA DE POBLACIONES

Esta exposición es sólo un breve resumen acerca de los hechos fundamentales de la genética y la biología del desarrollo en los que se basan los modelos de la genética de poblaciones, en particular el tratamiento que se hace en esta tesis y que es de los más sencillos y elementales. He tomado el caso particular de la genética de la clase de los mamíferos por tratarse de un caso sencillo, suficientemente estudiado y muy cercano a nosotros. No obstante, la mayoría de los hechos que se refieren aquí son casi idénticos a los que suceden en la mayoría de los eucariotes - aquéllos cuyas células poseen núcleo - con reproducción sexual, esto es, la mayoría de los animales y las plantas; las diferencias que existen para todos estos organismos son hasta cierto punto irrelevantes, al menos para el trabajo teórico de un matemático. Para una exposición un poco más amplia y accesible para el lector no familiarizado con la ciencia de la biología, recomiendo el magnífico libro de S. E. Luria [21].

Existen dos caminos para la presentación o el estudio de la genética y la biología del desarrollo: el histórico cronológico, que sigue paso a paso el desenvolvimiento de esta ciencia estudiando los hechos y descubrimientos desde los elementales hasta los pasmosos conocimientos de la actualidad, camino recomendado para el estudiante de

biología o el lector paciente y sistemático, y el camino de la exposición fría y breve de los hechos y fundamentos tal y como los conocemos actualmente - por supuesto, evitando la complejidad de los estudios especializados -; este camino es el adecuado para un lector con cierta cultura científica y para aquél - como es el caso del físico y el matemático - que no requiere de una clase de disección o de microscopía para poder suponer que un cromosoma existe y se comporta como los especialistas dicen que lo hace. Yo he escogido este último. Hecha esta advertencia, espero ser suficientemente claro y breve.

Los seres vivos cuya genética nos interesa están compuestos por muchas células, todas con núcleo, en cuyo interior se hallan las estructuras conocidas como cromosomas. Estas estructuras son "arreglos" tridimensionales compuestos por una molécula gigantesca llamada ácido desoxirribonucleico o ADN. Esta molécula está constituida por un arreglo lineal de pares de bases en la forma de peldaños de una escalera que en el arreglo tridimensional gira en torno a su eje longitudinal formando la conocida doble hélice.

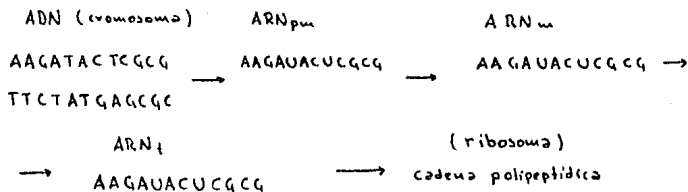
Los pares de bases son enlaces de dos moléculas llamadas bases, que pueden ser de dos tipos: púricas (adenina, A y guanina, G) o pirimidicas (timina, T y citosa, C). Cada uno de los "peldaños" está constituido por el enlace, mediante "puentes" de hidrógeno, de una base púrica con otra pirimidica

de acuerdo con las siguientes reglas: $A \equiv T$, $C \equiv G$; esto es, la adenina se "complementa" con timina mediante un doble enlace, mientras que la citosina se "complementa" con la guanina mediante un triple enlace. El uracilo, U es otra base pirimidica que, sustituyendo a la timina, toma el papel de ésta para formar otra molécula: ARN o ácido ribonucleico, de la que hablaremos en seguida. A su vez, cada uno de los pedacitos está unido a los adyacentes mediante enlaces con moléculas de azúcar y fosfatos, los que formarían el "pasamanos" de la escalera.

Debido a la complementariedad de las moléculas que forman los pares de bases, una de las mitades longitudinales de la escalera complementa a la otra, de tal modo que si, por ejemplo, de un lado se tuviese la secuencia AAGATACTCGCG, en el otro lado forzosamente deberá estar TTCTATGAGCGC.

Estas "letras" A, G, T, C, constituyen los elementos de un "alfabeto" (el código genético), el cual procede, muy esquemáticamente, de la siguiente manera: la información contenida en el ADN (en los cromosomas, dentro del núcleo de las células) se transcribe en una molécula análoga de ARN (llamada ARNm o pre mensajero) que "edita" dicha información separando exones (información relevante) de intrones (información irrelevante), que a su vez lo transcribe a otra molécula ARN (ARNm o mensajero) la cual viaja desde el núcleo hacia otro complejo estructural llamado ribosoma (conocido como "fábrica" de la célula),

entonces la información se "lee" (se transfiere mediante ARN_i o de transferencia) y su "lectura" se traduce en la construcción o elaboración de proteínas o enzimas o, en general, cadena polipeptídica, que serán los materiales de construcción o de instrucción y funcionalidad bioquímicos que la célula usa para su funcionamiento y el de todo el cuerpo. Esquemáticamente, con nuestro ejemplo, la cosa se ve así:



o más esquemáticamente aún: "a un gene, una cadena polipeptídica", pues son estas "frases" o conjunto de elementos del código genético los que contienen información relevante y a los cuales, en términos modernos, denominamos genes. Al total de la información genética de un individuo dado se le denomina genotipo, y es exactamente la misma (salvo mutaciones, de lo cual no hablaremos aquí) en todas y cada una de las células que lo constituyen.

Pues bien, en la mayoría de las células del cuerpo o soma de un individuo, - las células somáticas - los cromosomas aparecen por pares, denominados pares homólogos. Cada uno de los constituyentes

de dicho par homólogo contiene información parecida, del mismo tipo u homóloga; esto es, por ejemplo, puede contener información acerca del "color de los ojos", pero puede suceder que uno de los elementos contenga información específica que se traduzca como "color de los ojos: azules", mientras que el otro puede tener yacer la misma información exactamente - homocigosis - u otra un poco diferente - heterocigosis - , como "color de los ojos: café's". Por tanto, es en el lugar físico específico o locus de la molécula de ADN donde debe estar la frase o gene que contiene la información acerca del "color de los ojos", y puede haber información más específica de varios tipos - en el mismo locus -. A estas formas diferentes de un mismo gene se les denomina alelos. El por qué la información genética contenida en un determinado locus de un cromosoma dado pueda o no ser exactamente igual a la información contenida en el mismo locus del cromosoma homólogo complementario, se entenderá pronto.

Decimos que las células somáticas son diploides pues la información genética que contienen está dada en pares (por ejemplo, el humano tiene 23 pares cromosómicos, 46 cromosomas en total). Sin embargo, esto no es así para todas las células; existen otro tipo de células, llamadas células gaméticas o simplemente gametos que contienen sólo la mitad del número cromosómico de las células diploides: uno de cada par homólogo; a éstas se les denomina entonces haploides.

El proceso mediante el cual son generadas estas células haploides, los gametos, se denomina gametogénesis, fenómeno bastante complicado que involucra procesos de división celular conocidos como mitosis y meiosis, de los cuales no hablaremos mucho aquí por ser innecesario para lo que nos interesa; sin embargo hay que decir que en tratamientos que se alejen de la suposición de que se cumplen las leyes de Mendel, será indispensable detenerse con cierta atención en dichos procesos pues el fenómeno de entrecruzamiento, mediante el cual dos cromosomas homólogos intercambian información, sucede en una de las fases de la meiosis. Esto puede verse con detalle en la vasta obra de Ayala [3]. Repito, no obstante, que en este tratamiento dicho fenómeno se supone implícitamente que no sucede cuando afirmo que "se cumplen las leyes de Mendel". Esto se entenderá adelante.

El proceso meiótico mediante el cual una hembra produce las células genéticas llamadas óvulos se denomina ovogénesis, y el proceso mediante el cual un macho produce las células genéticas llamadas espermias se denomina espermatoogénesis.

Los espermias se producen en los testículos y desde el nacimiento estas glándulas contienen células (diploides) llamadas espermatoogonios, que son los predecesores de las células espermáticas. Hasta la pubertad los espermatoogonios se dividen muy lentamente y producen muy pocos o ningún espermia; después de esta edad y bajo la influencia de hormonas, los espermatoogonios comienzan a dividirse

muy rápidamente produciendo millones de células hijas. Algunas de estas células se dividen meióticamente dando origen cada espermatozouio a cuatro espermátides; los espermátides se convierten en espermias mediante un proceso de maduración que involucra la formación de una cola filamentosa que permitirá a los espermias nadar, y la producción de sustancias que hacen que los espermias sean atraídos hacia el óvulo femenino. El proceso meiótico, en fin de cuentas, consiste en la separación o segregación de los cromosomas de los pares homólogos del espermatozouio y en su distribución en el núcleo de los espermatozoides (espermias) de tal modo que a cada uno de éstos le corresponde un cromosoma de cada par.

En el caso del huevo u óvulo, la situación es un poco más complicada. Cuando una hembra nace, ya tiene todos los óvulos que estarán presentes en los ovarios durante toda su vida. La maduración desde los ovogonios (diploides) hasta los ovocitos (haploides) se lleva a cabo durante la vida fetal. En la pubertad ocurren cambios hormonales y comienzan los ciclos de ovulación; en un momento del ciclo un óvulo madura, se desprende del saco o folículo que lo contiene y mediante el cual se halla individualmente unido al ovario, y comienza a descender por los cuernos hacia el útero. Si en el viaje al útero el óvulo es fertilizado por un espermia, se completa el proceso de la meiosis y se desechan tres de los cuatro núcleos haploides en forma de cuerpos polares (cuya única función para el huevo fertilizado o zigoto será la de establecer un punto de simetría para

el desarrollo embrionario); el cuarto núcleo se convierte en el pronúcleo huevo. Cuando el espermatozoide penetra en el óvulo deja afuera la cola y su núcleo - pronúcleo espermático - se une al pronúcleo huevo formando entre ambos el núcleo del cigoto - de nuevo diploide - que comienza a reproducirse por mitosis en el desarrollo del embrión.

La fecundación consiste entonces en la fusión de dos células haploides: espermatozoide y óvulo, para formar una célula diploide: el cigoto. La meiosis en el caso femenino consiste entonces en la segregación de los cromosomas de los pares homólogos del ovogocito y su distribución en el núcleo de cuatro células haploides de tal modo que a cada una de éstas le corresponde un cromosoma de cada par.

En ambos casos, y muy esquemáticamente, la meiosis consiste en lo siguiente:

- i) una división celular mediante la cual una célula diploide ($2n$ cromosomas, n pares homólogos) produce dos células diploides ($2n$ cromosomas en cada una, $4n$ en total).
- ii) una división posterior de cada una de las células anteriores que resulta en la producción de cuatro células haploides (n cromosomas en cada una, uno de cada n par homólogo inicial).

Es en el paso ii) donde me interesa discutir las implicaciones que tienen estos hechos para las hipótesis cruciales del modelo H-W. Suponemos que la "asignación" de cromosomas del par homólogo en

esta guseta constituye un suceso independiente para cada par experimental; esto es, los elementos de cada par cromosómico, al seguirse, se distribuyen al azar en cada guseta. Dicha suposición no contradice, a primera aproximación - hasta donde se conoce al fenómeno - ninguna de las observaciones conocidas.

En el caso de los espermas tenemos entonces millones (un millón de cuatro) de ellos en cantidad suficiente como para que, debido al proceso azaroso que les da origen, haya un número igual de representantes de cada combinación específica de cromosomas posible para el genotipo diploide que les da origen, por lo que todos los tipos posibles tienen la misma oportunidad de ser escogidos en una elección al azar. (Llamamos genotipo haploide a cada uno de estas combinaciones específicas posibles).

En el caso de los óvulos sólo se tienen cuatro representantes de todas las posibles combinaciones que pueden tener lugar en el paso (ii); sin embargo, la suposición de aleatoriedad en el proceso de asignación de cromosomas a las células haploides nos lleva a concluir que cada combinación específica posible ha tenido la misma oportunidad de estar representada en uno de los cuatro posibles óvulos.

Por tanto, aunque los procesos genotípicos específicos para espermas y óvulos son muy diferentes, al final podemos suponer que, por la aleatoriedad de los procesos de segregación de cromosomas,

Cada una de las posibles combinaciones tiene la misma oportunidad de ser "escogida" en una "elección" al azar; dicha elección en el caso masculino consiste en saber qué espermia específico será el fecundante, y en el caso femenino cuál es el óvulo específico que será fecundado.

Otro hecho que necesitamos suponer es el siguiente:

No existe diferencia en capacidad de supervivencia, vigor, etc. de los espermias, en cuanto a su capacidad de fecundar, dada por sus diferencias genéticas; asimismo, no existe diferencia en vigor o vitalidad o capacidad de supervivencia de los óvulos, en cuanto a su susceptibilidad de ser fecundados, dada por sus diferencias genéticas.

Además, los espermias no eligen a qué óvulo habrá de fecundar ni los óvulos eligen qué espermia habrá de fecundarlos. Por tanto, el proceso de fecundación es un proceso aleatorio de encuentro; esto es, constituye la ocurrencia simultánea de eventos independientes.

De todo lo que hemos dicho hasta aquí extraemos las hipótesis que son centrales para esta tesis:

- la información genética de cada individuo (diploide) está dada en pares de cromosomas, siendo esta información dividida o segregada en cada uno de los elementos de dicho par, elementos que pasan - uno solo de ellos - a los gametos (haploides). Esta es la primera

Ley de Mendel.

cundo consideramos más de un par cromosómico (dos en esta tesis) la segregación de la información genética de un par cromosómico es independiente de la segregación en los otros pares cromosómicos considerados. Esta es la segunda Ley de Mendel.

el proceso de fecundación es un proceso completamente aleatorio -salvo por la forma de apareamiento, que aquí también se considera que sea al azar- dado por la ocurrencia simultánea de dos eventos independientes: el encuentro de un determinado espermia con un determinado óvulo, cuyas probabilidades de aparición dependen sólo de los genotipos de los individuos involucrados.

Tengo que aclarar que el hecho de que los experimentos de Mendel o los fenómenos observados y cuidadosamente registrados por él no hayan sido tratados aquí se debe a una deliberada omisión de mi parte por dos razones: en primer lugar porque hablo brevemente del famoso monje austriaco y sus trabajos en el Apéndice A y en segundo lugar porque he preferido la claridad y brevedad del conocimiento actual - evitando hablar de fenómenos muy complicados que afectan a todo cuanto he expuesto - que la más didáctica y fascinante, aunque muy larga,

historia de la Genética. Para un lector interesado en anécdotos muy interesantes y en análisis críticos agudos por parte de científicos como R.A. Fisher sobre los famosos experimentos de Mendel, recomiendo la obra de Sherwood y Stern [23].

Por último, mencionare' concierto detalle el comportamiento autogámico de un par cromosómico específico particularmente importante: los llamados cromosomas sexuales.

A diferencia de los otros pares cromosómicos, denominados autosómicos, el par de cromosomas sexuales no se comporta exactamente con el carácter de homología de los otros. Hablaré entonces brevemente acerca de algunas características de dicha homología: en el proceso de traducción ARNm \rightarrow proteínas existen variantes en un continuo respecto a la expresión de los genes en función de la heterosis en determinado locus; me estoy refiriendo al fenómeno de la recesividad - dominante. Explicaré esto mediante algunos ejemplos:

Puede darse el caso de que un alelo codifique para estatura alta y otro para estatura baja; en este caso el individuo heterocigo es de estatura intermedia. El otro caso extremo se da por ejemplo en el del color de los ojos: si un alelo codifica para color azul y otro para color café, el individuo heterocigo (café, azul) tiene los dos ojos café, exactamente del mismo modo que el homocigo (café, café) ... sólo el homocigo (azul, azul) tiene los ojos azules. En este caso se dice que el alelo "café" es dominante

(domina sobre el "azules") y el alelo "azules" es recesivo, pues "espera" a manifestarse hasta la segunda generación, cuando hay oportunidad de otro homocigoto (azules, azules). En el caso de la estatura no hay dominancia.

Hay que subrayar que este fenómeno no es el resultado de una "mezcla" de caracteres, sino que la expresión parece estar — según ya dijimos — en el proceso de traducción genes — proteínas: puede suceder que la traducción de un alelo específico no solo produzca determinada proteína, sino que inhiba la producción de aquella otra determinada por el otro alelo; en este caso el primero sería dominante y el segundo recesivo. Otro caso es aquel en que ninguno de los dos genes inhiba al otro y se produzca, ya bien sea una especie de mezcla, o bien que el proceso de traducción combinado produzca una tercera posibilidad híbrida.

Pues bien, en el caso de los cromosomas sexuales, estos son de dos tipos que difícilmente conservan el sentido de homología anterior: cromosomas X y cromosomas Y (la nomenclatura les viene de su forma, parecida a la de estas letras). Las hembras tienen el par homólogo (aquí sí) XX, mientras que los machos tienen el genotipo XY. Esto es sólo la manifestación de un hecho, pero podríamos pensar que se trata de una relación causal:

XX produce hembras, XY produce machos. Pero el hecho es así, por lo que la producción de óvulos, con base en lo que

hemos explicado, forzosamente dará todos con un cromosoma X; por las razones expuestas, la producción de espermias será forzosamente de tal modo que la mitad tendrán un cromosoma X y la otra mitad un cromosoma Y. Por tanto, en el momento de la fecundación, si el espermia fecundante es X el resultado será un cigoto que se desarrollará en una hembra, y si es Y, el cigoto resultante se convertirá en un macho. Como no hay posibilidad de que haya óvulos Y, no es posible la combinación diploide YY. Es esto, entre otras, la razón fundamental por la cual en el capítulo II separemos a la población en dos subpoblaciones: la femenina y la masculina, y a los procesos fecundativos en dos clases: los que originan hembras y los que originan machos.

Ahora bien, puede haber formas alélicas diversas en ciertos loci del cromosoma X, lo que da lugar a distintos genotipos. Si pensamos que hay dos posibilidades alélicas, digamos X y X', como en el caso de la hemofilia y del daltonismo en los humanos, tendremos que XX es una hembra normal, XX' es normal pero "portadora", esto es, puede haber una relación de dominancia-recesividad entre X y X' y el heterocigoto XX' es como XX en su manifestación externa o fenotipo, mientras que X'X' es una hembra afectada. En los machos: XY es un macho normal y X'Y es un macho afectado; sólo hay dos posibilidades, y la falta de homología entre X y Y hace que la información en el cromosoma Y no intervenga en ningún tipo de

dominancia o recesividad con respecto a la información codificada en el cromosoma X.

Por último, si bien es cierto que puede haber formas alélicas en loci situados en el cromosoma Y, la dinámica de estas variantes es tan sencilla que no tiene caso estudiarla a fondo pues no hay influencia - por la falta de homología - del cromosoma X en ellos (hasta donde se sabe): los hijos machos de un macho XY' serán XY' y los de XY serán XY pues, para ser machos, deben heredar del padre el cromosoma Y (o Y'), las hijas serán XX en todos los casos (salvo variantes alélicas en X). Sin la intervención de otros fenómenos (selección, mutación, migración) la población en cuestión estará siempre en equilibrio genético respecto a la distribución de probabilidad de los alelos Y, Y' .

Existen múltiples fenómenos - muchos no comprendidos todavía - bastante más complicados que los que acabo de esbozar aquí e incluso algunos de los hechos de los que me he servido como ejemplos nos ocurren en la forma tan sencilla y directa como pareciera, sino que dependen de una dinámica y de relaciones funcionales entre los genes mucho más complejas, pero que salen fuera del marco teórico de este trabajo y por lo tanto del objetivo de este apéndice, y que no invalidan la utilidad del modelo simple de herencia en el que se basa esta tesis.

APÉNDICE C

QUESTIONES TÉCNICAS

En el caso que nos ocupa (2 loci, n alelos) se tienen las relaciones que conectan a la distribución haploide con la distribución diploide para la generación t , relaciones que hemos denominado de gametogénesis:

$$P_{ik}(t) = P_{\delta(i)} \delta(k) (t) + \frac{1}{2} \left(\sum_{s \in L \cup R \cup H} P_{\delta(s)} s^{(t)} + \sum_{r \in L \cup R} P_{r \delta(r)} (t) \right) + \frac{1}{4} \left(\sum_{s \in L \cup R \cup H} \sum_{r \in L \cup R} P_{rs} (t) \right) \dots (IV.22)$$

y las relaciones entre la distribución diploide para $t+1$ y la distribución haploide para t , las relaciones de fecundación: (IV.23)

$$P_{rs}(t+1) = \begin{cases} P_{\sigma_1(r) \sigma_1(s)}^2 (t) & \text{si } \sigma_1(r) = \sigma_2(r), \sigma_1(s) = \sigma_2(s) \\ 2 P_{\sigma_1(r) \sigma_1(s)}^{(t)} P_{\sigma_1(r) \sigma_2(s)} (t) & \text{si } \sigma_1(r) = \sigma_2(r), \sigma_1(s) < \sigma_2(s) \\ 2 P_{\sigma_1(r) \sigma_1(s)}^{(t)} P_{\sigma_2(r) \sigma_1(s)} (t) & \text{si } \sigma_1(s) < \sigma_2(s), \sigma_1(r) = \sigma_2(r) \\ 2 (P_{\sigma_1(r) \sigma_1(s)}^{(t)} P_{\sigma_2(r) \sigma_2(s)}^{(t)} + P_{\sigma_1(r) \sigma_2(s)}^{(t)} P_{\sigma_2(r) \sigma_1(s)}^{(t)}) & \text{si } \sigma_1(r) < \sigma_2(r), \sigma_1(s) < \sigma_2(s) \end{cases}$$

Entonces, tomando (IV.22) para $t+1$ y (IV.23) para t , podemos llegar, por un camino completamente análogo al que nos llevó a (IV.24), a:

$$\begin{aligned}
 P_{ik}(t+1) = & P_{ik}(t) \left[P_{ik}(t) + \sum_{s \in L_k} P_{i\sigma_s(t)}^{(t)} + \sum_{s \in R_k} P_{i\sigma_s(t)}^{(t)} \right. \\
 & + \sum_{r \in L_i} P_{\sigma_r(t)k}^{(t)} + \sum_{r \in R_i} P_{\sigma_r(t)k}^{(t)} \\
 & + \sum_{r \in L_i} \sum_{s \in L_k} P_{\sigma_r(t)\sigma_s(t)}^{(t)} + \sum_{r \in L_i} \sum_{s \in R_k} P_{\sigma_r(t)\sigma_s(t)}^{(t)} \\
 & \left. + \sum_{r \in R_i} \sum_{s \in L_k} P_{\sigma_r(t)\sigma_s(t)}^{(t)} + \sum_{r \in R_i} \sum_{s \in R_k} P_{\sigma_r(t)\sigma_s(t)}^{(t)} \right] \\
 & + \frac{1}{2} \left[\sum_{r \in L_i} \sum_{s \in L_k} (P_{i\sigma_r(t)}^{(t)} P_{\sigma_s(t)k}^{(t)} - P_{ik}(t) P_{\sigma_r(t)\sigma_s(t)}^{(t)}) \right. \\
 & + \sum_{r \in L_i} \sum_{s \in R_k} (P_{i\sigma_r(t)}^{(t)} P_{\sigma_s(t)k}^{(t)} - P_{ik}(t) P_{\sigma_r(t)\sigma_s(t)}^{(t)}) \\
 & + \sum_{r \in R_i} \sum_{s \in L_k} (P_{\sigma_r(t)k}^{(t)} P_{i\sigma_s(t)}^{(t)} - P_{\sigma_r(t)\sigma_s(t)}^{(t)} P_{ik}(t)) \\
 & \left. + \sum_{r \in R_i} \sum_{s \in R_k} (P_{\sigma_r(t)k}^{(t)} P_{i\sigma_s(t)}^{(t)} - P_{\sigma_r(t)\sigma_s(t)}^{(t)} P_{ik}(t)) \right]
 \end{aligned}$$

resultado que, en vista de (IV.27 y siguientes), puede escribirse como:

$$\begin{aligned}
 P_{ih}(t+1) &= P_{ih}(t) \left[P_{ai}(t) + P_{\rho h}(t) - P_{ih}(t) + 1 - P_{\rho h}(t) \right. \\
 &\quad \left. - P_{ai}(t) + P_{ih}(t) \right] \\
 &\quad + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n \begin{vmatrix} P_{ih}(t) & P_{ih}(t) \\ P_{jh}(t) & P_{jh}(t) \end{vmatrix} \\
 &= P_{ih}(t) + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n \begin{vmatrix} P_{ih}(t) & P_{ih}(t) \\ P_{jh}(t) & P_{jh}(t) \end{vmatrix}
 \end{aligned}$$

resultado que muestra lo expuesto en (IV.5) para el caso particular $n=3$.

Pero también se tiene, siguiendo los pasos que llevan de (IV.27) a (IV.28), que:

$$\begin{aligned}
 P_{ih}(t+1) &= P_{ih}(t) \left[P_{ai}(t) + P_{\rho h}(t) - P_{ih}(t) \right. \\
 &\quad \left. + \sum_{j=1}^n P_{aj}(t) - \sum_{j=1}^n P_{jh}(t) - P_{ai}(t) + P_{ih}(t) \right] \\
 &\quad + \frac{1}{2} \left[P_{ai}(t) P_{\rho h}(t) - P_{ih}(t) \right] \\
 &= P_{ih}(t) \left(\sum_{j=1}^n P_{aj}(t) \right) + \frac{1}{2} \left[P_{ai}(t) P_{\rho h}(t) - P_{ih}(t) \right],
 \end{aligned}$$

donde se ve la necesidad de tener la condición de normalidad

$$\sum_{j=1}^n P_{aj}(t) = 1 \quad \text{para llegar a tener la relación}$$

buscada :

$$P_{ik}(t+1) = \frac{1}{2} P_{ik}(t) + \frac{1}{2} P_{ki}(t) P_{\beta k}(t)$$

Sin embargo, para asegurar que la condición de normalidad es compatible, requerimos a su vez de la expresión anterior, razón por la cual el texto sólo contiene dicha expresión para $t+1 = t$; la demostración de ambos hechos, simultáneamente, la tenemos en la Proposición (IV.29), que habremos de mostrar por inducción en bloque :

Proposición .- Para todo $t \in \mathbb{N}$ se cumple

$$a) \sum_{i=1}^n \sum_{k=1}^n P_{ik}(t) = 1$$

$$b) P_{ki}(t) = P_{ki} \quad , \quad P_{\beta k}(t) = P_{\beta k}$$

$$c) P_{ik}(t+1) = \frac{1}{2} P_{ki} P_{\beta k} + \frac{1}{2} P_{ik}(t)$$

Se tiene, como primera hipótesis, que $\sum_{i=1}^n \sum_{k=1}^n P_{ik} = 1$,

o bien, $\sum_{i=1}^n P_{ki} = \sum_{k=1}^n P_{\beta k} = 1$ de la primera generación,

lo cual nos ha llevado hasta el resultado (IV.28):

$$P_{ikh}(1) = \frac{1}{2} P_{\alpha i} P_{\beta h} + \frac{1}{2} P_{ik}$$

con ésto, arribamos fácilmente a (IV.28')

$$\sum_{i=1}^n \sum_{h=1}^n P_{ikh}(1) = 1 \quad (\text{que es a) para } t=1),$$

lo cual era de esperarse si $P(1)$ es la distribución de probabilidad para la generación 1, pero que habríamos de probar compatible con (IV.28). De aquí se sigue inmediatamente, como hemos visto, (IV.28'')

$$P_{\alpha i}(1) = P_{\alpha i}, \quad P_{\beta h}(1) = P_{\beta h}$$

Esto es, tenemos a) y b) para $t=1$; con esto demos tramos c) para $t=1$:

$$\begin{aligned} P_{ikh}(2) &= P_{ikh}(1) \left[P_{\alpha i}(1) + P_{\beta h}(1) - P_{ik}(1) \right. \\ &\quad \left. + \sum_{l=1}^n (P_{\alpha l}(1) - P_{il}(1)) - (P_{\alpha i}(1) - P_{ik}(1)) \right] \\ &\quad + \frac{1}{2} (P_{\alpha i}(1) P_{\beta h}(1) - P_{ikh}(1) \sum_{j=1}^n \sum_{z=1}^n P_{jz}(1)) \end{aligned}$$

usando (IV.27 y siguientes)

$$\begin{aligned}
 &= P_{ik}(t) \left[P_{ai} + P_{\beta k} - P_{ik}(t) + 1 - P_{\beta k} - P_{ai} + P_{ik}(t) \right] \\
 &\quad + \frac{1}{2} (P_{ai} P_{\beta k} - P_{ik}(t)) \\
 &= \frac{1}{2} P_{ai} P_{\beta k} + \frac{1}{2} P_{ik}(t) \quad \text{usando a) y b) para t.}
 \end{aligned}$$

Por tanto, tenemos a), b) y c) para $t=1$.

Hacemos ahora la inducción "en bloque" pues, para mostrar que a) y b) se cumplen para $t+1$ debemos tener que c) se cumple para t , y para demostrar que c) se cumple para $t+1$ debemos tener a) y b) en $t+1$; así pues, como hipótesis de inducción proponemos:

$$a) \sum_{i=1}^n \sum_{k=1}^n P_{ik}(t) = 1$$

$$b) P_{ai}(t) = P_{ai}, \quad P_{\beta k}(t) = P_{\beta k}$$

$$c) P_{ik}(t+1) = \frac{1}{2} P_{ai} P_{\beta k} + \frac{1}{2} P_{ik}(t)$$

Dem. para $t+1$:

$$a) \sum_{i=1}^n \sum_{k=1}^n P_{ik}(t+1) = \sum_{i=1}^n \sum_{k=1}^n \left(\frac{1}{2} P_{ik}(t) + \frac{1}{2} P_{ai} P_{\beta k} \right)$$

de la hipótesis c) para t

$$\begin{aligned}
 &= \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n P_{ik}(t) + \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n P_{\alpha i} P_{\beta k} \\
 &= \frac{1}{2} + \frac{1}{2} \quad \text{de hip. (a) para } t \\
 &= 1
 \end{aligned}$$

$$b) P_{\alpha i}(t+1) = \sum_{k=1}^n P_{ik}(t+1) = \sum_{k=1}^n \left(\frac{1}{2} P_{ik}(t) + \frac{1}{2} P_{\alpha i} P_{\beta k} \right)$$

$$\text{de c) para } t \quad = \frac{1}{2} P_{\alpha i}(t) + \frac{1}{2} P_{\alpha i}$$

$$\text{de b) para } t \quad = \frac{1}{2} P_{\alpha i} + \frac{1}{2} P_{\alpha i} = P_{\alpha i}$$

$$\text{y } P_{\beta k}(t+1) = P_{\beta k} \quad \text{por analogía.}$$

por último: c) :

$$\begin{aligned}
 P_{ik}(t+2) &= P_{ik}(t+1) \left[P_{\alpha i}(t+1) + P_{\beta k}(t+1) - P_{ik}(t+1) \right] \\
 &+ \sum_{\ell=1}^n \left(P_{\alpha \ell}(t+1) - P_{\ell k}(t+1) \right) - \left(P_{\alpha i}(t+1) - P_{ik}(t+1) \right) \\
 &+ \frac{1}{2} \left(P_{\alpha i}(t+1) P_{\beta k}(t+1) - P_{ik}(t+1) \sum_{\ell=1}^n \sum_{j=1}^n P_{j\ell}(t+1) \right)
 \end{aligned}$$

por (IV.27)

$$\begin{aligned}
 &= P_{i,h}(t+1) \left[P_{\alpha i} + P_{\beta h} - P_{i,h}(t+1) + 1 - P_{\beta h} - P_{\alpha i} + P_{i,h}(t+1) \right] \\
 &\quad + \frac{1}{2} (P_{\alpha i} P_{\beta h} - P_{i,h}(t+1)) \quad \text{por a) y b)} \\
 &\hspace{15em} \text{para } t+1 \\
 &= \frac{1}{2} P_{\alpha i} P_{\beta h} + \frac{1}{2} P_{i,h}(t+1)
 \end{aligned}$$

que es c) para $t+1$.

Por lo tanto, queda demostrada la proposición (IV.29) y la (IV.5) que es el caso particular para $n=3$.

Veremos enseguida lo relativo al método de identificación de gametos dado un zigoto particular, comenzando por la función definida en (IV.13)

La función $f: [-1, n] \rightarrow [0, \frac{n(n+1)}{2}]$

$$\text{con } f(-1) \equiv 0, \quad f(i) = \sum_{x=0}^i (n-x);$$

$$\text{se tiene, para } i \in [0, n]: \quad f(i) = \sum_{x=0}^i (n-x)$$

$$\begin{aligned}
 &= \sum_{x=0}^i n - \sum_{x=0}^i x = (i+1)n - \frac{i(i+1)}{2} = \frac{(i+1)(2n-i)}{2}
 \end{aligned}$$

que es lo que aparece en (IV.13); por tanto, los conjuntos

$$I_i \equiv (\rho(i-1), \rho(i)) = \{x \in \mathbb{Z} \mid \rho(i-1) < x \leq \rho(i)\}$$

son ajenos por construcción y $\bigcup_{x=1}^n I_x = (\rho(0), \rho(n)]$
 (ρ es creciente) $= (0, \frac{n(n+1)}{2}]$.

Esto hace que la función $\sigma_i: [1, \frac{n(n+1)}{2}] \rightarrow [1, n]$

está bien definida como $\sigma_i(r) = i$ así $r \in I_i$,

por lo que si $r \in I_{i-1} \Rightarrow r \in (\rho(i-2), \rho(i-1)]$

$$\Rightarrow \frac{(i-1)(2n - (i-2))}{2} < r \leq \frac{i(2n - (i-1))}{2}$$

$$\Rightarrow \frac{(i-1)(2n - (i-2))}{2} + 1 \leq r \leq \frac{i(2n - (i-1))}{2}$$

$$\Rightarrow \frac{(i-1)(2n - (i-2)) + 2}{2} \leq r \leq \frac{i(2n - (i-1))}{2}$$

Recordemos que I_0 representa al renglón 1 de Z_n :

$$\begin{array}{ccccccc} 1 & 2 & \dots & i & \dots & n \\ (1, 1) & (1, 2) & \dots & (1, i) & \dots & (1, n) \end{array}$$

I_1 representa al renglón 2 :

$$\begin{array}{cccc} n+1 & & n+2 & \dots & n+(n-1) \\ (2, 2) & , & (2, 3) & , \dots & , (2, n) \end{array}$$

etc. I_{i-1} representa al renglón i :

$$\begin{array}{cccc} f(i-2)+1 & & f(i-2)+2 & \dots & f(i-2)+n-(i-1) \\ (i, i) & , & (i, i+1) & , \dots & (i, n) \end{array}$$

por tanto, si $r \in I_{i-1}$, r debe ser de la forma

$$f(i-2)+1, f(i-2)+2, \dots, f(i-2)+n-(i-1) = f(i-1)$$

esto es, la primera componente asociada por σ ($\sigma_1(r)$) es i , y la segunda ($\sigma_2(r)$) debe ser de la forma $i+x$ con $x=0, 1, \dots, n-i$; por tanto, deberá de ser de la forma: la diferencia de r menos el extremo izquierdo del intervalo cerrado I_{i-1} (o sea $f(i-2)+1$) más el número i , por lo que $\sigma_2(r)$ debe ser:

$$\begin{aligned} \sigma_2(r) &= r - (f(i-2)+1) + i \\ &= r - \left(\frac{(i-1)(2n-(i-2)) + 2 + 2i}{2} \right) \\ &= r - \left(\frac{(i-1)(2n-i+2) + 2(1-i)}{2} \right) \end{aligned}$$

$$= r - \left(\frac{(i-1)2n - (i-1)(i-2) - 2(i-1)}{2} \right)$$

$$= r - \left(\frac{(i-1)2n - (i-1)(i-2+2)}{2} \right)$$

$$= r - \frac{(i-1)(2n-i)}{2} \quad \text{que es lo que aparece en (W.15)}$$

la función queda como $\sigma(r) = \left(i, r - \frac{(i-1)(2n-i)}{2} \right)$

ssi $r \in I_{i-1}$

demostramos que es inyectiva:

$$\text{si } r, s \in \left[1, \frac{n(n+1)}{2} \right] \Rightarrow r \in I_{i-1}, s \in I_{j-1}$$

si $r \neq s$ hay dos posibilidades:

$$1) \quad i \neq j \quad \text{en cuyo caso } \sigma_1(r) = i + j = \sigma_1(s) \\ \Rightarrow \sigma(r) \neq \sigma(s)$$

$$2) \quad i = j \quad \text{en cuyo caso } r = \rho(i-2) + x \\ s = \rho(i-2) + y$$

$$\text{con } x, y \in [1, n-(i-1)]$$

$$\text{como } r \neq s \Rightarrow x \neq y$$

$$\Rightarrow \sigma_2(r) = r - (\rho(i-2) + 1) + i = x - 1 + i$$

$$\sigma_2(s) = s - (\rho(i-2) + 1) + i = y - 1 + i$$

$$\Rightarrow \sigma_2(r) \neq \sigma_2(s) \quad \Rightarrow \quad \sigma(r) \neq \sigma(s)$$

El que σ sea suprayectiva se muestra como sigue:

$$\text{dado } (k, l) \in \left\{ (i, j) \mid i \in [1, n], j \in [1, n] \right\}$$

se tiene que $1 \leq k \leq l \leq n$, lo que implica

$$k + \frac{(k-1)(2n-k)}{2} \leq l + \frac{(k-1)(2n-k)}{2} \leq n + \frac{(k-1)(2n-k)}{2}$$

$$\Rightarrow \frac{2k + (k-1)(2n-k)}{2} \leq l + \frac{(k-1)(2n-k)}{2} \leq \frac{2n + (k-1)(2n-k)}{2}$$

$$\Rightarrow \frac{(k-1)2n + 2k - k^2 + k - 2 + 2}{2} \leq l + \frac{(k-1)(2n-k)}{2} \leq \frac{2n(1+k-1) - k(k-1)}{2}$$

$$\Rightarrow \frac{(k-1)2n - (k-1)(k-2) + 2}{2} \leq l + \frac{(k-1)(2n-k)}{2} \leq \frac{k(2n - (k-1))}{2}$$

$$\text{esto es: } \rho(k-2) + 1 \leq l + \frac{(k-1)(2n-k)}{2} \leq \rho(k-1)$$

o sea que el número $r = l + \frac{(k-1)(2n-k)}{2}$ es tal que

$$\rho(k-2) + 1 \leq r \leq \rho(k-1)$$

$$\Rightarrow r \in I_{k-1} \quad \Rightarrow \quad \sigma_1(r) = k$$

$$\sigma_2(r) = r - \frac{(k-1)(2n-k)}{2} = l$$

$\therefore \sigma$ es biyectiva y $\sigma^{-1}(k, \ell) = \ell + \frac{(k-1)(2n-k)}{2}$
 como se afirma en (IV.16)

Ahora, la función $\delta: [1, n] \rightarrow [1, \frac{n(n+1)}{2}]$ dada por

$$\delta(k) = \ell(k-2)+1$$

es creciente pues ℓ lo es, y recordando el papel de ℓ , tenemos que $\ell(k-2)+1$ es el extremo izquierdo del intervalo I_{k-1} , por lo que $r = \ell(k-2)+1 \Rightarrow \sigma_1(r) = \sigma_2(r) = k$; esto es, en $\ell(k-2)+1$ se representa a la pareja (k, k) como se afirma en (IV.18).

De (IV.19 y 20) se tiene que $L_i = I_{i-1} \setminus D_i$ consta de todos los números r tales que

$$\ell(k-2)+1 < r \leq \ell(k-1)$$

por lo que $\sigma_1(r) = i$ y $\sigma_2(r) = r - (\ell(k-2)+1) + i > i$

esto es, r representa a una pareja (i, j) con $j > i$; o sea que i aparece sencillito a la izquierda en los pares representados por elementos de L_i .

Los elementos de R_i son los que están en la columna i -ésima del arreglo triangular, excepto (i, i) que es el último.

				R_i			
$(1, 1)$	$(1, 2)$	$(1, 3)$...	$(1, i)$	$(1, i+1)$...	$(1, n)$
	$(2, 2)$	$(2, 3)$...	$(2, i)$	$(2, i+1)$...	$(2, n)$
		$(3, 3)$...	$(3, i)$	$(3, i+1)$...	$(3, n)$
				\vdots			
				$(i-1, i)$			
				(i, i)	L_i		
					$(i, i+1)$...	(i, n)
					$(i+1, i+1)$...	$(i+1, n)$
							\vdots
							(n, n)

el primer elemento de R_i está en el lugar i :

$$i = p(-1) + i - (-1+1)$$

el segundo elemento de R_i está en el lugar $(n+1) + (i-2)$

$$(n+1) + (i-2) = p(0) + i - (0+1)$$

el tercero en el lugar

$$n + (n-1) + 1 + (i-3)$$

$$n + (n-1) + 1 + (i-3) = p(1) + i - (1+1)$$

etc. el j -ésimo en el lugar

$$p\left(\frac{j-2}{2}+1\right) + (i-j)$$

$$= p(j-2) + i - (j-2+1)$$

el $i-1$ en el lugar $f(i-3) + i - (i-1)$
 $= f(i-3) + i - (i-3+1)$

esto es: $R_i = \left\{ f(x) + i - (x+1) \mid x \in [-1, i-3] \right\}$

como dice en (IV.21)

para el caso $i=1$ se tiene que $i-3 = -2$

y el intervalo $[-1, i-3] = [-1, -2] = \emptyset$

por lo que $R_i = \emptyset$ como es en efecto.

Escribo que si $r \in L_i \Rightarrow \sigma_1(r) = i$ por def.

ahora, si $r \in R_i$:

$$r = f(x) + i - (x+1) \quad x \in [-1, i-3]$$

$$\Rightarrow i \leq r \leq f(i-3) + i - (i-2)$$

$$\Rightarrow i \leq r \leq f(i-3) + 2$$

$$\Rightarrow 1 \leq \sigma_1(r) \leq i-2$$

además $\sigma_1(r) = x+2 \Rightarrow \sigma_2(x) = r - \frac{(x+2)-1}{2} (2n - (x+2))$

$$= f(x) + i - (x+1) - \frac{(x+1)(2n - (x+2))}{2}$$

$$= \frac{(x+1)(2n-x) + 2i - 2(x+1)}{2}$$

$$= \frac{(x+1)(2n - (x+2))}{2} = i$$

El hecho de que $L_i \cap R_i = \emptyset$ queda claro ahora:

$$r \in R_i \Rightarrow r = \rho(x) + i - (x+1) \quad x \in [-1, i-3]$$

$$\leq \rho(x) + n - (x+1)$$

$$= \rho(x+1) \leq \rho(i-3+1) = \rho(i-2)$$

$$\Rightarrow r \in \rho(i-2) \Rightarrow r \in I_{i-1} \Rightarrow r \notin L_i.$$

Por tanto, el conj. $L_i \cup R_i$ consta de tantos elementos $(n-1)$ como la suma de los elementos de L_i más los de R_i ; esto implica:

$$\sum_{s \in L_i \cup R_i} a_s = \sum_{s \in L_i} a_s + \sum_{s \in R_i} a_s$$

como usamos después de (IV.24)

BIBLIOGRAFÍA

- [1] Abercrombie, M.; Hickman, C.J.; Johnson, M.L. (1980)
The Penguin Dictionary of Biology. Penguin Books
- [2] Akin, E. (1979) The Geometry of Population Genetics
Springer-Verlag Berlin
- [3] Ayala, F.J.; Kiger, J.A. (1988) Genética Moderna
SITESA México
- [4] Breiman, L. (1969) Probability and Stochastic Processes
Houghton Mifflin Co. Boston
- [5] Darwin, C. (1969) El Origen de las Especies
UNAM México
- [6] Dobzhanski, T. (1950) Genetics of the Evolutionary Process
Columbia Univ. Press New York
- [7] Dubinin, N.P. (1984) Genética General
Ed. MIR Moscú

- [8] Ewens, W.J. (1979) *Mathematical Population Genetics*
Springer - Verlag Berlin
- [9] Feller, W. (1957) *An Introduction to Probability Theory and its Applications*
John Wiley and Sons New York
- [10] Fisher, R.A. (1949) *The Theory of Inbreeding*
Oliver & Boyd London
- [11] Fisher, R.A. (1958) *The Genetical Theory of Natural Selection*
Dover New York
- [12] Garden, E.J. (1982) *Principios de Genética*
LIMUSA México
- [13] Hofbauer, J.; Sigmund, K. (1988) *The Theory of Evolution and Dynamical Systems*
Cambridge Univ. Press Cambridge
- [14] Huxley, J. (1955) *Evolution: The Modern Synthesis*
George Allen & Unwin Ltd. London
- [15] Kimura, M.; Ohta, T. (1971) *Theoretical Aspects of Population Genetics*
Princeton Univ. Press Princeton N.J.

- [16] Kolmogorov, A.N. (1956) Foundations of the Theory of Probability
Chelsea Publ. Co. New York
- [17] Lehninger, A.L. (1970) Biochemistry
Worth Publishers Inc. New York
- [18] León de Garay, A. (1978) Genética de la Evolución
UNAM, México
- [19] Li, C.C. (1976) First Course in Population Genetics
The Boxwood Press Pacific Grove, Ca.
- [20] Li, H.W. (1977) Stochastic Models in Population Genetics
Dowden, Hutchinson, Ross Ltd. London
- [21] Luria, S.E. (1975) 36 Lectures in Biology
The MIT Press Cambridge, Mass.
- [22] Pearson, E.S. (1978) Pearson: Creador de la Estadística Aplicada
Espasa - Calpe Argentina Buenos Aires
- [23] Stern ; Sherwood (1980) El Origen de la Genética
Ed. Alhambra Barcelona

- [24] Watson, J. D. (1970) Molecular Biology of the Gene
W.A. Benjamin Menlo Park, Ca.
- [25] Wright, S. (1955) Evolution and the Genetics of
Populations Univ. of Chicago Press Chicago