

MODELS MATEMÀTICS I DE SIMULACIÓ DEL POLIMORFISME CROMOSÒMIC

PRODUÏT PER INVERSIONS

Memòria presentada per  
optar al grau de Doctor  
en Biologia per la Univer-  
sitat de Barcelona per  
en

JORDI OCAÑA I REBULL

Vist i Plau  
El Director de la Tesi

Dr. ANTONI PREVOSTI I PELEGRIN  
Catedràtic de Genètica  
Director del Departament de Genètica  
Facultat de Biologia  
Universitat de Barcelona

Barcelona, Setembre de 1981.



## 5. MACROMODELS

---

## 5. MACROMODELS

### 5.1. Model determinista d'una població diploïde amb aparellament a l'atzar, i selecció, pel cas de L loci.

Les equacions que defineixen el canvi de les freqüències gamètiques d'una generació a l'altra pel cas de dos loci dial·lèlics a una població diploïde, amb aparellament aleatori, sense coexistència de generacions i sota un règim selectiu d'eficàcies constants a l'espai i el temps varen ser definides per Lewontin i Kojima (1960). Ja abans, Kimura (1956) havia definit unes equacions equivalents però pel cas de temps continu (generacions intercalades).

Les equacions de Lewontin i Kojima es poden expressar de la manera següent: siguin dos loci A i B cada un dels quals pot presentar dos estats al·lèlics, que per raons de conveniència formal des d'ara indicarem com 0 i 1 (en lloc de l'habitual  $A_1$ ,  $A_2$  o A, a, pel primer locus, i  $B_1$ ,  $B_2$ , o B, b pel segon). Els gamets possibles són 00, 01, 10 i 11, corresponents a la representació binària dels nombres naturals representats en el sistema decimal 0, 1, 2 i 3, respectivament. Cada gamet està representat biunívocament per un nombre natural entre 0 i  $3 = 2^2 - 1$ . En general, si s'estudien L loci dial·lèlics i es representen aquests al·lèls pels dígit 0 i 1, hom pot indicar tot gamet

per una seqüència o vector de  $L$  zeros o uns el qual serà la representació binària d'un nombre natural entre 0 i  $2^L - 1$ , i, recíprocament, tot nombre entre aquests límits, expressat un sistema binari, representarà un gamet. Hi haurà, per tant,  $2^L$  gamets possibles als quals els correspondrà un i solament un número, el qual podrem fer servir per identificar-lo.

Pel cas de dos loci, indicarem per  $p_0, p_1, p_2, p_3$  les freqüències dels gamets 00, 01, 10 i 11 respectivament (0,1,2,3) de manera que es tindrà:

locus			
A	B		
gamets		número	freqüència
0	0	0	$p_0$
0	1	1	$p_1$
1	0	2	$p_2$
1	1	3	$p_3$

Taula 5.1.0.

Indiquem com  $W(x,y)=W_{xy}$ ,  $x,y = 0,1,2,3$  l'eficàcia del genotip constituït a partir dels gamets  $x,y$ . Com sabem  $W(x,y)$  és la probabilitat, o una magnitud proporcional, d'arribar a sobreviure i reproduir-se condicionat a tenir el genotip  $x,y$ :  $W(x,y) = \text{Prob}(S/x,y)$ , on  $S$  indica l'esdeveniment "sobreviu i es reproduceix (contribueix amb els seus gamets a la formació de la generació següent)".

Sigui  $R$  la fracció de recombinació entre els loci A i B. Si sumem respecte de tots els genotips que poden produir el gamet  $x$  ( $x=0,1,2,3$ ) tant per recombinació com per segregació, tenint en compte les probabilitats que passi això i les eficàcies (probabilitat que els individus amb aquell genotip arribin a produir gamets per la generació següent) i si suposem a més que  $W_{03} = W_{12}$ , és a dir que l'eficàcia depèn del contingut al·lèlic i no de la disposició d'aquests al·lels, s'obté amb algun treball el sistema d'equacions

$$p'_x = p_x - m_w^{-1} (p_x (W_x - m_w) - (-1)^x \cdot R \cdot DW_{12}) \quad (5.1.1.)$$

on  $p'_x$  indica la freqüència del gamet  $x$  a la generació següent,  $m_w$  l'eficàcia mitjana de la població parental

$$m_w = \sum_x \sum_y W(x,y) p_x p_y \quad (5.1.2.)$$

o mesura igual o proporcional a la probabilitat que un individu qualsevol arribi a reproduir-se,  $W(x)$  l'eficàcia marginal del gamet  $x$

$$W(x) = \sum_{y=0}^3 W(x,y) p_y \quad (5.1.3.)$$

$R$ , com hem dit, la fracció de recombinació entre els dos loci, i  $D$  el determinant de desequilibri al lligament o desequilibri gamètic

$$D = p_0 p_3 - p_1 p_2 \quad (5.1.4)$$

que expressa l'allunyament de les freqüències gamètiques reals de les freqüències gamètiques pel cas d'independència entre els dos loci. Efectivament, mitjançant la substitució  $p_0 = 1 - p_1 - p_2 - p_3$ , obtenim immediàtament

$$D = p_3 - p_1 p_3 - p_2 p_3 - p_3^2 - p_1 p_2 = p_3 - (p_1 + p_3)(p_2 + p_3)$$

però de les equivalències expressades a la taula 5.1.0 deduïm que  $p_1 + p_3 = p_{01} + p_{11} = p_{.1}$ , freqüència gènica de l'al·lel 1 del locus B, i igualment  $p_2 + p_3 = p_{10} + p_{11} = p_{1.}$ , freqüència gènica de l'al·lel 1 al locus A, de manera que

$$D = p_{11} - p_{1.} p_{.1} \quad (5.1.4\text{-bis})$$

D és la diferència entre la freqüència del gamet 3(11) i el producte de les freqüències gèniques de l'al·lel 1 a cada locus per separat.

Les propietats d'aquestes equacions han estat força estudiades sota diferents règims selectius i de recombinació. Igualment el model ha estat estès en diverses direccions, com ara la consideració de la dinàmica d'una inversió cromosòmica supressora de la recombinació, pel cas de 2 loci, (Deakin (1972), vegeu l'apartat 3) i la consideració de subdivisió del medi (recordem per exemple el treball de Christiansen, repassat a l'apartat

3.4.4. Una altra extensió interessant seria la consideració d'un nombre  $L \geq 2$  qualsevol de loci.

Lewontin (1964a) definí les equacions que serien l'extensió natural de (5.1.1) pel cas de  $L$  loci. Amb el mateix procediment de representar els dos al·lels per 0 i 1, cada gamet com una tira o vector de zeros o uns que representa un nombre natural entre 0 i  $2^L - 1$  (així per exemple, per  $L=3$ , els gamets possibles serien 000, 001, 010, 011, 100, 101, 110, 111 que són respectivament, la representació binària dels nombres naturals indicats 0,1,2,3,4,5,6,7 i 8 en representació decimal) obtingué l'expressió:

$$p'_x = p_x - m_w^{-1} (p_x (W(x) - m_w) - \rho(P, R))$$

(5.1.5)

per  $x = 0, 1, \dots, 2^L - 1$

summament satisfactòria si no fos per l'expressió de l'operador de recombinació  $\rho$

$$\rho(P, R) = \sum_{h=1}^{\frac{L(L-1)}{2}} \left[ R_h \sum_{j,k,l,m \in \Omega} \frac{(p_j p_k - p_l p_m) W(j, k)}{2^{H-2}} \right] +$$

$$+ \sum_{h=\frac{L(L-1)}{2} + 1}^{2^{L-1} - 1} \left[ O_h (-1)^{S_h} \sum_{j,k,l,m \in \Omega} \frac{p_j p_k W(j, k)}{2^{H-2}} \right] \quad (5.1.6)$$

on  $P = (p_0, p_1, \dots, p_{2^L - 1})$  és el vector de freqüències de gamets possibles,  $R$  és el vector de fraccions de recombinació,

$R = (R_1, R_2, \dots, R_{\frac{L(L-1)}{2}}, O_{\frac{L(L-1)}{2}+1}, \dots, O_{2L-1-1})$  essent les  $R_h$  fraccions de recombinació usuals, entre dos loci dels  $L$ , i les  $O_h$  fraccions de recombinació especials, paràmetres que cal afegir per definir totes les probabilitats de producció de gamets per part d'un genotip diploide,  $H$  és el nombre de loci heterozigòtics al genotip  $j,k$ ,  $S_h$  el nombre total d'intercanvis (esdeveniments de recombinació) associats a la fracció de recombinació  $O_h$  i  $\Omega, \omega$  són dos conjunts de parelles de genotips  $((j,k), (l,m))$  que renunciem a definir en poques paraules i d'una manera entenedora.

Tal com reconeix Lewontin, aquestes equacions són intractables analíticament i numèrica, de manera que proposà la utilització del mètode dels "operadors genètics", conjunt de transformacions algebraïques que permeten obtenir numèricament el vector de freqüències gamètiques d'una generació  $P^* = (p_0^*, p_1^*, \dots, p_n^*)$  (amb  $n=2^L-1$ ) a partir de les freqüències de la generació anterior  $P = (p_0, p_1, \dots, p_n)$ . Això s'obté segons els següents passos:

1) Es construeix la matriu de freqüències genotípiques a partir del producte extern  $P \otimes P' = G$

$$\begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{pmatrix} \otimes (p_0, p_1, \dots, p_n) = \begin{pmatrix} p_0^2 & p_0 p_1 & \dots & p_0 p_n \\ p_1 p_0 & p_1^2 & & \\ \vdots & & \ddots & \\ p_n p_0 & \dots & & p_n^2 \end{pmatrix} \quad (5.1.7-I)$$



2) La matriu  $G$  es multiplica terme a terme per la matriu  $W$  d'eficàcies

$$\begin{pmatrix} p_o^2 & p_o p_1 & \dots & p_o p_n \\ & \ddots & & \\ & & \ddots & \\ p_n p_o & \dots & \dots & p_n^2 \end{pmatrix} \begin{pmatrix} w_{oo} & \dots & w_{on} \\ \vdots & \ddots & \vdots \\ w_{no} & \dots & w_{nn} \end{pmatrix} = \\
 = \begin{pmatrix} p_o^2 w_{oo} & p_o p_1 w_{o1} & \dots & p_o p_n w_{on} \\ & p_1^2 w_{11} & & \vdots \\ \vdots & & \ddots & \vdots \\ p_n p_o w_{on} & \dots & \dots & p_n^2 w_{nn} \end{pmatrix} \quad (5.1.7-II)$$

obtenint la matriu de probabilitats de participar a la formació de la generació següent per part de cada genotip.

3) Es determinen tots els  $2^{L-1}$  gamets  $z$  que pot produir cada terme  $p_x p_y w_{xy}$  per les  $2^{L-1}$  probabilitats de transició  $P(xy \rightarrow z)$  que el genotip  $xy$  produeixi el gamet  $z$  (per recombinacions o per segregació sense recombinació) afegint  $p_x p_y w_{xy} P(xy \rightarrow z)$  a la freqüència de  $z$ . Això ens donaria el vector de freqüències de gamets a la generació següent (però sense normalitzar):  $\hat{P}$

4) Finalment es normalitza el vector de freqüències gamètiques obtingut  $P^* = \hat{P}/m_w$  per tal que sumin 1.

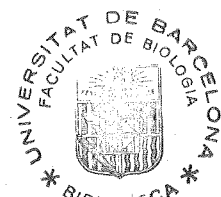
Les operacions 1,2 i 3 es poden fer en una sola iteració pels índexos  $x,y$  de tots els gamets possibles. Si es suposa simetria a les eficàcies  $w_{xy} = w_{yx}$  i a les probabilitats de recom-

binació, la iteració caldrà fer-la per  $x=1, \dots, 2^L$  i per  $y = 1, \dots, x$ , és a dir,  $\frac{2^L x 2^L - 2^L}{2} + 2^L = \frac{2^L x 2^L + 2^L}{2}$  vegades, cada una de les quals, a la seva vegada implicarà  $2^{L-1}$  operacions segons allò especificat al punt 3). Per tant caldrà fer  $\frac{2^L x 2^L + 2^L}{2} \times 2^{L-1} = 2^{2L-2} (2^L + 1)$  operacions bàsiques de determinar el zigot  $z$  i l'afegir  $p_x p_y w_{xy} P(xy \rightarrow z)$  a la seva freqüència. A continuació donem el nombre d'operacions d'aquestes segons alguns nombres de loci:

L	$2^{2L-2} (2^L + 1)$
2	20
3	144
4	1088
5	8192
6	65.536
7	528.384
8	4.210.688
9	33.619.968
10	$\sim 2.69 \times 10^8$
15	$\sim 8.79 \times 10^{12}$
20	$\sim 2.88 \times 10^{17}$
30	$\sim 3.09 \times 10^{26}$

Taula 5.1.2

Es poden definir unes altres equacions les quals sembla que són més adequades formalment que (5.1.5) i més eficaces pel que fa al càlcul numèric que el mètode dels operadors genètics. Men-



tre que el mètode dels operadors genètics es basa a avaluar tots els gamets que pot produir un genotip donat, la idea fonamental de les equacions que proposem als paràgrafs següents és la consideració de tots els genotips que poden produir un gamet.

Establim la següent notació: un gamet qualsevol, pel cas de  $L \geq 1$  loci, l'indiquem mitjançant una lletra sense subíndex:  $x, y, z$ , etc...,  $x$  és un vector  $x = (x_1, x_2, \dots, x_L)$  on cada component  $x_i$ ,  $i = 1, \dots, L$ , indica la constitució al·lèlica del locus  $i$ :  $x_i = \begin{cases} 0 \\ 1 \end{cases}$ . Podem dir, per tant, que  $x_i \in \{0, 1\}$  i que  $x \in G = \{0, 1\} \times \dots \times \{0, 1\}$ . Indiquem com  $G$  el conjunt dels gamets possibles. Sigui  $p_x = p(x)$  la freqüència (probabilitat, diem "freqüència" seguint la desafortunada terminologia de genètica de poblacions) del gamet  $x$ ,  $\sum_{x \in G} p_x = 1$ ,  $p_x \geq 0$ , i  $p'_x$  la freqüència a la generació següent. Una parella ordenada  $(y, z) \in G \times G$  indica un genotip:  $(y, z) = (y_1, \dots, y_L; z_1, \dots, z_L)$ .  $W_{yz} = W(y, z)$ , com abans, indica l'eficàcia del genotip  $(y, z)$ .  $R^*(x/yz)$  és la probabilitat que el genotip  $(y, z)$  produeixi el gamet  $x$ . Finalment, com abans,  $m_w$  indica l'eficàcia mitjana

$$m_w = \sum_{(y, z) \in G \times G} p_y p_z W_{yz}.$$

Les equacions més evidents en principi són

$$p'_x = m_w^{-1} \sum_{(y, z) \in G \times G} p_y p_z W_{yz} R^*(x/yz) \quad (5.1.8)$$

(5.1.8) es pot deduir dels següents raonaments: el pas d'una generació a l'altra es pot representar pels passos

zigots → adults reproductors → gamets de la generació següent

la generació següent la produeixen els adults reproductors, cal conèixer la freqüència (probabilitat) dels diferents genotips però condicionat que sobreviuen per poder reproduir-se  $\text{Prob}(xy/S)$ . Segons el teorema de Bayes, aquesta probabilitat és

$$\text{Prob}(yz/S) = \frac{\text{Prob}(yz) \cdot \text{Prob}(S/yz)}{\text{Prob}(S)}$$

on  $\text{Prob}(yz) = p_y p_z$  ja que l'aparellament es considera aleatori,  $\text{Prob}(S/yz) = w_{yz}$  i  $\text{Prob}(S) = \sum_{(y,z) \in G \times G} \text{Prob}(yz) \text{Prob}(S/yz) = \sum_{(y,z) \in G \times G} p_y p_z w_{yz} = m_w$ , segons la fórmula de les probabilitats totals, i per tant

$$\text{Prob}(yz/S) = m_w^{-1} p_y p_z w_{yz} .$$

Segons el pas b de l'esquema anterior, tot gamet x serà produït per algun individu de genotip (y,z) sempre que hagi sobreviscut i s'hagi reproduït, per tant, segons la fórmula de les probabilitats totals:

$$\begin{aligned} \text{Prob}(x) &= \sum_{(y,z) \in G \times G} \text{Prob}(yz/S) \text{Prob}(x/yz) = \\ &= \sum_{(y,z) \in G \times G} m_w^{-1} p_y p_z w_{yz} R^*(x/yz) = p'_x \end{aligned}$$

obtenim les equacions 5.1.8.

Les equacions (5.1.8) són inútils a la pràctica, ja que cada una d'elles té  $2^L \times 2^L$  sumands, amb paràmetres  $R^*(x/yz)$  en principi difícilment avaluable d'una manera general i concisa.

En absència de mutació, el terme  $R^*(x/yz)$  correspon a la probabilitat que el genotip  $(y,z)$  produeixi el gamet  $x$  per segregació o recombinació. Si  $(y,z)$  és homozigòtic, i amb al·lel diferent del corresponent a  $x$ , a algun locus, aleshores  $R^*(x/yz) = 0$ . És a dir, els genotips que poden produir  $x$  són aquells tals que a cada locus, o bé són heterozigòtics, o bé són homozigòtics amb el mateix al·lel que  $x$ .

Aquesta idea anterior la podem formalitzar definint el conjunt

$$\begin{aligned} H_x^* &= \{(y,z) \in G \times G \mid \forall i, i=1, \dots, L; y_i \neq z_i \vee y_i = z_i = x_i\} = \\ &= \{(y,z) \in G \times G \mid \forall i, i=1, \dots, L; (y_i = x_i) \vee (z_i = x_i)\} \end{aligned}$$

$H_x^*$  està definit segons la condició indicada en lletra negreta al paràgraf anterior. Si un genotip pertany a aquest conjunt, pot produir  $x$ , en cas contrari  $R^*(x/yz) = 0$ , per tant, en lloc de (5.1.8) podríem indicar

$$p_x' = m_w^{-1} \sum_{(y,z) \in H_x^*} p_y \cdot p_z \cdot w_{yz} R^*(x/yz) \quad (5.1.9)$$

amb menys sumands que (5.1.8).

Per exemple, si  $L=3$ , un gamet qualsevol, com ara  $x=101$  podrà ser produït pels genotips amb cada locus heterozigòtic o homozigòtic amb el mateix al·lel que presenta  $x$ . Per exemple el genotip  $101/101$  pot produir  $x$  ja que a tots els loci és homozigòtic amb el mateix al·lel que  $x$ , el genotip  $001/110$  també pot produir  $x$  ja que a tots els loci és heterozigòtic, el genotip  $100/001$  igualment el pot produir ja que al primer i al tercer locus és heterozigòtic i al segon és homozigòtic amb el mateix al·lel que  $x$  presenta l'al·lel 1. El conjunt  $H_{101}^*$  estarà format pels 27 genotips:

$$\begin{aligned} H_{101}^* &= \{(y,z) \mid \forall i, (y_i=x_i) \vee (z_i=x_i)\} = \\ &= \{101/101, 101/100, 101/111, 101/001, 101/110, 101/000, \\ &\quad 101/011, 101/010, 100/101, 100/111, 100/001, 100/011, \\ &\quad 111/101, 111/100, 111/001, 111/000, 001/101, 001/111, \\ &\quad 001/100, 001/110, 110/101, 110/001, 000/101, 000/111, \\ &\quad 011/101, 011/100, 010/101\} \end{aligned}$$

en lloc dels 64 genotips que componen  $G \times G$ . La freqüència de  $x$  a la generació següent serà la summa

$$\begin{aligned} P'_x &= m_w^{-1} (p_{101} p_{101} w(101,101) R(101/101,101) + \\ &\quad + p_{101} p_{100} w(101,100) R(101/101,100) + \dots \\ &\quad + p_{010} p_{101} w(010,101) R(101/010,101)) \end{aligned}$$

Si no trobem una manera general de determinar tots els elements de  $H_x^*$ , per qualsevol  $x$ , i una expressió convenient per

$R^*(x/yz)$ , les equacions (5.1.9) segueixen essent inutilitzables a la pràctica. Això es pot fer mitjançant els teoremes de-

mostrats als paràgrafs següents. Definim en primer lloc l'operació  $\oplus$  "summa lògica" al conjunt  $0,1$ , mitjançant la taula següent:

$\oplus$	0	1
0	0	1
1	1	0

Si  $a \in \{0,1\}$  i  $b \in \{0,1\}$ , podríem indicar  $a \oplus b = (a \vee b) \wedge \neg(a \wedge b)$ , on  $\vee, \wedge, \neg$  són les habituals operacions lògiques

$\wedge$	0	1
0	0	0
1	0	1

$\vee$	0	1
0	0	1
1	1	1

$\neg$	
0	1
1	0

L'operació  $\oplus$  es pot estendre immediatament a  $G$ :  $x \oplus y = (x_1 \oplus y_1, \dots, x_L \oplus y_L)$ , operant terme a terme tots els elements de les  $L$ -ples ordenades  $x$  i  $y$ . Fixem-nos que si  $a \in \{0,1\}$ , aleshores  $a \oplus 0 = 0 \oplus a = a$  i  $a \oplus 1 = 1 \oplus a = \neg(a)$ , és a dir, que l'operació amb 0 "no canvia l'al·lel" i amb 1 "canvia l'al·lel". Una propietat de  $\oplus$  que ens serà útil més endavant és la seva associativitat: per tot  $a, b, c \in G$ ,  $(a \oplus b) \oplus c = a \oplus (b \oplus c)$ , de demostració immediata.

Definim el conjunt

$$H^* = \{(h_1; h_2) \in G \times G \mid \forall i, i=1, \dots, L; h_{1i} = 0 \text{ o } h_{2i} = 0\} = \\ \{(h_1; h_2) \in G \times G \mid \forall i, i=1, \dots, L; h_{1i} \wedge h_{2i} = 0\}$$

format per aquelles parelles ordenades de gamets (o, millor dit, simplement seqüències de 0 i 1 de llargada L, ja que si bé sintàcticament podem dir  $h_1 \in G$ , semànticament  $h_1$  no és un gamet) tals que a cada posició (millor dit que no pas "locus")  $h_1$  o  $h_2$  (o ambdues alhora) presenta un 0.

El teorema següent indica que  $H^*$  caracteritza totalment el conjunt  $H_x^*$  dels genotips que podem produir el gamet  $x$ , per qualsevol  $x$ .

Teorema 5.1.1.

Per tot gamet  $x \in G$ , es compleix

a) Per tot  $(h_1; h_2) \in H^*$  la parella  $(h_1 \oplus x; h_2 \oplus x) = (x \oplus h_1; x \oplus h_2) \in H_x^*$ , és a dir, és un dels genotips que poden produir  $x$

b) Per tot  $(y, z) \in H_x^*$ , possible genotip productor de  $x$ , existeix un i solament un  $(h_1, h_2) \in H^*$  tal que  $(y, z) = (x \oplus h_1; x \oplus h_2) = (h_1 \oplus x; h_2 \oplus x)$ .

En particular això vol dir que existeix una bijecció entre  $H^*$  i  $H_x^*$ , per qualsevol  $x$ , de manera que ambdós conjunts finits tenen el mateix nombre d'elements,  $\forall x: \text{card}(H^*) = \text{card}(H_x^*)$  i que  $H_x^*$  pot ser determinat a partir de  $H^*$  segons l'expressió

$$H_x^* = \{(y, z) \in G \times G \mid y = x \oplus h_1, z = x \oplus h_2, (h_1, h_2) \in H^*\}$$



Efectivament:

a) Si  $(h_1, h_2) \in H^*$ , per tot locus  $i$ , existeix  $j=1,2$  tal que  $h_{ji} = 0$ , de manera que  $x_i \circ h_{ji} = x_i$  i per tant es compleix la condició  $(x_i \circ h_{1i} = x_i) \vee (x_i \circ h_{2i} = x_i)$  de pertinença a  $H_x^*$ .

b) Si  $(y, z) \in H_x^*$  podem definir  $h_1 = x \circ y$ ,  $h_2 = x \circ z$ . Aquest  $(h_1, h_2)$  compleix:

1)  $(h_1, h_2) \in H^*$  és a dir  $\forall i (h_{1i} = 0) \vee (h_{2i} = 0)$ . Efectivament: per tot locus  $i$ , si  $y_i \neq z_i$ , segons la definició de  $H_x^*$ , o bé  $y_i = x_i$  i aleshores  $h_{1i} = x_i \circ z_i = 0$  segons la taula de  $\circ$ , o bé  $z_i = x_i$  i aleshores  $h_{2i} = x_i \circ z_i = 0$ .

2)  $(x \circ h_1, x \circ h_2) = (y, z)$ . Efectivament, donada l'associativitat de  $x \circ h_1 = x \circ (x \circ y) = (x \circ x) \circ y = y$ , ja que  $\forall i, i=1, \dots, L$ ,  $x_i \circ x_i = 0$ . La demostració de  $x \circ h_2 = z$  és idèntica.

3) És únic. Efectivament, si  $(h_1, h_2) \in H^*$  compleixen  $x \circ h_1 = y$ ,  $x \circ h_2 = z$  ha de ser forçosament  $h_1 = x \circ y$ ,  $h_2 = x \circ z$ , ja que com  $\circ$  és

$$x \circ h_1 = y \quad \Rightarrow \quad x \circ (x \circ h_1) = x \circ y \quad \Rightarrow \quad h_1 = x \circ y$$

$$x \circ h_2 = z \quad \Rightarrow \quad x \circ (x \circ h_2) = x \circ z \quad \Rightarrow \quad h_2 = x \circ z$$

Com hem dit,  $H^*$  caracteritza totalment el conjunt de genotips que poden produir qualsevol gamet  $x$ . Podria ser anomenat "patró d'heterozigosis possibles" ja que donat qualsevol  $x$  ens permet definir tots els genotips,  $(x \circ h_1, x \circ h_2)$  que poden pro-

duir  $x$ . El nombre d'elements de  $H^*$  (i per tant de  $H_x^*$ ) es pot caracteritzar mitjançant el següent

Teorema 5.1.2

$H^*$  té  $3^L$  elements. On  $L$  recordem que és el nombre de loci.

Efectivament, els elements de  $H^*$  a cada posició, presenten sengles 0 o bé un 1 i un 0. De les  $L$  posicions, n'hi pot haver 0, 1, 2, ... fins a  $L$  amb un 0 i un 1 (0/1). Si hi ha  $k$  ( $k=0, 1, \dots, L$ ) posicions amb 0/1, aquestes poden presentar  $2^k$  configuracions possibles (d'acoblament-repulsió dels 0 i 1). Per altra banda hi ha  $\binom{L}{k}$  maneres de repartir  $k$  0/1 en  $L$  posicions, de manera que el total d'elements de  $H^*$  serà

$$\binom{L}{0}2^0 + \binom{L}{1}2^1 + \dots + \binom{L}{L}2^L = (2+1)^L = 3^L \quad .$$

El teorema 5.1.1 permet reescriure les equacions (5.1.9)

com

$$p'_x = m_w^{-1} \sum_{(h_1, h_2) \in H^*} p(x\ominus h_1) p(x\ominus h_2) w(x\ominus h_1, x\ominus h_2) R^*(x/x\ominus h_1, x\ominus h_2) \quad (5.1.10)$$

que serien fàcilment computables si coneguéssim  $R^*(x/x\ominus h_1, x\ominus h_2)$  ja que l'operació  $x\ominus h$  és fàcilment realitzable a un ordinador digital, i tant  $x\ominus h_1$  com  $x\ominus h_2$  són l'expressió en el sistema de numeració binari de sengles números naturals que especifiquen un element del vector  $(p_x)$  de freqüències gèniques o un element de la matriu  $(w_{xy})$  d'eficàcies.

Un avenç important el proporciona el següent

Teorema 5.1.3

Per  $(h_1; h_2) \in H^*$ ,  $x \in G$ ,  $R^*(x/x\phi h_1, x\phi h_2)$  és independent del gamet  $x$ , i podria ser indicat sense ambigüitat com  $R^*(h_1, h_2)$

Efectivament, definim la variable aleatòria  $C_i$  com

$$C_i = \begin{cases} 0 & \text{si, donat un genotip, es "llegeix" la prime-} \\ & \text{ra cromàtida, al locus } i, \text{ durant la forma-} \\ & \text{ció d'un gamet} \\ 1 & \text{idem, si es llegeix la segona cromàtida} \end{cases}$$

Si  $\xi_i$  indica un dels nombres 0 o 1, la seqüència  $(C_1 = \xi_1, C_2 = \xi_2, \dots, C_L = \xi_L)$  es pot interpretar com una "caminata aleatòria" o conjunt de canvis de cromàtida que, donat un genotip, genera un gamet.  $R^*(x/x\phi h_1, x\phi h_2)$ , la probabilitat que a partir del genotip  $(x\phi h_1, x\phi h_2)$  es generi  $x$ , és la summa de les probabilitats de totes aquelles "caminates aleatòries" que produeixen  $x$ :

$$R^*(x/x\phi h_1, x\phi h_2) = \sum_{\{\xi_1, \dots, \xi_L\}} \text{Prob}(C_1 = \xi_1, C_2 = \xi_2, \dots, C_L = \xi_L)$$

on  $\xi_i$  pot ser

$$\xi_i = \begin{cases} 0 \text{ o } 1 & \text{si } x_{i\phi h_{1i}} = x_{i\phi h_{2i}} \text{ (és a dir, si al lo-} \\ & \text{cus } i, \text{ el genotip és homoz. amb el ma-} \\ & \text{teix al.lel que } x). \\ 0 & \text{si } x_{i\phi h_{1i}} = x_{i\phi h_{2i}} \text{ (si al locus } i \text{ és he-} \\ & \text{terozigot, i la 1}^{\text{a}} \text{ cromàtida té el ma-} \\ & \text{teix al.lel que } x) \\ 1 & \text{si } x_{i\phi h_{1i}} \neq x_{i\phi h_{2i}} \text{ (si al locus } i \text{ és he-} \\ & \text{terozigot i la 2}^{\text{a}} \text{ cromàtida té el ma-} \\ & \text{teix al.lel que } x) \end{cases}$$

Les condicions anteriors són independents de  $x$ , ja que es podrien indicar

$$\xi_i = \begin{cases} 0 & \text{si } h_{1i} = h_{2i} = 0 \\ 0 & \text{si } h_{1i} = 0, h_{2i} = 1 \\ 1 & \text{si } h_{1i} = 1, h_{2i} = 0 \end{cases}$$

i per tant  $R^*(x/x\ominus h_1, x\ominus h_2)$  es podria indicar  $R^*(h_1, h_2)$  sense fer esment del gamet concret  $x$ .

A partir del teorema 5.1.3 podem escriure

$$p'_x = m_w^{-1} \sum_{(h_1, h_2) \in H^*} p(x\ominus h_1) p(x\ominus h_2) W(x\ominus h_1, x\ominus h_2) R^*(h_1, h_2) \quad (5.1.11)$$

Tornant a l'exemple anterior, pel cas  $L=3$ , podem comprovar que el conjunt de patrons d'heterozigosi

$$H^* = \{(h_1, h_2) \in G \times G \mid \forall i, h_{1i} = 0 \vee h_{2i} = 0\}$$

té  $3^3 = 27$  elements que són

$$H^* = \{000/000 \text{ (cap locus heterozigot)}, 000/001, 000/010, 000/100, \\ 000/011, 000/101, 000/110, 000/111, 001/000, 001/010, \\ 001/100, 001/110, 010/000, 010/001, 010/100, 010/101, \\ 100/000, 100/001, 100/010, 100/011, 011/000, 011/100, \\ 101/000, 101/010, 110/000, 110/001, 111/000\}$$

els que no tenen cap 1 produeixen homozigots a tots els loci, els que tenen un 1 (per exemple 000/001) produeixen genotips amb tots els loci homozigòtics llevat d'un, etc. Operant, mitjançant l'operació  $\ominus$ , un gamet  $x$  qualsevol amb tots els ele-

ments de  $H^*$  podem obtenir el conjunt  $H_x^*$  dels genotips que el poden produir per segregació o recombinació. Per exemple per  $x=101$  podem obtenir el conjunt  $H_{101}^*$  indicat alguns paràgrafs enrera d'aquesta secció:

$$101\oplus000 / 101\oplus000 = 101 / 101$$

$$101\oplus000 / 101\oplus001 = 101 / 100$$

$$101\oplus000 / 101\oplus010 = 101 / 111$$

$$101\oplus000 / 101\oplus100 = 101 / 001$$

etc.

fins a:

$$101\oplus111 / 101\oplus000 = 010 / 101$$

Tal com podem apreciar en aquest darrer exemple, el conjunt  $H^*$  (i per tant  $H_x^*$ ) té tots els elements "repetits", excepte  $(h,h)$  amb  $h=(0,0,\dots,0)$ , en el sentit que per tot  $(h_1,h_2)\in H^*$ , existeix un altre  $(h'_1,h'_2)\in H^*$  tal que  $h_1=h'_2$  i  $h_2=h'_1$ . Evidentment, per aquests elements  $p(x\oplus h_1)p(x\oplus h_2) = p(x\oplus h'_1)p(x\oplus h'_2)$ ,  $w(x\oplus h_1, x\oplus h_2) = w(x\oplus h'_1, x\oplus h'_2)$  (si, per exemple no es considera cap "efecte matern") i  $R(h_1, h_2) = R(h'_1, h'_2)$ , de manera que podríem definir un nou conjunt sense "repeticions", per exemple:

$$H = \{(h_1, h_2)\in H^* \mid \text{ent}(h_1) \leq \text{ent}(h_2)\}$$

on  $\text{ent}(h)$  vol dir "el nombre enter representat per la seqüència binària  $h$ ".  $H$  té  $(3^L-1)/2 + 1 = (3^L+1)/2$  elements. Si fem la convenció

$$R(h_1, h_2) = 2 \cdot R^*(h_1, h_2) \quad \text{si } h_1 \neq h_2$$

$$R(h_1, h_2) = R^*(h_1, h_2) \quad \text{si } h_1 = h_2$$

podem escriure finalment:

$$p'_x = m_w^{-1} \sum_{(h_1, h_2) \in H} p(x\phi h_1) p(x\phi h_2) w(x\phi h_1, x\phi h_2) R(h_1, h_2) \quad (5.1.12)$$

amb  $(3^L+1)/2$  sumands.

Com hi ha  $2^L$  gamets i cada equació del tipus (5.1.12) té  $(3^L+1)/2$  sumands, caldria fer un total de  $2^L(3^L+1)/2 = (3^L+1) \cdot 2^{L-1}$  operacions bàsiques de multiplicar  $p(x\phi h_1)p(x\phi h_2)w(x\phi h_1, x\phi h_2)R(h_1, h_2)$  i sumar-ho al terme corresponent, això segons el nombre de loci, implica les següents quantitats d'operacions

L	$(3^L+1)2^{L-1}$
2	20
3	112
4	656
5	3904
6	23.360
7	140.032
8	839.936
9	5.039.104
10	30.233.600
15	$\sim 2.35 \times 10^{11}$
20	$\sim 1.83 \times 10^{15}$
30	$\sim 1.1 \times 10^{23}$

taula 5.1.2.

progressivament grans a mesura que augmenta  $L$  però també progressivament menors que les indicades a la taula 5.1.1 pel mètode dels operadors genètics. De tota manera, per nombres grans de loci es fa palesa la inefectivitat d'aquests mètodes enfront a la simulació finita.

El càlcul efectiu de les freqüències gamètiques d'una generació  $p'$ , donades les freqüències a la generació anterior  $\{p_y\}_{y \in G}$  i la matriu d'eficàcies  $W = (w_{yz})_{(y,z) \in G \times G}$ , es pot fer si podem trobar tots els elements de  $H$  i avaluar  $R(h_1, h_2)$  per tot  $(h_1, h_2) \in H$ . Aquestes dues qüestions les tractem tot seguit.

Els següent algorisme, expressat en un llenguatge de programació imaginari (però confiem que intel·ligible per la majoria de programadors) permet trobar tots els elements de  $H$ . Una versió en llenguatge BASIC, com és habitual molt menys intel·ligible, és troba a les sentències 10515 a 10525 del programa documentat a l'apèndix 4.

```

DEFINIM I,J,K,L : ENTERS; COMENTARI(L es el nombre de loci)
LLEGIM(L);
DEFINIM PARELL : AGREGAT DE (2) ENTERS
H : AGREGAT DE ((3L+1)/2) PARELLS;

K=∅;
PER I=∅ FINS 2L-1;
  PER J=I FINS 2L-1;
    FEM
      SI BIN(I) ^ BIN(J) = BIN(∅) ALESHORES
        FEM
          K=K+1;
          H(K)=(I,J);
        PROU

```

SEGÜENT J;  
 SEGÜENT I;  
 COMENTARI(BIN( ) indica la representació binària. L'operació es realitza dígit a dígit. Juguem amb el fet que cada enter I o J té una representació binària que es pot considerar un element de G i viceversa).

El tractament de les probabilitats de recombinació que indiquem a continuació és de caire més aviat estadístic, enfront al de Lewontin (1964a) més aviat de tipus algebràic.

No oblidem que  $R(h,h') = R(x/x\text{ò}h,x\text{ò}h')$ , la probabilitat que el genotip  $(x\text{ò}h,x\text{ò}h')$  produeixi per recombinació o segregació  $x$ , és independent del gamet  $x$  i només depèn del patró d'heterozigosi  $(h,h')$ . Deixarem de banda el cas trivial  $h=h'=(0,0,0,\dots,0,0)$  on  $R(h,h')=1$  i considerarem aquesta funció per aquells patrons d'heterozigosi que impliquem, com a mínim, un locus heterozigòtic.

La parella  $(h,h')$  és equivalent a  $(h,l)$  on  $l=(l_1,l_2,\dots,l_s)$ , essent  $s$  el nombre de loci heterozigòtics ( $1 \leq s \leq L$ ) i  $l_j$  la posició (número d'ordre) del  $j$ -éssim locus heterozigòtic ( $1 \leq j \leq s$ ,  $1 \leq l_j \leq L$ , llevat del cas  $s=0$ ). Sigui  $NQ_{jk}$  la variable aleatòria 'nombre de quiasmes entre el locus  $l_j$  (exclòs) i el  $l_k$  (inclòs)' i  $NQ_{01}$  la variable aleatòria 'nombre de quiasmes entre l'origen del gamet i el primer locus heterozigòtic  $l_1$ '. Definim  $I=\{1,3,5,7,\dots\}$  i  $P=\{0,2,4,6,\dots\}$ . L'esdeveniment  $NQ_{jk} \in I$  implica un canvi de cromàtida des del locus  $h_j$  al  $h_k$  a partir de les quals es copia el gamet. L'esdeveni-



ment  $NQ_{jk} \in I$  implica un canvi de cromàtida des del locus  $h_j$  al  $h_k$  a partir de les quals es copia el gamet. L'esdeveniment  $NQ_{jk} \in P$  no implica cap canvi. Definim finalment una variable aleatòria  $C$  tal que

$$C = \begin{cases} 0 & \text{si el gamet s'inicia pel primer gamet } x \oplus h \\ 1 & \text{si el gamet s'inicia pel segon gamet } x \oplus h' \end{cases}$$

Si  $C=0$  (s'ha iniciat la rèplica pel primer gamet) i  $h_{1_1}=0$  (no és el primer gamet el que té l'al·lel diferent del de  $x$  al primer locus heterozigòtic) aleshores cal que  $NQ_{01} \in P$  (no hi hagi canvi de cromàtida) des de l'origen al locus  $1_1$ , perquè es copiï  $x$ . Si  $C=0$  i  $h_{1_1}=1$  (s'ha començat pel primer cromosoma i al primer locus heterozigòtic aquest primer cromosoma té un al·lel diferent del de  $x$ ) caldrà que  $NQ_{01} \in I$ . Igualment, si  $C=1$  i  $h_{1_1}=0$  aleshores s'ha de complir  $NQ_{01} \in I$  i si  $C=1$  i  $h_{1_1}=1$   $NQ_{01} \in P$ . La condició  $NQ_{01} \in P$  es pot expressar com  $NQ_{01} = 2i_1$  amb  $i_1=0,1,2,3,4,5,\dots$  i la condició  $NQ_{01} \in I$  com  $NQ_{01}=2i_1+1$ . En forma general podem enunciar que, perquè es copiï  $x$  fins a  $1_1$  cal que es compleixi l'esdeveniment  $NQ_{01} = 2i_1 + (C \oplus h_{1_1})$ , per  $i_1=0,1,2,\dots$ . Més generalment, per qualsevol parella de loci heterozigòtics en posicions  $1_{j-1}$  i  $1_j$  ( $j=2,3,\dots,s$ ), la condició per la rèplica de  $x$  entre ells (exclòs  $1_{j-1}$ , inclòs  $1_j$ ) equival a

$$NQ_{j-1,j} = 2i_j + (h_{1_{j-1}} \oplus h_{1_j}) \quad \text{per } i_j = 0,1,2,3,\dots$$

(nombre parell de quiasmes si  $h_{1_{j-1}} = h_{1_j}$ , no cal canviar de cromàtida replicada, i senar si  $h_{1_{j-1}} \neq h_{1_j}$ , cal canviar).

Per tant, la probabilitat  $R(h, h')$  és igual a la probabilitat de la unió de tots els esdeveniments del tipus

$$[C=c, NQ_{01}=2i_1+(c\oplus h_{1_1}), NQ_{12} = 2i_2+(h_{1_1} \oplus h_{1_2}), \dots, NQ_{s-1,s} = 2i_s+(h_{1_{s-1}} \oplus h_{1_s})]$$

per tots els  $i_j=0,1,2,3,\dots$  i  $j=1,2,\dots,s$ , és a dir

$$R(h, h') = \sum_{c=0}^1 \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \dots \sum_{i_s=0}^{\infty} \text{Prob}[C=c, NQ_{01}=2i_1+(c\oplus h_{1_1}), \dots, NQ_{s-1,s} = 2i_s + (h_{1_{s-1}} \oplus h_{1_s})] \quad (5.1.13)$$

Si  $\text{Prob}[C=0] = \text{Prob}[C=1] = 0.5$ , és a dir que no hi ha cap mena d'impuls meiótic, i el nombre de quiasmes és independent del cromosoma inicial, aleshores

$$R(h, h') = 0.5 \sum_{c=0}^1 \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \dots \sum_{i_j=0}^{\infty} \text{Prob}[NQ_{01}=2i_1+(c\oplus h_{1_1}), \dots, NQ_{s-1,s} = 2i_s + (h_{1_{s-1}} \oplus h_{1_s})] \quad (5.1.13\text{-bis})$$

i com la summa  $\sum_{c=0}^1 \sum_{i_1=0}^{\infty}$  s'extén a tots els possibles valors que pot prendre  $NQ_{01}$ , l'expressió (5.1.13) equival a

$$R(h, h') = 0.5 \sum_{i_2=0}^{\infty} \sum_{i_3=0}^{\infty} \dots \sum_{i_s=0}^{\infty} \text{Prob}[NQ_{12}=2i_2+(h_{1_1} \oplus h_{1_2}), \dots, \dots, NQ_{s-1,s} = 2i_s + (h_{1_{s-1}} \oplus h_{1_s})] \quad (5.1.14)$$

Pel que s'ha dit en introduir les equacions (5.1.12), en realitat  $R(h, h')$  ha d'esser el doble, per tant s'ha d'eliminar el factor 0.5 de l'expressió 5.1.14.

Cal conèixer la distribució conjunta de les variables aleatòries  $NQ_{12}, NQ_{23}, \dots, NQ_{s-1,s}$  les quals, per separat podriem tenir, per exemple, distribució de Poisson amb paràmetre  $\lambda_{j-1,j} = \lambda \Delta(1_{j-1}, 1_j)$  proporcional a la distància genètica entre els loci  $1_{j-1}$  i  $1_j$ ; o distribució binomial negativa, etc. Aleshores es tractaria d'un problema de construcció d'una distribució conjunta donades unes marginals.

Si en una primera aproximació considerem independència estocàstica entre els nombres de quiasmes als diferents segments del cromosoma, obtenim

$$R(h, h') = \prod_{j=2}^s \left( \sum_{i_j=0}^{\infty} \text{Prob}[NQ_{j-1,j}=2i_j+(h_{1_{j-1}} \oplus h_{1_j})] \right) \quad (5.1.15)$$

una funció de les fraccions de recombinació  $r(1_{j-1}, 1_j)$  entre els loci  $1_{j-1}$  i  $1_j$  per  $j=2, 3, \dots, s$ , ja que

$$\sum_{i_j=0}^{\infty} \text{Prob}[NQ_{j-1,j}=2i_j+(h_{1_{j-1}} \oplus h_{1_j})] = \begin{cases} r(1_{j-1}, 1_j) & \text{si } h_{1_{j-1}} = h_{1_j} \\ 1-r(1_{j-1}, 1_j) & \text{en cas contrari} \end{cases} \quad (5.1.16)$$

en general

$$R(h,h') = \prod_{j=2}^s r(l_{j-1}, l_j)^{h_1 j-1 \oplus h_1 j} (1-r(l_{j-1}, l_j))^{1-h_1 j-1 \oplus h_1 j} \quad (5.1.17)$$

El coneixement de tots els  $R(h,h')$  implicarà el coneixement de  $(L-1)!$  paràmetres  $r(l_{j-1}, l_j)$ . Si a més, es considera uniformitat de totes les zones cromosòmiques quant a probabilitats de recombinació i loci veïns tots a la mateixa distància, tots els paràmetres del tipus  $r(l, l+k)$  per tot  $1 \leq l \leq L$  i  $0 \leq k \leq L$  tal que  $1 \leq l+k \leq L$  es redueixen a un sol  $r(k)$  i tots els  $R(h,h')$  depenen de  $L-1$  paràmetres,  $r(1), r(2), \dots, r(L-1)$  corresponents a les fraccions de recombinació entre loci a  $1, 2, \dots, L-1$  intervals interloci de distància, respectivament

$$R(h,h') = \prod_{j=2}^s r(l_{j-1} - l_{j-1})^{h_1 j-1 \oplus h_1 j} (1-r(l_{j-1} - l_{j-1}))^{1-h_1 j-1 \oplus h_1 j} \quad (5.1.18)$$

aquesta és l'expressió que utilitzen a les simulacions deterministes descrites en aquesta secció.

Com a exemple final podríem deduir les equacions de Lewontin i Kojima pel cas de dos loci lligats, amb fracció de recombinació  $r$  entre ells, com un cas particular de les equacions 5.1.12

L'aplicació de l'algorisme d'obtenció de tots els elements de  $H$  ens donaria el conjunt

$$H = \{(0,0),(0,1),(0,2),(0,3),(1,2)\}$$

o, millor dit,

$$H = \{(00,00),(00,01),(00,10),(00,11),(01,10)\}$$

amb  $(3^2+1)/2 = 5$  elements o patrons d'heterozigosi possibles.

$R(00,00) = 1$  ja que a partir de l'homozigot és segur que es produirà el gamet amb la mateixa constitució.  $R(00,01) = 2(\frac{1}{2}(1-r) + \frac{1}{2}r) = 1$ , si el segon locus és heterozigòtic, es pot iniciar el gamet pel cromosoma "correcte" i no recombinar  $\frac{1}{2}(1-r)$  o bé per l'"incorrecte" i recombinar  $\frac{1}{2}r$ , per produir el gamet. Multipliquem per 2 ja que tractant-se de dos patrons d'heterozigosi diferents caldria considerar el cas 00,01 i el 01,00. Igualment  $R(00,10) = 2\frac{1}{2} = 1$ ,  $R(00,11) = 2\frac{1}{2}(1-r) = 1-r$  i  $R(01,10) = 2\frac{1}{2}r = r$ .

$$\begin{aligned} p'_{00} &= (p_{00\oplus00}p_{00\oplus00} W(00\oplus00,00\oplus00) R(00,00) + \\ &+ p_{00\oplus00}p_{00\oplus01} W(00\oplus00,00\oplus01) R(00,01) + \\ &+ p_{00\oplus00}p_{00\oplus10} W(00\oplus00,00\oplus10) R(00,10) + \\ &+ p_{00\oplus00}p_{00\oplus11} W(00\oplus00,00\oplus11) R(00,11) + \\ &+ p_{00\oplus01}p_{00\oplus10} W(00\oplus01,00\oplus10) R(01,10)) m_w^{-1} = \\ &= (p_{00}^2 W(00,00) + p_{00}p_{01} W(00,01) + p_{00}p_{10} W(00,10) + \\ &+ p_{00}p_{11} W(00,11)(1-r) + p_{01}p_{10} W(01,10)r) m_w^{-1} = \\ &= (p_{00}[p_{00} W(00,00) + p_{01} W(00,01) + p_{10} W(00,10) + p_{11} W(00,11)] - \\ &- p_{00}p_{11} W(00,11)r + p_{01}p_{10} W(01,10)r) m_w^{-1} = \end{aligned}$$

$$= (p_{00}W_0 - DW_{12}r) m_w^{-1}$$

$$p'_0 = p_0 - m_w^{-1}(p_0(w_0 - m_w) - DW_{12}) \text{ fent } w_0 = p_{00}W(00,00) + \\ + p_{01}W(00,01) + p_{10}W(00,10) + p_{11}W(00,11)$$

$$W(00,11) = W(01,10) = W_{12}$$

$$D = p_{00}p_{11} - p_{01}p_{10}$$

igualment

$$p'_{01} = (p_{0100}p_{0100}W(0100,0100)R(00,00) + \\ + p_{0100}p_{0101}W(0100,0101)R(00,01) + \\ + p_{0100}p_{0110}W(0100,0110)R(00,10) + \\ + p_{0100}p_{0111}W(0100,0111)R(00,11) + \\ + p_{0101}p_{0110}W(0101,0110)R(01,10))m_w^{-1} = \\ = (p_{01}p_{01}W(01,01) + p_{01}p_{00}W(01,00) + p_{01}p_{11}W(01,11) + \\ + p_{01}p_{10}W(01,10)(1 - ) + p_{00}p_{11}W(00,11))m_w^{-1} = \\ = (p_{01}(p_{01}w(01,01) + p_{00}w(01,00) + p_{11}w(01,11) + \\ + p_{10}w(01,10)) - p_{01}p_{10}w(01,10) + p_{00}p_{11}w(00,11) \\ r)m_w^{-1} = \\ = (p_{01}W_1 + DW_{12}r)m_w^{-1}$$

$$p'_1 = p_1 + m_w^{-1}(p_1(w_1 - m_w) + DW_{12}r)$$

i igualment s'obtidrien les expressions corresponents pels cromosomes 10(2) i 11(3).

5.2. Inversions cromosòmiques i subdivisió del medi al model determinista d'una població mendeliana.

Les equacions (5.1.12) es poden modificar fàcilment per estudiar la dinàmica de canvi d'una inversió cromosòmica apareguda de nou, i les característiques dels possibles estats polimòrfics d'equilibri.

Considerarem l'existència d'un tipus gamètic invertit I, amb la inversió incloent tots els loci estudiats. Caldrà considerar que hi ha  $2^L - 1 + 1 = 2^L$  gamets possibles,  $2^L - 1$  estàndard ( $x=0,1,\dots,2^L-1$ ) i un d'invertit I amb composició gènica constant ja que considerarem la inversió produïda un sol cop (no recurrent) i despreciarem (no considerarem al model) els efectes de la mutació i les dobles recombinacions dins de la zona invertida.

Sigui  $x$  un gamet estàndard qualsevol. Les equacions (5.1.12) es poden reescriure, pel cas d'inversions:

$$p'_x = m_w^{-1} \left\{ \left( \sum_{(h,h') \in H} p_{x \oplus h} p_{x \oplus h'} W(x \oplus h, x \oplus h') R(h, h') \right) + p_x q W(x, I) \right\}$$

per  $x=0,1,\dots,2^L-1$

$$q' = m_w^{-1} \left\{ \left( \sum_x p_x q W(x, I) \right) + q^2 W(I, I) \right\} \quad (5.2.1)$$

on  $q$  indica la freqüència gamètica de l'invertit.

Tot gamet estàndard  $x$  el pot produir qualsevol genotip homocariotípic per estàndard, del tipus  $(x\bar{h}, x\bar{h}')$ , per recombinació o segregació (primer terme de la dreta de la primer igualtat de 5.2.1) o el genotip heterocariotípic format per l'invertit i l'estàndard  $x$  (segon terme de la dreta de la primera igualtat de 5.2.1), solament per segregació, ja que exclouem tota possibilitat de recombinació als heterocariotips. Igualment, el gamet invertit podrà ésser produït només per tot heterocariotip, per segregació (primer terme de la dreta de la segona igualtat de 5.2.1), o per l'homocariotip invertit, per segregació amb probabilitat 1.

En principi, als heterocariotips  $xI$ , la probabilitat que un gamet concret sigui  $I$  (o  $x$ , és el mateix) és 0.5, però com la probabilitat de formar un heterocariotip  $xI$  és  $2p_x q$ , el terme  $2p_x q \cdot 0.5W(x, I)$  l'indiquem a les equacions (5.2.1) com  $p_x q W(x, I)$ .

Com suposem que la inversió no té cap efecte selectiu en si (per exemple efecte de posició o reducció de la fecunditat), totes les expressions derivades a l'apartat anterior i on hi intervenia l'eficàcia mitjana  $m_w$  són igualment vàlides, amb l'eficàcia mitjana calculada tenint en compte que hi ha un gamet més

$$m_w = \sum_{(x,y) \in G \times G} W(x,y) p_x^* p_y^*$$



$$\text{amb } \begin{cases} p_{\xi}^* = p_{\xi} & \text{si } \xi \neq I \\ p_{\xi}^* = p_{\xi} + q & \text{si } \xi = I, \text{ per } \xi = x, y \end{cases}$$

si volguéssim considerar la possibilitat d'una reducció de fecunditat als heterozigots estructurals o d'un efecte de posició, la situació es complicaria més, podríem per exemple considerar l'existència d'un factor d'eficàcia  $w^*(\xi, \eta)$  amb  $w^*(x, y) = 1$ , pels genotips estàndard cromosòmicament i  $w^*(x, I) = w^*(I, x)$ ,  $w^*(I, I)$  no necessàriament 1. L'eficàcia mitjana s'hauria de calcular

$$m_w = \sum_{(x,y) \in G \times G} w(x,y) p_x p_y + 2 \sum_{x \in G} w(x,I) w^*(x,I) p_x q + w(I,I) w^*(I,I) q^2$$

i les equacions de recursió serien:

$$p'_x = m_w^{-1} \left\{ \left( \sum_{(h,h') \in H} p_{x\emptyset h} p_{x\emptyset h'} w(x\emptyset h, x\emptyset h') R(h, h') \right) + p_x q w(x, I) w^*(x, I) \right\}$$

per  $x=0, 1, \dots, 2^L - 1$

$$q' = m_w^{-1} \left\{ \left( \sum_{x \in G} p_x q w(x, I) w^*(x, I) \right) + q^2 w(I, I) w^*(I, I) \right\} \quad (6.2.1\text{-bis})$$

Si al model determinista estudiat a l'apartat anterior (població diploide,  $L$  loci lligats, una inversió, aparellament a l'atzar, medi uniforme quant a selecció) relaxem la suposició d'uniformitat del medi i la substituïm per l'existència de  $n$  "clapes" o "taques" on un mateix genotip  $xy$  pot tenir diferent

eficàcia ( $w_i(x,y)$ , per  $i=1,\dots,n$ , indicarà l'eficàcia del genotip  $xy$  quan és a la clapa  $i$  de les  $n$  en què està subdividit el medi) i amb els zigots, després de produïts a l'aparellament aleatori (sense cap subdivisió quant a l'aparellament), repartits a l'atzar per les  $n$  clapas, on experimentaran la selecció, tindrem la situació plantejada per Levene (1953) però per un nombre indeterminat ( $L \geq 1$ ) de loci i una inversió. Això permetrà estudiar alguns aspectes de la dinàmica determinista de canvi de la freqüència d'una ordenació cromosòmica apareguda de nou, així com l'existència i les característiques dels possibles equilibris polimòrfics, sota condicions de subdivisió del medi.

Suposem que la clapa  $i$  té una "capacitat"  $c_i$ , en el sentit que una proporció  $c_i$  dels zigots hi va a parar i la mateixa proporció dels adults reproductors provenen d'ella. Evidentment  $\sum_{i=1}^n c_i = 1$ . Indicarem com  $p_i^1(x)$  la proporció de gamets "estàndard"  $x$  a la clapa  $i$  després de la selecció i com  $q_i^1$  la proporció de gamets "invertits" a la mateixa clapa  $i$  també després de la selecció. Com la dispersió dels zigots és totalment aleatòria, cada clapa rebrà les mateixes freqüències gamètiques  $p_x$ , de manera que podrem escriure, directament a partir de (5.3.1)

$$p_i^1(x) = m_w^{-1} \left\{ \left( \sum_{h,h' \in H} p_{x\emptyset h} p_{x\emptyset h'} w_i(x\emptyset h, x\emptyset h') R(h,h') \right) + p_x q w_i(x,I) \right\}$$

per  $x=0,1,\dots,2^L-1$

$$q_i^1 = m_w^{-1} \left\{ \left( \sum_x p_x q w_i(x,I) \right) + q^2 w_i(I,I) \right\} \quad (5.2.2)$$

on  $m_{w_i}$  indica l'eficàcia mitjana a la clapa  $i$ . Aplicant la fórmula de les probabilitats totals podem calcular la freqüència de cada gamet a la població global d'adults reproductors:

$$p_x = \sum_{i=1}^n c_i p_i(x), \quad \text{per } x=0, \dots, 2^L-1$$

$$q = \sum_{i=1}^n c_i q_i$$

és a dir:

$$p_x = \sum_{i=1}^n c_i m_{w_i}^{-1} \left\{ \left( \sum_{h, h' \in H} p_{x\emptyset h} p_{x\emptyset h'} w_i(x\emptyset h, x\emptyset h') R(h, h') \right) + p_x q w_i(x, I) \right\}$$

per  $x=0, \dots, 2^L-1$

$$q = \sum_{i=1}^n c_i m_{w_i}^{-1} \left\{ \left( \sum_x p_x q w_i(x, I) \right) + q^2 w_i(I, I) \right\} \quad (5.2.3)$$

que, intercanviant els sumatoris  $i$  i traient factor comú es pot convertir en

$$p_x = \sum_{h, h' \in H} p_{x\emptyset h} p_{x\emptyset h'} R(h, h') \left\{ \sum_{i=1}^n c_i m_{w_i}^{-1} w_i(x\emptyset h, x\emptyset h') \right\} +$$

$$+ p_x q \left\{ \sum_{i=1}^n c_i m_{w_i}^{-1} w_i(x, I) \right\}$$

$$q' = \sum_x p_x q \left\{ \sum_{i=1}^n c_i m_{w_i}^{-1} w_i(x, I) \right\} + q^2 \left\{ \sum_{i=1}^n c_i m_{w_i}^{-1} w_i(I, I) \right\} \quad (5.2.4)$$

Aquesta darrera expressió ajuda a comprendre la diferència entre les formulacions de Levene (1953) i de Dempster (1955) (recordem l'apartat 3.4.4) i perquè han estat qualificades de "selecció

tova" i de "selecció dura" respectivament (la discussió que se-  
gueix es podria realitzar en absència d'inversions i per un sol  
locus), efectivament:

Indiquem  $m_w = \sum_{i=1}^n c_i m_{w_i}$ , l'eficàcia mitjana global de  
la població. Aquesta definició és correcta si tenim en compte  
que  $m_{w_i}$  és una mesura proporcional a la probabilitat de sobre-  
viure i arribar a reproduir-se d'un zigot qualsevol, condicionat  
que és a la clapa  $i$ , i  $c_i$  és la probabilitat que un zigot hi  
sigui. Dividint i multiplicant per  $m_w$  la part de la dreta de  
les equacions (6.4.3) obtindrem

$$p'_x = m_w^{-1} \left( \sum_{h, h' \in H} p_{x\bar{e}h} p_{x\bar{e}h'} R(h, h') w(x\bar{e}h, x\bar{e}h') + p_x q_w(x, I) \right) \quad \text{per } x=0, \dots, 2^L-1$$

$$q' = m_w^{-1} \left( \sum_x p_x q_w(x, I) + q^2 w(I, I) \right) \quad (5.2.5)$$

amb

$$w(\xi, \eta) = \sum_{i=1}^n c_i m_{w_i}^{-1} w_i(\xi, \eta) m_w \quad (5.2.6)$$

(on  $\xi$  i  $\eta$  poden indicar  $x\bar{e}h$ ,  $x\bar{e}h'$ ,  $x$  o  $I$ ) és a dir, que el  
model amb medi subdividit és equivalent a un model amb medi uni-  
forme al qual definíssim les eficàcies segons (5.2.6). Però  
(5.2.6) es pot indicar com

$$w(\xi, \eta) = m_w \sum_{i=1}^n c_i \frac{w_i(\xi, \eta)}{m_{w_i}} \quad (5.2.7)$$

Una ullada a l'expressió 5.2.7 permet treure algunes conclu-

sions, que caldrà comprovar i precisar estudiant més a fons les equacions de canvi de les freqüències gamètiques.

Primerament, les eficàcies són dependents de les freqüències genotípiques, ja que depenen dels valors  $m_{w_i} = \sum_x \sum_y p_x p_y w_i(x,y)$ . La dependència entre eficàcies i freqüències sembla que ha de tendir a la uniformització dels valors selectius dels diferents tipus, ja que desafavoreix els que assoleixen més freqüència gràcies al seu valor selectiu inicial superior. Efectivament, considerem per exemple l'evolució del valor selectiu global d'un genotip homozigòtic format a partir de dos gamets invertits I. Si  $\frac{w_i(I,I)}{m_{w_i}}$  és gran (en comparació amb altres valors d'altres genotips) a prou clapes perquè  $\sum_{i=1}^n c_i \frac{w_i(I,I)}{m_{w_i}}$  sigui prou gran, és d'esperar que aquest avantatge selectiu impliqui un increment de la freqüència del gamet I i per tant del genotip (I,I). L'increment de la freqüència del genotip favorable augmentarà l'eficàcia mitjana  $m_{w_i}$  a les mateixes clapes i per tant disminuirà el pes dels termes  $\frac{w_i(I,I)}{m_{w_i}}$  a l'avaluació de l'eficàcia global  $W(I,I) = \sum_{i=1}^n c_i \frac{w_i(I,I)}{m_{w_i}}$  la qual cosa implicarà una disminució relativa de l'eficàcia de (I,I). Tot allò dit anteriorment és vàlid per a qualsevol altre genotip.

Una conseqüència derivada de l'anterior és que sembla més fàcil l'augment ràpid d'un gamet invertit -produït de nou i favorable- a una població no coadaptada, en el sentit que si l'eficàcia mitjana  $m_{w_i}$  és baixa a aquelles clapes on la inversió és favorable, els termes  $\frac{w_i(I, )}{m_{w_i}}$  tindran més pes a l'avalua--

ció de l'eficàcia global dels genotips portadors del gamet invertit. Es podria fer la conjectura que una inversió "favorable" apareguda de nou a una població subdividida generada també aquell moment, amb totes les freqüències gamètiques iguals i en total equilibri gamètic, possiblement incrementarà més de pressa que la mateixa inversió, al mateix tipus de gamet, apareguda de nou amb la mateixa freqüència inicial a una població iniciada en les mateixes condicions que abans però que ja ha experimentat uns cents de generacions de selecció segons un determinat tipus d'ambient, el mateix en el qual es produirà la inversió. Com hem dit, el model és equivalent a un model amb selecció dependent de les freqüències. (Llevat del cas en què  $m_{w_i} = m_{w_j}$  per tota parella de clapes i tota generació, aleshores

$$W(\xi, n) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \frac{m_{w_j}}{m_{w_i}} \quad w_i(\xi, n) = \sum_{i=1}^n c_i w_i(\xi, n) \sum_{j=1}^n c_j =$$

$$= \sum_{i=1}^n c_i w_i(\xi, n)$$

el model és equivalent a un amb medi no subdividit amb eficàcies definides segons l'expressió anterior, igual que el model de Dempster discutit tot seguit) és lògic que si no es tenen en compte les precisions de Wallace (1975) sobre el concepte de selecció tova, en el sentit que és aquella dependent ahora de les freqüències i de la densitat, hom hagi qualificat aquest model (pel cas més simple d'un sol locus o dos, sense inversions) com a model de selecció tova, sense tenir en compte que no depèn de la densitat.

Si, com fa Dempster (1955), la contribució en adults de la clapa  $i$  no es fa constant sinó igual a

$$c_i^* = \frac{c_i^m w_i}{\sum_{j=1}^n c_j^m w_j} \quad \left( \sum_{i=1}^n c_i^* = 1 \right) \quad (5.2.8)$$

és a dir, proporcional a l'eficàcia mitjana a la clapa  $i$ , de manera que

$$p'_x = \sum_{i=1}^n c_i^* p_i(x)$$

$$q' = \sum_{i=1}^n c_i^* q_i$$

mantenint tota la resta igual, l'expressió (5.2.5) segueix essent vàlida, però substituint (5.2.6) per

$$W(\xi, \eta) = \sum_{i=1}^n c_i^* m_{w_i}^{-1} w_i(\xi, \eta) m_w \quad (5.2.9)$$

que desenvolupada és

$$\begin{aligned} W(\xi, \eta) &= \sum_{i=1}^n \frac{c_i^m w_i}{\sum_{j=1}^n c_j^m w_j} m_{w_i}^{-1} w_i(\xi, \eta) \sum_{j=1}^n c_j^m w_j = \\ &= \sum_{i=1}^n c_i w_i(\xi, \eta) \end{aligned} \quad (5.2.10)$$

de manera que el model és equivalent a un model amb medi unifor-

me amb les eficàcies definides segons (5.2.8), totalment independents de la freqüència i de la densitat (selecció dura).

### 5.3. Model de selecció segons allunyament a un òptim i additivitat

Un dels esquemes selectius d'existència real més versemblant i al mateix temps menys estudiat, que en principi podria explicar el manteniment d'alts graus de polimorfisme a les poblacions, és allò que hom ha anomenat "selecció optimitzadora" o "models òptims" o "selecció per a valors intermedis". Segons la definició de Lewontin (1964b), entendrem per "selecció optimitzadora" aquella a la qual, d'un continu de fenotips possibles, n'hi ha un d'intermedi que, per un medi determinat, té la màxima eficàcia, i que tot fenotip desviat d'aquest òptim, és menys eficaç. Encara que a aquesta definició no es diu obertament, segons allò que segueix a la resta d'aquell treball, es sobreentén que l'eficàcia és una funció decreixent de la "distància" a l'òptim, suposant que es pot definir adequadament aquesta "distància" o "allunyament".

La justificació de la selecció segons allunyament a un òptim es pot basar en nombroses consideracions, per exemple termodinàmiques, suposant que hi ha un límit a l'energia de què pot



disposar un organisme per realitzar tots els processos necessaris per sobreviure i reproduir-se i que la màxima probabilitat d'això està associada a un òptim equilibrat d'utilització d'energia per a cada procés. Possiblement Levins (1968) pensa en quelcom semblat quan creu necessari afitar l'integral de la funció d'eficàcies. Relacionada amb l'explicació anterior hi ha l'existència de lligams funcionals entre les components de l'eficàcia, que fan impossible maximitzar d'una manera absoluta una d'elles sense disminuir excessivament les altres.

Una característica interessant dels models òptims és que considerem dos nivells, un, d'expressió fenotípica dels genotips i l'altre, d'eficàcia en funció del fenotip. Aquesta idea ha estat refinada per Levins (1973) qui considera el fenotip com una entitat multidimensional, amb diferents nivells d'integració que culminen en una funció que dóna el valor adaptatiu d'un genotip segons cada ambient.

Les primeres discussions de la relació entre quelcom semblant a la selecció optimitzadora i el manteniment de polimorfismes es troben a Fisher (1930), però amb una certa imprecisió a la definició dels conceptes. Mather (1941, 1942, 1943, 1949) va fer un estudi experimental i de discussió teòrica (però rarament d'anàlisi matemàtica) molt més ampli, donà èmfasi al paper del lligament, suggerí el concepte de "balanç relacional"

i algunes prediccions com ara la formació de gamets "equilibrats" associats a desequilibris gamètics negatius entre loci veïns, suggestió ja apuntada per Fisher.

Bodmer i Parsons (1962) han estudiat matemàticament la teoria de Mather pel cas de dos loci, amb un resultat en certa manera negatiu quant al manteniment del polimorfisme, ja que l'estabilitat dels equilibris polimòrfics s'associa a suposicions addicionals de superioritat dels heterozigòtics.

El model de Wright (1935), encara que per a un cas molt particular, s'acosta molt més a la definició de Lewontin (1964b). Wright considerà el fenotip definit per una única variable i l'eficàcia decreixent segons el quadrat de la diferència entre l'òptim i cada fenotip concret. El fenotip s'avalua a partir del genotip d'una manera totalment additiva o, si no, amb dominiància completa. Els loci estudiats segreguen amb independència total. La conclusió final és que no hi pot haver cap equilibri polimòrfic estable de totes les freqüències gèniques sinó que les configuracions d'equilibri estables són aquelles amb tots els loci fixats excepte, eventualment, un únic segregant en equilibri polimòrfic. Robertson (1956) també arribà a la mateixa conclusió estudiant el mateix model que Wright llevat de considerar altres funcions decreixents de la diferència entre l'òptim i cada fenotip. En canvi, Kojima (1959), pel model de Wright

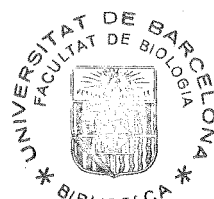
i també sense considerar els loci lligats, demostrà que es pot mantenir cert grau de polimorfisme gènic quan no hi ha additivitat ni dominància completa, sinó una situació intermèdia. Una altra conseqüència important d'aquest estudi és la demostració que, sense cap mena d'interacció entre loci a nivell d'expressió fenotípica, ni lligament, es generen intenses interaccions epistàtiques a nivell de l'eficàcia.

Els resultats de Kojima (1959), completats a Lewontin (1964b), indiquen que hi ha unes condicions bastant restrictives pel polimorfisme, de manera que per mantenir  $L$  loci polimòrfics és necessari que les relacions de dominància i l'òptim, si es representen en un sistema de coordenades cartesianes, defineixin punts dins d'una zona bastant reduïda. En particular, per cada nombre  $L$  de loci, hi ha un llindar que cal que sobrepassi l'òptim perquè puguin ser mantinguts els  $L$  loci en estat polimòrfic. Per sota d'aquest llindar només es podria mantenir un nombre menor de loci polimòrfics. Una altra característica del model òptim quadràtic en condicions de dominància intermèdia és la proximitat de les freqüències d'equilibri a la fixació, a mesura que augmenta el nombre de loci polimòrfics, la qual faria aquests equilibris especialment sensibles a la deriva.

Lewontin (1964b) estudià per primera vegada el paper del lligament en els "models òptims", emprant el mètode numèric dels

"operadors genètics", pels casos de 2 i 5 loci. La introducció de lligament al model quadràtic de Wright implica un increment el grau de polimorfisme dels equilibris, però atenuat en augmentar el nombre de loci, de manera que la situació es fa semblant a la del cas sense lligament.

Tots els resultats indicats anteriorment contradiuen la conjectura que els models òptims poden contribuir a explicar els polimorfismes observats a la natura. Els únics resultats del treball de Lewontin (1964b) en certa manera favorables a la validesa de la selecció segons allunyament a un òptim combinada amb lligament intens com a mecanisme explicatiu del polimorfisme es troben en considerar models de truncament doble, considerant un interval de valors fenotípics i "rebutjant" tot individu amb un fenotip per fora de l'interval. Com considera que el fenotip d'un individu segueix una normal centrada en el seu valor genotípic, per definició l'eficàcia és igual a l'àrea de probabilitat entre els límits de l'interval, sota una corba gaussiana amb mitjana els valor genotípic i desviació típica constant en alguns models i variable (variable segons els grau d'heterozigosi als models homeostàtics) a d'altres. A aquesta darrera classe de models, sense mantenir cap polimorfisme estable, es demostra que els temps fins a la fixació són molt llargs quan el lligament és intens.



Lewontin (1964 b) no realitzà cap mena d'anàlisi matemàtica del model òptim de Wright pel cas de loci lligats, al contrari d'allò realitzat per Kojima (1959) i ell mateix pel cas de loci no lligats. Aquesta anàlisi, encara que més complicada, és factible. Un primer pas seria la derivació d'expressions per calcular el valor fenotípic mitjà, la variància fenotípica, i l'eficàcia mitjana.

Plantejarem amb certa generalitat el model de selecció segons allunyament a un òptim per passar a continuació a fer algunes restriccions que ens permetin realitzar algunes anàlisis més concretes.

Sigui una població mendeliana diploide amb aparellament aleatori i de grandària infinita. L'expressió fenotípica de certs caràcters mètrics  $F_1, F_2, \dots, F_n$ , components del fenotip a partir de les quals s'avalua l'eficàcia, està controlada per  $L$  loci lligats cada un dels quals pot presentar dos estats al·lèlics que, com fem normalment, indicarem 0 i 1. Per tant, cada gamet  $x$  possible es podrà indicar com una tira de  $L$  dígit, des de 000 ... 0 a 111 ... 1. Suposarem que hi ha una funció  $f$  tal que a cada genotip  $(x,y)$  li fa correspondre el fenotip  $(F_1, F_2, \dots, F_n) = F$ :

$$(F_1, \dots, F_n) = f(x,y)$$

Usualment,  $f$  es podrà descompondre en  $n$  components  $f_1, f_2, \dots, f_n$   $F_i = f_i(x, y)$ . En realitat, el fenotip  $F_1, \dots, F_n$  s'ha de considerar el fenotip mitjà de tots els individus amb genotip  $(x, y)$ , ja que el fenotip real d'un individu concret s'hauria de considerar  $F = (F_1, \dots, F_n)$  més una perturbació aleatòria causada per l'ambient extern i la resta del genoma no estudiat.

Considerarem que la selecció s'ajusta a un model d'allunyament a un òptim, si l'eficàcia global d'un genotip  $xy$  és una funció decreixent  $g$  de la distància  $d(F, 0)$  entre el fenotip mitjà associat al genotip  $xy$  i l'òptim  $0 = (0_1, \dots, 0_n)$ :

$$w(xy) = g(d(f(xy), 0)) \quad (5.3.1)$$

$d$  no ha de ser forçosament una distància mètrica, sinó algun índex que indiqui l'allunyament de  $F$  de l'òptim  $0$ .

A partir d'aquesta idea general es poden fer algunes suposicions simplificadores addicionals.

Per una única component  $F$  del fenotip, si cada locus,  $i$ , té una contribució  $F^{(i)}$ , segons el genotip:

$F^{(i)}$	$x_i/y_i$	
$a_i$	per	1/1
$h_i a_i$	per	1/0    o    0/1
$-a_i$	per	0/0

(5.3.2)

considerant el valor de  $F$  com la simple summa per tot el genotip  $x_i/y_i$  dels valors  $(a_i, h_i a_i, -a_i)$  que corresponguin segons el genotip  $x_i/y_i$  al locus  $i$ ,  $F = \sum_{i=1}^L F^{(i)}$  i definint finalment les funcions  $d$  i  $g$  com

$$\begin{aligned} d(F,0) &= (F-0)^2 \\ g(d(F,0)) &= W_m - (F-0)^2 \end{aligned} \quad (5.3.3)$$

on  $W_m$  és una constant escollida adequadament, tindrem el model quadratic estudiat per Wright, Kojima i Lewontin.

La constant  $h$  indica les relacions de dominància al locus  $i$ . Si  $h_i=0$  tenim que  $h_i a_i=0$ , de manera que hi ha una total additivitat. Si  $h_i=1$  l'al·lel 1 és totalment dominant respecte del 0. Si  $h_i=-1$  passa la cosa contrària. Per a valors intermedis es donen situacions intermèdies de dominància. Si  $h_i > 1$  hi ha superdominància o superioritat dels heterozigòtics.

La contribució del locus  $i$  al fenotip es podria indicar en funció del genotip  $x_i/y_i$  al locus:

$$\begin{aligned} F^{(i)} &= a_i x_i y_i + h_i a_i x_i (1-y_i) + h_i a_i (1-x_i) y_i - a_i (1-x_i)(1-y_i) = \\ &= a_i (x_i + y_i - 1) + h_i a_i (x_i + y_i - 2x_i y_i) \end{aligned} \quad (5.3.4)$$

i el fenotip total com

$$F = \sum_{i=1}^L F^{(i)} = \sum_{i=1}^L a_i (x_i + y_i - 1) + \sum_{i=1}^L h_i a_i (x_i + y_i - 2x_i y_i) \quad (5.3.5)$$

el fenotip mitjà a una generació concreta és:

$$m_F = E(F) = \sum_{i=1}^L a_i (E x_i + E y_i - 1) + \sum_{i=1}^L h_i a_i (E x_i + E y_i - 2E(x_i y_i))$$

com  $x_i$  (i igualment  $y_i$ ) és la variable aleatòria indicador o de Bernouilli de la presència de l'al·lel 1 al locus  $i$  del gamet  $x$  (o del gamet  $y$ ), tindrem que  $E x_i = E y_i = p_i$ , la freqüència gènica de l'al·lel 1 al locus  $i$ , igualment  $E(x_i y_i) = E x_i E y_i = p_i p_i = p_i^2$  ja que es considera aparellament aleatori i per tant independència entre  $x_i$  i  $y_i$ , de manera que

$$\begin{aligned} m_F &= \sum_i a_i (2p_i - 1) + \sum_i h_i a_i (2p_i - 2p_i^2) = \\ &= \sum_i a_i (2p_i - 1) + 2 \sum_i h_i a_i p_i (1 - p_i) = \\ &= \sum_i a_i (2p_i - 1) + 2 \sum_i h_i a_i \sigma_i^2 \end{aligned} \quad (5.3.6)$$

on  $\sigma_i^2$  és la variància genètica al locus  $i$ , o  $2p_i(1-p_i) = 2\sigma_i^2$  la probabilitat d'heterozigosi al mateix locus.

L'esperança del quadrat del valor fenotípic és:

$$E(F^2) = E\left\{\sum_i a_i (x_i + y_i - 1) + \sum_i h_i a_i (x_i + y_i - 2x_i y_i)\right\}^2 =$$



$$\begin{aligned}
&= \sum_i \sum_j a_i a_j E\{(x_i + y_i - 1)(x_j + y_j - 1)\} + \\
&+ \sum_i \sum_j h_i h_j a_i a_j E\{(x_i + y_i - 2x_i y_i)(x_j + y_j - 2x_j y_j)\} + \\
&+ 2 \sum_i \sum_j a_i h_j a_j E\{(x_i + y_i - 1)(x_j + y_j - 2x_j y_j)\}
\end{aligned}$$

i tenint en compte que  $E(x_i x_j) = E(y_i y_j) = p_{ij}$  ja que es tracta de loci lligats (( $x_i, x_j$ ) no són variables aleatòries independents), tindrem:

$$\begin{aligned}
E(F^2) &= \sum_i \sum_j a_i a_j (2p_{ij} + 2p_i p_j - 2p_i - 2p_j + 1) + \\
&+ \sum_i \sum_j h_i a_i h_j a_j (4p_{ij}^2 - 4p_{ij} p_i - 4p_{ij} p_j + 2p_{ij} + 2p_i p_j) + \\
&+ 2 \sum_i \sum_j a_i h_j a_j (2p_{ij} - 4p_{ij} p_j + 2p_i p_j - 2p_j + 2p_j^2) \quad (5.3.7)
\end{aligned}$$

i per tant, la variància de  $F$  és:

$$\begin{aligned}
\sigma_F^2 &= E(F^2) - (EF)^2 = \sum_i \sum_j a_i a_j (2p_{ij} - 2p_i p_j) + \sum_i \sum_j h_i h_j a_i a_j (4p_{ij}^2 - \\
&- 4p_i^2 p_j^2 - 4p_{ij} p_i^2 + 4p_j p_i^2 - 4p_{ij} p_j^2 + 4p_i p_j^2 + 2p_{ij} - 2p_i p_j) + \\
&+ 2 \sum_i \sum_j a_i h_j a_j (2p_{ij} - 2p_i p_j - 4p_{ij} p_j + 4p_i p_j^2) \quad (5.3.8)
\end{aligned}$$

recordant que  $D_{ij} = p_{ij} - p_i p_j$  (fem la convenció  $D_{ii} = p_{ii} - p_i p_i = p_i - p_i^2$ ) i que  $p_{ij}^2 - p_i^2 p_j^2 = (p_{ij} - p_i p_j)(p_{ij} + p_i p_j) = D_{ij}(D_{ij} + 2p_i p_j)$

s'obté immediatament que

$$\begin{aligned} \sigma_F^2 &= 2 \sum_i \sum_j a_i a_j D_{ij} + \sum_i \sum_j h_i h_j a_i a_j (4D_{ij} (D_{ij} + 2p_i p_j) - 4p_i D_{ij} - 4p_j D_{ij} + \\ &+ 2D_{ij}) + 2 \sum_i \sum_j a_i h_j a_j (2D_{ij} - 4p_j D_{ij}) \\ \sigma_F^2 &= 2 \sum_{i=1}^L \sum_{j=1}^L a_i a_j D_{ij} + 2 \sum_{i=1}^L \sum_{j=1}^L h_i h_j a_i a_j D_{ij} \left\{ (1 - 2p_i + \frac{1}{h_i})(1 - 2p_j) + \right. \\ &+ \left. 2D_{ij} \right\} \end{aligned} \quad (5.3.9-I)$$

El primer terme de la dreta de (5.3.9-I) correspon clarament a la part additiva de la variància; si  $h_i = 0$  per a tot  $i$ , és a dir no hi ha cap mena de dominància, aleshores

$$\sigma_F^2 = 2 \sum_i \sum_j a_i a_j D_{ij}$$

El segon terme correspon a la part de variància associada a la dominància. Si considerem idèntica la contribució additiva de cada locus i el grau de dominància (per tot  $i, j$   $a_i = a_j = a$ ,  $h_i = h_j = h$ ), l'expressió 6.3.9 es simplifica considerablement:

$$\sigma_F^2 = 2a^2 \sum_{i=1}^L \sum_{j=1}^L D_{ij} \left\{ 1 + h^2 \left\{ (1 - 2p_i + \frac{1}{h})(1 - 2p_j) + 2D_{ij} \right\} \right\} \quad (5.3.9-II)$$

El lligament interacciona amb ambdues components de la variància. Si tots els loci fossin en equilibri gamètic, per tot  $i \neq j$ ;

$D_{ij}=0$ , de manera que, fent  $D_{ii}=p_i-p_i^2=p_i(1-p_i)$ , tindríem

$$\sigma_F^2 = 2 \sum_{i=1}^I a_i^2 p_i (1-p_i) + 2 \sum_{i=1}^I h_i^2 a_i^2 (p_i - p_i^2) (1-2p_i + 2p_i^2 + \frac{2}{h_i} - \frac{4p_i}{h_i}) \quad (5.3.10)$$

expressió que, desenrotllant-la, es comprova que és igual a l'expressió obtinguda per Kojima (1959)

$$\sigma_F^2 = \sum_{i=1}^I (p_i^2 a_i^2 + 2p_i(1-p_i)h_i^2 a_i^2 + (1-p_i)^2 a_i^2 - (E(F^{(i)}))^2) \quad (5.3.11)$$

també pel cas sense lligament i equilibri gamètic.

L'eficàcia mitjana de la població serà

$$m_W = E\{w_m - (O-F)^2\} = w_m - O^2 - E(F^2) + 2OEF \quad (5.3.12)$$

sumant i restant el quadrat de fenotip mitjà  $(EF)^2$  a (5.3.12), obtenim l'expressió

$$m_W = w_m - O^2 - (EF)^2 + 2OEF - E(F^2) + (EF)^2 = w_m - (O-m_F)^2 - \sigma_F^2 \quad (5.3.13)$$

de manera que l'eficàcia mitjana, que en certa manera expressaria el grau d'adaptació de la població al medi, és funció de la diferència al quadrat entre l'òptim i el fenotip mitjà  $(O-m_F)^2$  i de la variància fenotípica  $\sigma_F^2$ . Ambdós termes, positius,

es resten de la constant  $w_m$ . La minimització de l'un i de l'altre implicaria la maximització de  $m_w$ .

Les expressions anteriors han estat deduïdes pel cas sense inversions i amb medi no subdividit. L'extensió al cas amb un gamet invertit I és immediata i no requereix cap formulació addicional si no considerem cap component de l'eficàcia associada a l'estructura cromosòmica. Les fórmules per l'eficàcia mitjana, el fenotip mitjà i la variància segueixen essent vàlides ja que els paràmetres que intervenen (desequilibris gamètics, freqüències gèniques, valors additius...) tenen exactament el mateix significat, quant a fórmules per calcular-los, que en absència d'inversions - allò que possiblement serà diferent és la seva dinàmica de canvi.

Pel cas amb medi subdividit no tindrem pròpiament una matriu d'eficàcies  $(w_{xy})_{\substack{x=0, \dots, 2^L-1 \\ y=0, \dots, 2^L-1}}$   $w_{xy} = w(x,y) = w_m - (O - F_{xy})^2$ , si no un agregat tridimensional d'eficàcies

$$(w_{xyj})_{\substack{x=0, \dots, 2^L-1 \\ y=0, \dots, 2^L-1 \\ j=1, \dots, n}}$$

ja que cada genotip  $(x,y)$  tindrà diferent eficàcia segons la subdivisió del medi on sigui. Són vàlides les mateixes expressions que abans però considerant l'existència de  $n$  òptims diferents, segons la clapa del medi  $O_1, O_2, \dots, O_n$

$$w_{xyj} = w_j(x,y) = w_m - (O_j - F_{xy})^2 \quad (5.3.14)$$

amb el mateix fenotip  $F_{xy}$  pel genotip  $xy$  independentment de la clapa. A la fórmula 5.3.14 hi ha implícita la suposició que el fenotip no depèn de la clapa on s'experimenta la selecció, la qual és possiblement la mateixa on cada individu s'ha desenvolupat. Si volguéssim considerar una tal dependència, caldria per exemple assignar valors additius i dominants a cada locus  $a_{ij}$ ,  $h_{ij}$ , dependents no solament del locus  $i$  sinó, a més, de la subdivisió  $j$  del medi.

Com la dispersió és a l'atzar, cada clapa rep exactament les mateixes freqüències gamètiques i per tant exactament les mateixes freqüències gèniques, els mateixos desequilibris, etc., de manera que segons (5.3.6) i (5.3.9-I), el fenotip mitjà i la variància seran els mateixos per totes les subdivisions

$$\begin{aligned} m_{F_1} &= m_{F_2} = \dots = m_{F_n} = m_F \\ \sigma_{F_1}^2 &= \sigma_{F_2}^2 = \dots = \sigma_{F_n}^2 = \sigma_F^2 \end{aligned} \quad (5.3.15)$$

En canvi l'eficàcia mitjana pot ser diferent d'una clapa a l'altra, ja que està definida com:

$$m_{w_i} = E\{w_m - (O_i - F)^2\} = w_m - (O_i - m_F)^2 - \sigma_F^2 \quad (5.3.16)$$

#### 5.4. Anàlisi numèrica dels models anteriors

Mitjançant la utilització el programa descrit a l'apèndix 4 hem analitzat les propietats de les equacions definides als apartats 5.1, 5.2 i 5.3, per  $L=5$  loci, sota condicions de lligament -amb una probabilitat  $r=0.15$  de recombinació entre loci veïns, per tots ells i totes les simulacions- i d'additivitat total ( $h=0$ ) a tots els loci. A més, per fer els càlculs més ràpids, hem suposat que la contribució fenotípica de cada locus és

$$\begin{aligned} 2 & \text{ (en lloc de } a) \text{ pel genotip } 1/1 \\ 1 & \text{ (en lloc de } 0) \text{ pel genotip } 1/0 \text{ ò } 0/1 \\ 0 & \text{ (en lloc de } -a) \text{ pel genotip } 0/0 \end{aligned} \quad (5.4.1)$$

la mateixa per tots els loci. (5.4.1) és purament la transformació afí  $\hat{F}^{(i)} = \frac{1}{a} F + 1$ . Es tracta d'un canvi d'escala i origen de coordenades, que aplicat sobre les expressions (5.3.5), (5.3.6) (5.3.9) i (5.3.13) implica immediatament, si anomenen  $F$  al valor fenotípic calculat segons 5.4.1:

$$F = \sum_{i=1}^L (x_i + y_i)$$

$$m_F = 2 \sum_{i=1}^L p_i$$

$$\sigma_F^2 = 2 \sum_{i=1}^L \sum_{j=1}^L D_{ij}$$

$$m_w = w_m - (0 - m_F)^2 - \sigma_F^2 = w_m - (0 - 2 \sum_{i=1}^L p_i)^2 - 2 \sum_{i=1}^L \sum_{j=1}^L D_{ij} \quad (5.4.2)$$

on resulta evident que l'eficàcia mitjana té dues components, la primera funció de l'allunyament de les freqüències gèniques de l'òptim 0:  $(0 - 2 \sum_{i=1}^L p_i)^2$  i la segona de l'estructura gamètica

$$\sum_i \sum_j D_{ij} .$$

Totes les simulacions parteixen d'una població inicial amb freqüències gamètiques totes iguals, equilibri al lligament i freqüències gèniques de 0.5.

Pels casos amb medi no subdividit, considerem:

- 1) Un òptim  $O=5$  amb una constant  $w = 65$ , que impliquen els valors d'eficàcia, pels diferents genotips i fenotips, indicats a la taula 5.4.1
- 2) Un òptim  $O=4$  amb una constant  $w_m = 84$ , amb els valors d'eficàcia pels diferents fenotips (prescindint a partir d'ara d'indicar els genotips) indicats a la taula 5.4.2.

Pel casos amb medi subdividit en dues clapes, considerem

- 1) Uns òptims  $O_1=1$ ,  $O_2=9$ , una constant  $w = 81$  i unes grandàries  $c_1=c_2=0.5$  per cada clapa, que impliquen les eficàcies indicades a la taula 5.4.3. ( $w_1$  per la clapa 1,  $w_2$  per la clapa 2).

genotip xy $(x_1 x_2 x_3 x_4 x_5)$ $(y_1 y_2 y_3 y_4 y_5)$	fenotip $F_{xy}$	eficàcia $w(xy) = w_m - (0 - F_{xy})^2$ $= 65 - (5 - F_{xy})^2$	eficàcia rela- tiva referida a un màxim de 1 $w'(xy) = w(xy) / w_m$
00000 00000	0	40	0.6153846
00001 00010 00100 00000 00000 00000 01000 10000 00000 00000 (i els recíprocs 00000 00001 etc.)	1	49	0.7538461
00000 00000 00000 00011 00101···10001 00000 00000 10010 ·····11000 00001 00010 10000 00001 00001···00001 10000 ·····10000	2	56	0.8615384
00000 00000 00000 00111 01011···11100 00001 ·····00001 00011 11000 00010 00010 00011 ·····11000 10000 ·····11000	3	61	0.9384615
00000 00000 01111·····11110 00001 00001 00111·····11100 10000 ·····11100 00011 00011 00011 ·····11000	4	64	0.9846153
00000 00001 00001 11111 01111···11110 00011 11000 00111·····11100	5	65	1



continuació

00001 ..... 10000 11111            11111			
00011            11000 01111 ..... 11110  11100 .....            11100	6	64	0.9846153
00011            11000 11111 ..... 11111  11100 .....            11110	7	61	0.9384615
00111            11110 11111 ..... 11110	8	56	0.8615384
01111            11110 11111 ..... 11111	9	49	0.7538461
11111 11111	10	40	0.6153846

Taula 5.4.1.

F	$W = w_m - (0-F)^2$ $= 84 - (4-F)^2$	$W/w_m$
0	68	0.8095238
1	75	0.8928571
2	80	0.9523809
3	83	0.9880952
4	84	1
5	83	0.9880952
6	80	0.9523809
7	75	0.8928571
8	68	0.8095238
9	59	0.7023809
10	48	0.5714285

Taula 5.4.2.

F	$W_1 = w_m - (O_1 - F)^2$ $= 81 - (1 - F)^2$	$W_2 = w_m - (O_2 - F)^2$ $= 81 - (9 - F)^2$	$c_1 W_1 + c_2 W_2 =$ $\frac{1}{2}(W_1 + W_2) = W$	$W/w_m$
0	80	0	40	0.6153846
1	81	17	49	0.7538461
2	80	32	56	0.8615384
3	77	45	61	0.9384615
4	72	56	64	0.9846153
5	65	65	65	1
6	56	72	64	0.9846153
7	45	77	61	0.9384615
8	32	80	56	0.8615384
9	17	81	49	0.7538461
10	0	80	40	0.6153846

Taula 5.4.3.

F	$W_1 = w_m - (O_1 - F)^2$ $= 100 - (0 - F)^2$	$W_2 = w_m - (O_2 - F)^2$ $= 100 - (8 - F)^2$	$W = c_1 w_1 + c_2 w_2$ $= \frac{1}{2}(w_1 + w_2)$	$W/w_m$
0	100	36	68	0.8095238
1	99	51	75	0.8928571
2	96	64	80	0.9523809
3	91	75	83	0.9880952
4	84	84	84	1
5	75	91	83	0.9880952
6	64	96	80	0.9523809
7	51	99	75	0.8928571
8	36	100	68	0.8095238
9	19	99	59	0.7023809
10	0	96	48	0.5714285

Taula 5.4.4.

2) Uns òptims  $O_1=0$ ,  $O_2=8$ , una constant  $w_m=100$  i unes grandàries  $c_1=c_2=0.5$ , que impliquen les eficàcies de la taula 5.4.4.

Aquests valors fan que els dos casos 1), amb medi subdividit i no subdividit, siguin comparables i igualment els dos casos 2) ja que, tal com indiquem a les dues darreres columnes de les taules 5.4.3. i 5.4.4., l'eficàcia global de cada fenotip

$$w(xy) = \sum_{i=1}^2 c_i w_i(xy) = \frac{1}{2}(w_1(x,y) + w_2(x,y))$$

coincideix amb la indicada a les taules 5.4.1 i 5.4.2 respectivament. Però mentre que al cas 1) amb medi subdividit, les eficàcies mitjanes inicials són iguals a ambdues clapes

$$\begin{aligned} m_{w_1} &= w_m - (O_1 - 2 \sum_{i=1}^L p_i)^2 - 2 \sum_{i=1}^L \sum_{j=1}^L D_{ij} = \\ &= 81 - (1 - 2 \sum_{i=1}^L 0.5)^2 - 2 \sum_{i=1}^L (0.5 - 0.5^2) = 62.5 \end{aligned}$$

$$\begin{aligned} m_{w_2} &= w_m - (O_2 - 2 \sum_{i=1}^L p_i)^2 - 2 \sum_{i=1}^L \sum_{j=1}^L D_{ij} = \\ &= 81 - (9 - 2 \sum_{i=1}^L 0.5)^2 - 2 \sum_{i=1}^L (0.5 - 0.5^2) = 62.5 \end{aligned}$$

i iguals a l'eficàcia mitjana del cas amb medi no subdividit, pel cas 2) amb medi subdividit, les eficàcies mitjanes inicials a cada clapa no coincideixen

$$m_{w_1} = w_m - (O_1 - 2 \sum p_i)^2 - 2 \sum \sum D_{ij} =$$

$$= 100 - (0-2\sum 0.5)^2 - 2\sum(0.5-0.5^2) = 72.5$$

$$\begin{aligned} m_{w_2} &= w_m - (0-2\sum p_i)^2 - 2\sum\sum D_{ij} = \\ &= 100 - (8-2\sum 0.5)^2 - 2\sum(0.5-0.5^2) = 88.5 \end{aligned}$$

Presumiblement (i així es demostra seguidament a les simulacions) la igualtat de les eficàcies mitjanes del cas 1) es mantindrà en generacions succesives a causa de la simetria de les freqüències gamètiques inicials (tots els gamets en la mateixa freqüència 0.03125) mantinguda per la simetria del patró selectiu, amb òptims 1 i 9, entre uns fenotips possibles de 0 a 10. Això fa que aquest cas sigui equivalent a un cas sense subdivisió del medi (recordem l'apartat anterior on es comparen els enfocaments de Levene i Dempster) amb eficàcies definides per l'expressió (5.3.10)

$$W = \sum_{i=1}^2 c_i w_i = \frac{1}{2} (w_1 + w_2)$$

valors iguals, com hem dit, als de las eficàcies pel cas no subdividit. És previsible que en absència d'altres alteracions, l'evolució dels dos casos 1) (subdividit, no subdividit) sigui idèntica. Aquesta situació no es dóna pas al cas 2) al qual cal esperar una evolució diferent segons el grau de subdivisió del medi.

La introducció de gamets invertits, amb l'alteració de les probabilitats de recombinació que implica, ha d'alterar les simetries descrites als paràgrafs anteriors: l'increment (si es produeix) en la freqüència relativa del tipus invertit, a costa d'altres, trencarà la igualtat de les eficàcies mitjanes, de manera que és previsible que l'evolució de la freqüència d'inversions aparegudes de nou i les característiques dels possibles polimorfismes cromosòmics d'equilibri siguin diferents segons la subdivisió del medi, tant pel cas 1) com pel 2).

Pel cas 1) amb medi no subdividit (eficàcies segons la taula 5.4.1) en absència d'inversions, la població evoluciona de la manera usual de tots els models òptims amb lligament. En particular els resultats concorden amb els de Lewontin (1964b) i Fraser i Burnell (1970). La taula 5.4.5 indica les freqüències gamètiques a diferents generacions. En unes 80 generacions s'arriba a la configuració d'equilibri a les freqüències gamètiques indicada a la taula 5.2.2. Com podem apreciar els tipus gamètics més afavorits són els corresponents a composicions "equilibrades", amb 2 o 3 loci amb l'al·lel 1, d'acord amb les previsions i les anàlisis experimentals de Mather (1949 a,b), en particular el gamet 01010 i el seu complementari 10101 són els que assoleixen la màxima freqüència (0.035), seguits pels gamets 10010 i 01001 i els seus complementaris (0.034), pels gamets 10100 i 00101 i els seus complementaris (0.0337), mentre que els gamets

gamet	generació 27	generació 36	generació 50	equilibri
00000	.0222331869	.0222057278	.0221999788	.0221985426
10000	.0281321149	.0281157273	.02811206	.0281118082
01000	.0296058747	.0296043491	.0296039909	.0296039602
11000	.0310982232	.0310707676	.0310636309	.0310630021
00100	.0299317468	.0299254845	.0299233124	.029923049
10100	.0337332611	.0337514898	.0337556805	.0337559698
01100	.0331809954	.0331769469	.0331758018	.0331757069
11100	.0110982232	.0310707676	.0310636309	.0310630022
00010	.0296058747	.0296043491	.0296039909	.0296039602
10010	.0341749291	.0341982312	.0342042539	.0342047916
01010	.0351544656	.0352004693	.0352119303	.03521286
11010	.0337332611	.0337514898	.0337556804	.0337559698
00110	.0331809954	.0331769469	.0331758018	.0331757066
10110	.0341749292	.0341982312	.0342042539	.0342047913
01110	.0328298039	.0328332944	.0328348424	.0328350712
11110	.0281321149	.0281157273	.02811206	.0281118082
00001	.0281321148	.0281157273	.02811206	.0281118082
10001	.0328298039	.0328332944	.0328348424	.0328350712
01001	.0341749292	.0341982312	.034204254	.0342047913
11001	.0331809954	.0331769469	.0331758018	.0331757066
00101	.033733261	.0337514898	.0337556805	.0337559698
10101	.0351544657	.0352004693	.0352119303	.03521286
01101	.0341749291	.0341982312	.0342042539	.0342047916
11101	.0296059747	.0296043491	.0296039909	.0296039602
00011	.0310982232	.0310707666	.0310636309	.0310630022
10011	.0331809953	.0331769469	.0331758017	.0331757069
01011	.833733261	.0337514898	.0337556804	.0337559698
11011	.0299317468	.0299254845	.0299233124	.029923049
00111	.0310982232	.0310707676	.0310636309	.0310630021
10111	.0296058747	.0296043491	.0296039909	.0296039602
01111	.0281321149	.0281157273	.02811206	.0281118082
11111	.0222331869	.0222057278	.0221990788	.0221985426

Taula 5.4.5.



11000, 00011 i els seus complementaris són els menys freqüents dins de la classe dels de composició equilibrada, és a dir que a l'equilibri que s'estableix sota selecció i recombinació-segregació, les configuracions més afavorides són les que presenten el màxim grau d'alternància d'al·lels 0 i 1, i les menys freqüents a l'equilibri són les que presenten la màxima diferenciació en dues zones, una amb tots els loci 1 i l'altra amb tots els loci 0.

Aquesta tendència a la generació d'un cert "ordre" a la constitució dels gamets es reflecteix al patró de desequilibris al lligament indicat a la matriu de desequilibris  $D_{ij} = p_{ij} - p_i p_j$  de la generació 100:

	locus 1	locus 2	locus 3	locus 4
locus 2	-0.011			
locus 3	-0.006	-0.011		
locus 4	-0.005	-0.006	-0.011	
locus 5	-0.004	-0.005	-0.006	-0.011

amb valors negatius, associats a la preponderància dels gamets amb els tipus al·lèlics alternats de manera que la probabilitat  $p_{ij} = \text{Prob} [x_i=1, x_j=1]$  que als loci  $i, j$  es presenti alhora l'al·lel 1 és més petita que el producte  $p_i p_j = \text{Prob} [x_i=1] \cdot \text{Prob} [x_j=1]$  de les freqüències gèniques

per separat. Si calculem l'index

$$\rho_{ij} = \frac{D_{ij}}{\sqrt{p_i(1-p_i)p_j(1-p_j)}} = \frac{\text{cov}(x_i, x_j)}{\sqrt{\sigma^2(x_i)\sigma^2(x_j)}}$$

que es pot interpretar com un coeficient de correlació entre loci, ja que  $D_{ij}$  és la covariància entre  $x_i$  i  $x_j$  les variables que indiquen la constitució dels loci  $i$  i  $j$ , i  $p_i(1-p_i)$  és la variància de  $x_i$  (i igualment  $p_j(1-p_j)$  és la variància de  $x_j$ ) o bé la meitat de la probabilitat d'heterozigosi al locus  $j$ , obtenim la matriu

	locus 1	locus 2	locus 3	locus 4
locus 2	-0.059			
locus 3	-0.032	-0.059		
locus 4	-0.027	-0.032	-0.059	
locus 5	-0.021	-0.027	-0.032	-0.059

que té la mateixa interpretació que la matriu de desequilibris amb la informació addicional que les correlacions negatives, no són pas gaire properes a -1 sinó a 0, és a dir, molt febles, la qual cosa indica que la diferenciació de la població no és molt forta a causa de la selecció poc intensa i la recombinació relativament forta, que permet l'existència d'un nombre gran de gamets poc equilibrats (taula 5.4.5).

Per calcular  $D_{ij}$  i  $\rho_{ij}$  calen les freqüències gèniques. Aquestes es mantenen constants a 0.5 per tots els loci i totes les generacions.

L'eficàcia mitjana creix monotònicament des de 62.5 a 62.804 a l'equilibri.

Tots aquests resultats són coneguts des dels treballs precursors que ja hem citat. Ara analitzarem els dos aspectes nous: la subdivisió del medi i les inversions cromosòmiques.

Tal com és previsible, en absència d'inversions l'evolució de la població pel cas 1) amb medi subdividit (eficàcies segons la taula 5.4.3) és exactament igual a la del cas 1) amb medi no subdividit. Són vàlides les mateixes coses dites als paràgrafs anteriors i les taules de freqüències i desequilibris al lligament.

Si introduïm un gamet invertit a una freqüència inicial baixa, per exemple d'una mil·lèsima, i amb composició òptima, per exemple el gamet 11000 (indiquem els loci segons l'ordre estàndard) l'evolució dels dos casos 1), amb medi subdividit i amb medi no subdividit, és pràcticament la mateixa durant unes 500 generacions (figura 5.4.1.a), en particular calen

de l'ordre de 300 generacions perquè la freqüència del gamet invertit superi el valor 0.05 i unes 350 perquè superi 0.1; a partir de freqüències d'aquest ordre l'increment ja és més ràpid. Com es pot apreciar l'increment d'una inversió en freqüència inicial baixa, per les eficàcies expresades a les taules 5.4.1 i 5.4.3, a una població ja adaptada i en equilibri a les freqüències gamètiques (abans de produir-se la inversió) és extremadament lent pel cas determinista (per una grandària poblacional molt gran, teòricament infinita) i no és essencialment diferent segons la subdivisió del medi. Serà interessant estudiar l'efecte de les perturbacions aleatòries a aquest procés, per efecte de la deriva genètica associada a disminucions dràstiques de la grandària poblacional o "colls de botella", els quals sembla que, donada la durada del procés, es poden produir en un moment o altre. Precisament quan estudiem l'efecte de la deriva, discutim més detalladament els canvis de la freqüència del gamet invertit a les generacions inicials comparant els casos deterministes amb els casos amb deriva. Veieu més endavant.

A partir d'unues 500 generacions, el procés de canvi de la freqüència del gamet invertit ja és significativament diferent segons la subdivisió del medi (figura 5.4.1. a, a'). A ambdós casos s'arriba a un equilibri polimòrfic on coexistei-

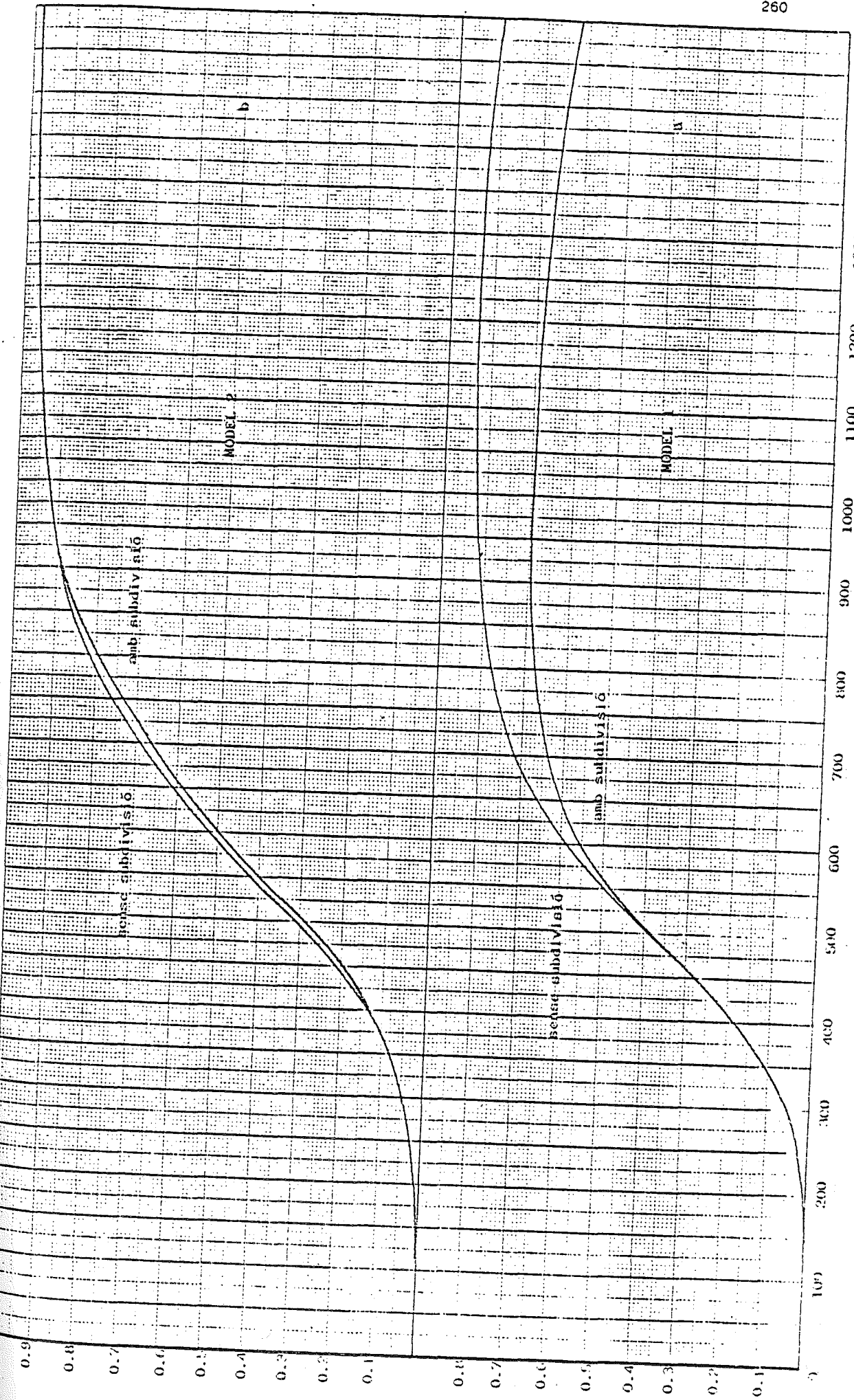


Figura 5.4-1.

"JST" "BARKER" "MAGNET" "RECORDING"

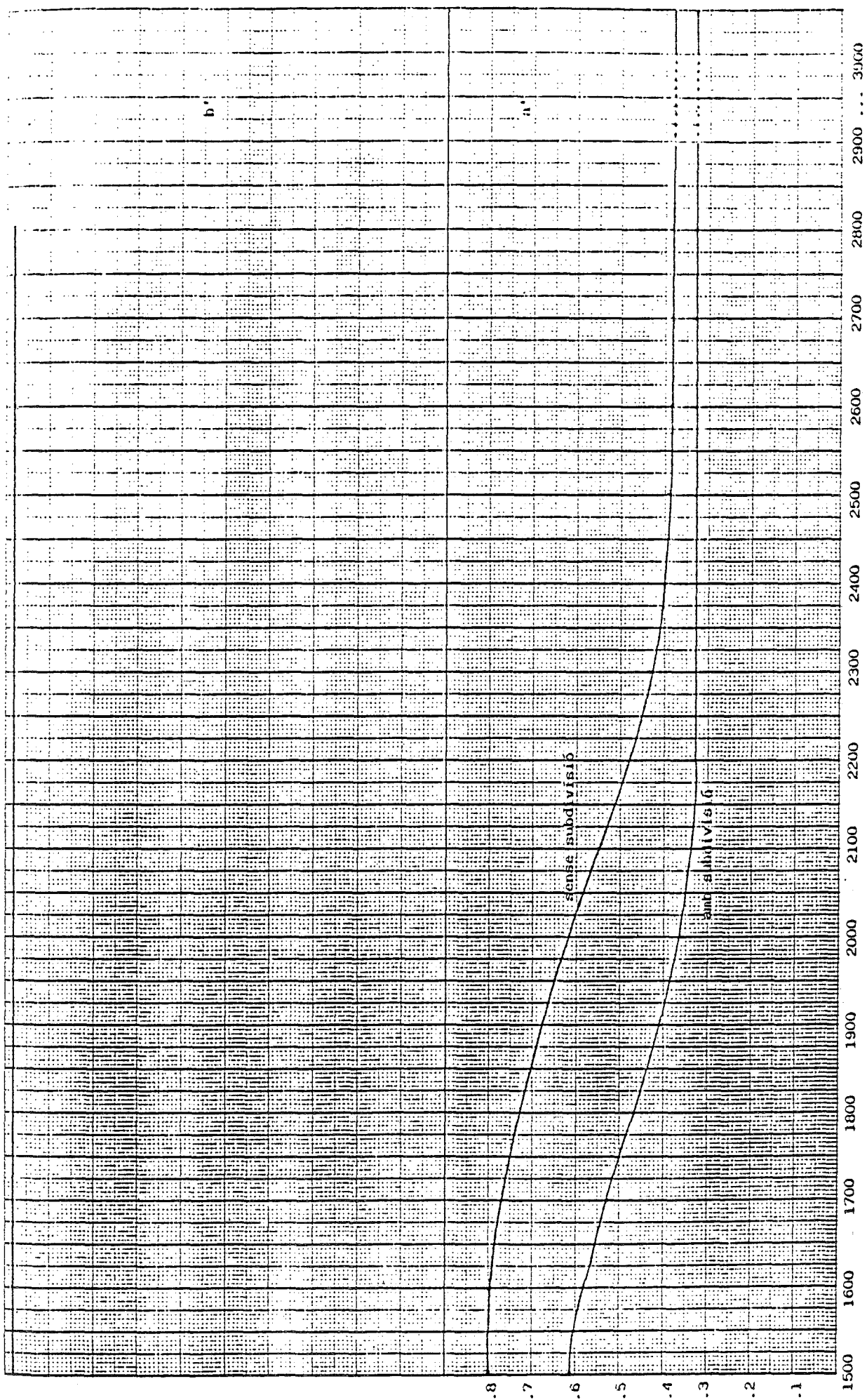


Figura 5.4.1

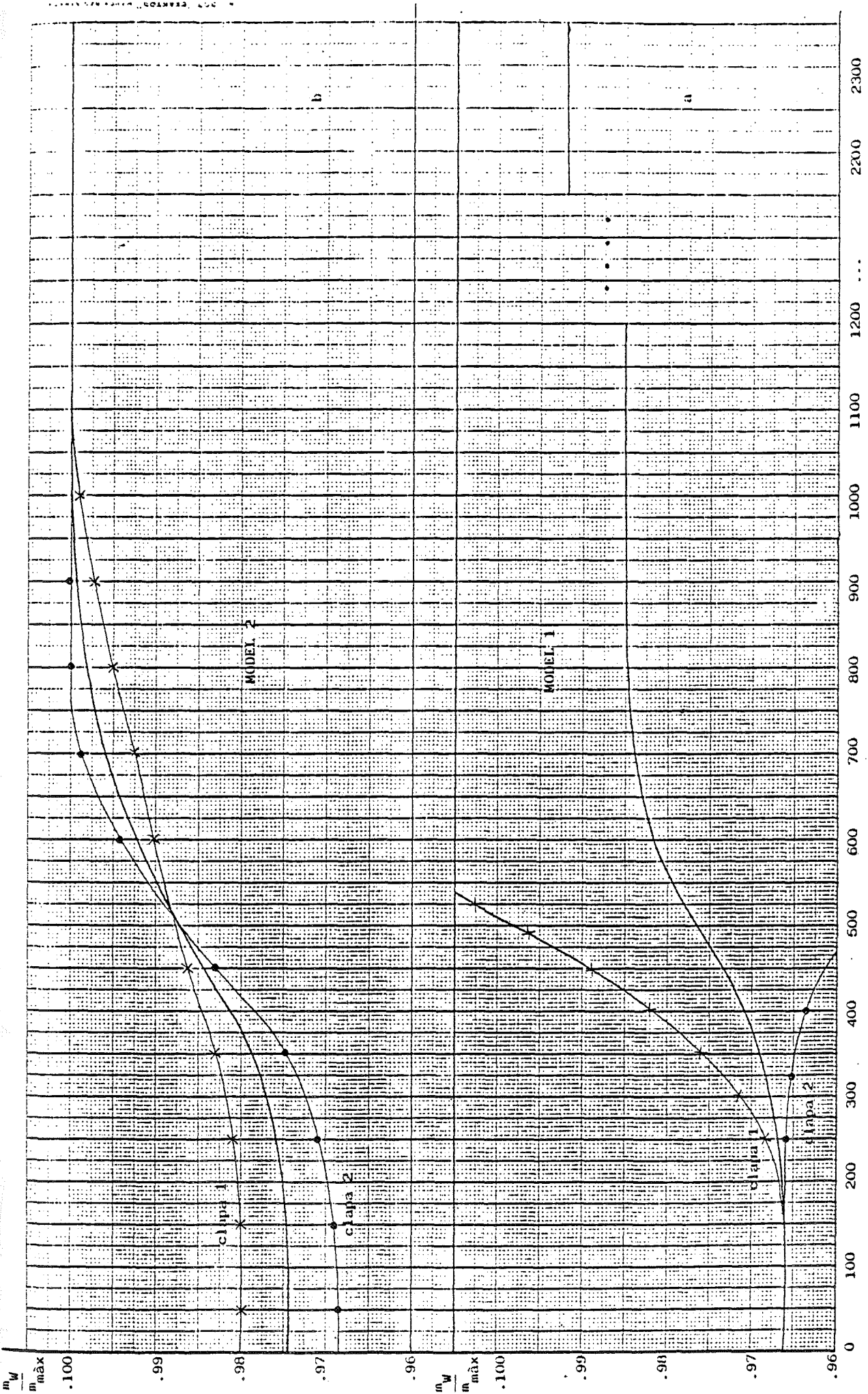


Figura 5.4.2.

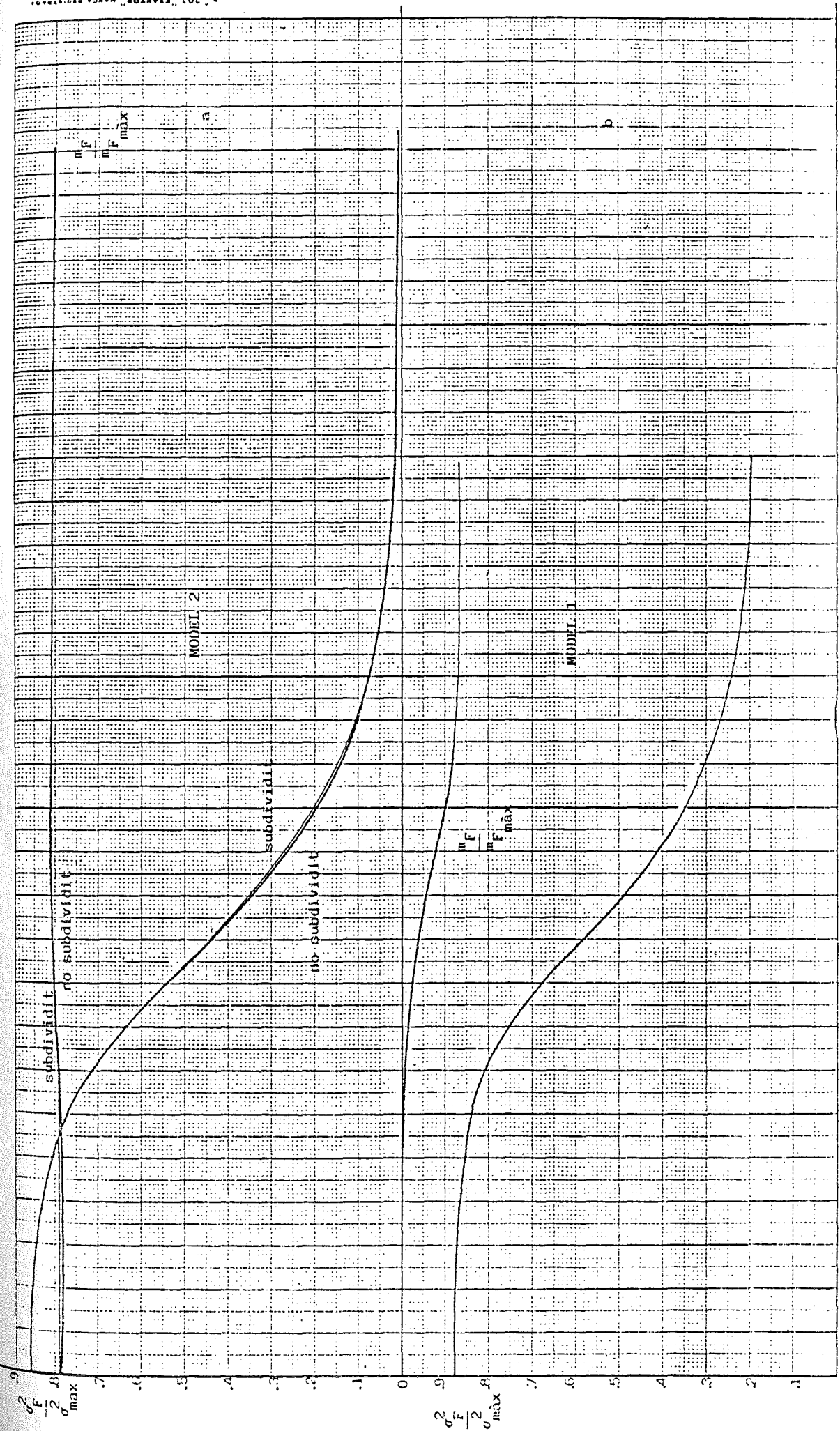


Figura 5.4.3.



xen gamets estàndard i gamets invertits, però les freqüències gamètiques i altres paràmetres són diferents segons el cas.

A ambdós casos, després d'una fase de creixement continuat, la freqüència del gamet invertit assoleix un màxim, als voltants de la generació 1000 pel cas amb subdivisió (amb una freqüència de 0.7) i als voltants de la generació 1200 pel cas sense subdivisió (amb una freqüència de 0.825). Si durant aquesta fase es produís una disminució forta de la grandària poblacional de manera que els efectes de la deriva fossin importants, en principi, sembla que seria més fàcil la fixació del gamet invertit pel cas sense subdivisió, donada la seva freqüència superior. Això ho estudiem més endavant.

Després d'assolir aquest màxim, la freqüència de la inversió decreix a ambdós casos fins a les freqüències d'equilibri al qual coexisteixen gamets estàndard i l'invertit.

Pel cas amb subdivisió coexisteixen els gamets estàndard 00111 amb una freqüència de 0.5 i 00101, amb una freqüència 0.16894, juntament amb l'invertit 11000 que presenta una freqüència de 0.33106. Com podem apreciar hi ha una freqüència 0.5 de gamets de potència 2 coexistint amb gamets de potència 3 també amb freqüència 0.5.

El valor de diversitat segons l'index de Shanon ( $H = -\sum_{i=1}^n p_i \log_2 p_i$ ,  $p_i$  és la freqüència del gamet  $i$  dels  $n=3$  presents) és de 1.43018 bits.

La situació d'equilibri del cas sense subdivisió és similar quant a la freqüència del gamet invertit però no quant als gamets estàndard que coexisteixen amb ell. La freqüència del gamet invertit és de 0.37912 i coexisteix amb els gamets estàndard 10101 amb freqüència 0.5 i 00101 amb freqüència 0.12088.

El valor de diversitat segons l'index de Shanon, associat a aquestes freqüències és de 1.39898 bits, menor que el corresponent al cas subdividit.

A les taules 5.4.6 i 5.4.7 hi ha les freqüències gamètiques estàndard corresponents a la fase amb màxim d'invertits i a una generació a les proximitats de l'equilibri. Desapareix totalment el patró de desequilibris al lligament propi de la selecció per un òptim intermedi, citat als paràgrafs anteriors, amb preponderància dels desequilibris negatius entre loci veïns, reflex de la predominància de les configuracions en repulsió. La matriu de desequilibris gamètics quan es produeix el màxim de la freqüència de la inversió, pel cas de medi no subdividit és

gamet	generació 1150	generació 2650
00000	6.16161148x10 <sup>-4</sup>	7.50973318x10 <sup>-12</sup>
10000	1.39662808x10 <sup>-3</sup>	1.98040522x10 <sup>-18</sup>
01000	7.71789588x10 <sup>-4</sup>	2.07353745x10 <sup>-22</sup>
11000	1.15863643x10 <sup>-3</sup>	4.26011388x10 <sup>-27</sup>
00100	2.55721586x10 <sup>-3</sup>	2.67265109x10 <sup>-7</sup>
10100	4.56132049x10 <sup>-3</sup>	5.84506349x10 <sup>-14</sup>
01100	2.16741842x10 <sup>-3</sup>	5.29822469x10 <sup>-16</sup>
11100	2.34644432x10 <sup>-3</sup>	8.88619693x10 <sup>-23</sup>
00010	2.08513132x10 <sup>-3</sup>	3.85459156x10 <sup>-11</sup>
10010	3.95065011x10 <sup>-3</sup>	8.79900341x10 <sup>-18</sup>
01010	2.09071883x10 <sup>-3</sup>	8.86274954x10 <sup>-20</sup>
11010	2.41397202x10 <sup>-3</sup>	1.54932959x10 <sup>-26</sup>
00110	5.90570522x10 <sup>-3</sup>	1.05387017x10 <sup>-6</sup>
10110	8.2033517 x10 <sup>-3</sup>	1.99243233x10 <sup>-13</sup>
01110	3.67993323x10 <sup>-3</sup>	1.73015543x10 <sup>-15</sup>
11110	2.98336836x10 <sup>-3</sup>	2.48005532x10 <sup>-22</sup>
00001	6.58411415x10 <sup>-3</sup>	5.52907682x10 <sup>-6</sup>
10001	.0127377102	1.29107577x10 <sup>-12</sup>
01001	6.93960755x10 <sup>-3</sup>	1.33386189x10 <sup>-14</sup>
11001	8.325594 x10 <sup>-3</sup>	2.39575366x10 <sup>-21</sup>
00101	.021933911	.168946719
10101	.031541977	3.25901794x10 <sup>-8</sup>
01101	.0145304276	2.89538775x10 <sup>-10</sup>
11101	.0120912469	4.22801123x10 <sup>-17</sup>
00011	.0150251342	2.18021032x10 <sup>-5</sup>
10011	.0228932519	4.39129611x10 <sup>-12</sup>
01011	.0117454048	4.33434062x10 <sup>-14</sup>
11011	.0104373829	6.61476183x10 <sup>-21</sup>
00111	.0308717772	.500004391
10111	.0332049738	8.32640271x10 <sup>-8</sup>
01111	.0144484454	7.07855028x10 <sup>-10</sup>
11111	9.25192965x10 <sup>-3</sup>	8.85851213x10 <sup>-7</sup>

Taula 5.4.6. Freqüències gamètiques estàndard corresponents al model 1, després de la introducció de la inversió, pel màxim del gamet invertit i per les proximitats de l'equilibri. Cas sense subdivisió

gamet	generació 1155	generació 3655
00000	$3.84245096 \times 10^{-4}$	$4.62818516 \times 10^{-17}$
10000	$1.10685271 \times 10^{-3}$	$2.60287188 \times 10^{-16}$
01000	$1.06784024 \times 10^{-3}$	$1.02556063 \times 10^{-33}$
11000	$1.80055934 \times 10^{-3}$	$4.23950393 \times 10^{-33}$
00100	$1.85976958 \times 10^{-3}$	$5.25949124 \times 10^{-10}$
10100	$4.00603733 \times 10^{-3}$	$2.47046288 \times 10^{-9}$
01100	$3.21933522 \times 10^{-3}$	$8.34438887 \times 10^{-27}$
11100	$3.32446563 \times 10^{-3}$	$2.76749225 \times 10^{-26}$
00010	$1.66533224 \times 10^{-3}$	$2.3715105 \times 10^{-33}$
10010	$3.89462601 \times 10^{-3}$	$1.15945053 \times 10^{-32}$
01010	$3.57512278 \times 10^{-3}$	0
11010	$4.04414108 \times 10^{-3}$	0
00110	$5.0535813 \times 10^{-3}$	$1.94920893 \times 10^{-26}$
10110	$7.40940381 \times 10^{-3}$	$7.91665852 \times 10^{-26}$
01110	$5.50012288 \times 10^{-3}$	0
11110	$3.53652413 \times 10^{-3}$	0
00001	$2.39413314 \times 10^{-3}$	$1.26680661 \times 10^{-8}$
10001	$5.7176678 \times 10^{-3}$	$6.30926962 \times 10^{-8}$
01001	$5.43463239 \times 10^{-3}$	$2.44283557 \times 10^{-25}$
11001	$6.48748493 \times 10^{-3}$	$8.80469221 \times 10^{-25}$
00101	$8.83339838 \times 10^{-3}$	.120883714
10101	.0136136806	.499999812
01101	.0104578758	$1.66709249 \times 10^{-18}$
11101	$6.93726181 \times 10^{-3}$	$4.76676879 \times 10^{-18}$
00011	$6.20023759 \times 10^{-3}$	$4.72232222 \times 10^{-25}$
10011	.0103206513	$2.01245633 \times 10^{-24}$
01011	$9.04711835 \times 10^{-3}$	0
11011	$6.57707236 \times 10^{-3}$	0
00111	.0115361668	$3.08544172 \times 10^{-18}$
10111	.0109212455	$1.08706943 \times 10^{-17}$
01111	$7.79748408 \times 10^{-3}$	$3.52336041 \times 10^{-35}$
11111	$3.62716949 \times 10^{-3}$	$8.54166437 \times 10^{-35}$

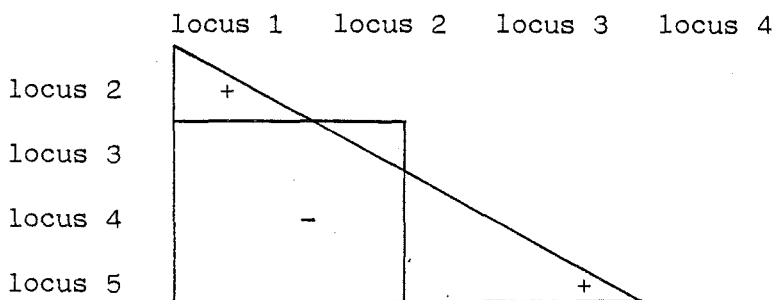
Taula 5.4.7. Freqüències gamètiques estàndard corresponents al model 1, després de la introducció de la inversió, pel màxim del gamet invertit i per les proximitats de l'equilibri. Cas amb subdivisió.

	locus 1	locus 2	locus 3	locus 4
locus 2	0.3			
locus 3	-0.045	-0.053		
locus 4	-0.042	-0.047	0.045	
locus 5	-0.051	-0.058	0.06	0.053

i la corresponent al cas amb medi subdividit

	locus 1	locus 2	locus 3	locus 4
locus 2	0.051			
locus 3	-0.065	-0.09		
locus 4	-0.06	-0.08	0.71	
locus 5	-0.081	-0.109	0.108	0.96

Ambdues reflecteixen el fet que el gamet predominant té constitució 11000, de manera que hi ha dos blocs de loci, amb repulsió entre blocs i acoblament dins de cada bloc, fet reflectit en els signes dels desequilibris:



A l'equilibri, els desequilibris gamètics són, pel cas amb manca de subdivisió

	locus 1	locus 2	locus 3	locus 4
locus 2	0.46			
locus 3	-0.46	-0.235		
locus 4	0	0	0	
locus 5	-0.46	-0.235	0.235	0

i pel cas amb subdivisió

	locus 1	locus 2	locus 3	locus 4
locus 2	0.221			
locus 3	-0.221	-0.221		
locus 4	-0.166	-0.166	0.166	
locus 5	-0.221	-0.221	0.221	0.166

La presència de la inversió també altera, naturalment, les freqüències gèniques, les quals en absència d'inversions en mantenen constants amb valor 0.5. Pel cas sense subdivisió, les freqüències gèniques pel màxim a la freqüència de la inversió són 0.915, 0.905, 0.108, 0.101, 0.126 (0.47629), mentre que pel cas amb subdivisió són 0.861, 0.812, 0.194, 0.178, 0.249 (0.69480). A l'equilibri, les freqüències gèniques són respectivament 0.879, 0.379, 0.621, 0, 0.621 (0.68085)

i 0.331, 0.331, 0.669, 0.5, 0.669 (0.93276). Els valors entre parèntesis corresponen al grau mitjà de polimorfisme segons la fórmula de Shanon ( $\bar{H} = \sum_{j=1}^L H_j/L$ ,  $H_j = -(p_j \log_2 p_j + (1-p_j) \log_2 (1-p_j))$ ), amb  $p_j$  igual a la freqüència gènica al locus  $j$ ).

Totes les observacions anteriors apunten cap al fet que en condicions de subdivisió del medi és superior el grau de polimorfisme cromosòmic mantingut a una població panmítica infinita. La inversió considerada té un contingut equilibrat, de manera que els individus que la porten són probablement generalistes, no especialment ben adaptats a una determinada clapa del medi sinó òptims globalment. Sembla que som davant d'un cas de selecció a favor dels heterozigots estructurals amb un haplotip invertit i un estàndard de potència 3 (que inicialment són els més nombrosos dins dels estàndard) com 11100, 10101, etc..., ja que aquests heterocariotips tenen un valor fenotípic de 5, òptim pel cas d'una clapa o òptim globalment pel cas de subdivisió del medi. La subdivisió únicament amplifica el grau de polimorfisme gràcies a la dependència de les freqüències que presenten les eficàcies.

Podem veure que a l'equilibri, els gamets estàndard més freqüents no són els màximament en repulsió (10101 i

01010), la qual cosa és explicable ja que ara tenim un cas amb manca de recombinació efectiva en coexistir invertits amb estàndard entre els quals no té cap efecte la recombinació.

Tot seguit estudiem l'efecte de la subdivisió del medi en un cas en què en principi no s'ha d'establir cap polimorfisme cromosòmic sinó la fixació del gamet invertit.

Pel cas amb eficàcies 2) (taules 5.4.2 i 5.4.4), tal com és previsible donada la diferència a les eficàcies mitjanes dins de cada clapa pel cas amb medi subdividit, l'evolució de les freqüències gamètiques és diferent segons la subdivisió del medi. A la taula 5.4.8 hi ha les freqüències gamètiques estàndard per diferents generacions abans de produir-se la inversió, pel cas amb medi subdividit i no subdividit. A la taula 5.4.9 indiquem les freqüències gèniques per ambdós casos. Associat amb el fet que l'òptim està desplaçat cap a valors menors que 5, de 4 exactament, la summa de freqüències gèniques disminueix ràpidament i es col·loca als voltants de 2, però les freqüències gèniques no es mantenen constants i totes elles iguals a 0.4 (valor que faria que sumessin 2) sinó que tendeixen a fer-se grans pels loci centrals i disminuir pels loci extrems, és a dir que aquest lleuger desplaçament de l'òptim implica una disminució de la capacitat



de manteniment del polimorfisme gènic, més gran que l'associada a un simple desplaçament de totes les freqüències cap a un altre equilibri.

Després de quatre-centes generacions de selecció, les matrius de desequilibris gamètics són indistingibles (amb tres xifres decimals) a ambdós processos i iguals a

	locus 1	locus 2	locus 3	locus 4
locus 2	$-8 \times 10^{-3}$			
locus 3	$-5 \times 10^{-3}$	$-8 \times 10^{-3}$		
locus 4	$-4 \times 10^{-3}$	$-5 \times 10^{-3}$	$-8 \times 10^{-3}$	
locus 5	$-3 \times 10^{-3}$	$-4 \times 10^{-3}$	$-5 \times 10^{-3}$	$-8 \times 10^{-3}$

tots negatius, és a dir, el mateix patró de desequilibris al lligament que als processos anteriors, també abans de la selecció.

La impressió general en vista de l'evolució de les freqüències gamètiques (taula 5.4.8), gèniques (taula 5.4.9) i del grau d'adaptació, expressat per l'eficàcia mitjana (taula 5.4.10) és que encara que l'estat final és essencialment semblant malgrat la subdivisió del medi, l'evolució és més ràpida pel cas sense subdivisió. També es pot dir que l'evolució en aquestes condicions (tant pel cas subdividit com no) és més lenta que al cas tractat primerament.

gamet	generació			
	100	200	400	800
00000	.0706400672	.0710545888	.0708177048	.0616514768
10000	.0519296607	.0512379005	.0476999974	.0214890728
01000	.0546726666	.0553330731	.0573655035	.0696966371
11000	.0346523818	.0344008427	.0333127424	.0209226969
00100	.0552682579	.0561908423	.0593125029	.084280495
10100	.0371211712	.0370190691	.0364970628	.0268129307
01100	.036934946	.0377804802	.0414767225	.0821225138
11100	.0214229229	.0214947535	.0220410152	.0225440407
00010	.0546726665	.055333073	.0573655036	.0696966453
10010	.0374364074	.0371642718	.0359872062	.022611473
01010	.0386799745	.0393868326	.0424723902	.071964286
11010	.0228510352	.022823882	.0229882625	.0201244816
00110	.0369349459	.0377804802	.0414767226	.0821225236
10110	.0231272244	.0232042059	.023793014	.0243473311
01110	.0225890393	.0232467287	.0265407512	.0731436836
11110	.0122152858	.0123311942	.0131498253	.0187155328
00001	.0519296606	.0512379	.0476999964	.0214890737
10001	.035880835	.0347263451	.0301959508	7.03636594E-03
01001	.0374364074	.0371642715	.0359872054	.0226114712
11001	.0223206264	.0217349724	.0196587914	6.3841285E-03
00101	.037121171	.0370190688	.0364970619	.0268129317
10101	.0234583236	.022946416	.0211301218	8.02496334E-03
01101	.0231272245	.0232042058	.0237930135	.0243473291
10101	.0126112011	.0124118618	.0118877285	6.2848038E-03
00011	.0346523817	.0344008424	.0333127418	.0209227003
10011	.0323206263	.0217349724	.0196587914	6.38412928E-03
01011	.0228510351	.0228238819	.022988262	.0201244823
11011	.012695637	.0124385635	.0117020501	5.29313679E-03
00111	.0214229228	.0214947533	.0220410148	.0225440443
10111	.012611201	.0124118618	.0118877285	6.28480453E-03
01111	.0122152858	.0123311941	.013149825	.0187155335
11111	6.19680703E-03	6.13667062E-03	6.11278991E-03	4.49428112E-03

Taula 5.4.8. Freqüències gamètiques estàndard abans de la inversió. Model 2, amb manca de subdivisió.

generació	locus 1		locus 2		locus 3		locus 4		locus 5	
100	.389	.391	.393	.395	.394	.396	.393	.395	.389	.391
200	.384	.387	.395	.398	.397	.399	.395	.398	.384	.387
300	.378	.381	.399	.401	.402	.404	.399	.401	.378	.381
400	.368	.371	.405	.407	.411	.413	.405	.407	.368	.371
500	.355	.355	.412	.416	.421	.426	.412	.416	.355	.355

Taula 5.4.9. Freqüències gèniques a diferents generacions . Els valors de l'esquerra corresponen al cas amb 1 clapa, els de la dreta al cas amb 2 clapes.

generació	1 clapa, $W_m=84$ , $O=4$	2 clapes, $W_m=100$ , $O_1=0, O_2=8$		
	GLOBAL	CLAPA 1	CLAPA 2	GLOBAL
100	81.833	82.325	81.34	81.8325
200	81.834	82.332	81.334	81.833
300	81.835	82.333	81.334	81.8335
400	81.837	82.336	81.336	81.836
500	81.843	82.341	81.343	81.842

Taula 5.4.10. Eficàcies mitjanes a diferents generacions.

L'evolució de la freqüència del gamet 11000, invertit i introduït a una freqüència inicial de 0.001 està indicada a la figura 5.4.1-b. En principi, pel cas amb medi no subdividit, d'acord amb el treball de simulació d'altres autors principalment de Fraser i Burnell, com l'òptim és 4 i el gamet invertit té una potència de 2, de manera que l'homozigot invertit té una potència gamètica de 4, (potència en el sentit de Fraser i Burnell (1970), és a dir, senzillament el nombre de loci amb l'al·lel 1), és d'esperar la fixació de la inversió. Tal com indica la gràfica, efectivament això és el que passa. En un principi l'increment de la inversió és extremadament lent, igual que pels casos descrits a la figura 5.4.1.-a. De fet, encara que a la figura 5.4 és impossible distingir-ho, a les generacions inicials, en què la freqüència de l'invertit és molt baixa, l'increment és més petit que al cas del model 1. Això és lògic ja que les diferències d'eficàcies (taules 5.4.1 a 5.4.4) són menors per aquest segon model. Calen unes tres-centes generacions perquè la inversió arribi a 0.05 i unes tres-centes setanta-cinc perquè arribi a 0.1. A partir d'aquest moment l'evolució és més ràpida, s'arriba a una freqüència de 0.95 en unes mil cent generacions i es tendeix novament, molt lentament, cap a la fixació.

La subdivisió del medi no implica l'establiment d'un polimorfisme cromosòmic, tal com podem apreciar també a la figura 5.4.1-b, tan sols una suavització de la gràfica, que a partir d'unies 350 generacions està per sota de la corresponent al cas

no subdividit. Aquest efecte és més palès a la part intermèdia de la gràfica. En canvi és impossible diferenciar ambdues corbes a la fase inicial i a la fase de fixació.

Per un gamet invertit intermedi com és ara 11000, i pel règim selectiu de les taules 5.4.2 i 5.4.4, no hi ha una condició de polimorfisme associada a la subdivisió del medi. Si pel cas no subdividit es produeix fixació de la inversió, pel cas amb medi subdividit igualment es fixa. L'únic efecte és un alentiment del procés, ja de si lent.

A les figures 5.4.2 a,b es reflecteix l'evolució del grau d'adaptació expressat mitjançant l'eficàcia mitjana  $m_w$  dividida pel valor màxim  $w_m$  per tal que estigui entre 0 i 1, pel cas amb medi no subdividit (línia llisa) i la corresponent a cada una de les clapes del medi subdividit (línies  $\text{---}\times\text{---}\times\text{---}$  i  $\text{---}\bullet\text{---}\bullet\text{---}$ ). L'eficàcia global pel medi subdividit estaria exactament entremig d'aquestes dues darreres línies i lleugerament per sota de la línia llisa. L'eficàcia global és sempre creixent (figura 5.4.2 b) malgrat que pel cas subdividit del model 1, l'eficàcia a una clapa és decreixent en algunes fases (figura 5.4.2 a). El mateix fet (el creixement de  $m_w$ ) queda indicat en representar les dues components de l'eficàcia mitjana:  $m_F$  i  $\sigma_F^2$  (figura 5.4.3 a,b). Pels dos casos (no subdividit, subdividit) l'estat final és aproximadament el mateix.

La tendència a l'increment del gamet invertit estudiat fins ara és independent de la freqüència inicial. No hem trobat cap llindar mínim de freqüència com ara el discutit per Fraser i col·laboradors (vegeu la secció 2) que és de l'ordre de 0.12 a 0.15, freqüències extraordinàriament grans i que fan difícil d'explicar l'establiment d'un polimorfisme cromosòmic a partir d'un esdeveniment de mutació cromosòmica únic. Per totes les freqüències inicials provades hi ha un increment mantingut de la freqüència de la inversió. A la taula 5.4.11 hi ha uns quants valors de la freqüència de la inversió quan es parteix d'una freqüència inicial de 0.0001, corresponent al cas d'una població de 5000 individus a la qual un d'ells és portador del gamet invertit. A la taula, la primera generació produïda es compta com la 0. Una mica més endavant en aquesta secció, quan estudiem l'efecte de la deriva, hi ha les freqüències corresponents a un gamet invertit introduït a una freqüència inicial de 0.00125 i segons el model determinista, les quals es comparen amb les del model amb deriva.

Anteriorment hem comentat l'evolució d'un gamet invertit de composició intermèdia. La idea general que s'en dedueix, de la discussió anterior, és que es produeix una mena d'heterosi quan el genotip òptim solament s'assoleix quan es combina el gamet invertit amb un gamet estàndard adequat. Aleshores s'estableix un estat d'equilibri amb polimorfisme cromosòmic. En canvi

MODEL 1, SENSE SUBDIVISIÓ

generació	Freqüència
0	$1.01349907 \times 10^{-4}$
5	$1.08376878 \times 10^{-4}$
10	$1.15890708 \times 10^{-4}$
15	$1.23925405 \times 10^{-4}$
20	$1.32517056 \times 10^{-4}$
25	$1.41704231 \times 10^{-4}$
30	$1.51528186 \times 10^{-4}$
35	$1.62033032 \times 10^{-4}$
40	$1.73265941 \times 10^{-4}$
45	$1.85277347 \times 10^{-4}$

MODEL 1, AMB SUBDIVISIÓ

generació	Freqüència
0	$1.01349907 \times 10^{-4}$
5	$1.08376854 \times 10^{-4}$
10	$1.15890592 \times 10^{-4}$
15	$1.23925145 \times 10^{-4}$
20	$1.32516609 \times 10^{-4}$
25	$1.41703561 \times 10^{-4}$
30	$1.51527254 \times 10^{-4}$
35	$1.62031797 \times 10^{-4}$
40	$1.73264355 \times 10^{-4}$
45	$1.85275354 \times 10^{-4}$

MODEL 2, SENSE SUBDIVISIÓ

generació	Freqüència
0	$1.01316979 \times 10^{-4}$
5	$1.08164466 \times 10^{-4}$
10	$1.15471752 \times 10^{-4}$
15	$1.23269651 \times 10^{-4}$
20	$1.31590737 \times 10^{-4}$
25	$1.40469684 \times 10^{-4}$
30	$1.49943422 \times 10^{-4}$
35	$1.60051277 \times 10^{-4}$
40	$1.70835118 \times 10^{-4}$
45	$1.82339508 \times 10^{-4}$

MODEL 2, AMB SUBDIVISIÓ

generació	Freqüència
0	$1.01321278 \times 10^{-4}$
5	$1.08192502 \times 10^{-4}$
10	$1.15527621 \times 10^{-4}$
15	$1.23357998 \times 10^{-4}$
20	$1.31716822 \times 10^{-4}$
25	$1.4063946 \times 10^{-4}$
30	$1.50163610 \times 10^{-4}$
35	$1.60329488 \times 10^{-4}$
40	$1.71179915 \times 10^{-4}$
45	$1.82760555 \times 10^{-4}$

Taula 5.4.11. Freqüències del gamet invertit 11000, a les primeres generacions, introduït a una freqüència inicial de 0.0001.

quan l'òptim el pot produir la combinació homozigòtica dels dos gamets invertits, la població tendeix a la fixació de la inversió. La subdivisió del medi és irrellevant quant a aquest estat final, quan el medi està subdividit i hi ha selecció disruptiva però amb unes eficàcies que en mitjana siguin comparables amb les del cas no subdividit, hi ha polimorfisme quan el millor genotip generalista, que no és òptim a cap de les clapes però té unes eficàcies "compromís" òptimes, s'assoleix en combinació heterozigòtica d'estàndard i invertit, en cas contrari, quan l'homozigot invertit és òptim, es produeix la fixació de la inversió.

Aquesta selecció a favor dels generalistes queda palesa quan estudiem la dinàmica d'inversions "extremes". Concretament hem seguit l'evolució dels gamets invertits de composició gènica 00000 i 01111.

A la taula 5.4.12 hi ha les freqüències d'aquests dos gamets invertits, pel cas amb medi no subdividit, seguides durant 100 generacions

Com és d'esperar hi ha una marcada caiguda de la freqüència d'aquests gamets, summament desequilibrats donat l'òptim únic i intermedi.



GAMET 00000

generació	MODEL 1 (OPTIM 5)	MODEL 2 (OPTIM 4)
0	$9.17964358 \times 10^{-4}$	$9.62130148 \times 10^{-4}$
20	$1.65897095 \times 10^{-4}$	$4.44704286 \times 10^{-4}$
40	$2.99703424 \times 10^{-5}$	$2.05469379 \times 10^{-4}$
60	$5.4116766 \times 10^{-6}$	$9.48819847 \times 10^{-5}$
80	$9.7698335 \times 10^{-7}$	$4.37903519 \times 10^{-5}$
100	$1.76366038 \times 10^{-7}$	$2.01986579 \times 10^{-5}$

GAMET 01111

generació	MODEL 1 (OPTIM 5)	MODEL 2 (OPTIM 4)
0	$9.81654165 \times 10^{-4}$	$9.66467386 \times 10^{-4}$
20	$6.77900853 \times 10^{-4}$	$4.88670647 \times 10^{-4}$
40	$4.68109507 \times 10^{-4}$	$2.46960176 \times 10^{-4}$
60	$3.23232605 \times 10^{-4}$	$1.24712625 \times 10^{-4}$
80	$2.23192679 \times 10^{-4}$	$6.29249679 \times 10^{-5}$
100	$1.54114814 \times 10^{-4}$	$3.17184202 \times 10^{-5}$

Taula 5.4.12

Com és també lògic, al model 1 (òptim 5) el gamet 00000 baixa en freqüència més ràpidament que 01111 ja que aquest segon està menys allunyat de l'òptim intermedi. Pel model 2 passa quel com semblant, encara que la diferència entre l'evolució d'ambdós gamets és menor.

Allò que ja és més sorprenent és l'evolució dels mateixos gamets invertits però en condicions de subdivisió del medi (tau la 5.4.13)

Com podem apreciar l'evolució és molt semblant a la del cas sense subdivisió del medi, hi ha la mateixa caiguda a les freqüències dels gamets invertits, només lleugeríssimament esmorteïda, però de manera no apreciable si està actuant qualsevol efecte aleatori com ara la deriva.

L'homozigot invertit és òptim, pel cas de la inversió 00000 i el model 2, a la primera clapa i pel cas de la inversió 01111 i el model 2, a la segona. Pel cas del model 1, els homozigots invertits corresponents són quasi òptims a cada una de les clapes, respectivament. És evident que no es produeix cap mena de selecció a favor dels tipus extrems, adaptats a cada òptim extrem, sinó a favor d'un intermedi que és òptim globalment o en mitjana, però no en alguna de les clapes. Tenim un cas com el discutit per Da Cunha (1960) (recordem l'apartat 3.4.1) on es

GAMET 00000

generació	MODEL 1 (OPTIMS 1,9)	MODEL 2 (OPTIMS 0,8)
0	$9.17964358 \times 10^{-4}$	$9.61620205 \times 10^{-4}$
20	$1.65972796 \times 10^{-4}$	$4.39847195 \times 10^{-4}$
40	$2.99887307 \times 10^{-5}$	$2.01133321 \times 10^{-4}$
60	$5.41478733 \times 10^{-6}$	$9.19292079 \times 10^{-5}$
80	$9.77509972 \times 10^{-7}$	$4.19984374 \times 10^{-5}$
100	$1.76458381 \times 10^{-7}$	$1.91794446 \times 10^{-5}$

GAMET 01111

generació	MODEL 1 (OPTIMS 1,9)	MODEL 2 (OPTIMS 0,8)
0	$9.81654164 \times 10^{-4}$	$9.6705281 \times 10^{-4}$
20	$6.77855288 \times 10^{-4}$	$4.949801 \times 10^{-4}$
40	$4.68049324 \times 10^{-4}$	$2.53272047 \times 10^{-4}$
60	$3.23180528 \times 10^{-4}$	$1.29518288 \times 10^{-4}$
80	$2.23152494 \times 10^{-4}$	$6.61909936 \times 10^{-4}$
100	$1.54085109 \times 10^{-4}$	$3.38036092 \times 10^{-5}$

Taula 5.4.13.

selecciona a favor del genotip que té un marge més ampli de microníxols possibles o una major homeòstasi enfront a les alteracions ambientals. Com veurem més endavant, hi ha una diferència notable entre aquests resultats i els obtinguts en condicions de competició (selecció dependent alhora de la freqüència i de la densitat).

A partir dels resultats anteriors no es pot afirmar categòricament que sota les condicions d'allò que en línies molt generals podríem anomenar "model de Levene", sempre ha d'augmentar la freqüència dels genotips òptims globalment i no la dels adaptats a cada clapa, o en altres paraules, que l'estratègia adaptativa de la població sempre es basa en l'augment en freqüència de genotips homeostàtics enfront a totes les condicions i no en la diversificació de genotips. Podem afirmar que per tots els casos estudiats fins ara, amb el present model, passa això. Altres autors, com ara Felsenstein (1979) per altres models de selecció han deduït les condicions a les quals hi ha selecció per un genotip intermedi. A la secció següent estudiarem un model al qual es produeix una selecció evident a favor dels tipus extrems, en condicions de subdivisió del medi.

Finalitzarem l'estudi mitjançant el macromodel determinista amb l'anàlisi de dues conjectures fetes a l'apartat 5.2.

Certes raons de caire teòric fan suposar que en una població no coadaptada la progressió d'una inversió favorable selectivament ha d'esser més ràpida que en una població que ja ha experimentat bastants generacions de selecció. Els resultats indicats a la taula 5.4.14 confirmen això. Hi ha l'evolució de la freqüència de la inversió equilibrada 11000, estudiada anteriorment, per les primeres 46 generacions quan és introduïda a una població amb totes les freqüències gamètiques estàndard iguals a 0.03125, és a dir amb totes les freqüències gèniques iguals a 0.5 i absència de desequilibri al lligament. La freqüència inicial de la inversió és de 0.00125. Si comparem les freqüències de la taula 5.4.17, corresponents a la mateixa situació però per una població ja coadaptada, amb les de la taula 5.4.14 podem comprovar que en tot moment són superiors les segones (les dades d'un mateix model, naturalment).

També a l'apartat 5.2 comparàvem teòricament els models de Levene amb els models amb manca de subdivisió ambiental o amb subdivisió del tipus proposat per Dempster (que com hem dit és formalment equivalent a l'absència de subdivisió). Hem vist que pel cas dels models amb subdivisió ambiental del tipus de Levene, l'eficàcia biològica d'un genotip  $(\xi, \eta)$  (on  $\xi, \eta$  poden indicar qualsevol gamet estàndard o invertit) es pot indicar

$$w(\xi, \eta) = m_w \sum_{i=1}^n c_i \frac{w_i(\xi, \eta)}{m_{w_i}}$$

MODEL 1, SENSE SUBDIVISIÓ

generació	freqüència
0	$1.27 \times 10^{-3}$
5	$1.36556635 \times 10^{-3}$
10	$1.46242011 \times 10^{-3}$
15	$1.56448996 \times 10^{-3}$
20	$1.67308496 \times 10^{-3}$
25	$1.78896256 \times 10^{-3}$
30	$1.91274123 \times 10^{-3}$
35	$2.04502345 \times 10^{-3}$
40	$2.18637943 \times 10^{-3}$
45	$2.3374782 \times 10^{-3}$

MODEL 1, AMB SUBDIVISIÓ

generació	freqüència
0	$1.27 \times 10^{-3}$
5	$1.36556253 \times 10^{-3}$
10	$1.46240164 \times 10^{-3}$
15	$1.56444839 \times 10^{-3}$
20	$1.67301374 \times 10^{-3}$
25	$1.78885577 \times 10^{-3}$
30	$1.9125928 \times 10^{-3}$
35	$2.04481565 \times 10^{-3}$
40	$2.18612688 \times 10^{-3}$
45	$2.33715769 \times 10^{-3}$

MODEL 2, SENSE SUBDIVISIÓ

generació	freqüència
0	$1.2810559 \times 10^{-3}$
5	$1.42170826 \times 10^{-3}$
10	$1.55022192 \times 10^{-3}$
15	$1.67507481 \times 10^{-3}$
20	$1.80085117 \times 10^{-3}$
25	$1.93048007 \times 10^{-3}$
30	$2.06598836 \times 10^{-3}$
35	$2.20887485 \times 10^{-3}$
40	$2.36032763 \times 10^{-3}$
45	$2.52135709 \times 10^{-3}$

MODEL 2, AMB SUBDIVISIÓ

generació	freqüència
0	$1.28759985 \times 10^{-3}$
5	$1.44415339 \times 10^{-3}$
10	$1.5749959 \times 10^{-3}$
15	$1.69782432 \times 10^{-3}$
20	$1.82078182 \times 10^{-3}$
25	$1.94806528 \times 10^{-3}$
30	$2.08203058 \times 10^{-3}$
35	$2.22415082 \times 10^{-3}$
40	$2.37548051 \times 10^{-3}$
45	$2.53688641 \times 10^{-3}$

Taulla 5.4.14. Freqüència a les primeres generacions de la inversió 11000 introduïda a una freqüència inicial de  $1.25 \times 10^{-3}$  i sense coadaptació prèvia.

on  $n$  és el nombre de clapes o subdivisions ambientals,  $c_i$  és la capacitat de la clapa  $i$  ( $\sum_{i=1}^n c_i = 1$ ),  $w_i(\xi, \eta)$  és l'eficàcia del genotip a la clapa  $i$ ,  $m_{w_i}$  és l'eficàcia mitjana a la clapa  $i$  ( $m_{w_i} = \sum_x \sum_y p_x p_y w_i(x, y)$ ) i  $m_w$  és l'eficàcia mitjana global

$$(m_w = \sum_{i=1}^n c_i m_{w_i}) .$$

La presència dels factors  $m_{w_i}^{-1}$  a l'expressió de  $w(\xi, \eta)$ , els quals depenen de les freqüències genotípiques, fa que un model de Levene sigui formalment equivalent a un model amb manca de subdivisió ambiental però amb les eficàcies dependents de les freqüències. Quan les eficàcies a cada clapa  $m_{w_i}$  siguin totes iguals  $m_{w_1} = m_{w_n} = m_w$ , aleshores

$$w(\xi, \eta) = \sum_{i=1}^n c_i w_i(\xi, \eta),$$

i el model de Levene coincidirà amb el de Dempster o amb un model sense subdivisió. Quan les eficàcies mitjanes a cada clapa siguin molt diferents, l'evolució pel model de Levene serà maxime-ment diferent de la corresponent als altres casos i l'efecte de dependència de les eficàcies de les freqüències (via eficàcies mitjanes) més marcat.

Els següents resultats numèrics concorden amb les idees anteriors.

Igualment pel cas de  $L=5$  loci, a partir d'una població en equilibri al lligament, amb totes les freqüències gèniques 0.5 i gamètiques estàndard 0.03125, hem realitzat 100 generacions en condicions de selecció amb manca d'inversions i manca de subdivisió ambiental, segons els mateixos esquemes que fins ara, llevat que a l'expressió de l'eficàcia d'un genotip  $xy$ ,  $w(x,y) = w_m - (O - F_{xy})^2$ , hem pres  $w_m=100$  i  $O=10$ , és a dir, un esquema de selecció direccional, a favor dels fenotips extrems, associats a una majoria de loci amb l'al·lel 1. La població final com és d'esperar, es caracteritza per una preponderància de gamets amb 4 i 5 loci amb l'al·lel 1, freqüències gèniques properes a 1, una eficàcia mitjana de 99.577 i una matriu de desequilibris gamètics positius.

A partir d'aquesta darrera població hem estudiat l'evolució dels gamets invertits extrems 00000 i 11111 introduïts en una freqüència 0.00125 en condicions de subdivisió del medi quan els òptims són 0 i 10, és a dir el dos fenotips més extrems, i també en condicions de manca de subdivisió, per un òptim 5. Els dos casos són comparables ja que hem considerat uns valors de la constant  $w_m=100$  i  $w_m=75$  respectivament, els quals impliquen els valors d'eficàcia expressats a la taula 5.4.15, per cada possible fenotip F.



Com la població de partida havia estat seleccionada per l'òptim extrem 10, a les noves situacions simulades l'eficàcia mitjana inicial és molt menor (indicada a la taula 5.4.15-II) donada la manca de coadaptació de les poblacions al nou ambient simulat. Al cas subdividit, a més, l'eficàcia mitjana a la primera clapa és molt menor que a la segona, per la mateixa raó expressada abans. Aquesta gran diferència a les eficàcies mitjanes ha d'implicar un efecte de les freqüències molt fort a les poblacions subdividides i una evolució apreciablement diferent, més que als casos estudiats fins ara. Això és el que efectivament passa, com podem comprovar si observem la taula 5.4.15-II.

Tant al cas subdividit com al no subdividit el gamet 00000 experimenta un increment inicial de freqüència donada la composició majoritària de la població inicial en gamets amb excés de loci amb l'al·lel 1, de manera que els genotips portadors de 00000 ténen una "habilitat combinatòria" inicial alta (entendem habilitat combinatòria en el sentit utilitzat per Fraser i Burnell (1970), com a capacitat per produir gamets que combinin adequadament amb els de la resta de la població per produir genotips òptims, no en el sentit més habitual d'heterosi obtinguda en creuar línies consanguínies). Però aquest increment inicial és més ràpid i intens al cas amb medi subdividit (màxima freqüència de 00000, .0242437091, a la generació 15) que al cas sense

subdivisió (màxim de 0.120567129 a la generació 30). Pel cas subdividit l'eficàcia mitjana inicial és molt menor a la primera clapa, on precisament el gamet 00000 és òptim, que a la segona ( $m_{w_1} = 6.247$ ,  $m_{w_2} = 99.524$ , taula 5.4.15-II), la qual cosa, segons hem comentat als paràgrafs anteriors té un efecte, dependent de les freqüències, que afavorirà l'augment de la freqüència de 00000. Els dos casos, subdividit i no subdividit, tendeixen a ésser equivalents a mesura que es fan iguals les eficàcies mitjanes de cada clapa.

L'avantatge inicial del gamet 00000 desapareix a mesura que augmenta l'eficàcia mitjana, és a dir, que es produeixen més gamets intermedis de manera que l'habilitat combinatòria de 00000 disminueix.

Pel cas del gamet 11111, ni a un cas ni a l'altre es produeix cap increment donada la seva manca d'habilitat combinatòria inicial, a causa que la majoria de gamets tenen excés d'alels 1 i per tant 11111 no combina adequadament per produir genotips intermedis.

Per simular l'efecte de la deriva sobre aquests models hem considerat la situació a la qual el nombre de zigots produïts a cada clapa el podem considerar molt gran (infinit) però el nombre d'adults supervivents relativament petit.

F	$w_1=100-(0-F)^2$	$w_2=100-(10-F)^2$	$w=0.5w_1+0.5w_2$	w/75
0	100	0	50	0.6
1	99	19	59	0.786
2	96	35	66	0.88
3	91	51	71	0.946
4	84	64	74	0.986
5	75	75	75	1
6	64	84	74	
7	51	91	71	
8	36	96	66	
9	19	99	59	
10	0	100	50	

F	$w=75-(0-5)^2$	w/75
0	50	0.6
1	59	0.786
2	66	0.88
3	71	0.946
4	74	0.986
5	75	1
6	74	
7	71	
8	66	
9	59	
10	50	

Taula 5.4.15.

Medi subdividit, òptims 0 i 10,  $w_m=100$ ,  $m_{w_1}=6.247$ ,  $m_{w_2}=99.524$ ,  
Inversió 00000

gen	freq.	$m_{w_1}$	$m_{w_2}$
0	$8.10986399 \times 10^{-3}$	15,343	98.337
5	.153377456	49.114	84.28
10	.24301734	62.354	74.422
15	.242437091	65.756	72.265
20	.21728705	66.979	72.228
25	.187038928	67.722	72.605
30	.156860384	68.344	72.995
35	.128614433	68.921	73.299
40	.103220355	69.468	73.5
45	.0811551734	69.981	73.602
100	$2.35751901 \times 10^{-3}$	72.563	72.92

Medi subdividit, òptim 5,  $w_m=75$ ,  $m_w=52.885$ ,  
Inversió 00000

gen	freq.	$m_w$
0	$1.76774052 \times 10^{-3}$	53.318
5	$8.59294327 \times 10^{-3}$	56.279
10	0.0295245826	60.35
15	0.0662878552	64.355
20	0.101592673	67.134
25	0.119796577	68.767
30	0.120567129	69.775
35	0.110240748	70.482
40	0.0947227464	71.02
45	0.0779094109	71.445
100	$2.64726711 \times 10^{-3}$	72.74

Taula 5.4.15-II Continua

Medi subdividit, òptims 0 i 10,  $w_m = 100$ ,  $m_{w_1} = 6.064$ ,  $m_{w_2} = 99.59$   
 Inversió 11111

gen.	freq.	$m_{w_1}$	$m_{w_2}$
0	$9.41658844 \times 10^{-4}$	14.486	98.776
5	$2.56851755 \times 10^{-4}$	38.67	93.575
10	$8.64641963 \times 10^{-5}$	52.639	87.619
15	$3.58968154 \times 10^{-5}$	61.012	82.565
20	$1.7702492 \times 10^{-5}$	65.896	78.943
25	$9.88215027 \times 10^{-6}$	68.732	76.565
30	$5.9911061 \times 10^{-6}$	70.387	75.072
35	$3.82773343 \times 10^{-6}$	71.359	74.156
40	$2.5262217 \times 10^{-6}$	71.93	73.602
45	$1.70054612 \times 10^{-6}$	72.273	73.269
100	$2.90280122 \times 10^{-8}$	72.773	72.777

Medi no subdividit, òptim 5,  $w_m = 75$ ,  $m_w = 52.827$   
 Inversió 11111

gen	freq.	$m_w$
0	$1.21717273 \times 10^{-3}$	53.239
5	$1.00341474 \times 10^{-3}$	55.976
10	$7.33742079 \times 10^{-4}$	59.63
15	$4.76575358 \times 10^{-4}$	63.455
20	$2.84761899 \times 10^{-4}$	66.688
25	$1.63779272 \times 10^{-4}$	69.026
30	$9.39279073 \times 10^{-5}$	70.552
35	$5.48105306 \times 10^{-5}$	71.487
40	$3.28160628 \times 10^{-5}$	72.038
45	$2.01827768 \times 10^{-5}$	72.356
100	$2.4396538 \times 10^{-7}$	72.77

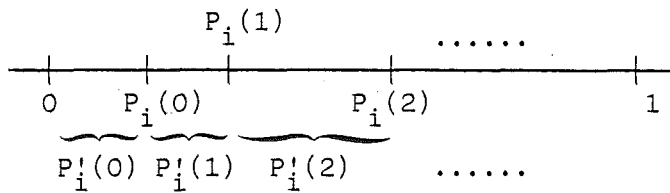
Taula 5.4.15-II

El procediment per simular aquest efecte de deriva és el següent: segons les expressions habituals pel càlcul de les freqüències gamètiques, calculem les freqüències corresponents a cada clapa  $p'_i(x)$  i  $q'_i$  (equacions 6.2.2). A continuació les freqüències gamètiques acumulatives, també a cada clapa

$$P_i(x) = \sum_{y \leq x} p'_i(y)$$

$$P_i(I) = P_i(2^L - 1) + q_i = 1$$

aquests valors defineixen una partició  $\{0, P_i(0), P_i(1), \dots, 1\}$  de l'interval  $[0, 1]$



amb la llargada del subinterval  $(P_i(x-1), P_i(x)]$  igual precisament al valor de freqüència gamètica  $P'_i(x)$ . Si produïm  $n$  nombres a l'atzar, segons una distribució uniforme dins de l'interval  $[0, 1]$ , la probabilitat que un d'ells,  $r$ , estigui comprès entre  $P_i(x-1)$  i  $P_i(x)$ :  $P_i(x-1) < r_j \leq P_i(x)$  és igual a la llargada  $P'_i(x) = P_i(x) - P_i(x-1)$ , de manera que contant la freqüència  $F_i(x)$  de vegades que es produeix l'esdeveniment  $P_i(x-1) < r_j \leq P_i(x)$  simulem l'aparició aleatòria de gamets segons les

frequències teòriques  $p_i'(x)$ . El nombre  $N_i$  de valors  $r_j$  generats segons una distribució uniforme, i utilitzats per produir les frequències gamètiques dels adults supervivents a la clapa  $i$ , cal que compleixi  $c_i = N_i/N$ , on  $c_i$  és la contribució relativa en adults de la clapa  $i$ , i  $N = N_1 + \dots + N_n$  és el total de nombres aleatoris generats, per totes les clapas juntes.

Després de repetir aquest procés per les  $n$  clapas, la frequència del gamet  $x$  a la població de la generació següent, formada per un nombre finit de  $N/2$  individus o  $N$  gamets, és

$$p_x = \frac{\sum_{i=1}^n F_i(x)}{N}$$

En una primera sèrie de simulacions hem estudiat l'evolució, durant 50 generacions, d'una inversió produïda de nou durant una fase en què la grandària poblacional es redueix a 400 individus. L'elecció d'aquest nombre és dictada bàsicament pel desig de comparar aquests resultats amb els de les simulacions finites tractades a la secció 6. La inversió introduïda és la produïda sobre el gamet 11000, és a dir, sobre el mateix considerat anteriorment i que augmenta en frequència a tots els casos deterministes. Hem fet 20 rèpliques en cada una de les quatre condicions estudiades també al cas determinista. Cada rèplica consisteix a suposar l'aparició d'un gamet invertit (per tant en frequència  $1/800=0.00125$ ) a les mateixes poblacions inicials que abans i seguir l'evolució de la seva frequència durant 50 generacions.

Pel cas 1) amb medi no subdividit ( $\delta$ ptim 5,  $w_m=65$ ) els resultats queden reflectits a la taula 5.4.16, on per cada rèplica i generació (comptades de 5 en 5 i comptant a partir de 0) hi ha la freqüència del gamet invertit.

Com podem veure els resultats clarament no concorden amb els de Fraser i Burnell (1970) ja que no hi ha la barrera de freqüència mínima inicial de la inversió per sota de la qual es perd amb probabilitat pràcticament 1. La freqüència inicial (0.00125) és clarament inferior al valor d'aproximadament 0.12 trobat pels autors esmentats com a freqüència mínima inicial. En 1 de les 20 rèpliques s'ha mantingut la inversió durant les 50 generacions de deriva, de manera que la freqüència relativa de vegades en què no s'ha perdut la inversió és  $1/20=0.05$ , aquest és un esdeveniment rar, però no impossible.

Encara més important és el fet que quan la inversió no es perd per atzar, la seva freqüència assoleix valors relativament grans, de 0.16875.

Cal recordar que pel cas determinista la progressió inicial de la inversió és extraordinàriament lenta. Quan la freqüència inicial és de 0.001 ja hem vist (figura 5.4.1) que calen unes 275 generacions per assolir una freqüència de l'ordre de 0.03. A la taula 5.4.17 hi ha el canvi de les freqüències d'inversió,



MODEL 1, MANCA DE SUBDIVISIO, DERIVA (800 GAMETS).

Rèplica	generac.	freq.
1	0	0
2	0	0
3	0	$1.25 \times 10^{-3}$
	5	0 (pèrdia 5)
4	0	$1.25 \times 10^{-3}$
	5	0 (pèrdia 2)
5	0	$1.25 \times 10^{-3}$
	5	0 (pèrdia 2)
6	0	$1.25 \times 10^{-3}$
	5	$5 \times 10^{-3}$
	10	$2.5 \times 10^{-3}$
	15	0.01
	20	0.02625
	25	0.03125
	30	0.05375
	35	0.0725
	40	0.0925
45	0.16875	
7	0	$2.5 \times 10^{-3}$
	5	0 (pèrdia 1)
8	0	$2.5 \times 10^{-3}$
	5	0 (pèrdia 1)
9	0	$1.25 \times 10^{-3}$
	5	0 (pèrdia 1)
10	0	$1.25 \times 10^{-3}$
	5	0 (pèrdia 2)
11	0	$3.75 \times 10^{-3}$
	5	0 (pèrdia 3)
12	0	$2.5 \times 10^{-3}$
	5	0.01
	10	$3.75 \times 10^{-3}$
	15	$5 \times 10^{-3}$
	20	$1.25 \times 10^{-3}$
	25	0 (pèrdia 21)

Taula 5.4.16 continua

continuació

Rèplica	generac.	freq.
13	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 3)
14	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 3)
15	0	0
16	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 2)
17	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 2)
18	0	$2.5 \times 10^{-3}$
	5	0.01125
	10	$7.5 \times 10^{-3}$
	15	$3.75 \times 10^{-3}$
	20	$6.25 \times 10^{-3}$
	25	0 (pèrdua 25)
19	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
20	0	0

Taula 5.4.16.

MODEL 1, SENSE SUBDIVISIÓ

generació	freqüència
0	$1.26687383 \times 10^{-3}$
5	$1.3545858 \times 10^{-3}$
10	$1.44831622 \times 10^{-3}$
15	$1.5485213 \times 10^{-3}$
20	$1.65564487 \times 10^{-3}$
25	$1.77015931 \times 10^{-3}$
30	$1.89257054 \times 10^{-3}$
35	$2.02341965 \times 10^{-3}$
40	$2.16328478 \times 10^{-3}$
45	$2.31278327 \times 10^{-3}$

MODEL 1, AMB SUBDIVISIÓ

generació	freqüència
0	$1.26687384 \times 10^{-3}$
5	$1.35458209 \times 10^{-3}$
10	$1.44829817 \times 10^{-3}$
15	$1.54848063 \times 10^{-3}$
20	$1.65557512 \times 10^{-3}$
25	$1.77005468 \times 10^{-3}$
30	$1.89242509 \times 10^{-3}$
35	$2.02322689 \times 10^{-3}$
40	$2.16302014 \times 10^{-3}$
45	$2.31245397 \times 10^{-3}$

MODEL 2, SENSE SUBDIVISIÓ

generació	freqüència
0	$1.26646224 \times 10^{-3}$
5	$1.3519519 \times 10^{-3}$
10	$1.4431329 \times 10^{-3}$
15	$1.54042077 \times 10^{-3}$
20	$1.64421881 \times 10^{-3}$
25	$1.75495333 \times 10^{-3}$
30	$1.87307876 \times 10^{-3}$
35	$1.99907918 \times 10^{-3}$
40	$2.13346976 \times 10^{-3}$
45	$2.27679844 \times 10^{-3}$

MODEL 2, AMB SUBDIVISIÓ

generació	freqüència
0	$1.26651598 \times 10^{-3}$
5	$1.35230273 \times 10^{-3}$
10	$1.44383232 \times 10^{-3}$
15	$1.54152687 \times 10^{-3}$
20	$1.64579729 \times 10^{-3}$
25	$1.75707853 \times 10^{-3}$
30	$1.87583463 \times 10^{-3}$
35	$2.00256047 \times 10^{-3}$
40	$2.13778331 \times 10^{-3}$
45	$2.28206458 \times 10^{-3}$

Taula 5.4.17. Freqüència del gamet invertit 11000, a les primeres generacions, introduït a una població coadaptada i amb freqüència 0.00125. Cas determinista.

pels diferents models i pel cas determinista, quan la freqüència inicial és de 0.00125, és a dir, per una situació comparable a l'estudiada pel cas amb deriva. Es pot comprovar que a la generació 45, la freqüència "determinista" és de l'ordre de  $2.3 \times 10^{-3}$ . És a dir, que les perturbacions aleatòries per deriva impliquen la pèrdua de la majoria de noves inversions, però al mateix temps poden implicar ràpides progressions per atzar d'algunes.

A la taula 5.4.18 hi ha els canvis de la freqüència de la inversió pel cas 2) amb medi no subdividit (òptim 8,  $w_m = 84$ ). Com és d'esperar, donada la semblança al canvi de les freqüències d'ambdós models malgrat les diferències en eficàcies, els resultats són similars i són vàlides les mateixes coses dites als paràgrafs anteriors.

No hi ha hagut cap rèplica a la qual s'hagi mantingut la inversió durant totes les generacions, però a diverses d'elles hi ha hagut increments per atzar de la freqüència, malgrat que posteriorment no han progressat a causa de canvis aleatoris en sentit contrari.

No es pot detectar cap diferència significativa entre ambdós casos amb deriva i ambient no subdividit. Per exemple si

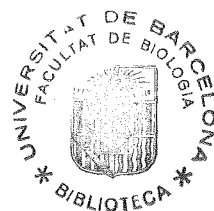
## MODEL 2, MANCA DE SUBDIVISIÓ, DERIVA (800 GAMETS).

Réplica	generació	freq.
1	0	$3.75 \times 10^{-3}$
	5	0 (perd. 5)
2	0	0
3	0	$1.25 \times 10^{-3}$
	5	$1.25 \times 10^{-3}$
	10	0 (perd. 7)
4	0	$5 \times 10^{-3}$
	5	$1.25 \times 10^{-3}$
	10	0 (perd. 6)
5	0	0
6	0	0
7	0	$1.25 \times 10^{-3}$
	5	0 (perd. 1)
8	0	0
9	0	0
10	0	$1.25 \times 10^{-3}$
	5	0 (perd. 2)
11	0	$1.25 \times 10^{-3}$
	5	$1.25 \times 10^{-3}$
	10	0.01625
	15	$8.75 \times 10^{-3}$
	20	$1.25 \times 10^{-3}$
25	0 (perd. 24)	
12	0	0
13	0	$3.75 \times 10^{-3}$
	5	0 (perd. 1)
14	0	0
15	0	$1.25 \times 10^{-3}$
	5	$7.5 \times 10^{-3}$
	10	$3.75 \times 10^{-3}$
	15	$7.5 \times 10^{-3}$
	20	$7.5 \times 10^{-3}$

Taula 5.4.18 continua

continuació	25	$1.25 \times 10^{-3}$
	30	0 (perd. 26)
<hr/>		
16	0	$2.5 \times 10^{-3}$
	5	0 (perd. 2)
<hr/>		
17	0	$1.25 \times 10^{-3}$
	5	0 (perd. 1)
<hr/>		
18	0	0
<hr/>		
19	0	$2.5 \times 10^{-3}$
	5	$1.25 \times 10^{-3}$
	10	0 (perd. 7)
<hr/>		
20	0	$1.25 \times 10^{-3}$
	5	$1.25 \times 10^{-3}$
	10	0 (perd. 9)
<hr/>		

Taula 5.4.18



comptem les rèpliques en què la inversió s'ha perdut a la primera generació, aquelles a les quals s'ha perdut en una generació posterior i les que l'han conservada tenim la següent taula de freqüències

	perd. a la 0	perd. posteriorment	no perd.
model 1	4	15	1
model 2	8	12	0

la qual té un valor de l'estadístic khi-quadrat (calculat ajuntant les dues darreres classes) de la prova d'homogeneïtat de 1.9048 amb 1 grau de llibertat, que és un valor no significatiu.

Cal fer notar també que tampoc no hi ha cap allunyament significatiu d'aquestes freqüències respecte d'allò que esperaríem en condicions de neutralitat: després de realitzar 10.000 rèpliques en condicions de neutralitat (xifra realitzable ja que aleshores el procés és molt més ràpid), les probabilitats dels tres esdeveniments anteriors (pèrdua a la primera generació, pèrdua en generació posterior i manteniment de la inversió) es poden estimar en 0.3491, 0.6095 i 0.0414 respectivament, valors que es poden prendre pràcticament com a paràmetres teòrics donada la gran grandària mostral. Els valors dels estadístics khi-quadrat per la prova d'ajust de les freqüències anteriors en aquestes probabilitats són respectivament, pel model 1) de

1.9567 i pel model 2) 0.2280, amb 1 grau de llibertat, clarament no significatius (calculats també ajuntant les dues darreres classes). De tota manera, a nivell purament descriptiu, podem dir que la diferència amb el cas neutre és més gran pel cas del model 1) al qual les diferències d'eficàcia són més marcades.

A la taula 5.4.19 hi ha els valors de freqüència corresponents al model 1) amb subdivisió del medi. Novament podem apreciar una freqüència relativa de  $2/20=0.1$  rèpliques a les quals no es perd la inversió, ans al contrari, arriba a unes freqüències de 0.035 i de 0.08375, superiors al valor de  $2.3 \times 10^{-3}$  esperable segons el model determinista (taula 5.4.17).

Si comparem els dos casos (no subdividit, subdividit) del model 1, a partir d'una taula de freqüències com l'anterior

	perduda a la 0	perduda posteriorment	no perduda
no subdiv.	7	11	2
subdiv.	4	15	1

obtenim un valor de l'estadístic khi-quadrat de la prova d'homogeneïtat de 1.7669 per dos graus de llibertat, no significatiu. Tampoc no és significatiu el valor khi-quadrat corresponent a la prova d'ajust a les probabilitats pel cas neutre (1.9567).



## MODEL 1, SUBDIVISIÓ, DERIVA (800 GAMETS).

Rèplica	generació	frequència
1	0	0
2	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
3	0	0
4	0	0
5	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
6	0	$3.75 \times 10^{-3}$
	5	$7.5 \times 10^{-3}$
	10	$1.25 \times 10^{-3}$
	15	0 (pèrdua 13)
7	0	0
8	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
9	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 3)
10	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
11	0	$5 \times 10^{-3}$
	5	0 (pèrdua 2)
12	0	$5 \times 10^{-3}$
	5	0 (pèrdua 5)
13	0	$1.25 \times 10^{-3}$
	5	$5 \times 10^{-3}$
	10	0.0175
	15	0.0225
	20	0.0375
	25	0.0325
	30	0.0175
	35	0.015
	40	0.02875
45	0.035	

Taula 5.4.19 continua.

continuació

14	0	0
15	0	$2.5 \times 10^{-3}$
	5	0.01125
	10	0.01875
	15	0.01875
	20	0.02625
	25	0.0125
	30	0.01625
	35	0.04375
	40	0.0475
	45	0.08375
16	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
17	0	$3.75 \times 10^{-3}$
	5	0 (pèrdua 2)
18	0	0
19	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
20	0	0

Taula 5.4.19.

## MODEL 2, SUBDIVISIÓ, DERIVA (800 GAMETS).

Rèplica	generació	freqüència
1	0	$2.5 \times 10^{-3}$
	5	0 (pèrdua 3)
2	0	0
3	0	0
4	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
5	0	$5 \times 10^{-3}$
	5	0.0125
	10	0.05375
	15	0.0375
	20	0.04125
	25	0.04
	30	0.0525
	35	0.0325
	40	0.02125
45	0.025	
6	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
7	0	0
8	0	0
9	0	0
10	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
11	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)
12	0	$6.25 \times 10^{-3}$
	5	0.01
	10	$2.5 \times 10^{-3}$
	15	0 (pèrdua 13)
13	0	$2.5 \times 10^{-3}$
	5	0 (pèrdua 5)
14	0	$1.25 \times 10^{-3}$
	5	0 (pèrdua 1)

continuació

13	0 5	$2.5 \times 10^{-3}$ 0 (pèrdua 5)
14	0 5	$1.25 \times 10^{-3}$ 0 (pèrdua 1)
15	0 5	$5 \times 10^{-3}$ 0 (pèrdua 5)
16	0	0
17	0 5	$1.25 \times 10^{-3}$ 0 (pèrdua 1)
18	0 5	$2.5 \times 10^{-3}$ 0 (pèrdua 1)
19	0	0
20	0	0

Taula 5.4.20.

Els resultats corresponents al model 2) pel cas amb subdivisió ambiental estan resumits a la taula 5.4.20. A una de les 20 rèpliques s'ha mantingut el gamet invertit. La freqüència final de la inversió per aquesta rèplica és de 0.025, superior al valor de l'ordre de  $2.28 \times 10^{-3}$  esperable segons el model determinista. Si comparem aquest cas amb el corresponent amb manca de subdivisió, la taula de freqüències és la següent

	Perduda a la 0	Perduda posteriorment	Mantinguda
No subdiv.	8	12	0
subd.	8	11	1

amb un valor de khi-quadrat de 0 (ajuntant les dues darreres classes per evitar les freqüències esperades de 0.5 corresponents a la darrera classe) evidentment no significatiu. La khi-quadrat d'ajust als valors de probabilitat del cas neutre és de 0.2280, no significativa.

Totes les khi-quadrats anteriors, d'associació entre factors, es poden resumir a la següent taula de contingència tridimensional, obtinguda quan considerem alhora els tres factors: 1) model (model 1, model 2), 2) subdivisió ambiental (manca de subdivisió, presència de subdivisió) i 3) sort de la inversió (pèrdua a la primera generació, pèrdua posterior, manteniment):

		Pèrdua a 0	Pèrdua post.	Manten.
Model 1	No sub.	7	11	2
	Sub.	4	15	1
Model 2	No Sub.	8	12	0
	Sub.	8	11	1

Si plantegem el contrast d'hipòtesis:

hipòtesi nul·la  $H_0$ : per tot  $i, j, k, p_{ijk} = p_i p_j p_k$

hipòtesi alternativa  $H_1$ : existeix  $i, j, k, p_{ijk} \neq p_i p_j p_k$

on  $p_{ijk}$  indica la probabilitat que es doni la combinació (intersecció) del nivell  $i$  del factor "model" amb el nivell  $j$  del factor "subdivisió" amb el nivell  $k$  del factor "sort de la inversió" mentre que  $p_i, p_j$  i  $p_k$  indiquen les probabilitats de cada un d'aquests nivells dels factors respectius per separat, és a dir, si contrastem l'hipòtesi nul·la de manca de cap dependència estocàstica entre els tres factors alhora amb l'alternativa d'existència de dependència, l'estadístic khi-quadrat

$$\chi^2 = \sum_i \sum_j \sum_k \frac{(O_{ijk} - E_{ijk})^2}{E_{ijk}}$$

té un valor 4.47, no significatiu per un nivell 0.05 i 2 graus de llibertat.  $O_{ijk}$  és la freqüència observada a la casella  $i, j, k$  mentre que  $E_{ijk} = \frac{O_i}{N} \times \frac{O_j}{N} \times \frac{O_k}{N} \times N = \frac{O_i \times O_j \times O_k}{N^2}$  és la freqüència esperada, essent

$$O_i = \sum_j \sum_k O_{ijk} \quad , \quad O_j = \sum_i \sum_k O_{ijk} \quad , \quad O_k = \sum_i \sum_j O_{ijk} \quad ,$$

les freqüències absolutes marginals observades i  $N = \sum_i \sum_j \sum_k O_{ijk}$   
el total de rèpliques realitzades.

Aquesta manca d'associació múltiple ja descarta la possibilitat d'existència de dependències parcials entre factors.

Dues conclusions són evidents a partir de les anàlisis anteriors. La primera, que no és possible detectar cap diferència significativa entre models ni entre graus de subdivisió del medi, ni entre aquests i allò que s'observa al cas neutre. El factor determinant és la deriva associada a la grandària poblacional de 400 individus.

La segona conclusió és que no hi ha una probabilitat 1 que es perdi per deriva una inversió produïda de nou a una població finita i que als casos als quals no es perd pot experimentar increments molt superiors als predits pel model determinista.