

APLICACION DE CADENAS ABSORBENTES DE MARKOV A UN MODELO DE
AUTOFECONDACION EN ALOHEXAPLOIDES.

N.J. BARTOLONI (1).

Recibido: 29-9-86

Aceptado: 27-3-87

RESUMEN

Ha sido considerado el caso de autofecundación para tres genes de un mismo locus en tres genomas distintos. Se han calculado la matriz de probabilidades de transición, las matrices espectrales de ella, los vectores correspondientes al número medio de generaciones que necesita un proceso para alcanzar la absorción, a la variancia y al momento centrado de tercer orden para dicha variable. También se han calculado la matriz promedio del número de pasajes a través de un estado determinado antes de absorción, el vector correspondiente al coeficiente de variación y el vector para el coeficiente de asimetría para cada estado transitorio. Se han considerado dos casos de selección (ventaja de heterocigotas y ventaja de homocigotas) y el caso de no selección, los tres con genomas distinguibles. Se efectúa una comparación final.

APPLICATION OF ABSORBING MARKOV CHAINS FOR A SELFING MODEL
IN ALOHEXAPLOIDS

SUMMARY

It has been considered the case of selfing for three genes in the same locus of three genomes. Are given the calculations of the transition probability matrix, the spectral matrices, the vectors of the number of generations that a process starting in a certain state needs to reach absorption, of the variance and the third order moments for the same variable. Also calculated the matrix of the number of times that a process passes through another before absorption, of the coefficients of variation and assymetry for each of the transient states. Three cases were considered: no selection, selective advantage for the heterozygous and selective advantage for the homozygous (all of three with distinguishable genomes). All final comparison is effectuated.

(1) Cátedra de Genética, Facultad de Agronomía, UBA.
Avda. San Martín 4453 (1417) Buenos Aires. Argentina.

INTRODUCCION

En problemas de genética de poblaciones, en estudios teóricos y en mejoramiento, es de vieja data el problema del cálculo de frecuencias genotípicas en generaciones avanzadas de endogamia. Han habido también muchos modelos y métodos de cálculo en respuesta. En este trabajo se ha considerado el caso particular de un gene triplicado en tres genomas con homeología completa y se han efectuado cálculos y comparaciones para tres casos bien definidos: a) no selección, b) ventaja de los heterocigotas para los tres genomas y, c) ventaja de ambos homocigotas, también para los tres genomas.

MATERIALES Y METODOS

Se ha utilizado una aplicación de la teoría de las cadenas finitas de Markov en sistemas genéticos de apareamiento. Se ha seguido la metodología sugerida por Bosso *et al.* (1969) y Feller (1980) para el cálculo de los vectores y matrices correspondientes al número de generaciones antes de absorción. Para el cálculo de las matrices espectrales se ha utilizado el método de las descomposiciones canónicas indicado por Máltsev (1978) de lo cual hay una aplicación elemental en Nelder (1952). Para la conformación de la matriz de probabilidades de transición (MPT) con selección se ha seguido el procedimiento indicado por Tan (1973). Para el caso de un gene en un hexaploide con dos alelos por cada genoma, se tiene la siguiente clasificación de estados:

- I) **absorbentes**; los tres genes al estado homocigota. Para tres genes: $(2) (2) (2) = 8$ estados posibles $(C_{3,0}; V_{2,3})$

- II) **transitorios de primer grado**: un gene al estado heterocigota y los otros dos al estado homocigota, o sea 12 estados posibles; $(C_{3,1}; V_{2,2})$

- III) **transitorios de segundo grado**: un gene al estado homocigota y los otros dos al estado heterocigota, es decir 6 estados $(C_{3,2}; V_{2,1})$

- IV) **transitorios de tercer grado**: los tres genes al estado heterocigota. Con genes dialélicos existe un solo estado posible para este grado, $(C_{3,3}; V_{2,0})$

La matriz de transición así como el resto de las matrices y vectores, aunque teniendo conformaciones similares varían en sus elementos según el caso. Se llamará **Caso I** al caso de no selección, **Caso II** al caso de ventaja de los heterocigotas y **Caso III** al de ventaja de los homocigotas. Considerando a las filas como determinando a la generación n y a las columnas como definiendo a la generación $n+1$ y disponiendo en primer término a los 8 estados absorbentes, en segundo término a los 12 estados transitorios de primer grado, en tercer término a los 6 transitorios de segundo grado y, por último, al de tercer grado tanto para filas como para columnas, la MPT es una matriz estocástica celular cuadrada triangular inferior (semidescompuesta) de cuarto orden. Las características de las 16 células matriciales son las siguientes:

- 1) las células $a_{12}, a_{13}, a_{14}, a_{23}, a_{24}$, y a_{34} son rectangulares nulas de orden $8 \times 12, 8 \times 6, 8 \times 1, 12 \times 6, 12 \times 1$ y 6×1 , respectivamente;

- 2) las células diagonales son matrices diagonales del tipo $(K) (E)$ donde E es una matriz unidad y donde K es un número real que varía según se trate de la MPT o de sus descomposiciones espectrales y varía dentro de

cada matriz de célula a célula, y cuyos órdenes son 8, 12, 6 y 1 para a_{11} , a_{22} , a_{33} , y a_{44} , respectivamente;

3) las células a_{41} , a_{42} y a_{43} son vectores fila de orden 8, 12 y 6, respectivamente, y del tipo k (E) donde E es un vector unidad y donde k varía entre MPT y sus espectrales y dentro de cada una de ellas, entre célula y célula;

4) las células a_{21} son rectangulares de orden 12×8 del tipo:

$$k \begin{pmatrix} A & O \\ B & B \\ O & A \end{pmatrix} \text{ donde } A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$B = E_{4 \times 4}$ y O es una matriz nula, y donde k varía según la matriz de que se trate.

5) las células a_{31} son de orden 6×8 del tipo:

$$K \begin{pmatrix} A & B & C & O \\ D & D & D & D \\ O & B & C & A \\ O & B & C & A \end{pmatrix} \text{ donde } A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}; B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix};$$

$$C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \text{ y } D = E \text{ y donde } k \text{ varía de matriz a matriz. En este caso } E \text{ es de orden } 2;$$

6) las células a_{32} son de orden 6×12 del tipo:

$$K \begin{pmatrix} A & B & C & O & D & O \\ F & D & E & E & F & D \\ I & F & I & B & C & G \end{pmatrix} \text{ donde } A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix};$$

$$C = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; D = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}; F = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; G = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix};$$

7) las células a_{22} , a_{23} , a_{24} , a_{33} , a_{34} , a_{42} , a_{43} , y a_{44} configuran la submatriz Q correspondiente a la MPT entre estados transitorios;

8) las células a_{21} , a_{31} y a_{41} configuran las submatrices R que corresponden a la MPT entre estados transitorios y estados ergódicos.

Con respecto a la matriz promedio N, cada elemento N_{ij} representa el número promedio de veces que un proceso que comienza en el estado S_i pasa a través de S_j antes de entrar en absorción. La configuración de la matriz N coincide con la de la matriz Q mencionada arriba, variando los valores de K de célula a célula y entre Caso y Caso de los que denominamos antes.

Los vectores columna T (cuyo elemento T_i representa el número medio de generaciones que un proceso que comienza en S_i necesita para llegar a un estado absorbente), T_2 (cuyo elemento T_{21} indica la variancia del número de etapas para alcanzar un estado absorbente comenzando en S_i) y T_3 (cuyo elemento T_{3i} señala el momento de tercer orden alrededor de la media T_i) tienen configuración:

$$\begin{pmatrix} kA \\ lB \\ mC \end{pmatrix} \text{ donde } A, B \text{ y } C \text{ son vectores unitarios de orden } 12, 6 \text{ y } 1, \text{ respectivamente, y donde } k, l \text{ y } m \text{ son diferentes números reales según de qué vector se trate. El número } k \text{ corresponde a los estados transitorios de primer grado, } l \text{ a los de segundo grado y } m \text{ a los de tercer grado.}$$

Los vectores correspondientes al coeficiente de variación (CV) y al coeficiente de asimetría (ASI) para cada estado, tienen una configuración similar a los T.

Valores selectivos

El cuadro general de valores selectivos para los tres genes es:

		GENOMA		
		A	B	C
GENOTIPO	A ₁ A ₁	x ₁	x ₂	x ₃
	A ₁ A ₂	1	1	1
	A ₂ A ₂	y ₁	y ₂	y ₃

Se considerarán para el **Caso I** $x_1=x_2=x_3=y_1=y_2=y_3$; λ_2 para el **Caso II** $x_1=x_2=x_3=y_1=y_2=y_3=0,5$ y, por último, los valores para el **Caso III** serán $x_1=x_2=x_3=y_1=y_2=y_3=2$.

Este esquema corresponde a genomas independientes en acción (valor selectivo agregado de los tres valores selectivos correspondientes a cada uno de los tres genes).

RESULTADOS

1) Ei en valores de la MPT

	λ_1	λ_2	λ_3	λ_4
Caso I	1	1/2	1/4	1/8
Caso II	1	4/7	5/16	1/6
Caso III	1	5/11	1/5	1/12

En cada caso el eigenvalor λ_1 es factor de la matriz espectral P_1 en la descomposición espectral; λ_2 es factor de P_2 ; λ_3 es factor de P_3 y λ_4 de P_4 . La matriz P_i , en todos los casos λ_i es las probabilidades de absorción final de cada estado transitorio.

2) Valores de k, 1 y m.

Caso I.

SUBMATRIZ

MATRIZ	a ₁₁	a ₂₁	a ₂₂	a ₃₁	a ₃₂	a ₃₃	a ₄₁	a ₄₂	a ₄₃	a ₄₄
MPT	1	1/4	1/2	1/16	1/8	1/4	1/64	1/32	1/16	1/8
P ₁	1	1/2	0	1/4	0	0	1/8	0	0	0
P ₂	0	-1/2	1	-1/2	1/2	0	-3/8	1/4	0	0
P ₃	0	0	0	1/4	-1/2	1	3/8	-1/2	1/2	0
P ₄	0	0	0	0	0	0	-1/8	1/4	-1/2	1
N	-	0	2	-	1/3	4/3	-	5/42	2/21	8/7

VECTOR	k	l	m
T	2,00	2,66	3,14
T ₂	2,00	2,66	2,93
T ₃	30,00	63,52	98,62
CV(%)	70,70	61,30	54,40
ASI	10,60	14,64	19,66

Caso II

SUBMATRIZ

MATRIZ	a ₁₁	a ₂₁	a ₂₂	a ₃₁	a ₃₂	a ₃₃	a ₄₁	a ₄₂	a ₄₃	a ₄₄
MPT	1	3/14	4/7	3/64	1/8	5/16	1/96	1/36	5/72	1/6
P ₁	1	1/2	0	1/4	0	0	1/8	0	0	0
P ₂	0	1/2	1	-14/29	14/29	0	-693/1972	231/986	0	0
P ₃	0	0	0	27/116	-14/29	1	135/406	40/87	10/21	0
P ₄	0	0	0	0	0	0	-101/952	23/102	10/21	1
N	-	-	7/3	-	14/33	16/11	-	49/330	4/33	6/5

VECTOR	k	l	m
T	2,33	3,15	3,71
T ₂	3,11	4,01	4,33
T ₃	49,46	106,29	166,06
CV(%)	75,57	63,57	56,08
ASI	9,01	13,23	18,43

CASO III

SUBMATRIZ

MATRIZ	a ₁₁	a ₂₁	a ₂₂	a ₃₁	a ₃₂	a ₃₃	a ₄₁	a ₄₂	a ₄₃	a ₄₄
MPT	1	3/11	5/11	3/40	1/8	1/5	1/48	5/144	1/18	1/12
P ₁	1	1/2	0	1/4	0	0	1/8	0	0	0
P ₂	0	-1/2	1	-55/112	55/112	0	-495/1372	165/686	0	0
P ₃	0	0	0	27/112	-55/112	1	135/392	-275/588	10/2	0
P ₄	0	0	0	0	0	0	-149/1372	935/4116	-10/21	1
N	-	-	11/6	-	55/192	5/4	-	5/48	5/66	12/11

VECTOR	k	l	m
T	1,83	2,39	2,83
T ₂	1,52	2,07	2,10
T ₃	22,51	45,36	73,57
CV(%)	67,37	60,19	51,20
ASI	12,01	15,23	24,17

CONSIDERACIONES

La MPT ha sido calculada según las leyes genéticas de segregación bajo un régimen de autofecundación y teniendo en cuenta valores selectivos de cada uno de los genotipos. La descomposición espectral permite el cálculo de las frecuencias genotípicas en cualquier generación de autofecundación mediante la expresión:

$$(MPT)^n = \lambda_1^n P_1 + \lambda_2^n P_2 + \lambda_3^n P_3 + \lambda_4^n P_4$$

La matriz espectral correspondiente al primer eigenvalor (P_1) brinda las probabilidades de que un proceso que comienza en un estado transitorio S_i ; termine en un estado absorbente S_j . Por esa razón sólo las células a_{11} , a_{21} , a_{31} , y a_{41} tienen valores positivos. Los restantes son iguales a cero. Los valores de probabilidades de absorción final también pueden ser calculados a partir de la matriz N como $(N)(R)$ (Bosso et al., 1969). Las expresiones para el cálculo de los distintos vectores (Bosso et al., 1969) se indicarán en los párrafos siguientes.

a) $N = (E - Q)^{-1}$ donde E es una matriz unidad. La matriz N ha sido calculada siguiendo el procedimiento habitual de inversión matricial a través del cálculo de las matrices adjuntas aprovechando el hecho de que la MPT es una matriz triangular lo que simplifica enormemente los cálculos. El i -ésimo componente del vector T es simplemente la suma de los elementos de la i -ésima fila de N .

b) $T_2 = (2N - E)T - T_{sq}$ donde $T_{sq} = T_i^2$

c) $T_3 = T'_3 - 3T_{21} + 2T^3$

siendo $T'_3 = 6N^2 T - 6NT + T$,

$$T_{21i} = (T_{21}) (T_i) \text{ y } T^3 = (T_i)^3$$

Los elementos T_{3i} del vector T_3 indican el momento centrado de tercer orden para el proceso que comienza en el estado transitorio S_i .

d) Para los coeficientes de variación y asimetría:

$$(CV)_i = \frac{(T_{21})^{0,5}}{T_i} \cdot 100$$

$$ASI_i = \frac{T_{3i}}{(T_{21})^{1,5}}$$

Las matrices P_1 son iguales para los tres casos dada la segregación simétrica que caracteriza a la autofecundación.

Los eigenvalores λ_2 , λ_3 y λ_4 son mayores para el Caso II, luego están los del Caso I y los menores son los del III y esto resulta fácilmente comprensible puesto que la solución del polinomio característico resulta coincidir con los elementos de la diagonal de la MPT y que son, justamente, las probabilidades de permanencia en los estados transitorios. Los valores de los elementos de T , T_2 y T_3 aumentan notoriamente al pasar del Caso I al III y disminuyen, también notoriamente, al pasar del Caso I al III como resulta lógico esperar. Pero los valores de variación y asimetría se mantienen bastante estables. A los fines de comodidad visual en la comparación se resumen en el Cuadro N° 1 los principales resultados.

Cuadro N° 1: Principales resultados.

		T	T ₂	T ₃	CV%	ASI
Caso I	k	2,00	2,00	30,00	70,70	10,60
	l	2,66	2,66	63,52	61,30	14,64
	m	3,14	2,93	98,62	54,40	19,66
Caso II	k	2,33	3,11	49,46	75,57	9,01
	l	3,15	4,00	106,29	63,57	13,23
	m	3,71	4,33	166,06	56,08	18,43
Caso III	k	1,83	1,52	22,51	67,37	12,01
	l	2,39	2,07	45,36	60,19	15,23
	m	2,83	2,10	73,57	51,20	24,17

CONCLUSIONES

- a) Los valores correspondientes al número medio de generaciones necesarios para alcanzar un estado absorbente habiendo partido de un estado transitorio aumentan cuando los heterocigotas tienen ventaja selectiva sobre los homocigotas, y disminuyen cuando los homocigotas tienen ventaja sobre los heterocigotas. Téngase en cuenta que se han considerado valores de aptitud similares pa-

ra los tres genes y que la disminución en aptitud se ha tomado de igual intensidad en los Casos I y III. Los valores de selección utilizados no implican una selección muy intensa.

- b) La forma de las distribuciones es bastante similar para los tres casos como lo demuestran los valores similares en los CV y en los ASI para cada uno de los grados de transición. Los valores de ASI positivos demuestran que las distribuciones en los tres casos tienen cola a la derecha lo que significa un nivel de heterocigosis relativamente superior al valor medio T_i después de varias generaciones de autofecundación para un proceso que comenzó en un estado transitorio ($\mu + \sigma$).

Dada la segregación simétrica característica de la autofecundación, los valores de probabilidad de fijación a partir de cada uno de los estados transitorios son independientes de los valores selectivos.

BIBLIOGRAFIA

- 1) BOSSO, J.A., O.M. SORRARAIN and E.A. FAVRET (1969). Application of finite Markov chains to sib mating populations with selection. *Biometrics*, 25:17-26.
- 2) FELLER, W.A. (1980). Introducción a la teoría de probabilidades y sus aplicaciones. Ed. Limusa, México, 504 págs. Volumen I.
- 3) TAN, E.Y. (1973). Applications of some finite Markov chain theories to two locus selfing model with selection. *Biometrics*, 29:331-346.
- 4) NELDER, J.A. (1952). Some genotypic frequencies and variance components occurring in biometrical genetics. *Heredity*, 6:387-394.
- 5) MALTSEV, A.I. (1978). Fundamentos de álgebra lineal. Ed. Mir, Moscú, 400 páginas.