IMA Journal of Mathematics Applied in Medicine and Biology (2001) **18**, 327–341

Genetic associations under mixed mating systems: the Bennett–Binet effect

J. A. VARGAS†
Independent scholar, Constitución 204, Oaxaca, Oaxaca 68000, México

AND

RAFAEL F. DEL CASTILLO‡
CIIDIR-IPN Oaxaca, Calle Hornos 1003, Xoxocotlán, Oaxaca 71230, México

[Received on 11 December 2000; revised on 19 December 2001]

Using elementary algebraic geometry and computational commutative algebra, supported by the program Macaulay2, we studied and developed operators that define the zygotic and gametic evolution under a mixed-mating system with parameters s selfing rate, r recombination rate, and g relative fitness of inbreeders, for any possible combination of initial zygotic or gametic frequencies with two alleles at each of two loci. We found that (i) the allelic frequencies are preserved in every generation; (ii) the gametic frequencies converge to values that depend exclusively on the allelic frequencies; (iii) every zygotic population converges to a population in equilibrium with double heterozygotes equally frequent; (iv) the rate of convergence decreases to arbitrary small values with sufficiently small values of r or with sufficiently large values of s and (v) as g decreases, the maximal 'association between the two loci' occurs with higher values of selfing. We also found generalizations for the case of several alleles at each locus.

Keywords: Bennett–Binet, two loci, mixed mating, recombination, projective space, rational map, zygotic algebra

1. Introduction

Outbreeding and self-fertilization are two basic sexual mechanisms present in plants and animals. The difference between them lies in the ancestry of the reduced gametes, or pronuclei, which fuse to form the zygote.

In selfing, both gametes are derived from the same organism, whereas outbreeding involves the fusion of gametes derived from two separate individuals.

Although selfing is widespread, it is well known that self-fertilized organisms may reproduce occasionally by outcrossing (Bell, 1982). Indeed, many hermaphroditic plants and animals reproduce by a combination of selfing and outcrossing, a method of sexual reproduction called mixed-mating (Aide, 1986). Such a system has interesting genetic consequences not present in pure selfing or in pure outcrossing, namely the development of nonrandom associations between loci, even if such loci are unlinked.

†Email: javargas1@excite.com
‡Email: rdelcast@yahoo.com

Haldane (1949) showed that, under partial inbreeding, excesses or deficiencies of doubly heterozygous and doubly homozygous zygotes became established, even in populations which are in linkage equilibrium. Bennett & Binet (1956) obtained mathematical expressions for the 'gametic and zygotic imbalances resulting from partial inbreeding', studying a case of two loci, each with two alleles. Later, Weir & Cockerham (1973) introduced two loci descent measures, and found genotypic frequencies and disequilibrium functions for every generation of a population reproducing with partial selfing with linked loci. In this paper, we generalize these results using elementary algebraic geometry and computational commutative algebra, powerful tools, yet underutilized in population genetics. For this purpose, we use the language of genetic algebras, with selfing introduced as indicated by Holgate (1978) and we construct operators that describe the zygotic and gametic evolution starting with any possible combination of zygotic or gametic frequencies for two loci, each with two alleles.

Our approach allows us to (*a*) find operators that produce the zygotic or gametic frequencies in any generation from information derived from the previous generation; (*b*) clarify the relation between the zygotic and gametic evolution through a commutative diagram; (*c*) give a clear account of the pure selfing case; (*d*) calculate the values of the Weir and Cockerham parameters at equilibrium; (*e*) prove that the parameters converge to those equilibrium values; (*f*) find expressions for the zygotic and gametic equilibrium states; (*g*) estimate the asymptotic dynamics of the process; (*h*) include a parameter expressing the possible reduced fitness of the progeny derived from inbreeding with respect to the progeny derived from outcrossing and (*i*) include a generalization for arbitrary many alleles at each locus.

2. The model

Consider a very large Mendelian population with no overlapping generations at two loci, each with two alleles: A_0, A_1 and B_0, B_1, whose respective frequencies are p_0, p_1, q_0, q_1, so that

$$0 \leqslant p_0, p_1, q_0, q_1 \leqslant 1, \;\; p_0 + p_1 = 1 \;\text{ and }\; q_0 + q_1 = 1. \tag{1}$$

The population reproduces by a mixed-mating system in which a fraction s of the progeny is derived from selfing and the rest, $1 - s$, is derived from random outcrossing. The fitness of the self progeny is a fraction g of the fitness of outbreeders, standarized to 1.

There exist ten zygotic genotypes:

$$\begin{gathered} A_0B_0/A_0B_0,\, A_0B_0/A_0B_1,\, A_0B_1/A_0B_1, \\ A_0B_0/A_1B_0,\, A_0B_0/A_1B_1,\, A_0B_1/A_1B_0,\, A_0B_1/A_1B_1, \\ A_1B_0/A_1B_0,\, A_1B_0/A_1B_1,\, A_1B_1/A_1B_1. \end{gathered} \tag{2}$$

We separate the two double heterozygotes: coupling A_0B_0/A_1B_1 and repulsion A_0B_1/A_1B_0, because their gametic outputs are different. These two genotypes recombine at a rate r. When $r = 0$, the two loci are completely linked; in the case $r = \frac{1}{2}$, they segregate independently.

The parameters r, s, g satisfy

$$0 \leqslant s, g \leqslant 1, \;\; 0 \leqslant r \leqslant \tfrac{1}{2}. \tag{3}$$

In the mathematical development, we can ignore conditions (1) and (3); we return to them only to derive biological consequences.

We have a ten-dimensional zygotic algebra $\mathfrak{Z}$, whose basis elements are labelled by these genotypes in the order given above. A population is identified with a point in projective space: $x = (x_0, \ldots, x_9) \in \mathbb{P}^9 = \mathbb{P}(\mathfrak{Z})$, see (Hartshorne, 1977), where each homogeneous coordinate x_i is the frequency of a genotype in x.

Right-hand multiplication with a basis element is given by right-hand multiplication with a matrix M_i, for $0 \leqslant i \leqslant 9$, like

$$M_0 = M_{A_0B_0/A_0B_0} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2}(1-r) & \frac{1}{2}r & 0 & \frac{1}{2}r & \frac{1}{2}(1-r) & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2}r & \frac{1}{2}(1-r) & 0 & \frac{1}{2}(1-r) & \frac{1}{2}r & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The result gives the structure of the population obtained from our population x outcrossing it with a population of fixed genotype.

We construct the outcrossing matrix $Q = x_0M_0 + \cdots + x_9M_9$, the selfing matrix S, whose ith row equals the ith row of M_i for $i = 0, \ldots, 9$; and the mixed-mating transformation $T = sghS + (1-s)Q$, where $h = x_0 + \cdots + x_9$ is a homogenizing factor that allows the use of projective space. This factor was first introduced by Holgate (1978).

The rational map, see (Hartshorne, 1977), $\varphi : \mathbb{P}^9 \to \mathbb{P}^9$, defined by $\varphi(x) = xT$ is the evolution operator. It gives the structure of the next generation. The construction of this model has followed steps similar to those in Vargas & del Castillo (1999) and Vargas (2000).

3. The gametic and allelic evolution

The four gametes A_0B_0, A_0B_1, A_1B_0, A_1B_1 occur with the following frequency in our population $x = (x_0, \ldots, x_9)$:

$$\begin{aligned} y_0 &= x_0 + \tfrac{1}{2}(x_1 + x_3 + x_4), & y_1 &= x_2 + \tfrac{1}{2}(x_1 + x_5 + x_6), \\ y_2 &= x_7 + \tfrac{1}{2}(x_3 + x_5 + x_8), & y_3 &= x_9 + \tfrac{1}{2}(x_4 + x_6 + x_8), \end{aligned} \tag{4}$$

where $A_0B_0 \leftrightarrow y_0$, $A_0B_1 \leftrightarrow y_1$, $A_1B_0 \leftrightarrow y_2$, $A_1B_1 \leftrightarrow y_3$.

There is a four-dimensional corresponding gametic algebra $\mathfrak{G}$ and an evolution operator $\psi : \mathbb{P}^3 \to \mathbb{P}^3$, where $\mathbb{P}^3 = \mathbb{P}(\mathfrak{G})$, so that the following is a commutative diagram:

$$\begin{array}{ccc} \mathbb{P}^9 & \xrightarrow{\rho} & \mathbb{P}^3 \\ \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \psi} \\ \mathbb{P}^9 & \xrightarrow{\rho} & \mathbb{P}^3, \end{array}$$

where the linear map $\rho(x) = (y_0, \dots, y_3)$ is given by (4).

If we write $y = (y_0, \dots, y_3) \in \mathbb{P}^3$ and $\psi(y) = (z_0, \dots, z_3)$, we find that

$$\begin{aligned} z_0 &= y_0 - \tfrac{1}{2}r(x_4 - x_5), & z_1 &= y_1 + \tfrac{1}{2}r(x_4 - x_5), \\ z_2 &= y_2 + \tfrac{1}{2}r(x_4 - x_5), & z_3 &= y_3 - \tfrac{1}{2}r(x_4 - x_5). \end{aligned} \tag{5}$$

These equations express the recombination phenomenon, as seen in Bennett & Binet (1956, p. 52). One cannot construct the evolution operator ψ intrinsically inside $\mathfrak{G}$, because the zygotic ingredient $x_4 - x_5$ is present; and it cannot be recovered from $\{y_0, \dots, y_3\}$. In fact, the matrix P has rank 5:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 1 \\ 0 & 1/2 & 1/2 & 0 & -1 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

and it satisfies

$$(y_0, y_1, y_2, y_3, x_4 - x_5) = (x_0, \dots, x_9)P. \tag{6}$$

The frequency of the alleles can be obtained from the x_i using

$$p_0 = (y_0 + y_1), \quad p_1 = (y_2 + y_3), \quad q_0 = (y_0 + y_2), \quad q_1 = (y_1 + y_3). \tag{7}$$

The allele frequencies are preserved from each generation to the next.

We write $w = (w_0, \dots, w_9) = \varphi(x)$, and we pass to the hyperplanes given by $x_0 + \cdots + x_9 = 1$ and $y_0 + \cdots + y_3 = 1$, which are subsets of $\mathbb{R}^{10}$ and $\mathbb{R}^4$ respectively, in the affine cones over $\mathbb{P}^9$ and $\mathbb{P}^3$, to obtain the following equations:

$$\begin{aligned} w_4 - w_5 &= 2(1-s)(y_0 - p_0q_0) - (srg - sr - \tfrac{1}{2}sg + r)(x_4 - x_5), \\ w_4 - w_5 &= -2(1-s)(y_1 - p_0q_1) - (srg - sr - \tfrac{1}{2}sg + r)(x_4 - x_5), \\ w_4 - w_5 &= -2(1-s)(y_2 - p_1q_0) - (srg - sr - \tfrac{1}{2}sg + r)(x_4 - x_5), \\ w_4 - w_5 &= 2(1-s)(y_3 - p_1q_1) - (srg - sr - \tfrac{1}{2}sg + r)(x_4 - x_5). \end{aligned} \tag{8}$$

They generalize the second equation of Bennett & Binet (1956, p. 52).

Equations (5) and (8) suffice to understand the gametic evolution and part of the zygotic evolution. Following Bennett and Binet, we observe that there are four two-dimensional vector spaces, contained in the algebra of polynomial functions over $\mathbb{R}^{10} \times \mathbb{R}^4$, invariant under an operator induced by both φ and ψ. These are generated by $x_4 - x_5$ and one of

$$y_0 - p_0q_0, \quad y_1 - p_0q_1, \quad y_2 - p_1q_0, \quad y_3 - p_1q_1.$$

These four quantities are called *linkage disequilibria*. It follows from (8) that they have the same absolute value for any given population x. Furthermore, the evolution operator in each invariant plane is given by the matrix

$$E = \begin{pmatrix} 1 & \mp\frac{1}{2}r \\ \pm 2(1-s) & -(srg - sr - \frac{1}{2}sg + r) \end{pmatrix}.$$

PROPOSITION 3.1 If $s < 1$ and $0 < r < \frac{1}{2}$, then E is a contraction: E is diagonalizable with eigenvalues of absolute value less than 1.

Proof. The polynomial $p(\lambda) = \det(\lambda - E)$ satisfies

$$\begin{aligned} p(0) = \det\ E &= sg(\tfrac{1}{2} - r) > 0, \\ p(s) &= -(1-s)s[1 - (\tfrac{1}{2} - r)g - r] < 0, \\ p(1-r) &= -sr[1 - (\tfrac{1}{2} - r)g - r] < 0, \\ p(1) &= (1-s)r > 0. \end{aligned} \tag{9}$$

Hence, $p(\lambda)$ has two positive real roots: one between 0 and s and the other one between s and 1. □

We observe that the biggest root is also bigger than $1 - r$.

Starting with a point $x = x^{(0)} \in \mathbb{R}^{10}$ over $\mathbb{P}^9$ and a corresponding point $y = y^{(0)} \in \mathbb{R}^4$ over $\mathbb{P}^3$, let $x^{(n)}$ and $y^{(n)}$ be the trajectories that evolve from x and y. The fact that E is a contraction implies that

$$\begin{aligned} x_4^{(n)} - x_5^{(n)} &\to 0, \\ y_0^{(n)} \to p_0q_0, \quad y_1^{(n)} \to p_0q_1, \quad y_2^{(n)} \to p_1q_0, \quad y_3^{(n)} &\to p_1q_1, \\ \text{as } n \to \infty. \end{aligned} \tag{10}$$

Equations (9), indicate that the rate of convergence for (10) decreases with larger values of s and with smaller values of r.

PROPOSITION 3.2 If $s < 1$ and $r = \frac{1}{2}$, then (10) holds.

Proof. In this case, the eigenvalues of E are 0 and $\frac{1}{2}(1 + s)$. □

4. Pure selfing case

The case $s = 1$, which occurs when all the progeny is derived from selfing, is extremely simple; and our language is fit to exhibit this simplicity.

The zygotic evolution operator reduces to $\varphi(x) = xS$. The matrix S is diagonalizable, with positive eigenvalues, the largest of which is 1, with corresponding eigenspace generated by the rows of the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Hence, $\lim_{n\to\infty}(x_0, \ldots, x_9)S^n = (x_0, 0, x_2, 0, 0, 0, 0, x_7, 0, x_9)$.

Thus, in this situation, only the double homozygotes persist, with respective zygotic equilibrium frequencies:

$$\overline{x}_0 = \frac{x_0}{x_0 + x_2 + x_7 + x_9}, \quad \overline{x}_2 = \frac{x_2}{x_0 + x_2 + x_7 + x_9},$$

$$\overline{x}_7 = \frac{x_7}{x_0 + x_2 + x_7 + x_9}, \quad \overline{x}_9 = \frac{x_9}{x_0 + x_2 + x_7 + x_9}.$$

The gametic equilibrium frequencies $y_0, \ldots, y_3$ satisfy

$$y_0 = \overline{x}_0, \quad y_1 = \overline{x}_2, \quad y_2 = \overline{x}_7, \quad y_3 = \overline{x}_9.$$

At equilibrium, we have $y_0 - p_0q_0 = y_3 - p_1q_1 = \Gamma$, $y_1 - p_0q_1 = y_2 - p_1q_0 = -\Gamma$, with $\Gamma = y_0y_3 - y_1y_2$.

5. The fixed points

From now on, we assume that $s \neq 1$.

The map φ has fixed points with closure given by the ideal I generated by the 2×2 minors of the 2×10 matrix with rows x and $\varphi(x)$.

At this point, one may try to guess the solutions for the system of equations I; alternatively, consider the polynomials in p_0, p_1, q_0, q_1 and the x_i obtained from

$$p_0(y_2 + y_3) - p_1(y_0 + y_1), \qquad q_0(y_1 + y_3) - q_1(y_0 + y_2) \tag{11}$$

by elimination of the y_i using (4), call them w_0, w_1. Form the ideal $J = \text{ideal}\,(w_0, w_1)$; and solve a system of linear equations contained in the ideal $K = \text{saturate}\,((I + J) : h)$, see (Eisenbud, 1995). The ideal J expresses the invariance of the alleles, while K discards points on the hyperplane $Z(h)$.

The solutions form the following irreducible surface $\mathbb{S}$ of degree 8:

$$\begin{aligned} x_u &= \sum_{3(i+k)+j+l=u} T_{ij,kl}, \text{ for } u = 0, 1, 2, 3, 4; \\ x_u &= \sum_{3(i+k)+j+l=u-1} T_{ij,kl}, \text{ for } u = 5, 6, 7, 8, 9; \end{aligned} \tag{12}$$

where

$$c = (1-r)^2 + r^2, \quad F = \frac{sg}{sg - 2s + 2},$$

$$\Phi = \frac{sg\{[sg - 2(1-s)]c - 2sg\}}{(sg - 2s + 2)[sgc - 2(sg - s + 1)]}, \quad \Delta = \Phi - F^2,$$

$$T_{ij,kl} = \left[F p_i \delta_{ik} \sum_{t=0}^{n} p_t + (1-F) p_i p_k \right] \left[F q_j \delta_{jl} \sum_{s=0}^{n} q_s + (1-F) q_j q_l \right]$$
$$+ \Delta p_i q_j \left[\left(\delta_{ik} \sum_{t=0}^{n} p_t \right) - p_k \right] \left[\left(\delta_{jl} \sum_{s=0}^{n} q_s \right) - q_l \right]. \qquad (13)$$

5.1 *Observations*

Regarding these expressions, we note:

(1) The symbols δ_{ik} and δ_{jl} are used for the Kronecker delta, so that $\delta_{ik} = 1$ if and only if $i = k$; and $\delta_{ik} = 0$, otherwise.
(2) The solutions are homogeneous polynomials in p_t, q_s. This is in order to qualify as valid projective geometry entities.
(3) We have considered only the case with two alleles ($n = 1$), but our result (13) allows arbitrary values for n. The next remark proves the validity of (13) for the case of more than two alleles at each locus.
(4) It turns out that we have calculated the parameters F, Φ and Δ, described by Weir and Cockerham as 'inbreeding coefficient', 'probability that an individual chosen at random is autozygous at both loci' and 'a measure of the association between the two loci' for the fixed points; and we will prove in the next section that every biologically relevant point converges to a fixed point.
(5) The parameters F, Φ and Δ have thus probability and biology content, although our methods have used no probability.
(6) Our solutions generalize those obtained by Haldane (1949), and also by Kimura (1958) and Weir & Cockerham (1973). We have adopted the compact notation of Nagylaki (1992). The substitution $g = 1$ produces the values obtained by these authors.
(7) The parameter F satisfies $0 \leqslant F \leqslant 1$. Hence, $0 \leqslant (1-F) \leqslant 1$.
(8) It is easy to see that $\Delta \geqslant 0$ with $\Delta = 0$ if and only if $s = 0, 1$.
(9) Equations (13) give the frequency of each 'ordered' zygotic genotype, while equations (12) give the same answer for the 'unordered' or true genotypes in case $n = 1$. The 'unordered' genotypes were the ones used to construct the algebra $\mathfrak{Z}$. There are 16 'ordered' zygotic genotypes for $n = 1$.
(10) The substitution $g = 1$ for two alleles, yields the fixed points given in Bennett & Binet (1956), validating the correction indicated in Nagylaki (1992).
(11) If $i \neq k$ and $j \neq l$, then

$$T_{ij,kl} = T_{kj,il} = T_{il,kj} = T_{kl,ij} = [(1-F)^2 + \Delta] p_i p_k q_j q_l.$$

The coefficient $(1-F)^2 + \Delta$ does not involve any allelic frequency, it only involves

the parameters r, s, g. Allowing a parametric coefficient, any product $p_i p_k q_j q_l$ with $i \neq k$ and $j \neq l$, occurs as $T_{ij,kl}$.

(12) If $i \neq k$ and $j \neq l$, then

$$T_{ij,il} = T_{il,ij},\ T_{ij,kj} = T_{kj,ij}.$$

(13) The formulae (13) also give answers in the case the loci A and B have $n+1$ and $m+1$ alleles respectively, with $n \neq m$: Assuming $m < n$, we write $q_{m+1} = \cdots = q_n = 0$, to obtain possibly non-zero values only for $T_{ij,kl}$ with $j, l \leqslant m$; and these values are given by (13) with the index of all sums involving the q_s going from 0 to m.

(14) The case $m = 0$ of the preceding observation is equivalent to the one-locus case. Hence, the frequency of the 'ordered' zygote $A_i A_j$ is

$$F p_i \delta_{ij} \sum_{t=0}^{n} p_t + (1 - F) p_i p_j,$$

while the frequency of the unordered zygote $A_i A_j$ is

$$F p_i \delta_{ij} \sum_{t=0}^{n} p_t + (2 - \delta_{ij})(1 - F) p_i p_j.$$

(15) We introduce the following notation for $i \neq k$, $j \neq l$:

$$\begin{aligned} t_{ij,kl} &= T_{ij,kl} + T_{kj,il} + T_{il,kj} + T_{kl,ij}; \\ t_{ij,il} = T_{ij,il} + T_{il,ij}, t_{ij,kj} &= T_{ij,kj} + T_{kj,ij}; \\ t_{ij,ij} &= T_{ij,ij}; \end{aligned} \tag{14}$$

in order to have for $0 \leqslant i, j, k, l \leqslant n$, the equation

$$\begin{aligned} t_{ij,kl} = &\left[F p_i \delta_{ik} \sum_{t=0}^{n} p_t + (2 - \delta_{ik})(1 - F) p_i p_k \right] \\ &\times \left[F q_j \delta_{jl} \sum_{s=0}^{n} q_s + (2 - \delta_{jl})(1 - F) q_j q_l \right] \\ +(-2)^{(2-\delta_{ik}-\delta_{jl})} \Delta p_i q_j &\left[\left(\delta_{ik} \sum_{t=0}^{n} p_t \right) - p_k \right] \left[\left(\delta_{jl} \sum_{s=0}^{n} q_s \right) - q_l \right]. \end{aligned} \tag{15}$$

The coefficient $\Delta = D(r, s, g)$, called *identity disequilibrium* by Weir & Cockerham (1973), is graphed as a function of the parameter s, with the indicated values for the other parameters, in Figs 1 and 2.

6. Geometric interpretation

There are $\binom{n+2}{2}$ monomials of degree 2 in $n + 1$ variables; and we associate a point in $\mathbb{P}^{\binom{n+2}{2}-1}$ to a zygotic population with one locus and $n + 1$ alleles. Observation 14 means

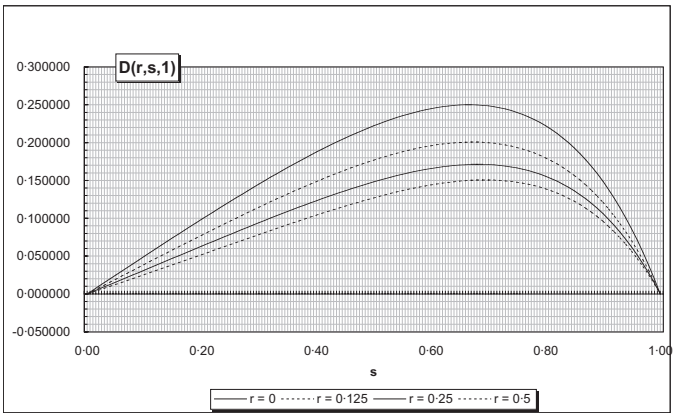

FIG. 1. $\Delta = D(r, s, 1)$.

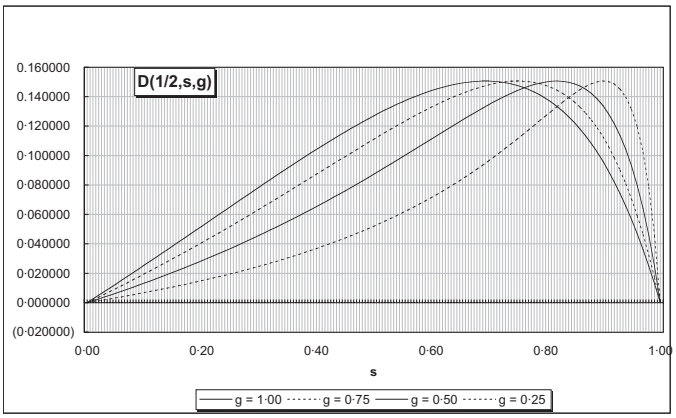

FIG. 2. $\Delta = D(\frac{1}{2}, s, g)$.

that the fixed points of the appropriate evolution operator acting on $\mathbb{P}^{\binom{n+2}{2}-1}$ have closure equal to the image of the Veronese embedding of degree two $\mathbb{P}^n \to \mathbb{P}^{\binom{n+2}{2}-1}$, followed by an automorphism of $\mathbb{P}^{\binom{n+2}{2}-1}$, the composition being $d : \mathbb{P}^n \to \mathbb{P}^{\binom{n+2}{2}-1}$ with

$$d(p_0, \ldots, p_n) = \left(\ldots, F p_i \delta_{ij} \sum_{t=0}^{n} p_t + (2 - \delta_{ij})(1 - F) p_i p_j, \ldots \right). \tag{16}$$

This parametrization generalizes our results in Vargas & del Castillo (1999), and the last equations of Weir & Cockerham (1973, p. 258).

The zygotic algebra $\mathfrak{Z}_n$ for two loci each with $n + 1$ alleles is obtained from the corresponding gametic algebra $\mathfrak{G}_n$ by commutative duplication.

The zygotic genotypes, which are used to construct $\mathfrak{Z}_n$ are all expressed as $A_i B_j / A_k B_l$ with $0 \leqslant i, j, k, l \leqslant n$, with the identifications $A_i B_j / A_k B_l = A_k B_l / A_i B_j$. If $i \neq$

k, $j \neq l$, then the genotypes A_iB_j/A_kB_l and A_iB_l/A_kB_j are different, and are capable of recombination.

The frequency of each A_iB_j/A_kB_l in a stable population with given allelic frequencies $p_0, \ldots, p_n$ and $q_0, \ldots, q_n$ is $T_{ij,kl}$, before any identification is made.

Starting with the vector space $\mathfrak{Z}_n$, we may further identify A_iB_j/A_kB_l with A_iB_l/A_kB_j. Formally, we construct the vector space $\mathcal{A}_n$ with basis $\{A_iA_kB_jB_l\}$, where each $A_iA_kB_jB_l$ is one representative for the class of A_iB_j/A_kB_l under all of the above identifications.

$\mathcal{A}_n$ is a homomorphic image of the vector space $\mathfrak{Z}_n$, and there is a natural projection $\mathbb{P}(\mathfrak{Z}_n) \to \mathbb{P}^{\binom{n+2}{2}^2-1} = \mathbb{P}(\mathcal{A}_n)$.

$\mathcal{A}_n$ classifies the zygotic genotypes up to their 'allelic content', (see Holgate, 1978). The frequency of $A_iA_kB_jB_l$ in a stable population is now $t_{ij,kl}$.

Observation 15 asserts that the projection in $\mathbb{P}^{\binom{n+2}{2}^2-1} = \mathbb{P}(\mathcal{A}_n)$ of the closure of the fixed points in $\mathbb{P}(\mathfrak{Z}_n)$ of the appropriate evolution map is the image of the composition $\beta \circ \sigma \circ (d \times d)$, where

$$\mathbb{P}^n \times \mathbb{P}^n \xrightarrow{d\times d} \mathbb{P}^{\binom{n+2}{2}-1} \times \mathbb{P}^{\binom{n+2}{2}-1} \xrightarrow{\sigma} \mathbb{P}^{\binom{n+2}{2}^2-1} \xrightarrow{\beta} \mathbb{P}^{\binom{n+2}{2}^2-1}, \tag{17}$$

d is given by (16), σ is the Segrè map: $\sigma((y_{ik}), (z_{jl})) = (x_{ij,kl})$ with $x_{ij,kl} = y_{ik}z_{jl}$, and $\beta \in \mathbb{PGL}_{\binom{n+2}{2}^2-1} = \mathrm{Aut}\,(\mathbb{P}^{\binom{n+2}{2}^2-1})$ is the *Bennett–Binet automorphism*.

We now order and partition the basis of $\mathcal{A}_n$ in three parts, so that the first part consists of the double heterozygotes and the last part contains all double homozygotes leaving in the middle part the remaining elements.

With respect to this partitioned basis, the automorphism β can be expressed as $I + B$, where I is the identity matrix and B is the partitioned matrix

$$B = \begin{pmatrix} C_0 & C_1 & C_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 = \end{pmatrix} \begin{pmatrix} + & - & + \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

such that $C_0 = 4bI$, with I an identity matrix of the appropriate size, C_1 a matrix with non-zero entries equal to $-2b$; and C_2 a matrix with non-zero entries equal to b, with

$$b = \frac{\Delta}{4(1-F)^2}. \tag{18}$$

The signs of the non-zero elements of C_0, C_1 and C_2 are as indicated in view of Observations (7), (8) and (15). These signs express the increased homozygosity and heterozygosity that we call the *Bennett–Binet effect*.

7. The convergence

In this section, we always have $n = 1$. Recall that $\mathfrak{Z} = \mathfrak{Z}_1$ is the zygotic algebra for two loci. We write $\mathcal{Z}$ for the one-locus zygotic algebra.

Let $\theta_1, \theta_2 : \mathbb{P}^9 = \mathbb{P}(\mathfrak{Z}) \to \mathbb{P}^2 = \mathbb{P}(\mathcal{Z})$ be the linear projections defined by

$$\theta_1(x) = (x_0 + x_1 + x_2, x_3 + x_4 + x_5 + x_6, x_7 + x_8 + x_9),$$
$$\theta_2(x) = (x_0 + x_3 + x_7, x_1 + x_4 + x_5 + x_8, x_2 + x_6 + x_9). \tag{19}$$

These maps describe the natural omission of one locus. They both factor through the linear map $\tau : \mathbb{P}^9 \to \mathbb{P}^8 = \mathbb{P}(\mathcal{A}_1)$ defined as

$$\tau(x_0, \ldots, x_9) = (x_0, x_1, x_2, x_3, x_4 + x_5, x_6, x_7, x_8, x_9). \tag{20}$$

The projection τ originates from the natural map $\mathfrak{Z}_1 \to \mathcal{A}_1$.

We have a commutative diagram

$$\begin{array}{ccccc} \mathbb{P}^2 & \xleftarrow{\theta_1} & \mathbb{P}^9 & \xrightarrow{\theta_2} & \mathbb{P}^2 \\ \downarrow{\scriptstyle \eta} & & \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \eta} \\ \mathbb{P}^2 & \xleftarrow{\theta_1} & \mathbb{P}^9 & \xrightarrow{\theta_2} & \mathbb{P}^2, \end{array}$$

where η is the evolution operator for one the locus model studied in Vargas & del Castillo (1999).

The convergence of the trajectories defined by each η is very fast, as seen in Vargas & del Castillo (1999), so that a fixed point is reached in each $\mathbb{P}^2$ before the convergence (10) is achieved for most points x with $x_4 \neq x_5$.

We claim that every point $x \in \mathbb{P}^9$, such that $x, \theta_1(x), \theta_2(x)$ and $\rho(x)$ are not fundamental points for φ, η and ψ, as applicable, converges to a point fixed under φ.

In order to prove this, we may assume that $\theta_1(x), \theta_2(x)$ and $\rho(x)$ are already fixed. Choose $a, b \in \mathbb{P}^2$ fixed under η, for each one-locus case, to have $\theta_1^{-1}(a) \cong \mathbb{A}^7 \cong \theta_2^{-1}(b)$; and assume $x \in (\theta_1^{-1}(a) \cap \theta_2^{-1}(b))$.

Then we choose the unique point $c \in \rho(\theta_1^{-1}(a) \cap \theta_2^{-1}(b)) \subseteq \mathbb{P}^3$ with $\psi(c) = c$, and write $K = Z(x_4 - x_5)$, to have

$$\mathcal{F}_p = (\theta_1^{-1}(a) \cap \theta_2^{-1}(b) \cap \rho^{-1}(c)) \cap K \cong \mathbb{A}^3,$$

where $\mathcal{F}_p$ contains a unique point $p \in \mathbb{P}^9$, fixed under φ; and identified with the origin of $\mathbb{A}^3$.

We may now assume $x \in \mathcal{F}_p$. The set $\mathcal{F}_p$ is invariant under translation by the row space of the matrix

$$\begin{pmatrix} 0 & -1 & 1 & 1 & 0 & 0 & -1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 & 0 & 0 & -1 & 0 & -1 & 1 \\ 1 & -2 & 1 & -2 & 2 & 2 & -2 & 1 & -2 & 1 \end{pmatrix}. \tag{21}$$

The action of the evolution operator φ on any such $\mathcal{F}_p$, with respect to the basis formed by the rows of (21), is given by the matrix

$$\frac{sg}{sg - s + 1} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & r^2 - r + \frac{1}{2} \end{pmatrix}.$$

Hence, φ induces a linear contraction on $\mathcal{F}_p$ with centre p. This implies that any point $x \in \mathbb{P}^9$, such that $x, \theta_1(x), \theta_2(x)$ and $\rho(x)$ are not fundamental points for φ, η and ψ, as applicable, converges to a fixed point.

8. The generalized gametic and allelic evolution

We fix two loci, each with $n+1$ alleles: $A_0, \ldots, A_n$ and $B_0, \ldots, B_n$. Then we have that a given population has associated to it two points $p, q \in \mathbb{P}^n$, each corresponding to a locus, where $p = (p_0, \ldots, p_n)$ and $q = (q_0, \ldots, q_n)$ have homogeneous coordinates such that p_i is the frequency of A_i and q_j is the frequency of B_j.

There are $(n+1)^2$ gametic genotypes A_iB_j for $i, j = 0, \ldots, n$.

We write y_{ij} for the frequency of A_iB_j in our population, to study the point $y = (y_{00}, \ldots, y_{nn}) \in \mathbb{P}^{(n+1)^2-1}$.

Every biologically relevant point converges to a fixed point, and in the case y is a fixed population, we have that

$$y_{ij} = \sum_{k,l=0}^{n} T_{ij,kl} = p_i q_j \Big(\sum_{k=0}^{n} p_k\Big)\Big(\sum_{l=0}^{n} q_l\Big). \tag{22}$$

We can use $\sum p_k = 1 = \sum q_l$, to see that at equilibrium, the gametic frequencies are the products of corresponding allelic frequencies, and the closure of the set of fixed points in $\mathbb{P}^{(n+1)^2-1}$ of the appropriate gametic evolution operator is given by the homogeneous ideal generated by all quadratic equations

$$y_{ij}y_{kl} = y_{il}y_{kj}; \qquad \text{with} \quad i, j, k, l = 0, \ldots, n. \tag{23}$$

It is easy to see, using a one-locus model, that the allelic frequencies are preserved from each generation to the next. This is basically because recombination only rearranges alleles, and the selection parameter g acts symmetrically on them.

9. Discussion

We have found that convergence towards the fixed points is like the convergence of power series with base

$$w_1 = \frac{sg}{2(sg - s + 1)} \tag{24}$$

or w_2 with

$$p(w_2) = w_2^2 - [sr(1-g) + \tfrac{1}{2}sg - r + 1]w_2 + sg(\tfrac{1}{2} - r) = 0, \tag{25}$$

which is the characteristic polynomial of matrix E in Section 3.

At any one single locus, the convergence is fast, like a power series with base w_1. This is also the case for the convergence after gametic equilibrium. The inclusion of inbreeding depression accelerates this convergence, because $g < 1$ implies

$$w_1 = \frac{sg}{2(sg - s + 1)} < \frac{s}{2}. \tag{26}$$

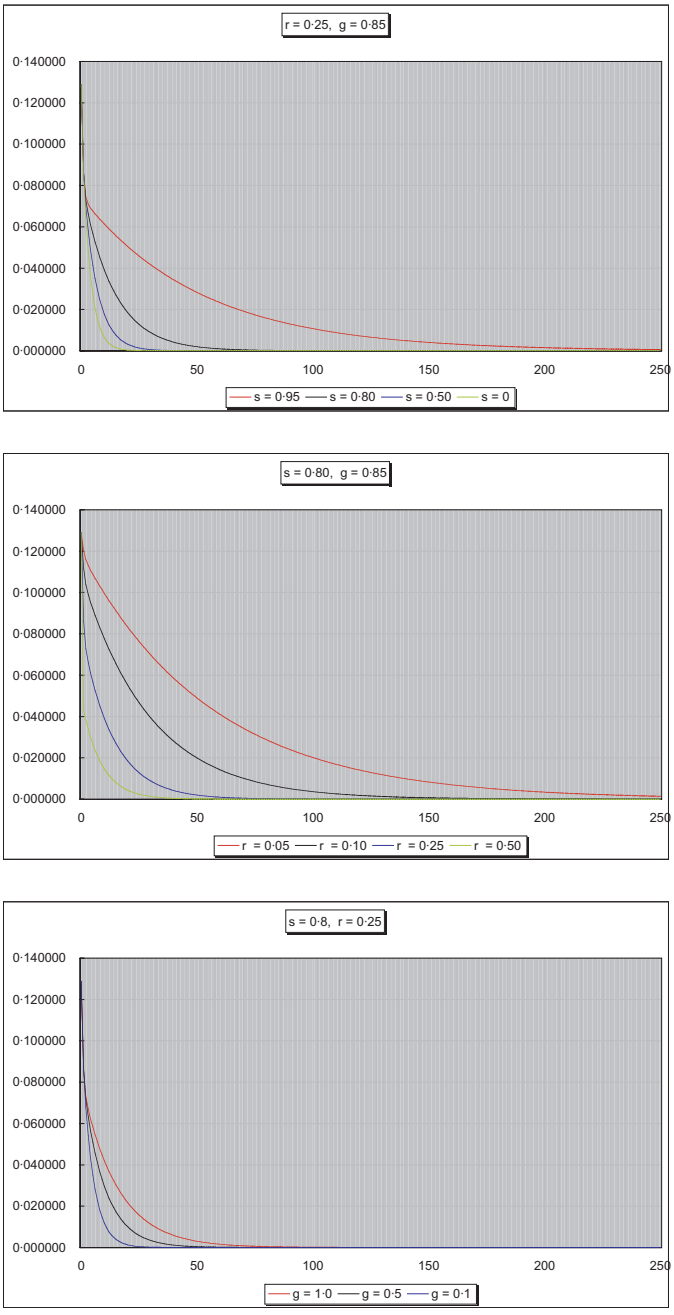

FIG. 3. Linkage disequilibria.

The potentially slow convergence takes place at the gametic level, the 'linkage disequilibria' $y_0 - p_0q_0$, etc. are intertwined with the behaviour of the zygotic frequency difference $x_4 - x_5$.

Our base w_2 is close to 1, slowing the gametic convergence, whenever s is close to 1 or r is close to zero. Thus, we are in agreement with Bennett & Binet (1956).

Weir & Cockerham (1973) claim that slow convergence, comparable with that of a power series with base w_3 satisfying

$$w_3^2 - \frac{\lambda + s + 1}{2} w_3 + \frac{\lambda s}{2} = 0 \tag{27}$$

occurs, symmetrically, for λ and s close to unity, for a parameter λ that they introduced. Their result agrees with ours because their parameter is related to recombination r by the equation

$$\lambda = 1 - 2r. \tag{28}$$

The effect of a fitness of inbreeders $g < 1$ on the gametic convergence is a marginal increase in the speed of this convergence.

Figure 3 shows the behaviour through time of the common absolute value of all linkage disequilibria for the indicated parameters. Here, in all graphs, we start we the same initial population

(0·005369, 0·016107, 0·268456, 0·013423, 0·033557,

0·375839, 0·107383, 0·026846, 0·080537, 0·072483).

The keen scrutiny on two-loci models, derived from the descent measures of Weir & Cockerham (1973), has allowed us to pass from two alleles at each loci to an arbitrary number of alleles. We have valid results for the fixed points and the convergence towards them in the general case, basically because we could prove the convergence to zero of all linkage disequilibria; then the remaining parameters F and Φ or Δ converge to values calculated using two alleles per locus.

Acknowledgements

Rafael del Castillo is grateful for the support of COFAA and the Darwin Initiative for the Survival of Species during the development of this work.

References

AIDE, T. M. (1986) The influence of wind and animal pollination on variations in outcrossing rates. *Evolution*, **40**, 434–435.

BELL, G. (1982) *The Masterpiece of Nature. The Evolution and Genetics of Sexuality*. Berkeley, CA: University of California Press.

BENNETT, J. H. & BINET, F. E. (1956) Association between Mendelian factors with mixed selfing and random mating. *Heredity*, **10**, 51–55.

COCKERHAM, C. C. & WEIR, B. S. (1973) Descent measures for two loci with some applications. *Theor. Pop. Biol.*, **4**, 300–330.

EISENBUD, D. (1995) *Commutative Algebra with a View toward Algebraic Geometry*, Graduate Texts in Mathematics 150. New York: Springer.

HALDANE, J. B. S. (1949) The association of characters as a result of inbreeding and linkage. *Ann. Eugenics*, **15**, 15–23.

HARTSHORNE, R. (1977) *Algebraic Geometry*, Graduate Texts in Mathematics 52. New York: Springer.

HOLGATE, P. (1978) Selfing in genetic algebras. *J. Math. Biol.*, **6**, 197–206.

KIMURA, M. (1958) Zygotic frequencies in a partially self-fertilizing population. *Ann. Rept. Natl Inst. Genet.*, **8**, 104–105.

NAGYLAKI, T. (1992) *Introduction to Theoretical Population Genetics*. New York: Springer.

VARGAS, J. A. & DEL CASTILLO, R. F. (1999) Inbreeding depression in a zygotic algebra. *Comm. Alg.*, **27**, 4425–4432.

VARGAS, J. A. (2000) Hardy–Weinberg theory for tetraploidy with mixed mating. *Adv. Appl. Math.*, **24**, 369–383.

WEIR, B. S. & COCKERHAM, C. C. (1973) Mixed self and random mating at two loci. *Genet. Res.*, **21**, 247–262.