*Journal of Difference Equations and Applications*
Vol. 15, No. 4, April 2009, 371–385

# Dynamical behaviour of a discrete selection-migration model with arbitrary dominance

James F. Selgrade[a1]*, Jordan West Bostic[b2] and James H. Roberds[c3]

[a]*Department of Mathematics and Biomathematics Program, North Carolina State University, Raleigh, NC, USA;* [b]*Department of Mathematics, North Carolina State University, Raleigh, NC, USA;* [c]*USDA Forest Service, Southern Research Station, Southern Institute of Forest Genetics, Saucier, MS, USA*

(*Received 21 March 2008; final version received 1 August 2008*)

In honor of Jim Cushing on the occasion of his 65th birthday

To study the effects of immigration of genes (possibly transgenic) into a natural population, a one-island selection-migration model with density-dependent regulation is used to track allele frequency and population size. The existence and uniqueness of a polymorphic genetic equilibrium is proved under a general assumption about dominance in fitnesses. Also, conditions are found which guarantee the existence of and determine the location of the global attractor for this model. The rate at which solutions approach the attractor is approximated. A measure of allelic diversity is introduced.

**Keywords:** natural selection; immigration; transgene; dominance; attractor

**AMS 2000 Mathematics Subject Classification:** 92D25; 39A11

## 1. Introduction

It is likely that alleles of genes carried by migration into a population will differ in frequency from alleles of the same genes in the recipient population. For cases in which these genes affect fitness, their allele frequencies in the receiving population will be shaped by the action of natural selection, as well as the rate of migration [2,14]. Here, we explore dynamics produced by the one-island model (also referred to as the continent-island model) which is designed to investigate the effects of natural selection together with one-way migration [6,10–12] in a single population. We consider a diploid population with two alleles, $A$ and $a$, segregating at a single gene locus. In each generation, a migrating population with constant allele frequency is incorporated into the gene pool of the island population through random mating. Our model is a discrete-time system of two nonlinear difference equations that describes changes in frequency for the $A$ allele, $p$, and population size, $x$, over generations. Density-dependent selection and migration are assumed to be the two evolutionary forces affecting change in allele frequency and population size in the island population.

This system of equations may be used to study the fate of transgenes that have migrated into non-targeted populations. In forest biology, there are concerns about

---

*Corresponding author. Email: selgrade@math.ncsu.edu

ISSN 1023-6198 print/ISSN 1563-5120 online

DOI: 10.1080/10236190802400741
http://www.informaworld.com

ecological consequences (see Williams [13]) that could arise because of transfer of such genetically engineered genes into natural tree populations. Understanding the dynamical behaviour that results from the combined action of selection and migration in the setting of the one-island model has potential to provide information about how transgene movement impacts genetic variability in these populations. Other discrete models have been investigated pertaining to the interaction of transgenic and natural populations. For instance, to study the mixing of transgenic mosquitoes resistant to malaria and wild type non-resistant mosquitoes, Li [3,4] introduced a model that tracks genotype numbers rather than allele frequency and total population size. He demonstrated that chaotic behaviour can result for his three-dimensional system [4].

In this paper, Section 2 discusses model background and Section 3 establishes properties for polymorphic equilibria. In Section 4, we discuss the classical genetic notion of *dominance* in fitness and present biologically reasonable conditions that guarantee the uniqueness of a polymorphic equilibrium for the cases of partial and complete dominance. Furthermore, we determine how allele frequency changes as the degree of dominance changes. Section 5 proves the existence of global attractors under general dominance assumptions and calculates attractor approach rates. If $q$ denotes a constant frequency for the allele $A$ in the immigrating population then the global attractor is contained within either the region where $p < q$ or the region where $p > q$, depending on which homozygote fitness is greater. This result generalizes a similar result in Ref. [11] for the case of complete dominance. Section 6 introduces a measure of allelic diversity for certain invariant sets and illustrates how this measure might change as the dominance parameter changes. Effects on the genetic composition of the recipient population are discussed.

## 2. Model background

For the one-island model, the island population consists of individual zygotes with one of three genotypes, $AA$, $Aa$ or $aa$. Let $x$ denote the island population size or density and let $p$ denote the frequency of the $A$ allele in the island population, where $0 \leq p \leq 1$. Hence, $1 - p$ is the frequency of the $a$ allele. The effects of density-dependent natural selection determine an average *per capita* replacement rate or *fitness* $f_{ij}(x)$ for the $ij$-genotype, where $i,j = A,a$, which measures fertility and viability of that genotype. Allele fitnesses $f_A$ and $f_a$ are linear combinations of genotype fitnesses weighted by allele frequency and are defined by $f_A \equiv pf_{AA} + (1 - p)f_{Aa}$ and $f_a \equiv pf_{Aa} + (1 - p)f_{aa}$. The population mean fitness $f$ is given by $f \equiv pf_A + (1 - p)f_a$.

Following selection in each generation, assume a constant number of gametes are contributed to the island population by immigration from a population with constant allele frequency $q$ where $0 \leq q \leq 1$. Random mating occurs following migration, so that the number of additional zygotes in the next generation due to immigration is denoted by the constant $y$ (Figure 1). The following system of difference equations describes changes in allele frequency and population size between generations $n$ and $n + 1$, see Refs. [6,10]:

$$p_{n+1} = \frac{p_n x_n f_A(p_n, x_n) + qy}{x_n f(p_n, x_n) + y}, \quad x_{n+1} = x_n f(p_n, x_n) + y. \tag{2.1}$$

System (2.1) depicts generational changes due to post-selection migration. When $y = 0$, this system is identical to the system previously studied for density-dependent selection, e.g. Ref. [9].

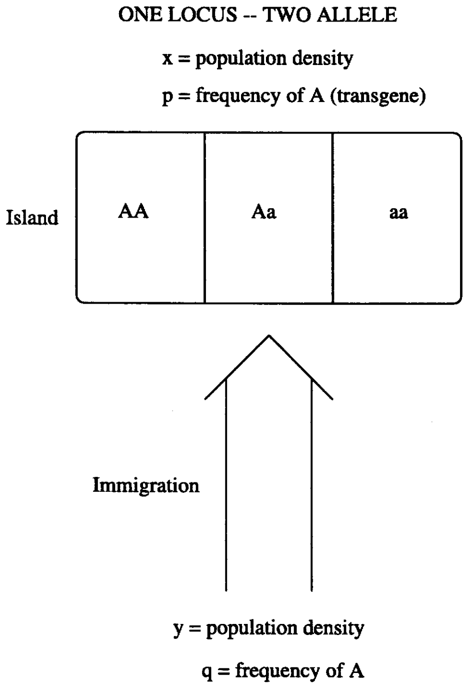

Figure 1. Immigration into an island population consisting of three genotypes.

It is convenient both mathematically and biologically to introduce a *per capita* migration rate for $x > 0$ given by

$$h(x) \equiv \frac{y}{x},$$

which measures a constant *per capita* migration per generation relative to the island population size $x$. Since $y$ is constant, $h(x)$ is a decreasing function of population size, in particular, $h'(x) = -y/x^2 < 0$. After replacing $y$ by $xh(x)$ in (2.1), the transition equations become

$$p_{n+1} = \frac{p_n f_A(p_n, x_n) + qh(x_n)}{f(p_n, x_n) + h(x_n)}, \quad x_{n+1} = x_n(f(p_n, x_n) + h(x_n)). \tag{2.2}$$

In (2.2), $f + h$ denotes the *per capita* transition function for the island population. Repeated iteration of (2.2) yields an orbit $\{(p_n, x_n) : n = 0, 1, 2, \ldots\}$ for this two-dimensional, dynamical system which is equivalent to (2.1) for $x > 0$. The phase space for system (2.2) is the slot in the $(p,x)$-plane designated by

$$\mathcal{S} \equiv \{(p, x) : 0 \leq p \leq 1, 0 < x\}.$$

When $y = 0$ (i.e. $h = 0$), the boundary lines of $\mathcal{S}$, $\{p = 0\}$ and $\{p = 1\}$, represent allele fixation and therefore, are invariant. If $y > 0$ and $0 < q < 1$, points on these vertical boundaries of $\mathcal{S}$ are mapped into the interior of $\mathcal{S}$ under the dynamical system.

## 3. Equilibria and isoclines

An equilibrium $E$ is an allele frequency $\bar{p}$, $0 \le \bar{p} \le 1$ and a population density $\bar{x} > 0$ which remain constant across generations, i.e. $p_n = \bar{p}$ and $x_n = \bar{x}$ for all $n$. Such an $E$ is said to be *polymorphic* if $0 < \bar{p} < 1$. From (2.2), an equilibrium $E = (\bar{p}, \bar{x})$ must satisfy the following system:

$$\bar{p} = \bar{p} f_A(\bar{p}, \bar{x}) + q h(\bar{x}), \quad 1 = f(\bar{p}, \bar{x}) + h(\bar{x}). \tag{3.1}$$

Since the frequency of the allele $a$ is $1 - p$ and is constant at equilibrium, the equation $1 - p_{n+1} = 1 - p_n$ implies that the following equation must also be satisfied at equilibrium:

$$1 - \bar{p} = (1 - \bar{p}) f_a(\bar{p}, \bar{x}) + (1 - q) h(\bar{x}). \tag{3.2}$$

Hence, a polymorphic equilibrium is a point of intersection of the three isocline curves:

$$\begin{aligned} \mathcal{C} &\equiv \{(p,x) : f(p,x) + h(x) = 1\}, \\ \mathcal{C}_A &\equiv \{(p,x) : p[f_A(p,x) - 1] + q h(x) = 0\} \text{ and} \\ \mathcal{C}_a &\equiv \{(p,x) : (1-p)[f_a(p,x) - 1] + (1-q) h(x) = 0\}. \end{aligned} \tag{3.3}$$

Any pair of equations in (3.3) will determine $E$.

Because of the detrimental effects of crowding, it is often assumed that genotype fitness $f_{ij}(x)$ decreases as population size increases. In addition, we may assume that each $f_{ij}(0) > 1$ and $f_{ij}(x) \to 0$ as $x \to \infty$. This guarantees that for each fixed $p$, $0 \le p \le 1$, there is a population density such that the population equilibrates, i.e. for each $p$ there is an $x > 0$, so that $f(p,x) = 1$. For $i,j = A,a$, we label this assumption as follows:

$$f'_{ij}(x) < 0 \text{ for all } x > 0, \; f_{ij}(0) > 1 \text{ and } f_{ij}(x) \to 0 \text{ as } x \to \infty. \tag{A1}$$

From (A1) it follows that $\partial f_A/\partial x < 0$, $\partial f_a/\partial x < 0$ and $\partial f/\partial x < 0$. Hence, the implicit function theorem gives that the curves defined in (3.3) may be considered as the graphs of $x$ as functions of $p$, which will be denoted by $\tilde{x}(p)$, $\tilde{x}_A(p)$ and $\tilde{x}_a(p)$, respectively. Since $\partial f/\partial p = 2(f_A - f_a)$, we have

$$\frac{\mathrm{d}\tilde{x}}{\mathrm{d}p} = \frac{-2(f_A - f_a)}{(\partial f/\partial x) + h'}, \tag{3.4a}$$

$$\frac{\mathrm{d}\tilde{x}_A}{\mathrm{d}p} = \frac{1 - f_A - p(\partial f_A/\partial p)}{p(\partial f_A/\partial x) + q h'}. \tag{3.4b}$$

Using (A1) and the fact that $h(x) \searrow 0$ as $x \to \infty$, it follows that the function $\tilde{x}(p)$ exists for all $p \in [0,1]$ and the curve $\mathcal{C}$ separates $\mathcal{S}$ into two subsets. Also, the function $\tilde{x}_A(p)$ has a vertical asymptote at $p = 0$ and exists for all $p \in (0,1]$. Appealing to these properties, Roberds and Selgrade [6] showed that the isoclines must cross at least once and so there is at least one polymorphism.

THEOREM 3.1. EXISTENCE OF A POLYMORPHIC EQUILIBRIUM. *Fix* $0 < q < 1$ *and* $y > 0$ *and assume* (A1). *Then,* (2.2) *has at least one polymorphic equilibrium* $E = (\bar{p}, \bar{x})$, *i.e.* $0 < \bar{p} < 1$ *and* $\bar{x} > 0$.

## 4. Dominance in fitness

The level or degree of dominance with regard to genetic control of fitness has been shown to affect substantially the properties of genetic equilibria influenced by migration and selection, e.g. Hedrick [2] and Nagylaki [5]. Here, we study dynamical behaviour when selection is density-dependent and where the genotypic fitness of the heterozygote is a linear combination of the homozygote fitnesses, i.e. *dominance*. Thus, for a real parameter $\delta$, we take

$$f_{Aa}(x) = \delta f_{AA}(x) + (1-\delta)f_{aa}(x) = f_{aa}(x) + \delta[f_{AA}(x) - f_{aa}(x)], \tag{4.1}$$

for all $x > 0$ or all $x$ within an invariant set. The concept of degree of dominance presupposes that heterozygote fitness can be expressed in terms of homozygote fitnesses. Selgrade and Roberds [11,12] studied the special cases, where $\delta = 1$ (*complete dominance*) and where $\delta = 0.5$. *Partial dominance* refers to the case, when $0 < \delta < 1$. If $\delta > 1$, then the heterozygote is said to be *overdominant*. See Falconer and Mackay [1] or Roughgarden [8] for application of this concept to selection theory in population genetics. With dominance, allele and mean fitnesses become

$$\begin{aligned} f_A &= f_{aa} + (p + \delta(1-p))[f_{AA} - f_{aa}], \\ f_a &= f_{aa} + \delta p[f_{AA} - f_{aa}] \\ f &= f_{aa} + p(p + 2\delta(1-p))[f_{AA} - f_{aa}]. \end{aligned} \tag{4.2}$$

If allele $A$ represents a transgene, $a$ represents a null allele and the $AA$ genotype is conferred with a certain fitness advantage then the heterozygote may express this advantage at some level. On the other hand, the natural population (the $aa$ genotype) may provide more drought resistance than the $AA$ genotype and the heterozygote may also partly manifest this. The degree of dominance $\delta$ scales the heterozygote fitness relative to the fitnesses of the homozygotes.

Theorem 3.1 establishes the existence of a polymorphic equilibrium but with dominance in fitness a uniqueness result also holds. From (4.2) notice that

$$f_A(p,x) - f_a(p,x) = (\delta(1-p) + p(1-\delta))[f_{AA}(x) - f_{aa}(x)]. \tag{4.3}$$

Hence, if $f_{AA}(x) > f_{aa}(x)$, then (3.4a) implies that the function $\bar{x}(p)$ is increasing. Selgrade and Roberds [12] proved uniqueness in the case of complete dominance ($\delta = 1$) by showing that the function $\bar{x}_A(p)$ is also decreasing, when $f_{AA}(x) > f_{aa}(x)$. Here, we obtain a general result for partial dominance where the monotonicity of $C_A$ does not hold. Because of technicalities in our proof, we assume that $0.5 \le q < 1$, but numerical simulations indicate that the result also holds for $0 < q < 0.5$.

THEOREM 4.1. UNIQUENESS OF POLYMORPHISM. *Assume that each genotype fitness $f_{ij}$ satisfies* (A1), *that* (4.1) *holds with $0 < \delta \le 1$ and that $f_{AA}(x) > f_{aa}(x)$, for all $x > 0$. If $0.5 \le q < 1$ or $\delta = 1$, then the polymorphic equilibrium $E = (\bar{p}, \bar{x})$ is unique. In addition, $q < \bar{p}$.*

*Proof.* To establish the result we show that where the isoclines $\mathcal{C}$ and $\mathcal{C}_A$ cross that

$$\frac{\mathrm{d}\tilde{x}_A}{\mathrm{d}p} < \frac{\mathrm{d}\tilde{x}}{\mathrm{d}p}. \tag{4.4}$$

Hence, $\mathcal{C}$ crosses $\mathcal{C}_A$ from below to above as $p$ increases and this can happen at most once for $0 < p < 1$.

From (3.4a) and (3.4b) for $0 < p < 1$ observe that (4.4) is equivalent to

$$\frac{p(1 - f_A - p(\partial f_A/\partial p))}{p(\partial f_A/\partial x) + qh'} = p\frac{\mathrm{d}\tilde{x}_A}{\mathrm{d}p} < p\frac{\mathrm{d}\tilde{x}}{\mathrm{d}p} < \frac{\mathrm{d}\tilde{x}}{\mathrm{d}p} = \frac{-2(f_A - f_a)}{(\partial f/\partial x) + h'}. \tag{4.5}$$

By cross multiplying the first and last terms in (4.5) and moving all terms to one side, we obtain the equivalent inequality

$$\left\{p\left(1 - f_A - p\frac{\partial f_A}{\partial p}\right)\right\}\left(\frac{\partial f}{\partial x} + h'\right) + \{2(f_A - f_a)\}\left(p\frac{\partial f_A}{\partial x} + qh'\right) < 0. \tag{4.6}$$

We need to verify the inequality (4.6) at each equilibrium $E = (\bar{p}, \bar{x})$. Use the first equation in (3.1) to substitute $qh(\bar{x})$ for $\bar{p}(1 - f_A)$ in the first bracketed term in (4.6). Then use equation (4.2) to write allele fitnesses and their derivatives in terms of genotypic fitnesses and their derivatives at $E$. Finally, rearrange terms by factors of $f'_{AA}(\bar{x}), f'_{aa}(\bar{x})$ and $h'(\bar{x})$ to get the inequality

$$\begin{aligned}
&f'_{AA}(\bar{x})[f_{AA}(\bar{x}) - f_{aa}(\bar{x})]\bar{p}\{(\bar{p} + 2\delta(1 - \bar{p}))(1 - \delta)(\bar{p}(1 - \bar{p}) + \bar{p})) + 2\delta^2(1 - \bar{p})\} \\
&\quad + f'_{AA}(\bar{x})\bar{p}qh(\bar{x})(\bar{p} + 2\delta(1 - \bar{p})) + f'_{aa}(\bar{x})qh(\bar{x})(1 - \bar{p})(1 - \delta\bar{p} + \bar{p}(1 - \delta)) \\
&\quad + f'_{aa}(\bar{x})[f_{AA}(\bar{x}) - f_{aa}(\bar{x})]\bar{p}(1 - \bar{p})(1 - \delta)(\bar{p}(1 - \bar{p}) + 2\delta(1 - \bar{p})^2) \\
&\quad + h'(\bar{x})\{qh(\bar{x}) + [f_{AA}(\bar{x}) - f_{aa}(\bar{x})](2\delta q(1 - \bar{p}) + (1 - \delta)\bar{p}(1 - \bar{p}))\} \\
&\quad + h'(\bar{x})[f_{AA}(\bar{x}) - f_{aa}(\bar{x})](1 - \delta)\bar{p}(2q - 1) \quad < 0.
\end{aligned} \tag{4.7}$$

Notice that each addend in (4.7) is negative or zero except possibly the last one. And since $0 < q$, the fourth line in (4.7) is strictly negative. The last addend is less than or equal to zero if $\delta = 1$ or if $0.5 \leq q$. Hence, inequality (4.7) holds and uniqueness follows.

To see the last assertion of Theorem 4.1 multiply the second equation in (3.1) by $\bar{p}$ and subtract from the first equation in (3.1) to obtain

$$(\bar{p} - q)h(\bar{x}) = \bar{p}[f_{AA}(\bar{x}) - f_{aa}(\bar{x})](1 - \bar{p})\{\bar{p}(1 - \delta) + \delta(1 - \bar{p})\}. \tag{4.8}$$

Since the right side of (4.8) is positive, we have $q < \bar{p}$. □

We will investigate how the subtle interplay of selection and dominance affect allele frequency. For example, if the *AA* genotype fitness is superior to the *aa* fitness, then one might think that increasing $\delta$ would favour the *A* allele and cause its frequency $p$ to increase. But increasing $\delta$ also increases the heterozygote fitness which should lead to more *Aa* genotypes in the population and a possible decrease in $p$.

First, we investigate how the position of a polymorphism $E = (\bar{p}, \bar{x})$ changes with dominance parameter $\delta$. To solve (3.1) for $\bar{p}$ and $\bar{x}$ as functions of $\delta$, we use (3.1) to define functions $F$ and $G$ as follows:

$$F(p,x,\delta) \equiv pf_A + qh - p, \quad G(p,x,\delta) \equiv f + h - 1. \tag{4.9}$$

The implicit function theorem may be used to solve simultaneously $F = 0$ and $G = 0$ for $\bar{p}$ and $\bar{x}$ as functions of $\delta$ where the denominator in (4.10) below is nonzero. In addition, the derivative of $\bar{p}$ with respect to $\delta$ is given by the following quotient:

$$\frac{\mathrm{d}\bar{p}}{\mathrm{d}\delta} = -\det\begin{bmatrix} \frac{\partial F}{\partial \delta} & \frac{\partial F}{\partial x} \\ \frac{\partial G}{\partial \delta} & \frac{\partial G}{\partial x} \end{bmatrix} \Big/ \det\begin{bmatrix} \frac{\partial F}{\partial p} & \frac{\partial F}{\partial x} \\ \frac{\partial G}{\partial p} & \frac{\partial G}{\partial x} \end{bmatrix}. \tag{4.10}$$

From (4.2) and (4.9), the numerator of (4.10) is computed to be

$$\begin{aligned} &\bar{p}(1-\bar{p})[f_{AA}(\bar{x}) - f_{aa}(\bar{x})]\left\{2\bar{p}\frac{\partial f_A}{\partial x} - \frac{\partial f}{\partial x} + (2q-1)h'(\bar{x})\right\} \\ &\quad = \bar{p}(1-\bar{p})[f_{AA}(\bar{x}) - f_{aa}(\bar{x})]\left\{\bar{p}^2 f'_{AA}(\bar{x}) - (1-\bar{p})^2 f'_{aa}(\bar{x}) + (2q-1)h'(\bar{x})\right\}. \end{aligned} \tag{4.11}$$

In terms of allele fitnesses, the denominator of (4.10) may be written as

$$\left[f_A + \bar{p}\frac{\partial f_A}{\partial p} - 1\right]\left[\frac{\partial f}{\partial x} + h'(\bar{x})\right] - \left[\bar{p}\frac{\partial f_A}{\partial x} + qh'(\bar{x})\right]\frac{\partial f}{\partial p}, \tag{4.12}$$

where

$$\begin{aligned} \frac{\partial f_A}{\partial p} &= (1-\delta)[f_{AA}(\bar{x}) - f_{aa}(\bar{x})], \\ \frac{\partial f_A}{\partial x} &= f'_{aa}(\bar{x}) + (\bar{p} + \delta - \delta\bar{p})[f'_{AA}(\bar{x}) - f'_{aa}(\bar{x})], \\ \frac{\partial f}{\partial p} &= 2(\bar{p} + \delta - 2\delta\bar{p})[f_{AA}(\bar{x}) - f_{aa}(\bar{x})] \text{ and} \\ \frac{\partial f}{\partial x} &= f'_{aa}(\bar{x}) + \bar{p}(\bar{p} + 2\delta - 2\delta\bar{p})[f'_{AA}(\bar{x}) - f'_{aa}(\bar{x})]. \end{aligned} \tag{4.13}$$

Using (4.13) to write (4.12) in terms of genotype fitnesses gives the cumbersome expression

$$\begin{aligned} &\delta\bar{p}^2[f_{AA}(\bar{x}) - f_{aa}(\bar{x})][f'_{AA}(\bar{x}) - f'_{aa}(\bar{x})] + \{2\bar{p}(1-\bar{p})(1-2\delta) + \delta\}[f_{AA}(\bar{x}) \\ &\quad - f_{aa}(\bar{x})]f'_{aa}(\bar{x}) + \{\bar{p}(\bar{p} - 2\delta\bar{p} + 2\delta)[f'_{AA}(\bar{x}) - f'_{aa}(\bar{x})] + f'_{aa}(\bar{x}) + h'(\bar{x})\}[f_{aa}(\bar{x}) \\ &\quad - 1] + h'(\bar{x})\{\bar{p} + (1-2q)(\bar{p} + \delta - 2\delta\bar{p})\}[f_{AA}(\bar{x}) - f_{aa}(\bar{x})]. \end{aligned} \tag{4.14}$$

Formula (4.14) is too complicated for controlling its sign by varying parameters, but the numerator of (4.10) given by (4.11) is much simpler. In fact, if the *AA* genotype is superior in fitness, i.e. $f_{AA}(x) > f_{aa}(x)$ for all $x > 0$, and if (A1) is assumed then the sign of (4.11) may be determined by whether $q > 0.5$ or $q < 0.5$. The following example illustrates both possibilities.

*Example* 4.1. Take $y = 1$ and genotype fitnesses

$$f_{AA}(x) = \mathrm{e}^{1-x} \text{ and } f_{aa}(x) = \mathrm{e}^{1-3x}. \tag{4.15}$$

Clearly $f_{AA}(x) > f_{aa}(x)$ for all $x > 0$. For $q = 0.9$, the first and third terms within the curly brackets in (4.11) are negative and the middle term is positive for any $\bar{x}$. When $\delta = 0$, then $E = (\bar{p}, \bar{x}) \approx (0.94, 1.74)$. We compute (4.11) to be approximately $-0.0128$ and (4.14) to be approximately 0.6234. Hence, $\mathrm{d}\bar{p}/\mathrm{d}\delta < 0$, so $\bar{p}$ decreases as $\delta$ increases. In fact, numerical simulations indicate that $\bar{p}$ decreases monotonically from 0.942 to 0.906 as $\delta$ increases from 0 to 1 (Figure 2(a)). Thus, as $\delta$ increases, the $p$-coordinate of the equilibrium moves toward $p = 0.5$, where both alleles are equally frequent and hence, where the population exhibits maximal allelic diversity. The following explanation may be a biological justification for this behaviour. When $\delta = 0$, the heterozygote fitness agrees with the homozygote *aa* which is inferior to the *AA* fitness. With a high immigrating frequency $q$ for *A* allele, the *AA* genotype dominates and equilibrium is attained at a high *A* frequency. However, as $\delta$ increases, then the heterozygote fitness improves which leads to an increase in *Aa* genotype numbers and an increase in the frequency of *a*, i.e. a decrease in *p*. In this case, an increase in the selective advantage of the heterozygote results in greater allelic diversity.

On the other hand, if $q = 0.2$ and $\delta = 0$, then $E \approx (0.239, 1.15)$. The last term within the curly brackets in (4.11) is positive and this changes the signs of (4.11) and of $\mathrm{d}\bar{p}/\mathrm{d}\delta$ to positive. Numerical simulations indicate that $\bar{p}$ increases monotonically from 0.239 to 0.328 as $\delta$ increases from 0 to 1 (Figure 2(b)). In this case, allelic diversity also increases

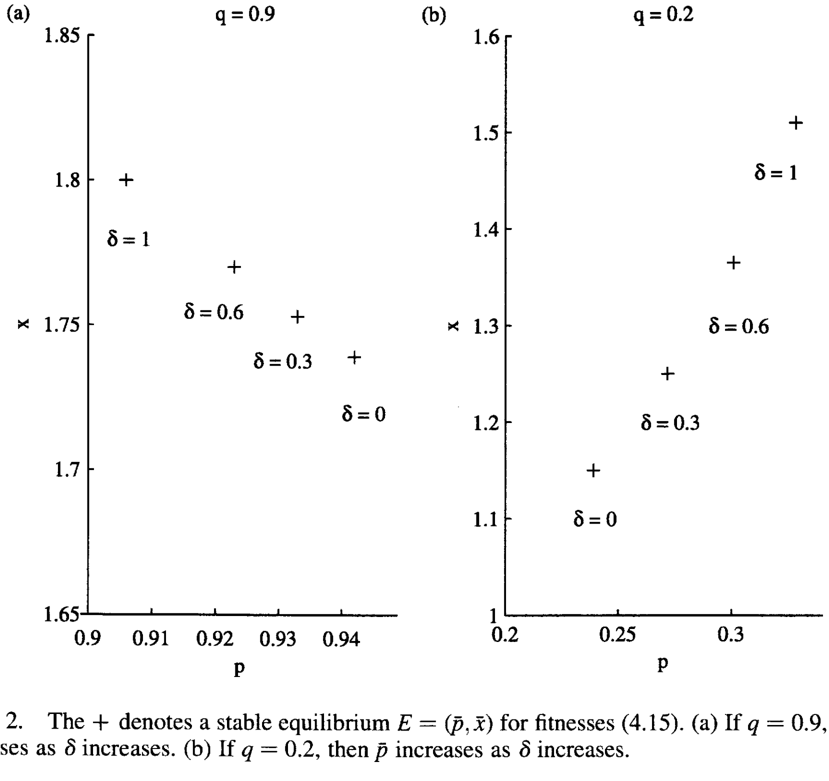

Figure 2. The + denotes a stable equilibrium $E = (\bar{p}, \bar{x})$ for fitnesses (4.15). (a) If $q = 0.9$, then $\bar{p}$ decreases as $\delta$ increases. (b) If $q = 0.2$, then $\bar{p}$ increases as $\delta$ increases.

as $\delta$ increases because the augmented *Aa* genotype fitness causes the frequency of the *A* allele to increase from its low levels.

## 5. Attracting regions

A globally attracting polymorphism is the simplest behaviour for the selection-migration model (2.2). Much more complicated dynamical behaviour has been observed for (2.2), including bistability and strange attractors [6] and in the case where $\delta = 0.5$, an apparent chaotic attractor [11]. Here, we study the existence and location of attractors for general dominance, i.e (4.1) holds, and we obtain results similar to those found by Selgrade and Roberds [11] for complete dominance ($\delta = 1$) and no dominance ($\delta = 0.5$). Also, we estimate the rate at which solution orbits approach a global attractor. We assume that $f_{AA} \geq f_{aa}$ or $f_{aa} \geq f_{AA}$, since such homozygote fitness relationships are expected when comparing a transgenic population with a natural population. For the results in this section, (A1) need not be assumed.

Note that the line $\{p = q\}$ divides the phase space $\mathcal{S}$ into two subregions:

$$\mathcal{S}^+ \equiv \{(p,x) : q \leq p \leq 1, 0 < x\} \text{ and } \mathcal{S}^- \equiv \{(p,x) : 0 \leq p \leq q, 0 < x\}.$$

For any set $\Lambda$, the topological *interior* and *closure* of $\Lambda$ are denoted Int$\Lambda$ and Cl$\Lambda$, respectively. A set $\Lambda$ is *invariant*, if for each $(p_0,x_0) \in \Lambda$, then the solution orbit $(p_n, x_n) \in \Lambda$, for all $n \geq 0$. Equilibria and periodic solutions are examples of invariant sets. First, we illustrate certain invariant subsets of $\mathcal{S}$.

LEMMA 5.1. *Assume* $0 < q < 1$ *and* (4.1) *where* $0 \leq \delta \leq 1$. *If* $f_{AA}(x) \geq f_{aa}(x)$ *for all* $x > 0$, *then* Int$\mathcal{S}^+$ *and* $\mathcal{S}^+$ *are invariant regions. If* $f_{aa}(x) \geq f_{AA}(x)$ *for all* $x > 0$, *then* Int$\mathcal{S}^-$ *and* $\mathcal{S}^-$ *are invariant sets.*

*Proof.* For the assertion about the Int$\mathcal{S}^+$, we show that if $p_n > q$, then $p_{n+1} > q$ for $n \geq 0$. Since

$$p_{n+1} = \frac{p_n f_A + qh}{f + h} = q\left[\frac{(p_n/q)f_A + h}{f + h}\right], \tag{5.1}$$

we need the bracketed term in (5.1) to be greater than 1. Hence, we should show $(p_n/q)f_A > f$, which is equivalent to showing $(p_n/q)f_A - f > 0$. Since $p_n > q$, then $(p_n/q)f_A - f > f_A - f$. From (4.2), it follows that

$$f_A(p_n,x_n) - f(p_n,x_n) = [f_{AA}(x_n) - f_{aa}(x_n)](1 - p_n)\{p_n(1 - \delta) + \delta(1 - p_n)\}. \tag{5.2}$$

Hence, $f_A - f \geq 0$ because the terms on the right in (5.2) are nonnegative for $0 \leq \delta \leq 1$. Also, if $p_n < 1$, then clearly $p_{n+1} < 1$. Thus, Int$\mathcal{S}^+$ is invariant. If we assume that $p_n \geq q$, then the preceding strict inequalities are weakened, which gives that $p_{n+1} \geq q$, i.e. that $\mathcal{S}^+$ is invariant.

To prove the assertions about $\mathcal{S}^-$, we reverse the preceding inequalities using the assumptions that $p_n < q$ and $f_{aa}(x) \geq f_{AA}(x)$. □

Notice that $x_n \geq y$ for $n \geq 1$ because of the immigration. Hence, after the first iterate, the inequalities on fitnesses in Lemma 5.1 need hold only for $x \geq y$ to obtain inequalities in the following corollary.

COROLLARY 5.2. *Assume* $0 < q < 1$ *and* (4.1) *where* $0 \le \delta \le 1$. *If* $f_{AA}(x) \ge f_{aa}(x)$ *for all* $x \ge y$ *and* $q < p_n < 1$, *then* $q < p_{n+1}$ *for all* $n \ge 1$. *If* $f_{aa}(x) \ge f_{AA}(x)$ *for all* $x \ge y$ *and* $0 < p_n < q$, *then* $p_{n+1} < q$ *for all* $n \ge 1$.

Monotone behaviour may be established for $p_n$ in the complement of $\mathcal{S}^+$ or $\mathcal{S}^-$, which shows that solution orbits are iterating toward $\mathcal{S}^+$ or $\mathcal{S}^-$.

LEMMA 5.3. *Assume* $0 < q < 1$ *and* (4.1) *where* $0 \le \delta \le 1$. *If* $f_{AA}(x) \ge f_{aa}(x)$ *for all* $x \ge y$ *and* $p_n < q$ *for* $n \ge 1$, *then* $p_n < p_{n+1}$. *If* $f_{aa}(x) \ge f_{AA}(x)$ *for all* $x \ge y$ *and* $p_n > q$ *for* $n \ge 1$, *then* $p_n > p_{n+1}$.

*Proof.* For $f_{AA}(x) \ge f_{aa}(x)$, to show

$$p_{n+1} = p_n \left[ \frac{f_A + (q/p_n)h}{f + h} \right] > p_n \tag{5.3}$$

we need to show the bracketed term in equation (5.3) is greater than 1 or equivalently that:

$$f_A(p_n, x_n) - f(p_n, x_n) + \left( \frac{q}{p_n} - 1 \right) h(x_n) > 0. \tag{5.4}$$

From (5.2), $f_A(p_n, x_n) - f(p_n, x_n) \ge 0$, when $f_{AA}(x) \ge f_{aa}(x)$. Also, since $q > p_n$, then $(q/p_n) - 1 > 0$. Thus, (5.4) holds and $p_{n+1} > p_n$.

Reversing the inequalities in (5.3) and (5.4) proves the assertion for $f_{aa}(x) \ge f_{AA}(x)$ and $p_n > q$. ☐

From a biological viewpoint, it is reasonable to assume that the population size $x$ is bounded for all generations. To guarantee this we make the following assumption on genotype fitnesses.

$$\text{There exists } B > 0 \text{ so that } x f_{AA}(x) < B \text{ and } x f_{aa}(x) < B \text{ for all } x > 0. \tag{A2}$$

From (A2), it follows that the genotype fitnesses are bounded and approach zero as $x \to \infty$. In addition, (A2) implies that $x_n f(p_n, x_n) + y \le B + y$ for all $n$, so all solutions to (2.2) are bounded. If the closed interval $\mathcal{J} \equiv [y, y + B]$, then it easily follows that:

LEMMA 5.4. *Assume* (A2), $0 < q < 1$ *and* (4.1) *where* $0 \le \delta \le 1$. *Then, each solution* $(p_n, x_n)$ *to* (2.2) *is contained in the rectangle* $\mathcal{R} \equiv [0,1] \times \mathcal{J}$ *for all* $n \ge 1$.

Since solution orbits are bounded, we now are able to show that the monotone behaviour of $p_n$ in the complement of Int$\mathcal{S}^+$ or Int$\mathcal{S}^-$ is exponential. This is the principal lemma in this section. It shows that if $f_{AA} > f_{aa}$, then $p_n$ grows exponentially as long as $p_n \le q$ and if $f_{aa} > f_{AA}$, then $p_n$ decays exponentially as long as $p_n \ge q$. Since $0 \le p_n \le 1$, neither may occur for all $n \ge 1$, so solution orbits must enter Int$\mathcal{S}^+$ or Int$\mathcal{S}^-$.

LEMMA 5.5. *Assume* (A2), $0 < q < 1$ *and* (4.1) *where* $0 \le \delta \le 1$. *If* $f_{AA}(x) > f_{aa}(x)$ *for all* $x \in \mathcal{J}$, *then there is* $r > 0$, *so that* $p_n \le q$ *implies that* $p_{n+1} \ge p_n(1 + r)$ *for any* $n \ge 1$. *If* $f_{aa}(x) > f_{AA}(x)$ *for all* $x \in \mathcal{J}$, *then there is* $s > 0$, *so that* $p_n \ge q$ *implies that* $1 - p_{n+1} \ge (1 - p_n)(1 + s)$ *for any* $n \ge 1$.

*Proof.* Assume that $f_{AA}(x) > f_{aa}(x)$ for all $x \in \mathcal{J}$. From (2.2), we have

$$p_{n+1} = p_n\left[1 + \frac{f_A - f + ((q/p_n) - 1)h}{f + h}\right]. \tag{5.5}$$

We want to minimize the fraction within the brackets in (5.5), so we minimize the numerator and maximize the denominator.

For the numerator let $\min_1$ denote $\min_{x\in\mathcal{J}}[f_{AA}(x) - f_{aa}(x)]$. Since $(p_n, x_n) \in \mathcal{R}$ for $n \geq 1$ and $p_n \leq q$, for $0 < \delta \leq 1$, we ignore the nonnegative $h$ term and use (5.2) to obtain

$$\begin{aligned} f_A - f + \left(\frac{q}{p_n} - 1\right)h &\geq (1 - p_n)\{p_n(1 - \delta) + \delta(1 - p_n)\}[f_{AA}(x_n) - f_{aa}(x_n)] \\ &\geq (1 - p_n)\{p_n(1 - \delta) + \delta(1 - p_n)\}\min_1 \\ &\geq \delta(1 - p_n)^2\min_1 \\ &\geq \delta(1 - q)^2\min_1. \end{aligned} \tag{5.6}$$

When $\delta = 0$, since $p_n \geq qy/(B + y)$ for $n \geq 1$, we have

$$\begin{aligned} f_A - f + \left(\frac{q}{p_n} - 1\right)h &\geq (1 - p_n)p_n[f_{AA}(x_n) - f_{aa}(x_n)] \\ &\geq p_n(1 - q)\min_1 \\ &\geq \frac{yq(1 - q)}{B + y}\min_1. \end{aligned} \tag{5.7}$$

For the denominator note that $x_n \geq y$ for $n \geq 1$, so $h(x_n) \leq 1$. Also from (4.1), we have $f_{Aa}(x_n) \leq f_{AA}(x_n)$, since $f_{AA}(x_n) > f_{aa}(x_n)$ for $n \geq 1$. Thus, for $(p_n, x_n) \in \mathcal{R}$, it follows that

$$f(p_n, x_n) + h(x_n) \leq f_{AA}(x_n) + h(x_n) \leq 1 + \max_{x\in\mathcal{J}} f_{AA}(x). \tag{5.8}$$

Thus, for $0 < \delta \leq 1$, define $r > 0$ by

$$r = \frac{\delta(1 - q)^2\min_1}{1 + \max_{x\in\mathcal{J}} f_{AA}(x)}. \tag{5.9}$$

For $\delta = 0$, define $r > 0$ by

$$r = \frac{yq(1 - q)\min_1}{(B + y)(1 + \max_{x\in\mathcal{J}} f_{AA}(x))}. \tag{5.10}$$

Hence, we have

$$p_{n+1} \geq p_n(1 + r). \tag{5.11}$$

For the second assertion of the lemma, assume that $f_{aa}(x) > f_{AA}(x)$ for all $x \in \mathcal{J}$. Define $v_n = 1 - p_n$, which is the frequency of $a$. The difference equation for $v_n$ is

$$v_{n+1} = \frac{v_n f_a + (1 - q)h}{f + h}. \tag{5.12}$$

Clearly, (5.12) is analogous to the difference equation for $p_n$ in (2.2), where $f_a$ replaces $f_A$

dominance, Adv. Stud. Pure Math. (AMS) (2007), to appear.

[13] C.G. Williams, *Framing the issues on transgenic forests*, Nat. Biotechnol. 23 (2005), pp. 530–532.

[14] S. Wright, *Evolution and the Genetics of Populations, II: The Theory of Gene Frequencies*, University of Chicago Press, Chicago, 1969.


(a)
δ = 0
16
(b)
δ = 1
16