BAKER'S LAW REVISITED: REPRODUCTIVE ASSURANCE IN A METAPOPULATION

JOHN R. PANNELL \textsuperscript{1} AND SPENCER C. H. BARRETT

Department of Botany, University of Toronto, 25 Willcocks Street, Toronto, Ontario, M5S 3B2, Canada

Abstract.—Baker's Law states that it is more likely for self-compatible than for self-incompatible individuals to establish sexually reproducing colonies after long-distance dispersal, because only the former can do so with a single individual. This hypothesis, proposed by H. O. Baker 40 years ago is based largely on the observation that self-compatibility is particularly frequent among colonists of oceanic islands. Here we argue that the principle of Baker's Law applies equally in the context of a metapopulation in which frequent local extinction is balanced by recolonization of sites by seed dispersal: metapopulation dynamics will select for an ability to self-fertilize. We review several studies that support this hypothesis and present a metapopulation model in which the seed productivity required by obligate outcrossers for their maintenance in a metapopulation is compared with that of selfers. Our model also estimates the reduction in the advantage of reproductive assurance to selfers as a result of perenniality and seed dormancy. In general, selection for reproductive assurance is greatest when the colony occupancy rate, $p$, is low and is much reduced when $p$ approaches its maximum. This provides an explanation for the observation that many highly successful colonizers, in which $p$ is often high, are self-incompatible. The basic model we present also lends itself to comparisons of metapopulation effects between unisexuality and coexistence and between different modes of self-incompatibility.

Key words.—Colonization, extinction, life history, outcrossing, perenniality, seed bank, self-fertilization.

Received June 16, 1997. Accepted February 24, 1998.

It has been 30 years since Baker (1967) reemphasized in these pages the significance of what Stebbins (1957) had termed "Baker's Law," namely, that it is more likely for self-compatible than for self-incompatible individuals to establish a sexually reproducing colony after "long-distance" dispersal, because only the former can do so with a single individual (Baker 1955, 1967). Although writing as a botanist, Baker's attention was drawn to the principle that now bears his name after noting similarities in colonization strategy between plants and the Notostracan invertebrates described by Longhurst (1955). The Notostraca are freshwater shrimps that inhabit ephemeral pools and whose dormant eggs are dispersed by wind or animals in much the same way as seeds. They are ancestrally dioecious, with derived hermaphroditism hypothesized to have been selected through the winnowing effects of long-distance dispersal. Baker (1967) noted that in animals "autogamous hermaphroditism provides the mechanism for ready establishment," whereas in plants "apomixis and purely vegetative reproduction are also available to replace self-incompatible hermaphroditism (and monoeism) or dioecism."

Evidence for Baker's Law has been drawn largely from comparative studies documenting a higher frequency of self-compatible relative to self-incompatible plant species on oceanic islands (Baker 1955, 1967; McMullen 1987; Webb and Kelly 1993; but see Carr et al. 1986), and the term has perhaps been most widely used in this context. To illustrate the process envisaged by underlying this floristic trend, Baker (1967) cited intraspecific variation in Armeria maritima, a self-incompatible plant in Europe that presumably lost its heteromorphic self-incompatibility system during migration to the New World. Other instances of this process include the evolution of selfing via homostyle formation in island populations of heterostylous Turnera ulmifolia and Eichhornia pacifica (Barrett and Shore 1987; Barrett et al. 1989) and the geographical distribution of heterostylous and homostylous forms in Amsinckia spectabilis (Ganders 1975). Common to these and other similar examples (for references, see Barrett 1996) has been the hypothesized selection of selfers as a means to assure reproductive success in the absence of mating partners (i.e., a corollary of the Allee effect; Allee et al. 1949). Darwin (1876) proposed that such reproductive assurance was the main selective force for the evolution of selfing in plants.

Functionally, both unisexual and self-incompatible colonists are similarly constrained in their inability to found colonies. In addition to the contrast between unisexuality and hermaphroditism in the Notostraca that Baker (1955) cited in support of his hypothesis, he also noted several cases from the plant kingdom (and see Cox 1989). For example, populations of Fragaria chiloensis are dioecious along the Pacific Coast of North and South America but are hermaphroditic in Hawaii (Baker 1967). A similar shift from gynodioecy to coexistence accompanied long-distance colonization of Tahiti from New Zealand in Fuchsia cyrtandraeoides (Sytsma and Smith 1991). The ruderal Mercurialis annua is another dioecious species in which self-compatible bisexuality (monoeism) has evolved in regions where selection for reproductive assurance is highest (Pannell 1997a). Significantly, in each of these cases coexistence seems derived from sexual dimorphism, which is the reverse evolutionary pathway to what is commonly assumed or modeled for angiosperms (Charlesworth and Charlesworth 1978; Bawa 1980; Bull and Charnov 1985; but see Lloyd 1972).

Although Baker emphasized long-distance dispersal to oceanic islands, the distance over which the dispersal event occurs is essentially unimportant. Indeed, Baker's Law more generally invokes the importance of reproductive assurance in selecting for self-fertilization in colonizing plants and animals. Solitary selfers will always be more successful than obligate outcrossers in founding new colonies, and thus we should generally expect selfing to be selected in species that

\textsuperscript{1} Present address: Institute of Cell, Animal, and Population Biology, Division of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, United Kingdom; E-mail: john.pannell@ed.ac.uk.

© 1998 The Society for the Study of Evolution. All rights reserved.
frequently colonize new sites (Lloyd 1980). Despite widespread reference to the significance of reproductive assurance (e.g., Lloyd 1979; Schoen et al. 1996), there has been little attempt to formalize the concept or quantify the selective advantage invoked by Baker’s Law. This lack of attention contrasts with other well-established predictions regarding mating-system evolution, such as the 3/2 transmission advantage enjoyed by selfers (Fisher 1941), the 1/2 threshold in the value of inbreeding depression below which self-fertilization should be selected (Lloyd 1979), or the relations derived for the effects of pollen discounting in partially selfing plants (Holsinger et al. 1984).

Baker (1955) stated his hypothesis verbally, suggesting that successful colonization of new sites by selfers would be “much more likely” than that by outcrossers. His subsequent reemphasis of the significance of his hypothesis a decade later was largely in response to Carlquist’s (1966) assertion that the disadvantage suffered by unisexual or self-incompatible colonists was “probably more imaginary than real.” How significant is Baker’s Law? Here, we quantify the advantage of reproductive assurance implied by Baker’s Law by asking how fertile an obligate outcrosser must be relative to a selfer to compensate for its inability to found new colonies by single seeds. Specifically, we compare the number of seeds that outcrossers must produce and disperse relative to selfers for their maintenance in a metapopulation at the same site-occupancy levels.

There has been much recent interest in the ecological and genetic consequences of frequent colony extinction and recolonization (e.g., McCauley 1993; Hastings and Harrison 1994; Husband and Barrett 1996; Hanski and Gilpin 1997; Ingvarsson 1997; Ronde and Olivieri 1997; Whitlock and Barton 1997), but surprisingly little attention has been focused on the effects of metapopulation dynamics on mating-system evolution (but see Barrett and Pannell 1998). The few studies that have addressed this issue include a metapopulation model developed to explain variation in female frequencies in gynodioecious populations of *Thymus vulgaris* in terms of recurrent extinction and recolonization, focusing on the mismatching of cytoplasmic male sterility factors with nuclear fertility restorers (Couv et al. 1986; Gouyon and Couvet 1987; Belhassen et al. 1989). In two different models, Holsinger (1986) and Ronfort and Couvet (1995) studied the maintenance of intermediate outcrossing rates in structured populations with seed dispersal among subpopulations, and J. R. Pannell and S. C. H. Barrett (unpubl. data) have considered the maintenance of a selfing-outcrossing polymorphism under the influence of drift and metapopulation dynamics. However, none of these studies has dealt with the effects of reproductive assurance.

Pannell (1997b) demonstrated the effects of local extinction and recolonization on the maintenance of cosexuality versus unisexual in a model that explicitly recognized the advantage enjoyed by hermaphrodites over unisexuals in colonizing unoccupied sites. While he drew attention to the selective advantage of self-compatible relative to self-incompatible cosexuals in a metapopulation, his models confined themselves to comparisons between unisexuals and these two reproductive modes (Pannell 1997b). Here, following Baker’s emphasis, we focus specifically on the difference in reproductive assurance between selfing and outcrossing cosexuals. Our approach is to consider the effect of recurrent colony extinction and recolonization on the maintenance of outcrossers in a metapopulation and to compare it with that on a metapopulation of selfers. The model thus differs fundamentally from those cited above in that we are not concerned with the maintenance of selfers and outcrossers together in a metapopulation, nor do we seek conditions for the invasion of one mating strategy into a metapopulation comprised of the other. Such considerations would require complicating assumptions regarding the inheritance of the mating system and rates of gene transfer between outcrossers and selfers and are beyond the scope of the present model; they have been dealt with, in part, in a related model elsewhere (J. R. Pannell and S. C. H. Barrett, unpubl. data; and see Discussion).

We first present a basic model of a metapopulation comprising annual plants without a seed bank. We then extend this model to address the effects of a perennial life history and seed dormancy on the advantages of reproductive assurance possessed by selfers relative to outcrossers. Perrenniality is likely to be important in the colonization of new sites, because colonists that fail to reproduce during the first year through lack of suitable mates may mate and reproduce in subsequent years with the accumulation of further immigrants. Such a process has been implicated in the colonization of islands by long-distance dispersal of outcrossing perennial hermaphrodites (Baker and Cox 1984; Böhle et al. 1996). The accumulation of dormant individuals in seed banks, which tend to be particularly common in plants inhabiting ephemeral environments (Cohen 1966, 1967), ensure the survival of individuals during years unsuitable for colony growth. Both perenniaility and seed dormancy thus reduce the extinction rate of colonies across the metapopulation, as well as increase the proportion of occupied sites (Kalisz and McPeek 1993). As our models show, levels of site occupancy, colony extinction rates, and colonization rates act together in determining the selective regime at the metapopulation level and are jointly important in influencing the relative advantage of reproductive assurance.

**The Model**

Consider a metapopulation consisting of an infinite array of occupied and vacant sites, and assume that its dynamics occur as follows. (1) Seeds immigrate into individual sites across the metapopulation. We assume an “island model” of migration, such that immigrant seeds to each site are drawn randomly from the metapopulation as a whole; that is, we adopt a “migrant pool” model of immigration (Slatkin 1977). Let $B$ be the rate of colonization occurring as a result of immigration into unoccupied sites, and let us assume that once a site has been colonized, further immigrants into it will have negligible effect on the colony’s size. (2) Following seed dispersal across the metapopulation, a proportion $E$ of the sites experiences conditions unsuitable for the maintenance of a colony. If we first assume that seeds are unable to persist in a seed bank during such years, then $E$ can be regarded as the colony extinction rate. (3) Finally, mature
plants in extant colonies mate and produce seeds. A fixed proportion of these seeds contributes to the migrant pool. Let \( p \) be the proportion of occupied sites at this stage in the metapopulation cycle, that is, after extinction events have occurred and during the colony growth, reproduction, and dispersal phases of the colony.

At metapopulation equilibrium, \( p \) will be static from one year to the next. This means that the proportion of occupied sites at the end of stage 2 will not differ from year to year. If \( p \) and \( p' \) are the proportion of occupied sites in years \( t \) and \( t + 1 \), respectively, then, for a metapopulation with no seed bank,

\[
p' = \frac{[p + (1 - p)B](1 - E)}{(1 - E)B + E}.
\]

Setting \( p = p' \) for equilibrium gives us

\[
p = \frac{(1 - E)B}{(1 - E)B + E}.
\]

(Note that Pannell [1997b] incorrectly assumed that \( E = B \) at metapopulation equilibrium; the correct expression for that model is \( E = B[1 - p]/p \).) We consider a site to be colonized when the number of founding individuals (i.e., immigrant seeds) is sufficient for reproduction to occur. This condition allows us to contrast the implications of self-fertilization versus outcrossing. In a metapopulation of selfers, a single individual is sufficient to establish a new colony, whereas at least two outcrossers are required for colonization. In general, let the minimum number of immigrants for colony establishment be \( m \); \( m \) might be greater than two if mating success is density-dependent, that is, \( m \) can be taken to quantify an Allee effect. Let \( I \) be the mean number of immigrants to each site per generation, and let their actual number in any instance follow a Poisson distribution with this mean. Then, because \( m \) or more seeds will be sufficient to establish a new colony in the metapopulation, the colonization rate in a metapopulation of plants with an annual life history and no seed bank is given as

\[
B = \sum_{n=m}^{\infty} P(I, n) = \sum_{n=m}^{\infty} \frac{I^n}{n!e^I}.
\]

As implied above, \( I \) will be proportional to the number of seeds produced per site, averaged across the metapopulation. This is simply the expected number of seeds at the end of the reproductive season in any extant site, \( \tilde{N} \), multiplied by the proportion of sites with reproductive individuals, \( p \). Thus

\[
I = Rp\tilde{N},
\]

where \( R \) is a constant of proportionality describing the fraction of seeds dispersed and the level of connectedness between sites. Let \( \tilde{N}(t, n) \) be the number of seeds produced in a colony at time \( t \) years after its establishment by \( n \) seeds, and let \( Q(n) \) be the probability that exactly \( n \) seeds first established the colony. If \( \tilde{N}_t \) is the expected number of seeds in an extant colony aged \( t \) generations, having been founded by \( n \geq m \) seeds at \( t = 0 \), then

\[
\tilde{N}_t = \sum_{n=m}^{\infty} \tilde{N}(t, n)Q(n).
\]

Summing over the distribution of colony ages, \( E(1 - E)' \), gives us \( \tilde{N} \). Thus

\[
\tilde{N} = \sum_{t=1}^{\infty} \tilde{N}_t(1 - E)' = \sum_{t=1}^{\infty} \sum_{n=m}^{\infty} \tilde{N}(t, n)Q(n)E(1 - E)'.
\]

Notice that we sum over \( t \geq 1 \) because only colonies that have survived to their first reproductive season contribute any seeds to the metapopulation.

Let us assume that \( \tilde{N}(t, n) \) is given by the population growth equation

\[
\tilde{N}(t, n) = \frac{K}{1 + Ce^{-At}},
\]

where \( K \) is the site's maximum carrying capacity, \( C \) and \( A \) are functions of \( n \) and \( r \), and \( r \) is the intrinsic rate of increase of a population after its establishment but before crowding affects its growth (i.e., when \( \tilde{N} \) is still small relative to \( K \)); \( r \) can be taken to be an index of individual fitness. The curve of \( \tilde{N}(t, n) \) is sigmoidal, with an initial value, when \( t = 0 \), of \( n \), and an approach to \( K \) with increasing \( t \). By substituting \( t = 0 \) and \( \tilde{N}(0, n) = n \) into equation (7), we obtain

\[
C = \frac{K}{n} - 1.
\]

\( A \) is found by letting \( \tilde{N}(1, n) = rn \) and rearranging terms after substitution into equation (7); thus

\[
A = -\ln\left(\frac{K - rn}{Crn}\right).
\]

Notice that we have not taken explicit account in these equations of the fraction of seeds dispersed from each colony to the migrant pool. Instead, we assume that the fraction of seeds dispersed from each colony is in excess of, and in proportion to, \( \tilde{N} \), and that this fraction is accounted for in the constant \( R \) in equation (4).

Finally, \( Q(n) \) is the truncated Poisson probability that exactly \( n \) seeds establish a site when the mean number of immigrants is \( I \) and the Poisson distribution is truncated at \( m - 1 \). \( Q(n) \) sums to unity over \( n \geq m \) and has the distribution

\[
Q(n) = \frac{P(I, n)}{B} = \frac{P(I, n)}{\sum_{j=m}^{\infty} P(I, j)}.
\]

Baker's Law contrasts the ability of selfers to found new colonies after a dispersal event, relative to that of outcrossers. In terms of our model parameters, this is analogous to the contrast between a species for which \( m = 1 \) (a selfer) and one for which \( m > 1 \) (an outcrosser). It will be intuitive that for outcrossers and selfers to be maintained in a metapopulation at equivalent site occupancy levels, \( p \) (and thus with equivalent colonization rates, \( B \)), the mean number of immigrants, \( I \), will have to be higher for outcrossers than for selfers. This is simply because in a metapopulation of outcrossers colonization will be unsuccessful whenever the num-
ber of immigrants is less than $m$, and this effect must be compensated for by increased $I$.

As equations (4) and (6) make clear, for any given $E$, $p$, and $R$ (which we will keep constant for our comparisons between selfer and outcrosser metapopulations), $I$ depends on both $N(t, n)$ and $Q(n)$. Note that because $Q(n)$ is a function of $I$ itself, equation (4) is a recursion. Note also that $N(t, n)$ is a function of $r$, the number of seeds produced by an individual before density effects are felt (eqs. 7 and 9), and so $r$ too must be found by solving a recursion numerically. We used a recursive algorithm to find values of $r$ and $I$ for a range of $E, p, R, K$, and $m$, such that equations (2), (3), and (4) were satisfied. Comparing these values between $m = 1$ and values of $m > 1$ gives us an estimate of the advantage of reproductive assurance. In particular, the ratio $r(m > 1)/r(m = 1)$ tells us how much more fecund an outcrosser must be than a selfer to compensate for the disadvantage of not being able to colonize new sites with single seeds.

The Effect of Life History

Consider now a metapopulation comprising plants that survive for more than one growing season. We might expect such perenniality to mitigate against the lack of reproductive assurance in outcrossers, because plants that fail to reproduce during the season in which they arrive at a new site might still mate with immigrants that reach the site in subsequent years. In this section, we seek to quantify the extent to which perenniality reduces the advantage of reproductive assurance in a metapopulation. Let $g$ be the number of seasons that a plant survives after it has established at a new site ($g = 1$ for the annual life history modeled above). We assume that once established at a site at the beginning of the colony’s first year (i.e., when $t = 0$), a plant has $g$ opportunities to reproduce and that the plant dies when $t = g$. As before, once mated, each plant produces $r$ viable seeds in recently established colonies. We further assume that a failure to reproduce during one season, due to an absence of mating opportunities, does not affect the value of $r$ over subsequent seasons of the plant’s life.

The colonization of an empty site requires the immigration of $n \geq m$ plants, either simultaneously when $t = 0$ or cumulatively over a maximum of $g$ years. Let $k_i$ be the number of seeds to reach a site at $t = i$ (i.e., at $i$ years prior to the ultimate establishment of a reproducively viable colony). Then the probability of colonization, $B$, is just the probability that $k_0 \geq m$ seeds reach a site at $t = 0$ (as before), plus the probability that $u = \sum_{i=1}^{g} k_i < m$ seeds immigrate during previous seasons, given that these $u$ (reproductively unsuccessful) immigrants do not die through colony extinction prior to $t = 0$ and that $k_0 > m - u$ seeds immigrate at $t = 0$. Note that we constrain $u$ to be less than $m$ because $u \geq m$ would mean that a colony had become reproducibly viable at $t < 0$, contrary to our adopted convention. Using the notation introduced in the previous section, we thus have

$$B = \sum_{k_0=m} P(I, k_0) + \sum_{k_0=m-k} P(I, k_0) \prod_{k_0=m-k} P(I, 1-E)P(I, k_0)$$

$$+ \sum_{k_0=m-k} P(I, k_2)P(I, k_1)(1-E)^2P(I, k_0)$$

$$+ \sum_{k_0=m-k} P(I, k_3)P(I, k_2)P(I, k_1) \times (1-E)^3P(I, k_0) + \ldots,$$

where the first, second, third, fourth . . . terms account for the probability that the first contributors to the colony, which becomes reproducively viable at $t = 0$, immigrated at $t = 0, -1, -2, -3, \ldots$, respectively. Because immigrants that arrive at a site before $t = 1 - g$ will have died before the colony becomes reproducibly viable, we include only the first $g$ terms in our summation.

As in the previous section, $I = RpN$ (eq. 4). We assume that colonies grow as described above in equations (7) to (9), noting that $n$ now equals $u + k_0$. In particular, we retain $r$ as the intrinsic rate of increase of a colony before density effects are felt. For $r$ greater than about two, $r$ will approximate the seed productivity of reproducing individuals and the effects of colony age structure can be reliably ignored. Retention of $r$ and the colony growth equations also allows us to estimate the effect of perenniality on reproductive assurance itself, without the confounding influences of colony growth dynamics (see Discussion).

To complete our model, we need to calculate $Q(n)$. In the previous section, $Q(n)$ was simply the truncated Poisson probability that exactly $n$ seeds reach a site simultaneously, with $n \geq m$ and the Poisson distribution truncated at $m - 1$ (equation 10). Allowing for perenniality, we must account not only for the possibility that $n \geq m$ seeds immigrate simultaneously, but also that some seeds immigrated in previous seasons and have survived to mate until the colony becomes reproducively viable. Therefore, following equation (11) and by analogy with equation (10), we see that $Q(n)$ will have the distribution

$$Q(n) = \frac{1}{B} P(I, n) + \frac{1}{B} \sum_{k_0=1}^{m-1} P(I, k_1)(1-E)P(I, n-k_1)$$

$$+ \frac{1}{B} \sum_{k_2=1}^{m-k_2} P(I, k_2)P(I, k_1)(1-E)^2$$

$$\times P(I, n-k_1-k_2) + \frac{1}{B} \sum_{k_3=1}^{m-k_3} \sum_{k_2=0}^{k_3-k_2} P(I, k_3)$$

$$\times P(I, k_2)P(I, k_1)(1-E)^3P(I, n-k_1-k_2-k_3) + \ldots;$$

(12)

$Q(n)$ sums to one over $n \geq m$, as required.

We used the $B$ and $Q(n)$, as modified above, and employed the same recursive calculations used as in the previous section to find values of $r$ and $I$ for a range of $E, p, R, K, m$, and $g$, such that equations (2), (4), and (11) were satisfied.

The Effect of a Seed Bank

We finally consider an annual life history in which seeds that do not germinate at the beginning of one growing season
may survive to germinate in later seasons. The case of a perennial life history with a seed bank is not considered explicitly, though see the Discussion section. Assume that $E$ is the probability that the site experiences a “bad” year, and that therefore the probability of a “good” year is $1 - E$. In good years, all seeds in the seed bank germinate and reach reproductive maturity, while in bad years, no seeds germinate. Seeds survive for a maximum of $d$ years in the seed bank, such that after $d$ bad years the colony goes extinct. Local colony extinction is thus the result of at least $d$ consecutive bad years, and the probability of extinction is thus $E^d$.

As before, we choose to count extant colonies during the phase of plant growth and reproduction, except that now some sites will be occupied by a seed bank. If $p$ and $p'$ are the proportion of sites occupied in years $t$ and $t + 1$, respectively, then

$$p' = p + (1 - p)B(1 - E) - pE^d,$$  \hspace{1cm} (13)

and at metapopulation equilibrium

$$p = \frac{(1 - E)B}{(1 - E)B + E^d}.$$  \hspace{1cm} (14)

Here, $p$ includes sites in which colonies exist as a dormant seed bank and that therefore do not contribute seeds to the migrant pool. Let $q$ be the proportion of sites with reproductive colonies only (i.e., colonies experiencing good years), then

$$q = \frac{(1 - E)B}{(1 - E)B + E}.$$  \hspace{1cm} (15)

$q$ will be the proportion of sites that contribute seeds to the migrant pool, so that the mean number of immigrants to each site, $I$, is now given by

$$I = Rq\tilde{N}.$$  \hspace{1cm} (16)

The accumulation of a seed bank will increase the probability of colonization, $B$, in a way similar to the effect of perenniality. In fact, equation (11) for the probability of colonization of a perennial colony can be applied for our present purposes by modifying it in two ways. First, we replace the term $(1 - E)$ for the probability of colony survival with $E$, the probability of a bad year. This is because the seed bank can only accumulate through immigration over a sequence of locally bad years; good years would cause the seeds to germinate, but with $n < m$ all plants would die without reproducing. We need also to alter the ranges over which terms are summed in the calculation of $B$. For the case of perenniality, we considered only the sum of probabilities that fewer than $m$ seeds had reached a colony prior to its establishment at $t = 0$; that is, the immigration of a total of at least $m$ seeds was required for colonization at $t = 0$, but not before. In contrast, where seeds can accumulate in a seed bank, we permit any number of seeds to reach a site prior to its year of colonization, provided that these immigrations coincide with bad years so that the colony does not yet comprise mature plants. Thus we have

$$B = \sum_{n=m}^{\infty} P(I, n) + \sum_{n=m}^{\infty} \sum_{k_1=1}^{n-1} P(I, k_1)bP(I, n - k_1)$$

$$+ \sum_{n=m}^{\infty} \sum_{k_2=1}^{n-k_1} \sum_{k_3=1}^{n-k_2} P(I, k_3)P(I, k_2)b^2P(I, n - k_1 - k_2)$$

$$+ \sum_{n=m}^{\infty} \sum_{k_2=1}^{n-k_1} \sum_{k_3=1}^{n-k_2} \sum_{k_4=1}^{n-k_3} P(I, k_4)P(I, k_3)P(I, k_2)b^3P(I, n - k_1 - k_2 - k_3) + \ldots,$$  \hspace{1cm} (17)

where the first, second, third, fourth, . . . terms are analogous to those in equation (11), and $n$ is the total number of seeds arriving at a site during a given year. Note that we include only the first $d$ terms in our summation. $Q(n)$ is modified similarly.

What is now the expected number of seeds produced by an extant colony? It will be convenient to consider colonies here not in terms of their absolute age, $t$, but in terms of the number of good years they have experienced since they were established, $v$. We therefore have

$$\bar{N}_v = \sum_{n=m}^{\infty} N(v, n)Q(n),$$  \hspace{1cm} (18)

where $N(v, n)$ has the same shape as $N(t, n)$ above. To calculate $\bar{N}$, we need $Q(n)$ as well as the probability that a site actually survives for $v$ good years and then goes extinct. The latter probability is given by

$$\left(1 - b\sum_{i=0}^{d-1} b^i\right)^v = (1 - b^d)^v b^d.$$  \hspace{1cm} (19)

This is just the age distribution of colonies, with “age” counted as the number of good years a colony experiences and during which it actually grows (a colony aged $v$ good years may be up to a maximum of $vd$ years of age in total). Summing $\bar{N}_v$ over this distribution gives us $\bar{N}$:

$$\bar{N} = \sum_{v=1}^{\infty} \bar{N}_v(1 - b^d)^vr^d.$$  \hspace{1cm} (20)

Once again, we calculated $I$ and $r$ for a range of values of $E$, $p$, $R$, $K$, $m$, and $d$ such that, this time, equations (2), (14), and (17) were satisfied.

**RESULTS**

**Annual Life History**

The maximum proportion of occupied sites in a metapopulation equals $1 - E$ when $B = 1$. Thus maximum site occupancy is independent of the mean number of immigrants per generation (Fig. 1). $I$ must increase rapidly for this maximum to be approached, though the increase is substantially more rapid for outcrossers than for selfers (cf. Figs. 1a and 1b). Increases in $E$ reduce maximum $p$ (cf. eq. 2) and increase the value that $I$ must attain for the maintenance of the metapopulation at $p$. Of particular interest is the value of $I$ as $p$ approaches zero, that is, as the metapopulation is brought closer and closer to the brink of global extinction. In a selfing metapopulation ($m = 1$; Fig. 1a) $I$ converges rapidly toward zero, whereas in an outcrossing metapopulation ($m = 2$; Fig.
Fig. 1. Model of an annual life history with no seed bank: the mean number of immigrants per site, I, required for the maintenance of a metapopulation over a range of site occupancy levels, p. Results are given for a metapopulation (a) of selfers, where a single seed is sufficient for colony establishment (m = 1) and (b) of outcrossers, where at least two seeds are required for colony establishment (m = 2). E denotes the probability of colony extinction. K = 10,000 and R = 0.01.

Fig. 2. Model of an annual life history with no seed bank: the intrinsic rate of increase of colonies before the negative effects of density are felt, r, required for the maintenance of a metapopulation with proportion p of sites occupied. (a) Metapopulation of selfers; m = 1. (b) Metapopulation of outcrossers; m = 2. E denotes the probability of colony extinction. Except for the curves indicated, K = 10,000 and R = 0.01.

1b) I remains relatively high and does not converge to zero for finite r.

This qualitative difference in the behavior of the model for selfers and outcrossers is more clearly illustrated in Figure 2. The most striking difference between selfers and outcrossers is that as p approaches zero, r converges to a finite value for selfers, whereas for outcrossers r approaches infinity. This means that whereas selfers can be maintained in a metapopulation at very low site-occupancy levels, a metapopulation of outcrossers is much more likely to go extinct if site-occupancy levels fall too low. With increasing E, r must increase for both selfers and outcrossers, especially as p approaches its maximum (Fig. 2).

Most of the curves in Figure 2 were plotted for R = 0.01 and K = 10,000. However, increasing R, the connectivity between sites, or K, the site carrying capacity, reduces r (see annotated curves in Fig. 2). This is because both R and K regulate the numbers of seeds that will be available for migration between sites; R does so directly and K does so by altering the total number of individuals in the metapopulation. r is more sensitive to changes in K for low E, because high extinction rates reduce the probability that a colony will ever reach the carrying capacity of a site (results not shown).

As noted above, the ratio r(m > 1)/r(m = 1) quantifies the advantage of reproductive assurance implied by Baker’s Law. This ratio increases with increasing E and m (Fig. 3). That is, the importance of reproductive assurance increases with increasing colony turnover and as the Allee effect becomes more marked. Significantly, both these effects are enhanced with smaller p, although with larger E the ratio always remains relatively high. As p approaches zero, this ratio approaches infinity. Thus the advantage of reproductive assurance intensifies dramatically when metapopulation site occupancy levels are low, even when the colony turnover rate (or E), is low.

Perennial Life History

Because perenniality increases the colonization rate by allowing colonists to accumulate over successive years, outcrossing annuals must produce and disperse more seeds than outcrossing perennials to persist in a metapopulation at equivalent site-occupancy levels (Fig. 4). This is particularly true when E and m are high and when p is low. The advantage of greater perenniality (i.e., increases in g) is most apparent when p is low and vanishes when site-occupancy levels approach their maximum (Fig. 4). Finally, when m = 1 (i.e., for selfers), there is no advantage at all in being perennial;
Proportion of occupied sites ($p$)

Fig. 3. Model of an annual life history with no seed bank: the seed productivity required by outcrossers for their maintenance in a metapopulation at site occupancy, $p$, relative to that of selfers; the ratio of $r$ for outcrossers ($m > 1$) to $r$ for selfers ($m = 1$). Curves for $m = 2, 3, 4,$ and $5$ are given. $E$ denotes the probability of colony extinction. $K = 10,000$ and $R = 0.01$.

The relative seed productivity is unity for all possible values of $p$, $E$, and $g$ (results not shown).

**The Effect of a Seed Bank**

The presence of a seed bank reduces the advantage of reproductive assurance in a metapopulation. This reduction is in general much greater than that brought about through perenniality (Table 1). When $p$ is close to its maximum, increases in $d$ strongly reduce the seed productivity required by outcrossers for their maintenance in the metapopulation relative to selfers (Fig. 5). This is particularly true when $m$ is high. When $p$ is low and $m$ is high, it becomes increasingly difficult for outcrossers to be maintained in the metapopulation (note that the curve for $m = 5$ crosses that for $m = 2$ in Fig. 5b). This is because with low $p$, very few sites will be occupied by reproductively mature plants capable of contributing plants to the migrant pool. As a result, the seed output of these few growing colonies must be that much higher if the metapopulation is to be maintained. The results presented in Figure 5 are for $E = 0.5$. We should expect qualitatively similar curves for lower $p$, but these values were not calculated because the presence of seed dormancy dramatically reduces the extinction rate and thus skews the colony age distribution to the right, thereby increasing computation time for the summations in equations for $B$, $Q$, and $N$.

The proportion of occupied sites in the metapopulation, $p$, generally increases when individuals are able to persist in a seed bank during bad years (i.e., when $d > 1$) particularly when $B$ and $E$ are low and when $d$ is high (Fig. 6a–c). With $d = 5$, almost all sites will be occupied in a metapopulation, even when $E$ is as high as 0.4 (or greater if colonies are recolonized frequently) (Fig. 6b). With very long-lived seed banks ($d = 50$), the metapopulation is fully occupied unless $E$ is near unity, a pattern that is relatively insensitive to the value of $B$ (Fig. 6c). The proportion of sites that are occupied only by a viable seed bank (i.e., $p - q$) tends generally to be highest for intermediate values of $E$. This peak is skewed
Table 1. Comparison of models for life history and seed bank: the relative seed productivity required for the maintenance of colonies at two levels of site occupancy, \( p = 0.40 \), and \( p = 0.01 \). Results are given for four levels of perenniaity \( (g = 2, 3, 4, \) and \( 5 \) ) and four levels of seed bank survivorship \( (d = 2, 3, 4, \) and \( 5 \) ) and are calculated relative to the seed productivity required by plants with an annual life history with no seed bank \( (^*) \), both for selfers \( (m = 1) \) and for outcrossers with different intensities of the Allee effect \( (m = 2, 3, 4, \) and \( 5 \) ). In all calculations, \( K = 10,000 \) and \( R = 0.01 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>(*)</th>
<th>( g )</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p = 0.40 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.84</td>
<td>0.83</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.76</td>
<td>0.74</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>0.69</td>
<td>0.67</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.63</td>
<td>0.60</td>
</tr>
<tr>
<td>( p = 0.01 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.50</td>
<td>0.42</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.72</td>
<td>0.67</td>
</tr>
</tbody>
</table>

toward lower \( E \) when \( B \) is low and toward higher \( E \) when \( d \) is high (Fig. 6d–f).

**DISCUSSION**

The models presented in this paper make predictions regarding the advantage of reproductive assurance enjoyed by selfers in a metapopulation under a variety of demographic and life-history conditions. In general, the strength of selection for reproductive assurance intensifies as the colony extinction rate, \( E \), increases, as the mean number of immigrants, \( I \), decreases, and as the proportion of occupied sites, \( p \), decreases. For any combination of these parameters, selection for reproductive assurance is diminished in perennial plants, which have more than one opportunity to reproduce, and especially in plants that may survive in a dormant seed bank when conditions are not suitable for the growth and maintenance of a mature colony.

Our models have highlighted the importance of the metapopulation site-occupancy rate, \( p \). In any metapopulation, the maximum proportion of habitat patches that are occupied at any point in time is determined directly by the local colony extinction rate and is approached as the colonization rate, \( B \), increases to unity. Thus, for maximum \( p \) to be approached, the mean number of seeds immigrating into each colony must increase, and in our models this occurs through enhanced seed productivity, \( r \). The increase in \( r \) required to maintain a metapopulation at high values of \( p \) applies to selfers as much as it does to outcrossers, except for high \( E \) when \( p \) is close to its maximum. By contrast, at low \( p \) the advantage of reproductive assurance increases greatly, so that outcrossers must produce and disperse increasing more seeds as \( p \) approaches zero, whereas the seed productivity of selfers converges to a constant for decreasing \( p \).

The importance of reproductive assurance when \( p \) is low is apparent for both annual and perennial life histories, as well as for plants with a seed bank, and helps to explain the otherwise puzzling observation that many successful sexual colonizing species may in fact be self-incompatible (e.g., *Centaurea solstitialis*, Sun and Ritland 1998; *Papaver rhoesas*, Campbell and Lawrence 1981; *Senecio squallidus*, Abbott and Forbes 1993; *Turnera ulmifolia*, Barrett 1978) or dioecious (e.g., *Mercurialis annua*, Pannell 1997a; and see below), in apparent contradiction to Baker’s Law. Our models indicate that the advantage of reproductive assurance will be smallest when most of the available sites in the metapopulation are occupied, even for high colony extinction rates, and greatest when a species is uncommon across the landscape, for example at the periphery of its distribution. This is in accord with empirical observations of the geographical distribution of selfing, which indicate that selfing occurs at otherwise out-

![Diagram](image-url)
crossing species commonly occur at the margins of the range (reviewed in Lloyd 1980; Barrett 1989). It also supports Baker's (1967) assertion that examples of self-incompatible colonizers do not refute his hypothesis, as claimed by Carlquist (1966).

In our model of the effects of a seed bank, we discriminate between \( q \), the proportion of sites occupied by reproductively mature individuals, and \( p \), the proportion of sites occupied either by mature plants or by plants as dormant seeds in the seed bank. Only individuals in reproductive colonies (i.e., those colonies experiencing good years) contribute seeds to the metapopulation seed pool, and hence \( q \) will directly influence the level of seed migration among sites. Unfortunately, whereas it may be relatively straightforward to assess the total number of reproductive colonies in a particular landscape, measuring \( q \) in the field is made difficult because we also need to know the number of unoccupied but available sites in the metapopulation. Measuring the value of \( p \) is even more of a challenge because dormant seeds are typically effectively invisible. Nonetheless, metapopulation studies of plant species need to account for the possibility of sites occupied by a seed bank (Husband and Barrett 1996). This is because seed dormancy not only dramatically reduces local colony extinction rates and decreases the chance of global extinction across the entire metapopulation; it also increases effective population size, thereby reducing genetic drift (Templeton and Levin 1979).

The assessment of seed bank dynamics in plant metapopulation studies is likely to be most important in the very species in which metapopulation dynamics play the greatest role. These are typically short-lived species that occupy seasonal and ephemeral habitats and in which apparent extinction rates may be high. In the aquatic annual *Eichhornia paniculata*, for example, an average of 34% of the colonies observed during one season were absent during the next, although some of these may have persisted as a dormant seed bank (B. C. Husband and S. C. H. Barrett, unpubl. data). This is in accord with the observation that species with an \( r \)-selected life-history strategy frequently possess seed dormancy (Harper 1977), particularly those that inhabit uncer-
tain environments (Cohen 1966, 1967). It is clear that un-
derstanding the persistence of many species at the metapopulation level will require a knowledge of both seed bank dynamics as well as patterns of dispersal and local extinction events (e.g., Valverde and Silvertown 1997).

Our seed bank model assumed an annual life history and was not generalized to include perenniality. In fact, the introduction of perenniality into the model treating seed dormancy would not alter our results substantially. If it were assumed that mature perennial plants died during bad years, then the reestablishment of a reproductive colony would occur during the next good year from the seed bank, while colony extinction would result after $d$ consecutive bad years, as for the annual case. If mature perennials were assumed to survive a bad year, without reproducing for example, then this would leave the extinction rate unaltered (as long as $g \leq d$) but would increase the colonization rate, $B$, for outcrossers because $m$ or more immigrants could accumulate at vacant sites over both good and bad years.

The reason why introducing perenniality would not unduly alter the results of our seed dormancy model lies in the assumptions we made in calculating $N(t, n)$, the number of plants in a colony founded by $n$ individuals $t$ years previously. We assumed that the seed productivity of perennials was constant and did not increase with age. We also assumed that the size of a colony in one year equals the number of seeds produced by plants in that colony the previous year. Both these assumptions are applicable to an annual life history but are unlikely to be accurate descriptions of the population dynamics of many perennials. In reality, many perennials increase seed production with increasing age and size. Moreover, because population size in year $t$ will include both the progeny of plants that reproduced during year $t - 1$, as well as all plants from previous generations aged less than $g$ years, we might expect perenniality to increase the colony growth rate directly. This expectation is complicated by the possibility that perenniality carries a fertility cost, with perennials producing fewer seeds per season in early years than their annual counterparts; their population growth rate may thus actually be lower than that of an annual. Indeed, it is expected on theoretical grounds (e.g., Stearns 1976; Lessells 1991) and it is well established empirically (e.g., Harper 1977) that the reproductive effort of short-lived species ($r$-strategists) is higher than that of longer-lived species. We chose to overlook these complications, as retention of $r$ and the use of the same colony growth equations for both annual and perennial life histories allowed us to estimate the effect of perenniality on reproductive assurance itself, without the confounding influences of colony growth dynamics. Nevertheless, our basic model could be modified to take these additional factors into account for cases where the population dynamics are understood and are of particular interest.

It is clear that processes occurring at the metapopulation level may affect simultaneously several different characters influencing a plant’s life history and reproductive system, and that trade-off relations and trait interactions are likely to be important. Thus, plants will be selected to allocate a certain proportion of their limited resources to dispersal, depending on local competition and the probability of colony extinction (Hamilton and May 1977; Olivieri and Gouyon 1997; Ronce and Olivieri 1997). Moreover, the allocation of resources to reproduction versus survival, which largely determines the life history of a plant, is expected to interact with patterns of allocation to male and female functions (Zhang and Wang 1994) and the mating system (Morgan and Schoen 1997), just as the mating system influences the evolutionarily stable sex allocation (Charlesworth and Charlesworth 1981).

Our models predict that selection for reproductive assurance will be greatest in annuals with no seed bank and should be diminished in perennial species and in species with a seed bank. One implication of this is that perenniality and seed dormancy may allow the maintenance of outcrossing in colonizers, despite selection for reproductive assurance, if the deleterious effects of inbreeding are high. Perenniality appears to have played this role in the maintenance of outcrossing in island species of Echium (Böhle et al. 1996), and its high incidence on islands generally (Carlquist 1974) may be due to similar causes. Nevertheless, in self-fertilizing species we might expect metapopulation dynamics (the frequent recurrence of local colony extinction, or bad years) to select also for perenniality and seed dormancy. It is thus interesting that whereas colonizing plants of disturbed and ephemeral habitats often possess a dormant seed bank, these species are often annual rather than perennial (Baker 1974). This apparent contradiction is likely due to the fact that high reproductive effort tends to be selected at the population level in disturbed ruderal situations and illustrates the principle that selection may act in different directions at the population and metapopulation levels (Olivieri and Gouyon 1997; Barrett and Pannell 1998). A comparison of ruderal and island colonists may shed further light on this notion. Because habitats available for colonization on islands are not necessarily prone to recurrent disturbance, the process of colonization tends to be uncoupled from the ruderal habit commonly found among mainland colonists. It is perhaps noteworthy in this context that an annual life history is more frequent in the colonists of the Galápagos Islands, which have a particularly high incidence of pioneer habitats, than on forested Hawaii (McMullen 1987).

It is intuitive that our model will be more robust to relaxing some of its assumptions than others. Most critically, allowing a certain amount of selfing by outcrossers (e.g., through “leaky” self-incompatibility or dioecy) would remove the advantage enjoyed by selfers in a metapopulation. It is also noteworthy that an ability of individuals in dioecious species to set some seed by selfing has been invoked to explain the successful colonization of oceanic islands by outcrossers (e.g., Baker and Cox 1984). Such leakage may have been significant in island colonization, where immigration is a very rare event, but it is possibly of less importance in the context of a metapopulation in which more frequent colonization should select for a strategy producing greater numbers of selfed seeds, especially in species where mating success is influenced by the Allee effect. Moreover, predominant outcrossing with an ability to produce a substantial number of seeds by selfing when mates are absent may in fact be a superior strategy to obligate or predominant selfing if selfed seeds suffer inbreeding depression (J. R. Pannell and S. C.
H. Barrett, unpubl. data), and we might thus expect delayed selfing to be selected under such conditions.

Our model makes no attempt to assess the evolution of mating-system traits in a metapopulation in which both selfers and outcrossers coexist. A model treating such a situation would ideally be able to identify threshold values of the rates of extinction and immigration on either side of which selfers or outcrossers would spread. However, as noted earlier, modeling the coexistence of selfers and outcrossers in a metapopulation requires assumptions regarding the genetics governing the mating system as well as the possibility of pollen transfer between mating-system morphs. These factors, which are beyond the scope of the present paper, have been considered in part elsewhere (J. R. Pannell and S. C. H. Barrett, unpubl. data).

Baker’s Law provides an explanation for the frequent capacity of colonizing species to produce viable progeny by self-fertilization. Baker (1955) was impressed by the generality of this principle, noting that conclusions drawn from studies of plants and animals were “virtually identical” in this respect. Whereas he drew a link specifically between self-compatibility and establishment after long-distance dispersal in the title of his paper, we wish to emphasize here that the principle of Baker’s Law may apply equally in the context of a metapopulation in which frequent colony extinction is balanced by the recolonization of available sites at a landscape level. Thus the propensity of oceanic island floras to contain many self-compatible species is a result of essentially the same process, albeit at a different spatial scale. The colonization of oceanic islands probably involves very low values of $I$, whereas reproductive assurance may be selected in a metapopulation with higher $I$ if $E$ is also sufficiently high.

By taking a landscape perspective, our metapopulation model provides a novel approach to understanding Baker’s Law; it allows us to distinguish quantitatively between the effects of extinction, colonization, density dependence, and seed productivity as well as the effects of life history and seed dormancy. While the model assumes simplified demographic and mating behavior, our comparison between an obligate outcrosser and a complete selfer gives an estimate of the upper bound to the advantage of reproductive assurance. The basic model we have presented here introduces a concept that might easily be modified to other circumstances. In particular, a similar approach could assess the relative likelihood of establishment of colonizing species with different types of self-incompatibility or sexual system. It would be interesting, for example, to compare the establishment likelihood of species with heteromorphic relative to various homomorphic incompatibility systems. Thus self-incompatibility may occur more commonly in a homomorphic than a heteromorphic colonizer, because the former typically have many more incompatibility alleles and would be more likely to coexist at a site with a cross-compatible mate. Similarly, metapopulation models have shown that the value of reproductive assurance may differ between sexual systems such as androecy and gynodioecy (Pannell 1997b). We suggest that contrasts in behavior of different reproductive systems and life histories in the context of a metapopulation may provide further insights into the joint evolution of traits influencing dispersal and mating.

Acknowledgments

We thank K. Holsinger, D. Charlesworth, L. Harder, C. Eckert, and two anonymous referees for comments on an earlier version of this manuscript. JRP was supported as a postdoctoral fellow by a research grant to SCBH from the Natural Sciences and Engineering Research Council of Canada.

Literature Cited


———. 1967. Optimizing reproduction in a randomly varying environment when a correlation may exist between the conditions at a time a choice has to be made and the subsequent outcome. J. Theor. Biol. 16:1–14.

Corresponding Editor: D. Piñero