The probability of fixation of a favoured allele in a subdivided population

N. H. BARTON

Division of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh EG9 3JT, U.K.

Summary

In a stably subdivided population with symmetric migration, the chance that a favoured allele will be fixed is independent of population structure. However, random extinction introduces an extra component of sampling drift, and reduces the probability of fixation. In this paper, the fixation probability is calculated using the diffusion approximation; comparison with exact solution of the discrete model shows this to be accurate. The key parameters are the rates of selection, migration and extinction, scaled relative to population size ($S = 4Ns$, $M = 4Nm$, $\Lambda = 4N\lambda$); results apply to a haploid model, or to diploids with additive selection. If new colonies derive from many demes, the fixation probability cannot be reduced by more than half. However, if colonies are initially homogeneous, fixation probability can be much reduced. In the limit of low migration and extinction rates ($M, \Lambda \ll 1$), it is $2s/\{1 + (\Lambda/MS)(1 - \exp(-S))\}$, whilst in the opposite limit ($M, \Lambda \gg 1$), it is $4sM/\{\Lambda(\Lambda + M)\}$. In the limit of weak selection ($S \ll 1$), it is $4sM/\{(\Lambda + 2)(\Lambda + M)\}$. These factors are not the same as the reduction in effective population size (N_e/N), showing that the effects of population structure on selected alleles cannot be understood from the behaviour of neutral markers.

1. Introduction

The simplest view of evolution is that it consists of the accumulation of favourable mutations. The rate of adaptation is then given by the product of the number of genes in the population, the mutation rate to alleles that increase fitness in the current environment and genetic background, and the average probability that each is fixed (P). If the fitness of individual genes follows a Poisson distribution, with the favoured allele having a mean fitness of $W = 1 + s$, then P is given by the solution to $P = 1 - \exp(-WP)$, with $P \approx 2s$ for small s (Fisher, 1930). It is usually assumed, following Maruyama (1970), that the fixation probability is independent of population subdivision, and that subdivision therefore has no influence on the accumulation of alleles that are everywhere favourable. In this note, I show that random extinction and recolonization can greatly reduce the chance that a weakly selected allele will be fixed.

2. The model

For simplicity, I will base the exact derivations on the haploid island model. However, the approximate results will also apply to the diploid case. There are infinitely many demes, each containing $2N$ genes. In each generation, every gene produces a very large number of offspring. (The fitness of the advantageous allele is $(1 + s)$, relative to the ancestral type. The results also apply with additive selection.) A fraction m of these are exchanged with a common pool of migrants (or equivalently, migration is equally likely between any pair of demes). The $2N$ genes that make up the next generation are then sampled at random (i.e. the Wright–Fisher model; Crow & Kimura, 1970). In deriving results for large N, it is convenient to define the scaled parameters, $S = 4Ns$, $M = 4Nm$, $\Lambda = 4N\lambda$, where m, λ are the migration and extinction rates.

In each generation, there is a chance λ that a deme will go extinct, and be immediately recolonized. The $2N$ colonists can either be drawn independently from the current migrant pool, or can all be derived from a single randomly chosen gene, taken after selection. These two forms of recolonization should be seen as the extremes: colonization by one or a few diploid individuals, from one or several demes, would lie between them. Colonization from many demes corresponds to Slatkin's (1977) 'migrant pool' model, where the number of haploid colonists equals $2N$. Colonization from one gene corresponds either to the

'migrant pool' or to the 'propagule pool' model with a single colonizing gene.

The assumption that the newly founded colony increases to its carrying capacity of $2N$ genes in a single generation is unrealistic. However, most of the results will be derived for the diffusion limit, when $2N \gg 1$, $m, \lambda, s \ll 1$. Then, one expects that the population will grow to its carrying capacity from one individual in $\approx \log(2N)$ generations, which is much smaller than the characteristic timescale of evolutionary change, $\approx 2N$ generations. Similarly, the delay before a colonist arrives should be $\approx 1/(Nm)$ generations, which will also be small relative to the characteristic timescale.

3. Approximate argument for small Nm, $N\lambda$

The fixation probability will be calculated numerically using both exact solutions for the discrete Wright–Fisher model, and the diffusion approximation. First, however, I will give an approximate argument that relies on separating the process of fixation into two parts: the establishment of the favourable allele in a single deme, and then spread from that one deme throughout the whole population. This kind of argument was introduced by Slatkin (1981) and Lande (1979, 1985) to follow the spread of underdominant alleles, and has recently been used by Tachida & Iizuka (1991). It will be accurate if, on average, the time taken for the favoured allele to spread through a single deme is much shorter than the time between extinctions, or between 'infections' of new demes, so
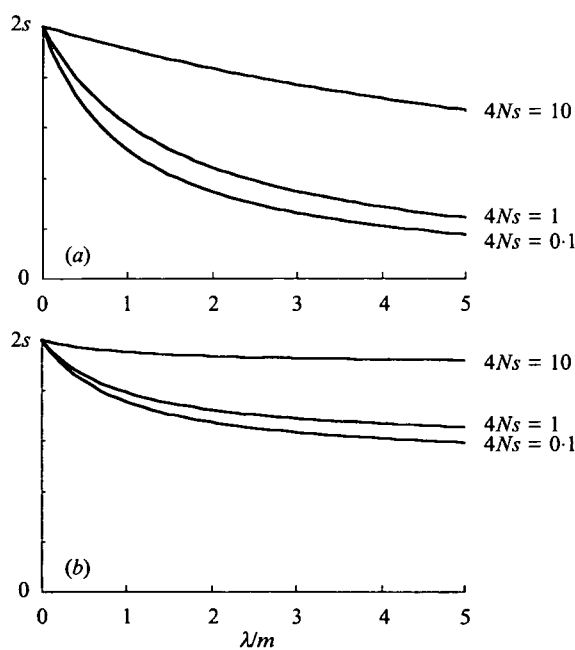

Fig. 1. The probability of fixation, as a function of the ratio between the rates of extinction and migration (λ/m), for $4Ns = S = 0.1$, 1, 10, in the limit of low migration and extinction rates $[\Lambda, M = 4N\lambda, 4Nm \ll 1$; eqn (1)]. (a) Colonists come from a single deme. (b) Colonists come from many demes.

that demes are usually fixed for one or other allele. This requires $m, \lambda \ll s$.

Suppose that the favoured allele is introduced into a single deme at frequency u. Neglecting migration and extinction, the probability that it will be fixed in that deme is

$$P_1(u) = (1 - \exp(-Su))/(1 - \exp(-S)).$$

Here, Kimura's (1962) diffusion approximation has been used; this is a good approximation for weak selection. The chance of fixation of a single allele is

$$P_1(1/2N) = (1 - \exp(-2s))/(1 - \exp(-S))$$
$$\approx 2s/(1 - \exp(-S)).$$

The chance that this one deme will successfully infect all the other demes depends on the probability per generation that a fixed deme will lose the new allele by chance (f_-), and on the expected number of other demes which it infects per generation (f_+). The chance that the wild-type allele carried by a single immigrant will displace the fitter allele is

$$2s/(\exp(S) - 1).$$

In every generation, the expected number of immigrants is $2Nm$, and so for small Nm, the probability of loss by drift is $4Nm/\exp(S) - 1)$. Allowing for the chance of extinction,

$$f_- = \{\lambda + (1 - \lambda)4Nms/(\exp(S) - 1)\}.$$

Assuming that m, s, λ are all small, this is approximately $f_- = \{\lambda + 4Nms/(\exp(S) - 1)\}$.

The number of fixed demes may increase through emigration or recolonization. A fixed deme sends out $2Nm$ emigrants on average, each of which has a chance $2s/(1 - \exp(-S))$ of successful invasion. If we assumed instead that migration occurs after selection, but before density regulation ('hard selection'), then a deme fixed for the fitter allele would on average send out $2Nm(1 + s)$ emigrants. However, this only increases the rate of infection, f_+, by a factor s, which is assumed to be negligible.

A fraction λ of demes becomes empty in every generation. If colonization is from a single deme, the expected number of colonizations by the new allele is just λ. If, on the other hand, colonists are drawn independently, the new allele will almost always be present in a single copy when it is present at all, and will then have a chance

$$P_1(1/2N) \approx 2s/(1 - \exp(-S))$$

of fixing. The expected number of successful invasions is thus $2N\lambda P_1(1/2N) = \lambda S/(1 - \exp(-S))$.

Summing, and approximating $(1 - \lambda)$ by 1, $f_+ = \{\lambda + Sm/(1 - \exp(-S))\}$ for colonization from one individual, or $f_+^* = S(\lambda + m)/(1 - \exp(-S))$ for colonization from a mixed propagule pool.

Slatkin (1981) showed that in this kind of model, where there is a probability f_- of loss, plus a Poisson distribution of increase with expectation f_+, the net

probability of fixation through the whole set of demes is $(1 - f_-/f_+)$ (for $f_-, f_+ \ll 1$). Multiplying this by the probability of initial fixation in one deme, we have

$$P(u) = \frac{(1 - \exp(-Su))}{1 + (\Lambda/MS)(1 - \exp(-S))}$$

(colonists from a single gene) (1*a*)

$$P^*(u) = \frac{(1 - \exp(-Su))}{S(M + \Lambda)}\left(\frac{\Lambda S}{(1 - \exp(-S))} + (MS - \Lambda)\right)$$

(colonists from many demes). (1*b*)

(This argument is based on the assumption that only one deme changes at a time, which is satisfied with small m, λ.) These approximations are illustrated in Fig. 1, which shows how the net probability of fixation decreases as the rate of extinctions increases. Note that only the rate of extinction *relative* to the rate of migration $(\Lambda/M = \lambda/m)$ is relevant. When extinction of demes is much more frequent than migration of individuals, and colonists come from one gene, the fixation probability declines to zero

$$(P(1/2N) \approx 2s(SM/\Lambda)/(1 - \exp(-S))).$$

When selection is strong relative to drift, the chance of fixation is proportional to the square of the selective advantage $(S = 4Ns \gg 1, P(1/2N) \approx 8Ns^2 m/\lambda)$.

In contrast, when colonists come from many demes, the fixation probability declines by a factor of not less than half $(P^*(u) \approx s$ when $\lambda \gg m, S \ll 1$, since

$$\lim_{S \to 0}(1/(1 - \exp(-S)) - (1/S)) = 1/2).$$

In general, the probability of fixation with multiple colonists is always greater than that with one colonist. This is because in the former case the advantageous allele gains an advantage during the competition between genes that occurs during the colonization process $(f_+^* > f_+)$.

4. Discrete solution

I now derive exact solutions for the discrete model. To develop the method, examine first the case with no extinction. Consider the chance Q_i that all i copies of the fitter allele in a deme are ultimately lost $(Q_i = 1 - P_i)$. After selection and immigration of wild-type alleles, the gene frequency in the gamete pool is

$$u^* = i(1 + s)(1 - m)/(2N + si(1 - m)).$$

The next generation is sampled from this pool to give j genes with binomial probability. The chance that all these j genes are ultimately lost is Q_j. The expected number of good genes that land (almost certainly, alone) in other demes is $mi(1 + s)$. (The model assumes hard selection, in that the fitter the deme, the more migrants it produces. As noted above, this makes little difference when selection is weak.) The actual number

of infected demes, k, follows a Poisson distribution, and the chance that the allele is lost from all these is Q_1^k. Multiplying these factors, and summing over the distribution of j and k gives

$$Q_1 = \left(\exp(-im(1 + s))\sum_{k=0}^{\infty}\frac{(im(1 + s))^k}{k!}Q_1^k\right)$$

$$\times\left(\sum_{j=0}^{2N}\binom{2N}{j}u^{*j}v^{*2N-j}Q_j\right)$$

$$= \exp(-im(1 + s)P_1)\left(\sum_{j=0}^{2N}\binom{2N}{j}u^{*j}v^{*2N-j}Q_j\right), \quad (2)$$

where $P_i = 1 - Q_i$.

It is tempting to suppose that the different alleles in each deme are lost independently of each other, as would be the case for a single large population. Then $Q_j = Q_1^j$, and eqn (2) could be summed to give a transcendental equation for P_1:

$$Q_i = \exp(-im(1 + s)P_1)(1 - u^*P_1)^{2N}. \quad (3)$$

However, since u^* depends on i, eqn (3) does not have the form Q_1^i, and so the argument is inconsistent: with Wright–Fisher sampling, the fitnesses of the different genes in a deme are negatively correlated, and so their chances of fixation are not independent. For the same reason, the solution to eqn (2) is not precisely independent of population structure (m), even with hard selection. Maruyama's (1970) result was derived using the diffusion approximation, and so does not hold exactly in small populations, and with strong selection. However, eqn (3) does reduce to Fisher's (1930) equation for weak selection $(s \ll 1)$, and is then independent of migration. Table 1*a* shows that the diffusion approximatioin is accurate even with selection as strong as $s = 0.2$, and populations of as few as five individuals.

Equation (2) can readily be extended to include extinction:

$$Q_i = \exp(-im(1 + s)P_1)\exp(-\lambda(1 + s)(i/2N)P_{2N})$$

$$\times\left\{(1 - \lambda)\left(\sum_{j=0}^{2N}\binom{2N}{j}u^{*j}v^{*2N-j}Q_j\right) + \lambda\right\}$$

(colonization from a single gene) (4*a*)

$$Q_i^* = \exp(-i(m + \lambda)(1 + s)P_1^*)$$

$$\times\left\{(1 - \lambda)\left(\sum_{j=0}^{2N}\binom{2N}{j}u^{*j}v^{*2N-j}Q_j^*\right) + \lambda\right\}$$

(colonists from many demes). (4*b*)

The right-hand term in each equation gives the chance that all the i alleles in a deme are ultimately lost from that deme. This is the chance that the deme immediately goes extinct (λ), plus the chance that the deme does not go extinct immediately, but that the j copies produced by binomial sampling are ultimately lost. Spread by migration gives a factor

$$\exp(-im(1 + s)P_1)$$

as before. If colonists are drawn from many demes, then the expected number of newly colonized demes which contain a single copy of the fitter allele is $\lambda i(1+s)$ (assuming that colonists are drawn from the gene pool after selection, giving the factor $(1+s)$). This leads to the factor $\exp(-i\lambda(1+s)P_1^*)$, in the same way as for spread by migration. If new colonies are derived from a single gene, then the expected number of colonies seeded by one of the i copies of the fitter allele is $E(k) = \lambda i(1+s)/2N$. The fitter allele is initially fixed in the new colonies, and so the chance that it is ultimately lost from all k colonies is Q_{2N}^k. This leads to the factor $\exp(-\lambda(1+s)(i/2N)P_{2N})$ in eqn (4a).

Equation (4) defines a set of simultaneous equations for $P_1, \dots, P_{2N}$, which can be solved numerically. In the following section, a diffusion approximation is derived, and is shown to be a good approximation to the discrete model for all but very small deme sizes.

5. The diffusion limit

We take the limit of large N, keeping the scaled parameters $S = 4Ns$, $M = 4Nm$, $\Lambda = 4N\lambda$ constant. The fixation probability P_i is then approximated by $P(u)$, where $u = i/2N$; the probability of fixation of a single gene is now written $P(1/2N)$, instead of P_1. The sum over the binomial distribution is replaced by an integral over a normal distribution, $\psi(\epsilon)$, with mean $(suv-mu)$, and variance $uv/2N$. Equation (4a), for colonization from a single gene, gives

$$1 - P(u) = \exp[-2Nmu(1+s)P(1/2N)]$$
$$\times \exp[-\lambda(1+s)uP(1)]$$
$$\times \left\{(1-(1-\lambda)\int_0^1 \psi(x-u)P(x)\,dx\right\}. \quad (5)$$

Expanding $P(x)$ in a Taylor's series around u, replacing m, s, λ by the corresponding scaled parameters, and keeping only leading terms gives:

$$0 = uv\frac{\partial^2 P}{\partial u^2} + (Suv - Mu)\frac{\partial P}{\partial u}$$
$$+ uMS\Pi(1-P) + u\Lambda\Xi(1-P) - \Lambda P, \quad (6a)$$

where $\Xi = P(1)$ for clarity. The factors of $(1+s)$, which represent the extra contribution of demes carrying the fitter allele with hard selection, make a negligible contribution. $P(1/2N)$ has been written as $2s\Pi$, leading, after scaling, to the factor $S\Pi$ in the third term. $P(0) = 0$, and in the continuous limit, $2s\Pi = P(1/2N) = (1/2N)(\partial P/\partial u)_{u=0}$. Hence,

$$(\partial P/\partial u)_{u=0} = 4ns\Pi = S\Pi.$$

Similarly, for colonization from many demes:

$$0 = uv\frac{\partial^2 P^*}{\partial u^2} + (Suv - Mu)\frac{\partial P^*}{\partial u}$$
$$+ u(M+\Lambda)S\Pi(1-P^*) - \Lambda P^*. \quad (6b)$$

The boundary conditions for eqn $(6a, b)$ are $P(0) = 0$, $(\partial P/\partial u)_{u=0} = S\Pi$. The parameter Π is then determined by the requirement that both P and $(\partial P/\partial u)$ be finite at $u = 1$. Consider first colonization from a single gene [eqn $(6a)$]. Near $u = 1$, $v \ll 1$, and so the diffusion term $uv(\partial^2 P/\partial u^2)$ becomes small. The equation therefore admits an unacceptable solution of the form $P = v^{1-M}$. For this singular component to be zero, we must have

$$-M\frac{\partial P}{\partial u} + MS\Pi(1-\Xi) - \Lambda\Xi^2 = 0,$$

where

$$\Xi = P(1), \quad S\Pi = (\partial P/\partial u)_{u=0}. \quad (7a)$$

Similarly, for colonization from many demes, we have the condition:

$$-M\frac{\partial P^*}{\partial u} + (M+\Lambda)S\Pi(1-\Xi^*) - \Lambda\Xi^* = 0,$$

where

$$\Xi^* = P^*(1), \quad S\Pi = (\partial P^*/\partial u)_{u=0}. \quad (7b)$$

Though I have not found any general analytic solution to eqn $(6a, b)$, it can be solved in the limit of rare migration and extinction. As M, Λ tend to zero, the terms in selection dominate, and the equations admit a family of solutions of the form $P(u) = \Pi(1-\exp(-Su))$. Substitution into eqn $(7a, b)$ then confirms the approximations of eqn $(1a, b)$.

An approximate solution can also be obtained in the opposite limit, where $M, \Lambda \gg S, 1$. In this case, the new allele is expected to move across many demes before selection appreciably raises its frequency. This suggests that its chance of fixation should be proportional to the number of copies ($P \approx u$), and be independent of their distribution across demes. This is the case where colonists come from one gene: substitution of this linear solution into eqn $(6a)$ confirms that this is a solution, to leading order in $1/M$. (Λ is assumed to be of the same order as M.) However, the coefficient of proportionality must be determined by considering higher-order terms. Substituting

$$P(u) = (\alpha u/M) - (\phi(u)/M^2) + O(1/M^3)$$

leads to

$$P(u) = \frac{2SMu}{\Lambda(\Lambda+M)}\left(1-\frac{Su}{\Lambda}\right) + O(1/M^3)$$

$$\text{(colonists from one gene)}. \quad (8a)$$

For colonists from many demes, the leading term is of order one, and does not have the expected simple linear form:

$$P^*(u) = (\gamma u)^{-(\Lambda/M)}e^{-\gamma u}\int_0^{\gamma u} x^{(\Lambda/M)}e^x\,dx + O(1/M)$$

$$\text{(colonists from many demes)}. \quad (8b)$$

where $\gamma = [S(\Lambda+2M)]/[2(\Lambda+M)]$.

Note that this is a function only of Λ/M and γu. The probability of fixation of a single gene, relative to that with no extinction, is

$$\frac{P(1/2N)}{2s} = \frac{2M}{\Lambda(\Lambda+M)} + O(1/M^2)$$

(colonists from one gene)   (9a)

$$\frac{P^*(1/2N)}{2s} = \frac{(\Lambda+2M)}{2(\Lambda+M)} + O(1/M)$$

(colonists from many demes).   (9b)

Note that in both cases, this is independent of the selection pressure, $S = 4Ns$. As migration and extinction become frequent ($M, \Lambda \to \infty$), the probability of fixation tends to zero with colonization from one gene [eqn (9a)], but remains of order $2s$ with colonization from many demes [eqn (9b)]. In the latter case, it tends to s as extinction becomes much more frequent than migration; this is also the minimum in the opposite limit of low migration and extinction [eqn (1a, b)].

The diffusion equation can also be solved in the limit of weak selection ($S = 4Ns \ll 1$). Taking terms of highest order in S, in the same way as above:

$$P(u) = \frac{2SMu}{(\Lambda+2)(\Lambda+M)}\left(1 - \frac{Su}{(\Lambda+2)}\right) + O(S^3)$$

(colonists from one gene)   (10a)

$$P^*(u) = \frac{S(\Lambda+2M)u}{2(\Lambda+M)}\left(1 - \frac{Su}{2}\right) + O(S^3)$$

(colonists from many demes).   (10b)

Equations (10) and (8a, b) are consistent, in that they reduce to the same form in the limit of small S and large Λ, M. The probability of fixation of a single gene, relative to that with no extinction, is:

$$\frac{P(1/2N)}{2s} = \frac{2M}{(\Lambda+2)(\Lambda+M)} + O(S)$$

(colonists from one gene)   (11a)

$$\frac{P^*(1/2N)}{2s} = \frac{(\Lambda+2M)}{2(\Lambda+M)} + O(S)$$

(colonists from many demes).   (11b)

In both cases, these equations are more accurate than eqn (9a, b) which was derived for the limit of large Λ, M. As selection becomes weaker ($S \to 0$), the probability of fixation tends to these limiting forms, which are shown by dotted lines in Fig. 2. With colonization from many demes, $P^*(1/2N)$ tends to s as extinction becomes much more frequent than migration, confirming that in this case, the fixation probability can never be reduced by more than half.

Numerical solutions to the diffusion equations were produced using the Runge–Kutta algorithm in *Mathematica* (Wolfram, 1991). For eqn (6b), where colonists
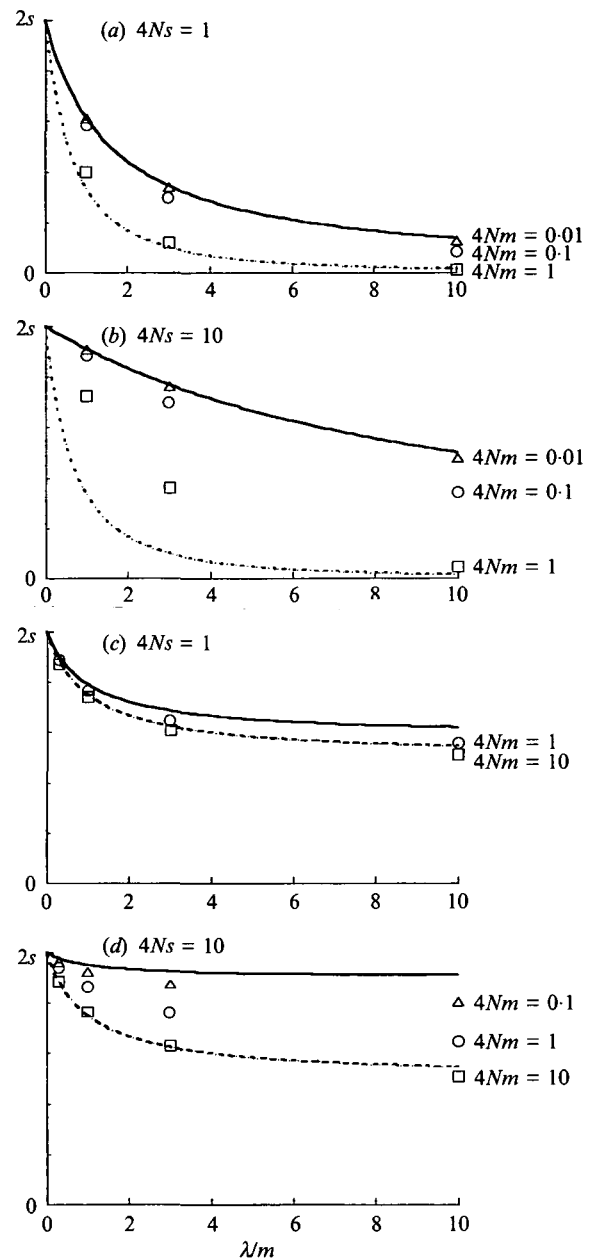

Fig. 2. For a given ratio between extinction and migration ($\lambda/m = \Lambda/M$), the fixation probability decreases as migration and extinction rates increase. The solid curves gave the limit of low migration and extinction rates [eqn (1)], as in Fig. 1. The dotted curves give the limit of weak selection [eqn (11)], which converges to the limit of high migration and extinction rates [eqn (9)] for large Λ, M. The symbols give numerical solutions to the diffusion equation [eqn (6)]. (a) Colonists from one gene; $S = 4Ns = 1$. (b) Colonists from one gene; $S = 4Ns = 10$. (c) Colonists from many demes; $S = 4Ns = 1$. (d) Colonists from many demes; $S = 4Ns = 10$.

come from many demes, one finds that value of the initial gradient ($S\Pi = (\partial P^*/\partial u)_{u=0}$) which satisfies eqn (7b), and hence gives a finite solution at $u = 1$. The solution to eqn (6a), where colonists derive from a single gene, is harder, because it depends on finding a value of $S\Pi = (\partial P^*/\partial u)_{u=0}$ which satisfies eqn (7a), and also a value of Ξ which gives $P(1) = \Xi$.

6. Results

Table 1a compares exact values for the discrete model [$s = 0.05, 0.025$; eqn (2)] with those from the diffusion approximation ($s = \epsilon$; $P(u) = (1 - \exp(-Su))$), in the

Table 1. *For each parameter combination, P(1/2N)/2s gives the probability of fixation of a single advantageous gene, relative to the expectation for weak selection in a large panmictic population (2s). P(1) gives the chance that a favoured gene that is fixed in one deme will ultimately spread through the whole population. All results are for S = 4Ns = 1. The first rows in each set give exact results for the discrete model [eqn (1)], for s = 0.05, N = 5, s = 0.025, N = 10, and (for some cases) s = 0.0125, N = 20. Rows headed by s = ϵ give results for the limit of weak selection, based on the diffusion approximation [eqn (6)]. The last row in each set, headed by 4Nm = ϵ', give results for the limit of low migration and extinction [eqn (1)]*

$4Ns$	$4Nm$	$4N\lambda$	s	$P(1/2N)/2s$	$P(1)$
		(a) No extinction			
1	1	0	0.05	0.9328	0.6242
			0.025	0.9654	0.6281
			ϵ	1.000	0.6321
1	0.1	0	0.05	0.9309	0.6233
			0.025	0.9644	0.6276
			ϵ	1	0.6321
1	ϵ'	0	ϵ	1	0.6321
		(b) Colonists from one gene			
1	1	10	0.05	0.1648	0.1456
			0.025	0.0816	0.0735
			ϵ	0.0122	0.0112
1	0.1	1	0.05	0.1433	0.1071
			0.025	0.1132	0.0833
			ϵ	0.0813	0.0588
1	ϵ'	$10\epsilon'$	ϵ	0.1366	0.0863
1	1	1	0.05	0.4720	0.3467
			0.025	0.4354	0.3167
			ϵ	0.3970	0.2863
1	0.1	0.1	0.05	0.5648	0.3837
			0.025	0.5723	0.3786
			ϵ	0.5805	0.3737
1	ϵ'	ϵ'	ϵ	0.6127	0.3870
		(c) Colonists from many demes			
1	1	10	0.05	1.0598	0.6219
			0.025	0.8005	0.4925
			0.0125	0.6803	0.4301
			ϵ	0.5673	0.3967
1	0.1	1	0.05	0.5828	0.3930
			0.025	0.5925	0.3889
			0.0125	0.5976	0.3868
			ϵ	0.6032	0.3849
1	ϵ'	$10\epsilon'$	ϵ	0.6200	0.2477
1	1	1	0.05	0.7979	0.5270
			0.025	0.7864	0.5099
			ϵ	0.7732	0.4919
1	0.1	0.1	0.05	0.7376	0.4943
			0.025	0.7620	0.4964
			ϵ	0.7881	0.4985
1	ϵ'	ϵ'	ϵ	0.7910	0.5000

absence of extinction. Agreement is close even when the population is small ($N = 5, s = 0.05$). Moreover, comparison of values for $4Nm = 1$ and $4Nm = 0.1$ shows that fixation probabilities depend only very slightly on migration rates. This is consistent with Maruyama's (1970) result in the limit of weak selection.

The last row in each part of Table 1, headed by $4Nm = \epsilon'$, gives values for the limit of rare extinction and migration [eqn (1)]. The diffusion approximation ($s = \epsilon$) does approach these values as both $4Nm$ and $4N\lambda$ become small. Figure 1 shows how the fixation probability in this limit depends on the relative rates of extinction and migration. Colonization from one gene can lead to a substantial reduction in fixation probability, provided that S is not too large (Fig. 1a). In the limit where drift dominates selection within demes ($S = 4Ns \ll 1$), the fixation probability tends to $2sM/(M + \Lambda)$ [from eqn (1a)]. On the other hand, if selection within demes is strong ($S \gg 1$), then the fixation probability tends to $2sMS/(\Lambda + MS)$. Now, the chance of fixation is proportional to s^2 when $1 \ll S \ll \Lambda/M$, so that extinction will disproportionally reduce the contribution of weakly advantageous genes. With founders from many demes, the effect of extinction is much smaller, and can never be larger than twofold ($2s > P(1/2N) > s$). This is because colonization from a mixed pool of immigrants gives an opportunity for selection which favours the fitter allele.

Figure 2 shows the effect of increasing migration and extinction rates, again plotted as a function of the ratio between them. This further reduces the fixation probability, for both models of colonization. Extinctions can now have a substantial effect even when selection within demes is strong ($S \gg 1$). The dotted curves show the analytic approximations of eqn (11a, b), for the limit of weak selection; this converges to the limit of frequent migration and extinction [eqn (9)] for large Λ, M. These curves are independent of selection, and show that with colonization from many demes, the reduction in fixation probability cannot be greater than twofold.

Figure 3 shows how the probability of fixation depends on the initial frequency of the gene within a single deme ($P(u)$). This is compared with the form expected in the limit of low migration and extinction, when the spread of the allele separates into establishment within one deme, followed by stochastic spread through the whole population: $P(u) = \Pi(1 - \exp(-Su))$ (dotted curves). Agreement is close for $4Nm = 0.1$, as expected. However, when migration is more frequent, and selection moderate ($4Ns = 1$, $4Nm = 1$), the curves approach the linear form predicted in the limit of high migration. [With colonization from many demes, $P(u)$ is not expected to be strictly linear in the limit of large M, Λ. However, for the parameters plotted in Figs 3(c, d), the function given by eqn (8b) is close to linear.]
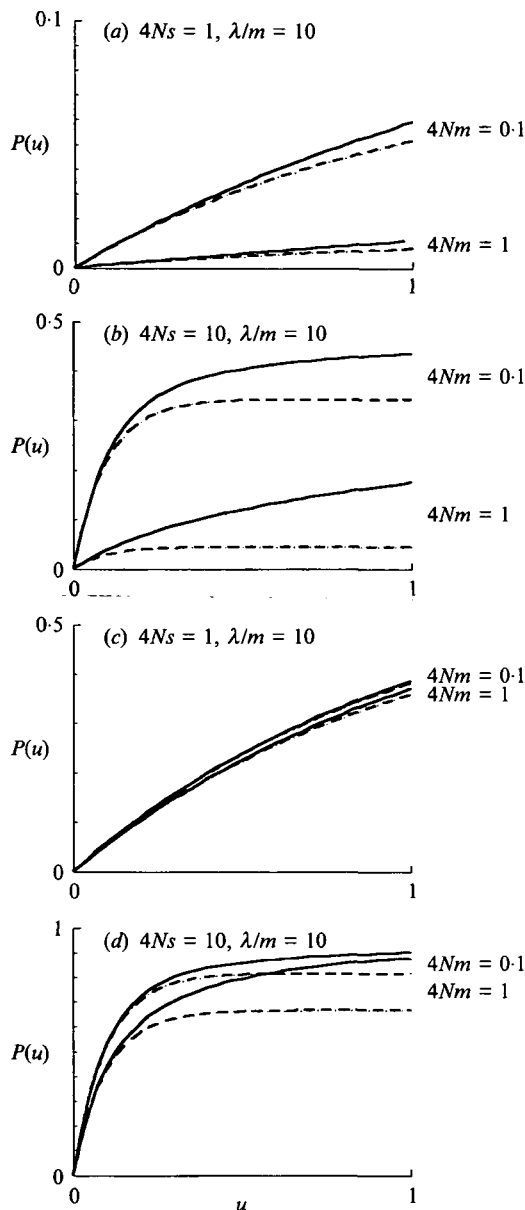
Fig. 3. The probability of fixation, $P(u)$, as a function of initial frequency in a deme, u. In each figure, the solid curves give $P(u)$, calculated from the diffusion approximation [eqn (6)]. The dotted curves give $P(u) = \Pi(1 - \exp(-Su))$, which is the form reached in the limit of low migration and extinction rates [eqn (1)].
(a) Colonists from one gene; $S = 4Ns = 1$, $\lambda/m = 10$.
(b) Colonists from one gene; $S = 4Ns = 10$, $\lambda/m = 10$.
(c) Colonists from many demes; $S = 4Ns = 1$, $\lambda/m = 10$.
(d) Colonists fromn many demes; $S = 4Ns = 10$, $\lambda/m = 10$.

7. Discussion

Random extinction introduces an extra component of sampling drift, and so reduces the probability of fixation of advantageous alleles. Unless demes are very small, and selection very strong, the fixation probability can be calculated using the diffusion approximation, and depends on the rates of selection, migration and extinction, scaled relative to population size ($S = 4Ns$, $M = 4Nm$, $\Lambda = 4N\lambda$). Though the diffusion equation can only be solved numerically,

analytic approximations can be found for the limits of low and high migration and extinction, and for weak selection [eqns (1), (9), and (11) respectively].

Both the analytic formulae, and numerical results for intermediate M, Λ, show that the effect of random extinction depends strongly on how new demes are colonized. If the genes in a new colony are drawn from many demes, then the fixation probability can be reduced by no more than half, from $2s$ to s. In contrast, when new colonies all derive from a single randomly chosen gene, the fixation probability can become very small. The difference between the two models is that when colonists are genetically variable, the colonization process involves selection as well as drift, and so favours the fitter allele. The importance of the diversity of new colonies has been emphasized in previous studies of the consequences of extinction for neutral alleles (e.g. Slatkin, 1977, 1981; Whitlock & McCauley, 1990).

When migration and extinction are infrequent (Λ, $M < 1$), the fixation probability can only be reduced substantially if colonies derive from a single gene, and if extinction is much more frequent than migration ($\Lambda \gg M$; Fig. 1a). Migration cannot be very much rarer than extinction, because new colonies must form from incoming migrants. However, the number of immigrants per generation is Nm, and so the delay before an immigrant arrives is approximately $1/(Nm)$ generations; this can be much smaller than the time between extinctions ($1/\lambda$) even if $\lambda \gg m$. Furthermore, if fitness is strongly density-dependent, immigrants may do well in an empty site, and yet rarely succeed in an established population.

Even if migration and extinction occur at similar rates, fixation probability can still be substantially reduced if colonies are founded from a single gene, and if extinction is frequent relative to selection (i.e. $\lambda \gg s$; Fig. 2a). Equation (11a), which applies in the limit of weak selection, shows that the chance that a single mutant will be fixed is

$$P(1/2N) = 2s/\{(1 + 2N\lambda)(1 + \lambda/m)\}.$$

Since $2N\lambda$ might well be large, especially if extinction is caused by environmental rather than demographic fluctuations, fixation probability could be much reduced.

Michalakis and Olivieri (1993) have recently used simulations to examine the probability that a chromosome rearrangement will fix throughout a two-dimensional population. They show that the probability of fixation of a karyotype favoured by meiotic drive can be substantially reduced, provided that migration is not too frequent ($Nm < 1$, $N\lambda$ 0·1 to 0·5; their Fig. 2). (Note that meiotic drive is equivalent to additive selection.) In these simulations, extinct demes were recolonized by migrants at a rate Nm, and so for small Nm, new colonies derive from one or a few individuals. Michalakis and Olivieri's results show a stronger effect than would be expected from the above

11-2

formulae, suggesting that extinction may have a stronger effect in two dimensions than in the island model.

Fisher's 'fundamental theorem' (Fisher, 1930) predicts that the increase in mean fitness due to natural selection should equal the additive genetic variance in relative fitness. Since weakly selected alleles ($4Ns \ll 1$) may have a much reduced chance of fixation in a structured population, they may produce less adaptation than would be expected from their contribution to the variance in fitness. The contribution of weakly selected alleles is reduced in a similar way by the 'hitch-hiking' effect of more strongly favoured substitutions (Barton, 1993), and by recurrent mutation (Keightley, 1991).

The reduction in fixation probability described here can be seen as being due to the random component of fitness introduced by extinction. Presumably, if extinction and/or recolonization were selective, as they are to a degree with a mixed founding population, the fixation probability would be greater. However, we have seen, in passing from the discrete model to the diffusion limit, that the difference between hard and soft selection is negligible. That is, the assumption that demes make a contribution to the pool of colonists proportional to their mean fitness has a negligible effect when selection is weak. This parallels the results of Rouhani & Barton (1993), who found that in models of Wright's 'shifting balance', group selection can only be important if the contribution to the migrant pool depends disproportionately on mean fitness. It also accords with the simulation results of Madalena & Hill (1972), who showed that for an additive quantitative trait, fixation probabilities are reduced even when there is strong selection among sublines.

In a single population, the fixation probability is given by $2s(N_e/N)$, where N_e is the effective population size (if $s \ll 1$; Kimura, 1962). The effects of extinction cannot generally be understood simply through a reduction in effective population size, because they depend on selection as well as on population structure. However, with high migration and extinction rates [eqn (9)], and with weak selection [eqn (11)], the fixation probability is reduced by a factor independent of selection, which suggests that this factor might just be (N_e/N).

To find (N_e/N), we solve Slatkin's (1977) equations in the limit of a large number (n) of demes, large N, and small mutation rate μ, and set the average heterozygosity to $\bar{f} = 1/(1 + 4nN_e\mu)$. (Slatkin's model differs slightly from that used here, since he assumes that drift occurs before migration. However, this has no effect in the limit of large N.) This definition of N_e is one of many alternatives. In a single panmictic population of size N, the probability $\bar{f}(t)$ that two randomly chosen genes will become identical by descent precisely t generations back declines geometrically at a rate $1/2N$. However, in a structured

population, $\bar{f}(t)$ is a more complicated function, and cannot be summarized by any one 'N_e'. The definition used here, which takes the limit of small μ, is equivalent to finding the asymptotic decline in $\bar{f}(t)$ for large t (see Ewens, 1979).

When colonies derive from many demes,

$$(N_e/N) = (1 + 2N\lambda + 4Nm)/(4N(m+\lambda))$$

[Slatkin, 1977, eqn (23)];

this differs from eqn (11b) by a factor

$$1/(1 + 1/(2N\lambda + 4Nm)).$$

When colonies are founded from a single gene,

$$(N_e/N) = (1 + 2N\lambda + 4Nm)/(4N(m+\lambda)(1 + 2N\lambda));$$

this differs from eqn (11a) by a factor

$$1/(1 + \lambda/2m + 1/4Nm),$$

implying that fixation probability may be reduced much more than neutral variability if extinction is frequent relative to migration. One can see the discrepancy more simply by considering a population with no extinction: then, the fixation probability is independent of Nm, whereas $(N_e/N) = (1 + 1/4Nm)$ [Maruyama, 1970; eqn (18) of Slatkin, 1977]. Thus, the effect of population structure on an advantageous gene cannot in general be understood from its effects on neutral alleles. The inadequacy of effective population size for understanding subdivided populations is already known from, for example, analyses of underdominance (Slatkin, 1981). However, it is somewhat more surprising that it also applies in a haploid model, or with additive selection.

References

Barton, N. H. (1993). Linkage and the limits to natural selection. *Genetics* (submitted).

Crow, J. F. & Kimura, M. (1970). *Introduction to Population Genetics Theory*. New York: Harper and Row.

Ewens, W. F. (1979). *Mathematical Population Genetics*. Berlin: Springer-Verlag.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

Keightley, P. D. (1991). Genetic variance and fixation probabilities at quantitative trait loci in mutation-selection balance. *Genetical Research* **58**, 139–144.

Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics* **47**, 713–719.

Lande, R. (1979). Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution* **33**, 234–251.

Lande, R. (1985). The fixation of chromosomal rearrangements in a subdivided population with local extinction and recolonization. *Heredity* **54**, 323–332.

Madalena, F. E. & Hill, W. G. (1972). Population structure in artificial selection programmes: simulation studies. *Genetical Research* **20**, 75–99.

Maruyama, T. (1970). Fixation probability of genes in a subdivided population. *Genetical Research* **15**, 221–225.

Maruyama, T. & Kimura, M. (1980). Genetic variability and effective population size when local extinction and recolonization of subpopulations are frequent. *Proceedings of the National Academy of Sciencess, USA* **77**, 6710–6714.

Maruyama, T. (1971). An invariant property of a geographically structured population. *Genetical Research* **18**, 81–84.

Michalakis, Y. and Olivieri, I. (1993). The influence of local extinctions on the probability of fixation of chromosomal rearrangements. *Journal of Evolutionary Biology* **6**, 153–170.

Rouhani, S. and Barton, N. H. (1993). Group selection and the 'shifting balance'. *Genetical Research* **61**: 127–135.

Slatkin, M. (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology* **12**, 253–262.

Slatkin, M. (1981). Fixation probabilities and fixation times in a subdivided population. *Evolution* **35**, 477–488.

Tachida, H. & Iizuka, M. (1991). Fixation probability in spatially changing environments. *Genetical Research* **58**, 243–251.

Whitlock, M. C. & McCauley, D. E. (1990). Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution* **44**, 1717–1724.

Wolfram, S. (1991). *Mathematica*. New York: Addison Wesley.