Stochastic Processes and their Applications 13 (1982) 249–261
North-Holland Publishing Company

A DIFFUSION PROCESS WITH KILLING: THE TIME TO FORMATION OF RECURRENT DELETERIOUS MUTANT GENES*

Samuel KARLIN
Department of Mathematics, Stanford University, Stanford, CA 94305, U.S.A.
Simon TAVARÉ
Department of Statistics, Colorado State University, Fort Collins, CO 80523, U.S.A.

Received 20 April 1981
Revised 26 October 1981

We analyze the role of recurrent mutation on the time to formation or detection of particular genotypes in finite populations. The traditional method of approximating Markov chain models by diffusion processes is used. However, the diffusion approximations that arise in this context are governed not only by infinitesimal drift and diffusion coefficients, but also by a state-dependent killing rate that arises from the formation events. The resulting processes are particularly tractable, and allow a comprehensive analysis of the role of mutation in this problem.

Markov chains	diffusion processes
population genetics	killing times

0. Introduction

Recently, some attention has been given to the following problem from population genetics. In a reproducing population of N diploid individuals, consider a single locus at which two alleles, A and a, are possible. The homozygote aa, referred to in this paper as the recessive genotype, is assumed to be visible as soon as it appears—corresponding, perhaps, to the a-allele being lethal in this form. Given that the population currently comprises only AA or Aa (heterozygote) genotypes, how long does it take to detect the first recessive aa-homozygote? Robertson [11] has analyzed some variations on this problem by discrete time methods, concentrating in particular on the effects of natural selection. These results were extended by a different approach in [7, 8]. Some of the interest in the problem derives from application of the results to artificial selection schemes and medical genetic screening. In the present context it is of interest for evolutionary theory, since it gives a more complete description of the process of gene substitution, and the formation times of genes.

* Supported in part by NIH Grant 5R01 GM10452-18 and NSF Grant MCS 79-24310.

Part of the aim of mathematical models in population genetics is to provide an analytical framework within which the relative importance of certain genetic factors can be assessed. One fundamental tool used here is the approximation of discrete Markov chains by diffusion processes whose properties may more readily be evaluated. A comprehensive account of the role of such diffusions in a genetic context is provided by Ewens [2, Chapters 3–5]. The novelty of the method we will describe hinges on the fact that the approximating diffusion processes are not conservative, but have a state-dependent killing rate. This killing rate corresponds to the appearance of population configurations that comprise *any* recessive homozygotes.

The resultant diffusion processes are particularly tractable, and give the natural setting for the analysis of problems involving gene formation. They also give a good illustration of the power of, and interest in, killing times for diffusion processes. We mention two other examples of such processes from the population genetics literature. One involves the calculation of the probability that a recombinant type appears before fixation [5], the other involves formation of high order mutants [4] (see also [12]).

In this note, we will assess the role of recurrent mutation from the A-allele to the a-allele on the behavior of the gene formation problem. Although the emphasis of the paper is on the effects of mutation, the results apply equally well to processes in which 'immigration' of a-alleles occurs into the system.

1. The model

Let X_n denote the number of heterozygotes in the population of fixed size N at times $n = 0, 1, 2, \ldots$. We suppose that recurrent mutation from the A-allele to the a-allele is possible; with probability m an A-allele mutates to an a-allele. We will also suppose that there are no selection effects operating on the system. We model the evolution of $\{X_n, n \geq 0\}$ by a Markov chain in the following way. Suppose that at time n, the population comprises $N-i$ AA-genes, i Aa-genes, and none of the recessive homozygotes. The proportion of A genes is then $1-i/2N$, and as a consequence of mutation the gene pool will have a fraction x_i of A-alleles, where

$$x_i = (1-i/2N)(1-m). \tag{1}$$

By considering the output of random matings in the population, we see that the probability of producing an Aa-type offspring is $\mathbf{P}[\mathrm{Aa}\,|\,i] = 2x_i(1-x_i)$. Similarly,

$$\mathbf{P}[\mathrm{AA}\,|\,i] = x_i^2, \qquad \mathbf{P}[\mathrm{aa}\,|\,i] = (1-x_i)^2.$$

To produce N individuals in the next generation, we take a multinomial sample of size N according to the three probabilities above. In order to continue the process of diploid formation, we require that no recessive homozygotes are sampled.

Hence the probability that $X_{n+1} = j$ and the process continues is

$$P_{ij} = \binom{N}{j}[2x_i(1-x_i)]^j x_i^{2(N-j)}, \quad i,j = 0, 1, \ldots, N. \tag{2}$$

Since the transition matrix in (2) is sub-stochastic, we add a fictitious state, H say, and then define

$$\begin{aligned} &P_{iH} = 1 - \sum_{j=0}^{N} P_{ij} = 1 - (1-(1-x_i)^2)^N, \\ &P_{HH} = 1, \quad P_{Hi} = 0, \qquad i = 0, 1, \ldots, N. \end{aligned} \tag{3}$$

The state H accounts for all population configurations in which at least one homozygous recessive genotype is found. Once the process has entered state H, we say that a killing or formation event has occurred.

To determine the parameters of the approximating diffusion process, define (for typographical convenience) $\delta_N = (2N)^{-1/3}$. Using the matrix determined by (1) and (2), explicit computation shows that as $N \to \infty$

$$\sum_{j=0}^{N} (j-i)P_{ij} \to \tfrac{1}{2}\nu, \tag{4a}$$

$$\delta_N \sum_{j=0}^{N} (j-i)^2 P_{ij} \to x, \tag{4b}$$

$$\delta_N^{-1}\left[1 - \sum_{j=0}^{N} P_{ij}\right] \to \tfrac{1}{2}x^2 \tag{4c}$$

and

$$\delta_N^3 \sum_{j=0}^{N} (j-i)^4 P_{ij} = O(\delta_N) \tag{4d}$$

with $i\delta_N \to x$ and $4Nm \to \nu > 0$ uniformly for x in compact subsets of $(0, \infty)$. If $f(x)$ is a function with two bounded continuous derivatives, a Taylor expansion and (4) lead to

$$\delta_N^{-1}\left[\sum_{j=0}^{N} P_{ij} f(j\delta_N) - f(i\delta_N)\right] \to \tfrac{1}{2}xf''(x) + \tfrac{1}{2}\nu f'(x) - \tfrac{1}{2}x^2 f(x) \tag{5}$$

as $N \to \infty$. Eq. (5) suggests that the sequence of processes $Y_N(t)$ defined by

$$Y_N(t) = (2N)^{-1/3} X_{[(2N)^{1/3}t]}, \quad t > 0, \tag{6}$$

converges to a diffusion process $\{Y(t), t \geq 0\}$ as $N \to \infty$, where $Y(\cdot)$ has state space $[0, \infty)$, infinitesimal mean and variance $\mu(x) = \frac{1}{2}\nu$, $\sigma^2(x) = x$ respectively, and killing rate $k(x) = \frac{1}{2}x^2$.

We pause to make some comments on the scalings used in this problem. As is common in population genetic problems, the mutation rate m has tc be of order

N^{-1} for the diffusion process to account for the effects of mutation (cf. [1, 2]). If m is of order $O(N^{-p})$, $0<p<1$, then the discrete process cannot be approximated by a diffusion process. If m is of order $O(N^{-p})$ for $p>1$, then the mutation rate is 'too small', and the approximating diffusion process has infinitesimal parameters $\mu(x)=0$, $\sigma^2(x)=x$, $k(x)=\frac{1}{2}x^2$; this latter process has been extensively analyzed [7, 18]. In the diffusion approximation, we keep track of heterozygote numbers to order $(2N)^{1/3}$, and the correct time scale is one in which one time unit in the diffusion corresponds to $(2N)^{1/3}$ generations in the discrete model. If the number of heterozygotes is of order much larger than $(2N)^{1/3}$, the detection of a recessive homozygote occurs effectively 'instantly'. The time scale $(2N)^{1/3}$ stands in marked contrast to the common scaling of $2N$ that arises in population genetic models.

From now on, we concentrate our attention on the diffusion process $\{Y(t), t\geq 0\}$ with state space $[0, \infty)$ and infinitesimal parameters given by

$$\mu(x)=\tfrac{1}{2}\nu, \qquad \sigma^2(x)=x, \qquad k(x)=\tfrac{1}{2}x^2. \tag{7}$$

2. Diffusion analysis

Of particular importance in the sequel is the behavior of the process Y near the boundaries 0 and ∞. In all cases, $\{\infty\}$ is a natural boundary, while $\{0\}$ is regular (if $0<\nu<1$) or entrance (if $\nu\geq 1$) (see, for example, [9, p. 50]). The first case reflects the fact that small mutation rates mean that state $\{0\}$ is accessible, which allows for realizations of the population comprising only A-alleles. The usual boundary condition corresponding to the discrete model is that of a reflecting barrier. In the second case the mutation rate is so large that the population can never comprise only A-alleles if it starts with a positive number of a-alleles, whereas if the Y process commences from $\{0\}$, then the process moves away from $\{0\}$. In all cases, the only possible outcome is that the process ends with a killing event. A variety of interesting probabilistic functionals is found by solving equations of the form

$$\mathrm{L}\,u=\tfrac{1}{2}xu''+\tfrac{1}{2}\nu u'-\tfrac{1}{2}x^2u=-f, \tag{8}$$

for appropriate functions f and boundary conditions on u.

Before proceeding to special cases we introduce some notation that will be used in the sequel. Let $I_\gamma(x)$ and $K_\gamma(x)$ be the modified Bessel functions defined by

$$I_\gamma(x)=(\tfrac{1}{2}x)^\gamma \sum_{k=0}^{\infty} \frac{(\frac{1}{2}x)^{2k}}{k!\Gamma(\gamma+k+1)}$$

and

$$K_\gamma(x)=\frac{\pi}{2\sin(\gamma\pi)}[I_{-\gamma}(x)-I_\gamma(x)],$$

$K_\gamma(\cdot)$ being defined by its limiting value if γ is an integer. We can now identify two linearly independent solutions $y_1(x)$ and $y_2(x)$ of the differential equation

L $y = 0$. These are defined by

$$y_1(x) = x^{(-\nu)/2} I_\gamma(\tfrac{2}{3}x^{3/2}) \tag{9a}$$

and

$$y_2(x) = x^{(1-\nu)/2} K_\gamma(\tfrac{2}{3}x^{3/2}) \tag{9b}$$

where $\gamma = \pm\frac{1}{3}(\nu - 1)$.

3. Large mutation rates $\nu \geq 1$

In this process the interest rests on the properties of the killing or formation time T. From standard theory for processes with an entrance boundary at $\{0\}$ the mean killing time $M(x) = \mathbf{E}(T \mid Y(0) = x)$ satisfies

$$\mathrm{L}\, M(x) = -1, \qquad M(\infty) = 0, \qquad \lim_{x \downarrow 0} x^{\nu} M'(x) = 0. \tag{10}$$

If we set

$$\gamma = \tfrac{1}{3}(\nu - 1) \geq 0, \tag{11}$$

then the Green's function for the problem is given by

$$G(x, \xi) = \begin{cases} \frac{4}{3}\xi^{\nu-1} y_1(\xi) y_2(x), & 0 < \xi \leq x, \\ \frac{4}{3}\xi^{\nu-1} y_1(x) y_2(\xi), & x \leq \xi < \infty. \end{cases} \tag{12}$$

The mean time to detection is then $M(x) = \int_0^\infty G(x, \xi)\, \mathrm{d}\xi$, or, from (12),

$$M(x) = \tfrac{4}{3} y_2(x) \int_0^x \xi^{\nu-1} y_1(\xi)\, \mathrm{d}\xi + \tfrac{4}{3} y_1(x) \int_x^\infty \xi^{\nu-1} y_2(\xi)\, \mathrm{d}\xi. \tag{13}$$

From a biological viewpoint we are most interested in the behavior of $M(x)$ near $x = 0$ (that is, when the number of heterozygotes in the initial generation is very small). Examining the form of y_1, y_2 near $x = 0$ shows that $y_2(x) \int_0^x \xi^{\nu-1} y_1(\xi)\, \mathrm{d}\xi = \mathrm{O}(x)$ as $x \downarrow 0$ $(\nu > 1)$ and $= \mathrm{O}(x \ln x)$ as $x \downarrow 0$ $(\nu = 1)$, so that

$$M(0+) = \tfrac{4}{3} y_1(0) \int_0^\infty \xi^{\nu-1} y_2(\xi)\, \mathrm{d}\xi.$$

The integral can be evaluated explicitly (cf. [3, p. 684, Eq. (16)]) to give

$$M(0+) = 3^{-4/3} \Gamma(\tfrac{1}{3}) 2\Gamma(\tfrac{1}{3}\nu) / \Gamma(\tfrac{1}{3}(\nu + 2)) \tag{14}$$

In terms of the underlying process of (2, 3) it follows that if the discrete model starts with very few heterozygotes (the population comprises almost all A-alleles), then it takes of order $(2N)^{1/3} M(0+)$ generations to produce the first homozygous

recessive genotype. For large values of ν we have

$$M(0+) = 1.24 \frac{\Gamma(\frac{1}{3}\nu)}{\Gamma(\frac{1}{3}(\nu+2))} \sim 2.48\nu^{-2/3}, \quad \nu \to \infty.$$

Another functional, related to the genetic cost of the deleterious mutant, is the expected cumulative number of heterozygotes that occur before the first recessive homozygote is formed. Denoting this function by $H(x)$ we find that H satisfies $\mathrm{L}\, H(x) = -x$, and hence, from (12),

$$H(x) = \tfrac{4}{3} y_2(x) \int_0^x \xi^\nu y_1(\xi)\, d\xi + \tfrac{4}{3} y_1(x) \int_x^\infty \xi^\nu y_2(\xi)\, d\xi. \tag{15}$$

For large values of x, $H(x)$ decreases like x^{-1}, whereas for small x we have

$$\begin{aligned} H(0+) &= \tfrac{4}{3} \frac{1}{3^\gamma \Gamma(1+\gamma)} \int_0^\infty \xi^\nu y_2(\xi)\, d\xi \\ &= 2\Gamma(\tfrac{1}{3}) 3^{-2/3} \Gamma(\tfrac{1}{3}(\nu+1))/\Gamma(\tfrac{1}{3}(\nu+2)). \end{aligned} \tag{16}$$

In terms of the discrete model (16) shows that if the detection process starts with very few heterozygotes, then the expected total number of heterozygotes produced before detection will be approximately $(2N)^{2/3} H(0+)$. For large values of ν, $H(0+) \sim 3.72\nu^{-1/3}$, $\nu \to \infty$.

The last functional we evaluate in this section is the distribution of the place of detection P. Define $w(x, I)$ to be the probability that the process $Y(\cdot)$ is killed in the set I, given that $Y(0) = x$. A simple probabilistic argument shows that $w(x, I)$ satisfies the differential equation

$$\tfrac{1}{2} x w''(x, I) + \tfrac{1}{2}\nu w'(x, I) + \tfrac{1}{2} x^2 w(x, I) = -k(x)\delta(x, I) \tag{17}$$

where

$$\delta(x, I) = \begin{cases} 1 & \text{if } x \in I, \\ 0 & \text{if } x \notin I. \end{cases}$$

The appropriate solution of (17) is given by

$$w(x, I) = \int_0^\infty G(x, \xi) k(\xi) \delta(\xi, I)\, d\xi = \int_I G(x, \xi) k(\xi)\, d\xi. \tag{18}$$

It follows from (18) that the density function of P for $Y(0) = x$ is given by

$$w(x, \xi) = G(x, \xi) k(\xi), \quad \xi > 0. \tag{19}$$

Again, we are particularly interested in the case in which the initial generation comprises a very small number of heterozygotes. Taking $x = 0$ in (19) and using

(9) and (12) we see that

$$w(0,\xi)=\tfrac{2}{3}y_1(0)\xi^{\nu+1}y_2(\xi)$$
$$=\frac{2\xi^{(3+\nu)/2}}{3^{\gamma+1}\Gamma(1+\gamma)}K_\gamma(\tfrac{2}{3}\xi^{3/2}),\quad \xi>0, \tag{20}$$

where $\gamma=\frac{1}{3}(\nu-1)$.

The class of densities specified in (20) is unimodal. To see this, differentiate $w(0,\xi)$ twice to get

$$w''(0,\xi)=\tfrac{2}{3}y_1(0)\xi^{\nu-1}[\nu(\nu+1)y_2(\xi)+2(\nu+1)\xi y_2'(\xi)+\xi^2 y_2''(\xi)].$$

Using the differential equation satisfied by $y_2(\xi)$ simplifies this to

$$w''(0,\xi)=\tfrac{2}{3}y_1(0)\xi^{\nu-1}[(\xi^3+\nu(\nu+1))y_2(\xi)+(\nu+2)\xi y_2'(\xi)].$$

Now at any turning point $\xi\geq 0$, we have $(\nu+1)y_2(\xi)+\xi y_2'(\xi)=0$, and hence, at such points,

$$w''(0,\xi)=\tfrac{2}{3}y_1(0)\xi^{\nu-1}y_2(\xi)(\xi^3-2(\nu+1)).$$

This function changes sign only once in the interval $\xi>0$, and $w(0,0)=0$, and hence the densities must be unimodal. Further, the moments of the detection position P can be computed explicitly:

$$\mathbf{E}(P^n\mid Y(0)=0)=\int_0^\infty \tfrac{2}{3}y_1(0)\xi^{n+\nu+1}y_2(\xi)\,d\xi$$
$$=3^{2n/3}\Gamma(\tfrac{1}{3}(n+\nu+2))\Gamma(1+\tfrac{1}{3}n)/\Gamma(\tfrac{1}{3}(\nu+2)),$$
$$n=0,1,2,\ldots. \tag{21}$$

For illustration we take $\nu=1$. In this case, the mode ξ_0 of the density can be computed numerically to give $\xi_0=1.229$, while the mean and variance of the detection position are given by 1.659 and 0.775 respectively. In terms of the original process we can deduce that if the population starts with one heterozygote, then the most likely detection position is at about $1.55N^{1/3}$ individuals, the mean detection position being about $2.09N^{1/3}$ individuals. The behavior of the mean and variance of the detection position as a function of ν is shown in Table 1.

4. Small mutation rates $0<\nu<1$

When the mutation rate is small $(0<\nu<1)$, the boundary at $\{0\}$ is accessible, and so a complete description of the process requires prescribing the boundary behavior there. In this section we concentrate on regular reflecting behavior at $\{0\}$. It is straightforward to check that if we set $\gamma=\frac{1}{3}(\nu-1)$, then the Green's function is given by (12) again.

It follows that the mean time to detection is given by (13), and, for x small, the result in (14) applies. We can now assess the effect of mutation on the time to detection. In Fig. 1 the values of $M(0+)$ are plotted as a function of $\nu = 4Nm$.

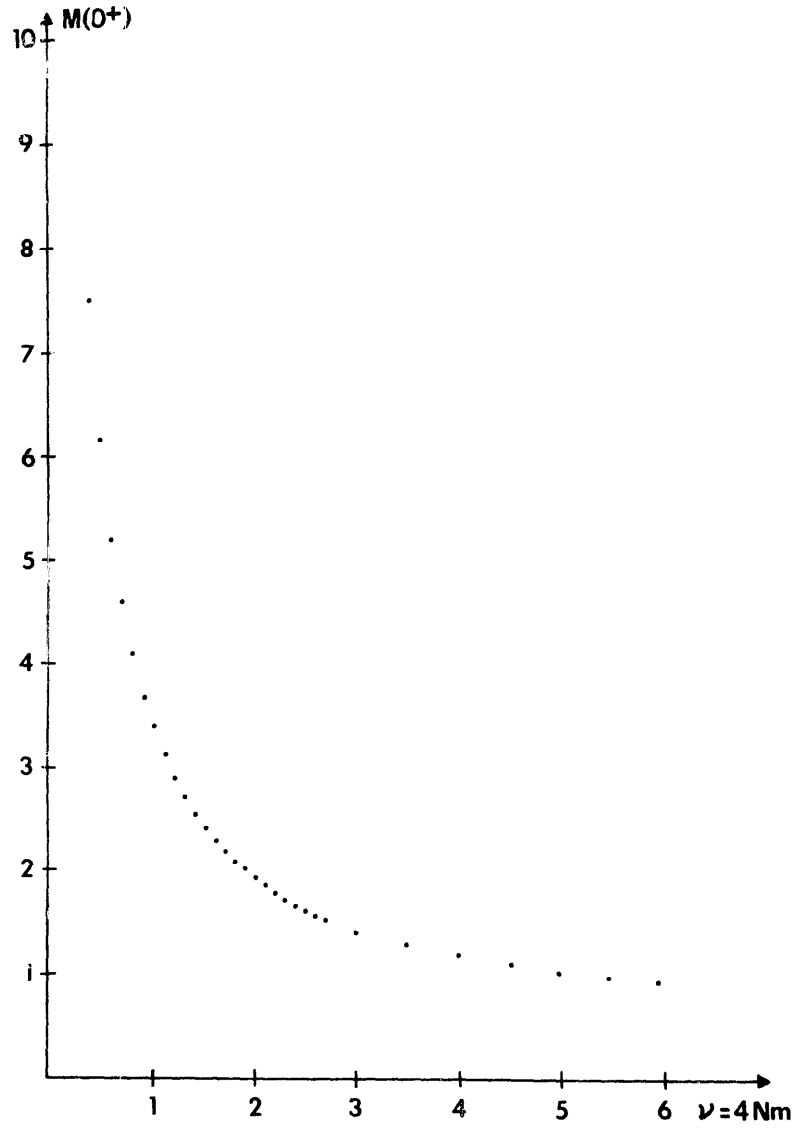

Fig. 1. Mean time to detection $M(0+)$ from (14) as a function of $\nu = 4Nm$. For the discrete process starting with $X_0 = 1$, the mean time for detection is approximately $(2N)^{1/3}M(0+)$ generations.

They confirm the intuitive result that as the mutation rate increases, the mean detection time decreases when the population starts with a very small number of heterozygotes.

Moments of the detection position are again given by (21). The effects of mutation on the mean and variance of this position are described in Table 1.

The results in Table 1 show that in the discrete model the mean detection position starting from 1 heterozygote, which is given by $(2N)^{1/3}\mathbf{E}(P)$ heterozygotes, increases slowly with ν. This corresponds to the production of larger numbers of heterozygotes as ν increases.

Table 1. Mean and variance of detection position P computed from (21), $\nu = 4Nm$.

ν	$\mathbf{E}(P)$	var(P)
0.100	1.405	0.703
0.300	1.468	0.720
0.500	1.527	0.744
0.700	1.582	0.752
0.900	1.634	0.767
1.000	1.659	0.775
1.500	1.774	0.812
2.000	1.878	0.848
2.500	1.972	0.883
3.000	2.058	0.918
3.500	2.137	0.952
4.000	2.212	0.986
4.500	2.281	1.018

5. Quasi-fixation problem

In this section we discuss the case $0 < \nu < 1$ in a different setting. We are now interested in the behavior of the process between visits to $\{0\}$. To model the behavior of the process after a visit to $\{0\}$, suppose the population remains fixed at 0 for a length of time W, at which time a new mutant deleterious gene is introduced into the population at a given frequency, from which state the process continues as before. Some problems of interest for this model include:

(a) What is the probability of a quasi-fixation event? That is, what is the probability that the mutant gene is not detected in homozygous form?

(b) What is the average number of new deleterious mutant types that arises before the first of them appears as a homozygote?

(c) What is the average time to detection of mutant homozygote in this model? What is the mean time to detection of a *particular* mutant, conditional on detection taking place?

To analyze these functionals we begin by computing the quasi-fixation probabilities. Let $u(x)$ be the probability that a quasi-fixation event occurs prior to detection, starting from $Y(0) = x$. Then $u(x)$ satisfies the differential equation $\mathrm{L}\, u = 0$, $u(0) = 1$, $u(\infty) = 0$. Examination of the functions in (9) reveals the required solution as

$$u(x) = \frac{y_2(x)}{y_2(0)}, \quad y_2(x) = x^{(1-\nu)/2} K_\gamma(\tfrac{2}{3}x^{3/2}), \ \gamma = \tfrac{1}{3}(1-\nu). \tag{22}$$

It is intuitively clear, and simple to prove, that the quasi-fixation probability $u(x)$ is monotone decreasing in ν as ν traverses 0 to 1 for fixed x. The probability that detection occurs first, is then $v(x) = 1 - u(x)$. For small values of x expansion of

(9) shows that

$$v(x)=\frac{\Gamma(1-\gamma)x^{1-\gamma}}{\Gamma(1+\gamma)3^{2\gamma}}-\frac{x^3}{9(1-\gamma)}+O(x^{4-\nu}),\quad x\downarrow 0. \tag{23}$$

To assess the role of the parameter ν we compute the probability of detection v in a population of size $N=500$ starting with one heterozygote. These are found from (23) with $x=(2N)^{-1/3}=0.1$. The results are presented in Table 2.

Table 2. Probability of detection, v, in population of size 500 with $X_0=1$ computed by taking $x=(2N)^{-1/3}=0.1$ in (23), $\nu=4Nm$.

ν	v
0.00	0.073
0.01	0.075
0.02	0.077
0.03	0.079
0.04	0.081
0.05	0.083
0.06	0.085
0.07	0.087
0.08	0.089
0.09	0.092
0.10	0.094
0.20	0.122
0.30	0.158
0.40	0.210
0.50	0.267

The moments of the killing/absorption times are found by identifying the appropriate Green's function; this is computed by taking the exist boundary condition at $\{0\}$. Setting $\gamma=\frac{1}{3}(1-\nu)$, we get

$$G(x,\xi)=\begin{cases}\frac{4}{3}\xi^{\nu-1}y_1(\xi)y_2(x), & 0<\xi\leqslant x,\\ \frac{4}{3}\xi^{\nu-1}y_1(x)y_2(\xi), & x\leqslant\xi,\end{cases} \tag{24}$$

where y_1, y_2 are the standard solutions exhibited in (9). If we define T to be the time to fixation *or* detection, then $M(x)=\mathbf{E}(T\mid Y(0)=x)$ satisfies $M(x)\sim C_0x^{1-\nu}$, as $x\downarrow 0$, where

$$C_0=\frac{2\Gamma(\frac{1}{3})\Gamma(\frac{1}{3}\nu)3^{2\nu/3}}{3(1-\nu)\Gamma(\frac{1}{3}(1-\nu))}. \tag{25}$$

For the discrete model, this result is interpreted as follows. If the large population comprises initially 1 heterozygote so that we may take $x=(2N)^{-1/3}$, then the mean time to absorption *or* detection is of order $C_0(2N)^{\nu/3}$ generations.

For further analysis of quasi-fixation we need to evaluate the means of the conditional absorption and detection times. The appropriate Green's function for

the process conditioned on absorption is given by

$$G_0(x,\xi) = G(x,\xi)u(\xi)/u(x), \tag{26}$$

where $G(\cdot,\cdot)$ and $u(\cdot)$ are specified by (22) and (24), respectively. Denoting the (conditioned) absorption time by T_0 we find that

$$M_0(x) = \mathbf{E}(T_0 \mid Y(0) = x) \sim C_1 x^{1-\nu}, \quad x \downarrow 0.$$

The constant C_1 can be evaluated explicitly using, for example, [3, p. 693, eq. (4)]:

$$C_1 = \frac{3^{2\nu/3}}{3(1-\nu)}\left(\frac{\Gamma(\frac{1}{3})}{\Gamma(\gamma)}\right)^2 \frac{2\Gamma(\frac{1}{3}\nu)\Gamma(\frac{1}{3}(2-\nu))}{\Gamma(\frac{2}{3})}. \tag{27}$$

The equivalent properties of the process conditioned on killing are computed using the conditioned Green's function $G_D(x,y) = G(x,y)(1-u(y))/(1-u(x))$.

The mean conditional time to killing is $M_D(x) = \int_0^\infty G_D(x,y)\,dy$, and its value for x small is given by

$$M_D(0+) = C_2 = \frac{2\Gamma(\frac{1}{3})\Gamma(\frac{1}{3}\nu)}{3^{4/3}\Gamma(\frac{1}{3}(\nu+2))}\left(1 - \frac{\Gamma(\frac{1}{3})\Gamma(\frac{1}{3}(2-\nu))}{\Gamma(\frac{1}{3}(1-\nu))\Gamma(\frac{2}{3})}\right). \tag{28}$$

It is in this case that we most expeditiously compare the effects of mutation ($0 < \nu < 1$) with the case of no mutation ($\nu = 0$), since the two processes have the same boundary behavior. We will focus on the conditional mean detection time starting with a very small number of heterozygotes, so that (28) applies. The mean time to detection is then given approximately by $(2N)^{1/3}C_2$ generations for the discrete process. The constant C_2 can be evaluated numerically for different values of ν, $0 < \nu < 1$. The limiting case $\nu = 0$ gives

$$C_2 = 2\pi\Gamma(\tfrac{1}{3})/3^{11/6}\Gamma(\tfrac{2}{3}) \approx 1.659.$$

Some numerical values are given in Table 3.

Table 3. Values of C_2 in (28), $\nu = 4Nm$.

ν	C_2
0.0	1.659
0.1	1.760
0.3	1.988
0.5	2.260
0.7	2.598
0.9	3.039

The following somewhat counterintuitive observation arises from Table 3. Conditional on eventual detection, the mean time to detection *increases* as the mutation rate increases. We point out that the same behavior occurs in the corresponding discrete process (when $0 < 4Nm < 1$), where numerical results can be found by matrix inversion.

This curious behavior is essentially due to the conditioning event. For another unintuitive result for conditioned processes, see the selection model discussed by Ewens [2, Chapter 5] and Maruyama [10].

We can now furnish a description of the simplest quasi-fixation model. We assume that initially $Y(0) = x_0$. We allow the process to evolve until either a quasi-fixation event occurs or until a detection event kills the process. If the former event occurs, we let the population remain fixed in an all-AA configuration for a length of time W. After W the process of detection restarts (with a new mutant 'a' gene introduced into the population) from x_0 and then continues as before. The process ends by detection of the first homozygous mutant.

If we let N denote the number of deleterious genes that appear before the first of them is detected in homozygous form, then clearly N has a geometric distribution with

$$\mathbf{P}[N = n] = u^{n-1}(1-u), \quad n \geq 1, \tag{29}$$

where $u = u(x_0)$ is the probability of a quasi-fixation starting from x_0. Hence the mean number of different deleterious genes that appears before the first is detected is

$$\bar{\mu} = \frac{1}{1-u(x_0)}. \tag{30}$$

Typically, a single new mutant will enter the (discrete) population. Setting $x_0 = (2N)^{-1/3}$ and using the approximation (23) we see that the average number of new mutants is given approximately by

$$\bar{\mu} \approx (2N)^{\gamma}\Gamma(1+\gamma)3^{2\gamma}/\Gamma(1-\gamma), \quad \gamma = \tfrac{1}{3}(1-\nu). \tag{31}$$

A simple compounding argument now shows that the mean length of time until the process ends by forming any recessive homozygotes is given by

$$\bar{T} = M_D(x_0) + (\{M_0(x_0) + \mathbf{E}[W]\}). \tag{32}$$

For small values of x_0 the estimates of (27) and (28) can be used to evaluate this expression.

We remark that many other functionals of this quasi-fixation process (for example, the average aggregate number of carriers and the distribution of the detection position) can be analyzed using methods similar to those developed here.

References

[1] J.F. Crow and M. Kimura, An Introduction to Population Genetics Theory (Harper and Row, New York, 1970).

[2] W.J. Ewens, Mathematical Population Genetics (Springer, New York, 1979).

[3] I.S. Gradshteyn and I.M. Ryzhik, Tables of Integrals, Series and Products (Academic Press, New York, 4th ed., 1965).

[4] S. Karlin, Sex and infinity, a mathematical analysis of the advantages and disadvantages of genetic recombination, in: M.S. Bartlett and R.W. Hiorns, eds., The Mathematical Theory of the Dynamics of Biological Populations (Academic Press, New York, 1973) pp. 155–194.

[5] S. Karlin, J.L. McGregor and W.F. Bodmer, The rate of production of recombinants between linked genes in finite populations, Proc. Fifth Berkeley Symp. on Math. Stat. Prob. Vol. IV (Univ. of Calif. Press, San Francisco, 1966) pp. 403–414.

[6] S. Karlin and S. Tavaré, The detection of a recessive visible gene in finite populations, Genetical Research, Cambridge 37 (1981) 33–46.

[7] S. Karlin and S. Tavaré, The detection of particular genotypes in finite populations Part I. Natural selection effects, Theor. Pop. Biol. 19 (1981) 187–214.

[8] S. Karlin and S. Tavaré, The detection of particular genotypes in finite populations Part II. The effects of partial penetrance and family structure, Theor. Pop. Biol. 19 (1981) 215–229.

[9] P. Mandl, Analytical Treatment of One-Dimensional Markov Processes (Springer, New York, 1968).

[10] T. Maruyama, The age of an allele in a finite population, Genetical Research, Cambridge 23 (1974) 137–143.

[11] A. Robertson, The time to detection of recessive visible genes in small populations, Genetical Research, Cambridge 31 (1978) 255–264.

[12] S. Sawyer, On the past history of an allele now known to have frequency p, J. Appl. Probab. 14 (1977) 439–450.