Implications of long tails in the distribution of mutant effects

D. Waxman[a],*, J. Feng[b]

[a] Centre for the Study of Evolution, School of Life Sciences, University of Sussex, Brighton BN1 9QG, Sussex, UK
[b] Centre for Scientific Computing, University of Warwick, Coventry CV4 7AL, UK

Abstract

Long-tailed distributions possess an infinite variance, yet a finite sample that is drawn from such a distribution has a finite variance. In this work we consider a model of a population subject to mutation, selection and drift. We investigate the implications of a long-tailed distribution of mutant allelic effects on the distribution of genotypic effects in a model with a continuum of allelic effects. While the analysis is confined to asexual populations, it does also have implications for sexual populations. We obtain analytical results for a selectively neutral population as well as one subject to selection. We supplement these analytical results with numerical simulations, to take into account genetic drift. We find that a long-tailed distribution of mutant effects may affect both the equilibrium and the evolutionary adaptive behaviour of a population.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

When genetic material is duplicated, during the production of offspring, copying errors—mutations—may occur. Non-mutated alleles are passed on identically to the next generation while mutated alleles differ from the parental genes. In the present paper, we concern ourselves with fundamental properties of the distribution of mutations and exclusively address the case of a population of asexual organisms.

Individuals are taken to be characterised by a single phenotypic trait that is controlled by many loci. We assume that the allelic mutation rate is sufficiently small that offspring are highly unlikely to differ from their parent by two or more mutations. In this case, the entire genome can be thought of as a single haploid locus. Thus, independent of the level of ploidy and the number of loci we shall treat individuals as consisting of a single haploid locus with a very large number of different possible alleles. We note that results for a single haploid locus may apply to a sexual population, if the neglect of linkage disequilibria is valid (which has to be established), since when it is, each individual can be viewed as a collection alleles at haploid asexual loci

in a genetic background consisting of the other alleles (see, e.g. [1]).

We adopt a model with a continuum of alleles [2] and in such models, alleles are typically labelled by a single, continuous, real parameter, x, where $\infty > x > -\infty$. Mutant values of x are randomly chosen from a continuous distribution and consequently can take a continuum of possible values. Since there is zero probability of obtaining precisely the same value twice, from a continuous probability distribution, every mutant allele differs from all alleles that were present in the population prior to the mutation. In the present work, the value of x is interpreted as the genotypic component of a phenotypic trait [1,2]. On account of this, we shall refer to x as the genotypic *effect* (or sometimes just the effect) of the allele.

If a parent, with effect x_p, undergoes a mutation, then the probability that the resulting mutant has an effect in the infinitesimal range $(x, x + \mathrm{d}x)$ is given by $M(x|x_p)\,\mathrm{d}x$ where $M(x|x_p)$ is the distribution of mutant effects. The function $M(x|x_p)$ is a probability density, and, as such, is non-negative and normalised to unity: $M(x|x_p) \geq 0$, $\int M(x|x_p)\mathrm{d}x = 1$ (here and elsewhere, all integrations with unspecified limits cover the range $-\infty$ to ∞). In general, $M(x|x_p)$ is a weighted average (equivalently, a mixture) over the distributions of mutant allelic effects at the different loci controlling the trait and as such, may have a very different shape from those of the underlying loci [3,4].

We shall follow most previous analyses by taking $M(x|x_p)$ to be a *unimodal* (single peaked) function of x. In contrast to previous analyses (see, e.g. [5]), we take $M(x|x_p)$ to possess long tails. The feature—or defining property—of long-tailed distributions (also called fat- or heavy-tailed distributions in the literature) is that while the distributions are necessarily non-negative and normalised to unity—as indicated above, such distributions always have an infinite variance and even their expected value may not be well defined [6]. We note, in passing, that discrete distributions can also be long-tailed, so there may also be applications of such distributions to meristic traits, such as offspring number or bristle number, as well as to the continuous traits considered here.

The fact that a probability distribution has an infinite variance does not mean the *measured* variance is infinite. Rather, we note that in a *sample* consisting of a *finite* number of terms taken from such a distribu-

tion, the variance of the sample is *finite*. Thus, from the viewpoint of summary statistics of *finite samples*, there is nothing manifestly pathological about such distributions. Some of the ways of viewing the significance of long tails in a distribution are discussed in Appendix A.

The objective of the present work is, principally, to re-examine the equilibrium properties of continuum of alleles models involving mutation and selection, without making the implicit assumption that the distribution of mutant effects is short-tailed. Long-tailed distributions have been extensively discussed in the past (see, e.g. [7]). However, to the best of our knowledge a long-tailed distribution has not been introduced into the model or the context we discuss in the current paper and therefore its functional role has not been addressed. Thus, the whole objective of the present paper may be summarised as looking for significant differences, for population genetics and evolution, between the outcome of the conventional, short-tailed, distributions of mutational effects, with the outcome arising from mutational distributions possessing long tails.

It seems to us that the objection that: because the genome is finite, the distribution of mutant effects has finite moments of all orders and hence is, necessarily, short-tailed, is not compelling. The number of different possible mutations is an astronomically large number, and the overwhelming proportion of mutations will never be observed. Thus, for practical purposes, the distribution of mutant effects could have all the appearances of a long-tailed distribution—out to a large but finite cutoff value—that is never likely to be even remotely approached and observed. Therefore, we view questions about the form of the distribution of mutant effects—such as whether it possesses long tails or not—as meriting further theoretical and experimental investigation, rather than simply being decided a priori. In the case where mutations are observed, whose effects are a number of standard deviations (say three or more) from the mean of the sampling distribution, one might suspect that a short-tailed distribution of mutant effects may not be an appropriate description. We note that there are general ways to assess whether a finite data sample collected from experiments is best described by a long-tailed distribution, for example, by using the Hurst index [8].

2. Form of the distribution of mutant effects

To perform an investigation of reasonable length into the implications of a distribution of mutant effects, $M(x|x_p)$, that has long-tails, we are forced to make restrictions on the form this distribution can take. The restrictions we adopt still allow, however, a variety of different forms of $M(x|x_p)$. In particular, we do not consider the most general long-tailed distribution, but base our analysis on a family of non-negative, normalised, symmetric and unimodal functions, $f(x)$, whose form we shall shortly give. There are several different forms that $M(x|x_p)$ can take even within this class, namely (i) $M(x|x_p) = f(x - x_p)$ [1,2]. Because of the dependence of $M(x|x_p)$ on $x - x_p$, this can be termed a "translationary" invariant distribution, since the same translation (i.e. shift) in both x and x_p leads to the distribution being unaltered; (ii) $M(x|x_p) = f(x)$ [9]. The final state of a mutation, x, is unrelated to the parental effect, x_p, and this is termed the "House of Cards" model of mutation, in analogy to the final, demolished state of a house of cards being unrelated to its initial, ordered, state; (iii)

$$M(x|x_p) = f(x - \gamma x_p) \tag{1}$$

[10] where $1 \geq \gamma \geq 0$. This is called the "regression" model of mutation, in view of the apparent connection of the argument of $f(\bullet)$ with linear regression. Clearly the regression model of mutation can interpolate between the House of Cards and translationary invariant models of mutation, by choosing the "regression" parameter γ to be 0 or 1 and we shall carry out most of our investigations in terms of this model for general values of γ.

In principle we could incorporate a systematic mutational bias into mutation, by incorporating a parameter b (where $\infty > b > -\infty$) into the argument of $f(\bullet)$, so that, e.g. in the regression mutation model, we would have $f(x - \gamma x_p - b)$. However, there have been some recent studies of the interplay between biased mutation and selection [11,12] and we shall not deal with the considerable added level of complexity associated with bias here, except as an aside, for the case of selective neutrality. We could also incorporate an asymmetry of the function $f(\bullet)$ about its maximum, that cannot be simply expressed as a mutational bias, but shall not do so here.
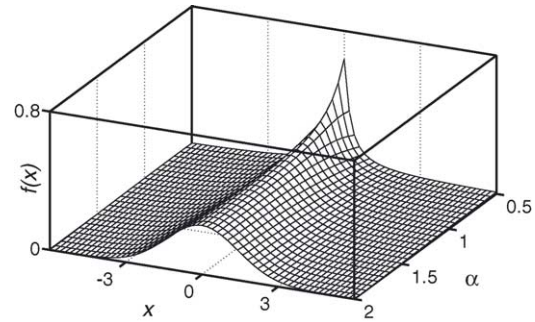

Fig. 1. The function $f(x)$ that is used in distribution of mutant effects, Eq. (2), is plotted as a function of genotypic effect, x, and "shape" parameter α. The parameter m, which is a measure of the width of the distribution, has been set to unity.

The form of $f(x)$ adopted here is characterised by two parameters, namely m and α, which appear in a Fourier integral representation of $f(x)$:

$$f(x) = \int \exp(-ikx)g(k)\frac{\mathrm{d}k}{2\pi}, \quad g(k) = \exp(-m^\alpha |k|^\alpha) \tag{2}$$

where $\infty > m > 0$ and $2 > \alpha > 0$. Since $g(-k) = g(k)$ and $g(0) = 1$, the function $f(x)$, as defined above, is manifestly real, symmetric and normalised to unity. However, it is by no means obvious that $f(x)$ is nonnegative, as it necessarily must be to be a probability density, and unimodal, as we have assumed. It has been established, however, that an $f(\bullet)$ of the form of Eq. (2) is indeed non-negative and unimodal [13] and is an example of a "Levy stable distribution".

The parameter α in Eq. (2) controls aspects of the shape of the distribution but only a few values of α lead to an integral in Eq. (2) that can be evaluated in a relatively simple form. The cases $\alpha = 1$ and 2 lead to straightforward integrations, and yield the well-known distributions associated with Cauchy and Gauss and are given in Appendix B. Some other values of α lead to integrals which may be evaluated in terms of special functions and examples of these may also be found in Appendix B (where a total of four representative forms for $f(x)$ are given). In Fig. 1, we plot the distribution, $f(x)$ of Eq. (2), as a function of x, for a range of α but with the parameter m set to unity.

The parameter α also determines the asymptotic behaviour of the family of distributions given

in Eq. (2):

$$f(x) \sim A_\alpha |x|^{-\alpha-1}, \quad \text{for } |x| \longrightarrow \infty, \quad (2 > \alpha > 0)$$

$$A_\alpha = \frac{1}{\pi} m^\alpha \sin\left(\frac{\pi\alpha}{2}\right) \Gamma(1+\alpha) \qquad (3)$$

[6] where $\Gamma(\bullet)$ denotes Euler's gamma function [14] and we shall make use of these results later. The small x properties of $f(x)$ are contained in the first few non-zero derivatives at the origin:

$$f(0) = \frac{\Gamma(1/\alpha)}{\pi\alpha m}, \qquad (4)$$

$$\left.\frac{d^2}{dx^2} f(x)\right|_{x=0} = -\frac{\Gamma(3/\alpha)}{\pi\alpha m^3} \equiv -\frac{1}{m^2}\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)} f(0).$$

$$(5)$$

All distributions of the form Eq. (2), with $\alpha < 2$, have an infinite variance, since we can write $\text{Var}(x) = \int x^2 f(x)\,dx = -[d^2 g(k)/dk^2]_{k=0}$ and this diverges. Equivalently, the absence of a finite variance for $\alpha < 2$ arises because of the slow power-law decrease of the distributions at large $|x|$, Eq. (3), so e.g. the large $|x|$ contribution is proportional to $\int^\infty x^2 x^{-\alpha-1}\,dx$, which diverges. It is precisely in this sense that the distributions of the type in Eq. (2) with $\alpha < 2$ have long tails.

The parameter m controls the range of x over which $f(x)$ changes appreciably. There are several, informative ways of seeing this. (i) The maximum value of $f(x)$ occurs at $x = 0$ and $\max_x f(x) = f(0)$ is given above in Eq. (4). The dependence of $\max_x f(x)$ on m^{-1} indicates that increasing m reduces the maximum value of $f(x)$, and because of normalisation and unimodality of $f(x)$, this decrease can only arise because of a resultant broadening of the tails of the distribution. (ii) From Eq. (5), it follows that in the vicinity of $x = 0$, we have $f(x) \simeq f(0) \times (1 - (1/2)(x/\ell)^2 + \cdots)$ where the characteristic scale of variation (i.e. the "range") of $f(x)$ is ℓ and $\ell = m\sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)}$. Note that ℓ is an increasing function of α and writing $\ell = \ell(\alpha)$ we have $\ell(1/2) \simeq 0.09\,m$, $\ell(1) \simeq 0.71\,m$ and $\ell(2) \simeq 1.41\,m$, i.e. for $\alpha \gtrsim 1$, ℓ is $O(m)$. (iii) From the analytical forms for $f(x)$ in Appendix B, we have the following results. When $\alpha = 1$, m equals the "half width at half height"—i.e. the value of x that results in the distribution having half its maximum value. For $\alpha = 2$, the variance of the distribution equals $2m^2$. (iv) Since,

Table 1
The range of x, around $x = 0$, where 50% of the normalisation of the function $f(x)$ of Eq. (2) resides can be written as $|x| \le c(\alpha)m$

α	$c(\alpha)$
0.50	1.28
0.75	1.07
1.00	1.00
1.25	0.98
1.50	0.97
1.75	0.96
2.00	0.95

The Table give the results of numerical calculation of $c(\alpha)$.

for $\alpha < 2$, we cannot classify the distributions of Eq. (2) by their variance, we could characterise them by the region around the maximum where an appreciable proportion of mutations lie. The range of x, around $x = 0$, where 50% of the normalisation of $f(x)$ resides can be written $|x| \le c(\alpha)m$. In Table 1, the results of numerical calculation for $c(\alpha)$ are presented.

Evidently, 50% of mutant effects lie within a range $\sim m$ of $x = 0$, again pointing to m as a useful indicator of the scale of x over which $f(x)$ changes appreciably.

A common choice made in the literature for the distribution of mutant effects is a Gaussian (corresponding, in Eq. (2), to $\alpha = 2$). The variance of the Gaussian is an empirically determined quantity, which varies from trait to trait and from species to species. A typical value of the variance is $2m^2 \sim 0.05$, i.e. $m \sim 0.2$ [15] and although we do not study a Gaussian distribution here, we shall use the value $m = 0.2$ for all of the numerical studies presented below.

3. Neutral case

A long-tailed distribution of mutant effects manifests itself most strongly in a situation where mutation is the only evolutionary force acting on a population. We therefore, consider a large (effectively infinite) population of haploid asexual organisms with one locus, where there is no genetic drift and individuals are not subject to selection. The lifecycle of the population, that takes place in discrete generations, begins with newly born individuals (juveniles) and is: (i) maturation of juveniles to adulthood, but with no selection operating, (ii) production of offspring by adults, followed shortly by the death of all adults. Mutation is taken to occur during the production of offspring.

We adopt the somewhat flexible model of mutation given by the regression model described above, and summarised in Eq. (1).

3.1. Equilibrium distribution of an infinite population

The distribution (probability density) of genotypic effects in one generation is denoted by $\varphi(x)$ and in the following generation, is denoted by $\varphi'(x)$. It satisfies

$$\varphi'(x) = (1 - u)\varphi(x) + u \int f(x - \gamma y)\varphi(y)\,\mathrm{d}y \qquad (6)$$

where the trait mutation rate is u. Eq. (6) follows directly from considerations of genes being perfectly transmitted to the next generation with probability $1 - u$ (first term on the right-hand-side) and being imperfectly transmitted—i.e. containing a mutation, with probability u (second term on the right-hand-side).

Let us define moments of $\varphi(x)$ by $\overline{x^n} = \int x^n \varphi(x)\,\mathrm{d}x$, and use a prime to denote the corresponding moment in the following generation. For any value of $\alpha < 2$, long-tailed distributions do not yield a dynamical equation for $\overline{x^2}$ that is meaningful. To see this, multiply Eq. (6) by x^2 and integrate over all x. This involves the quantity $\int \int x^2 f(x - \gamma y)\varphi(y)\,\mathrm{d}x\,\mathrm{d}y$ and this can be shown to diverge because $\int x^2 f(x)\,\mathrm{d}x$ diverges. By contrast, a meaningful equation for $\bar{x}$ can be defined for $\alpha > 1$: namely $\bar{x}' = [1 - u(1 - \gamma)]\bar{x}$. This equation coincides with previously found results [10] and $\bar{x}$ approaches an equilibrium value of 0 when $\gamma < 1$. However, the equation for $\bar{x}$ is generally not meaningful when the shape parameter, α, appearing in the distribution of mutant effects, Eq. (2), lies in the range $\alpha \leq 1$. The problem arises because we cannot give a definite value to $\int x f(x)\,\mathrm{d}x$ when $\alpha \leq 1$. It may be tempting, because of symmetry of $f(\bullet)$, to take $\int x f(x)\,\mathrm{d}x = 0$ but if $f(\bullet)$ is long-tailed, and possesses a non-zero level of asymmetry, then for $\alpha \leq 1$, $\int x f(x)\,\mathrm{d}x$ could actually diverge.

Despite the divergence of some or all of the moments arising from Eq. (6), we have established that in a continuous time approximation to Eq. (6), the distribution $\varphi(x)$ equilibrates for any $\gamma < 1$ (see Appendix C). The equilibrium distribution is independent of the trait mutation rate u—although the time taken to achieve equilibrium does depend on u. In Appendix C, it is shown that the exact equilibrium solution to Eq. (6), for the
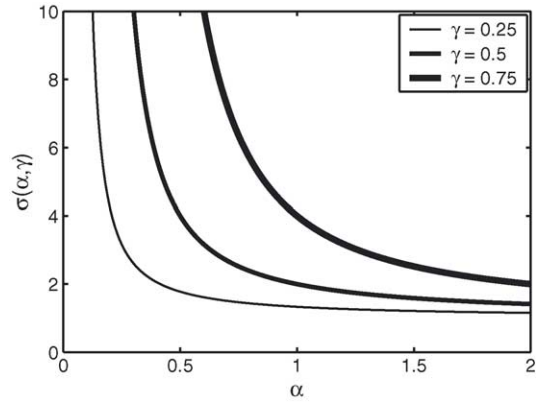

Fig. 2. The parameter $\sigma(\alpha, \gamma)$ of Eq. (8) is plotted as a function of the "shape" parameter α. The quantity $\sigma(\alpha, \gamma) \times m$ is a measure of the width of the equilibrium distribution of genotypic effects in the selectively neutral case.

family of $f(\bullet)$ given in Eq. (2), is

$$\hat{\varphi}(x) = \int \mathrm{e}^{ikx} \exp\left(-\frac{m^\alpha}{1 - \gamma^\alpha}|k|^\alpha\right)\frac{\mathrm{d}k}{2\pi}. \qquad (7)$$

Thus, the mutation distributions of Eq. (2) lead, at equilibrium, to a distribution of genotypic effects with the same shape parameter α, but with a "scale" parameter changed from m, in the mutation distribution, to $\sigma(\alpha, \gamma)m$ in the distribution of genotypic effects, where

$$\sigma(\alpha, \gamma) = \frac{1}{(1 - \gamma^\alpha)^{1/\alpha}}. \qquad (8)$$

Note that for any $\gamma > 0$ and any $\alpha > 0$, we have $\sigma(\alpha, \gamma) > 1$, thus, in accordance with intuition, the model of mutation adopted leads to an equilibrium distribution of genotypic effects that is *broader* than the distribution of mutant effects. Furthermore, because the equilibrium distribution of genotypic effects is characterised by the same shape parameter, it has the same asymptotic power-law behaviour as the distribution of mutant effects, Eq. (3). Fig. 2 contains a plot of the "width" of the equilibrium distribution of allelic effects, $\sigma(\alpha, \gamma)$, as a function of α, for several values of γ.

It is possible to include mutational bias in the selectively neutral case of this Section. A distribution of mutant effects of the form $f(x - \gamma y - b)$ with non-zero bias parameter, b, leads, using the method of Appendix C, to an equilibrium distribution of genotypes effects

given by $\hat{\varphi}(x - b/(1 - \gamma))$ where $\hat{\varphi}(x)$ is the distribution of an unbiased problem given in Eq. (7). The modified argument of $\hat{\varphi}(x)$ simply corresponds to a shift in position of the maximum of the distribution from $x = 0$ to $x = b/(1 - \gamma)$.

3.2. Equilibrium distribution of a finite population

The infinite population results illustrate some consequences of a long-tailed distribution of mutant effects. However, the resulting distribution of genotypic effects does have an infinite variance. This means the infinite population results cannot be directly used to make predictions for large but finite populations, in contrast to the results following from a short-tailed distributions of mutant effects. To understand the implications of long-tailed distributions in finite populations, we have investigated the behaviour of a large but finite population, from individually based numerical simulations. For the population size and mutation rates considered, the speed of the underlying dynamics are largely determined by the mutation rate. The population was thus simulated over a time interval that is large compared with the inverse of the mutation rate, u^{-1}. Over such a time interval the population approaches a highly stochastic "equilibrium" state associated with a finite population. This approach is illustrated in Fig. 3.

Although the population approaches a peaked distribution at long times, there are always a small fraction of outliers, as exemplified by the line of bins that rapidly develop at the extremities of the x range (at $x = \pm 5$). Increasing the range of x values, at fixed bin widths, is observed to not change the existence of outliers, but merely cause a reduction in their frequency.

4. Inclusion of selection

We now investigate the implications of a non-zero level of selection acting in the model considered above. Selection acts on viability, which depends on a single phenotypic trait characterising individuals and whose genetic component is determined by the allele at the single locus in question. Fitness is taken to be of a stabilising viability type. The viability, $w(x)$, of individuals of a particular trait value x arises from an average, over environmental effects, of viability as a function of
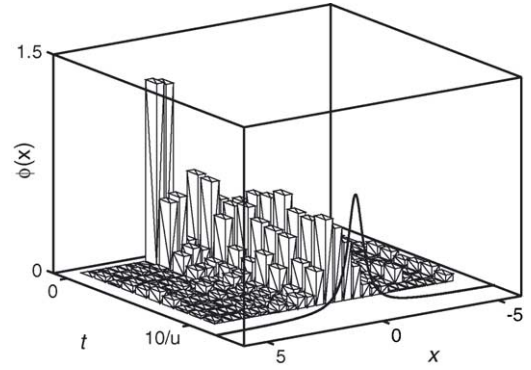

Fig. 3. A histogram of frequencies of different genotypic effects is plotted as a function of genotypic effect, x, and time t (measured in generations). This figure shows the behaviour of the distribution of genotypic effects, $\varphi(x)$, over a time interval that is long compared with the inverse of the trait mutation rate, u^{-1}, when there is no selection acting. The histogram was determined from a numerical simulation of a finite population of individuals and is normalised, at each time, so the area under the histogram is unity. Also plotted on the same figure is the distribution describing an infinite population at equilibrium (solid curve). The parameter values used were $u = 10^{-2}$, $m = 0.2$, $\gamma = 0.5$, $N = 10^4$ and the x values, at time $t = 0$, were drawn from a Gaussian distribution, with mean 2 and variance of 0.1.

phenotype (see, e.g. [5]). We take [16]

$$w(x) = \exp(-sx^2) \qquad (9)$$

where s is a positive parameter that is a measure of the intensity of selection and the optimal allelic value—the one leading to the maximum of $w(x)$—has been taken to lie at $x = 0$. Selection acts during the maturation stage (i) of the lifecycle given in Section 3.

The dynamical equation describing the evolution of the distribution of genotypic effects follows from Eq. (6) by replacing $\varphi(x)$ on the right-hand-side of this equation by the distribution after selection: $w(x)\varphi(x)/\int w(x)\varphi(x)\,\mathrm{d}x$, hence

$$\varphi'(x) = \frac{(1 - u)w(x)\varphi(x) + u\int f(x - \gamma y)w(y)\varphi(y)\,\mathrm{d}y}{\int w(x)\varphi(x)\,\mathrm{d}x}. \qquad (10)$$

We have investigated the behaviour of a finite population, from individually based numerical simulations and the results, at long times, are illustrated in Fig. 4 for the case $\alpha = 1$.

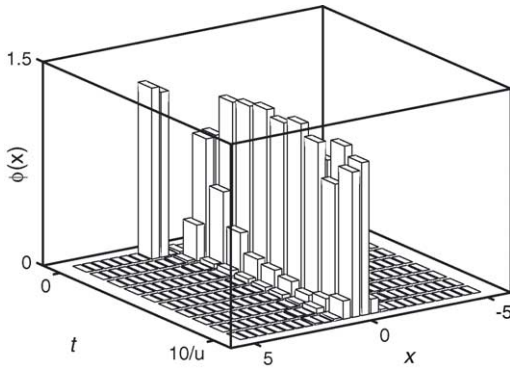The distribution is observed to be far more narrowly peaked than the corresponding distribution in the neu-

Fig. 4. A histogram of frequencies of different genotypic effects is plotted as a function of genotypic effect, x, and time t (measured in generations). This figure shows the behaviour of the distribution of genotypic effects, $\varphi(x)$, over a time interval that is long compared with the inverse of the trait mutation rate, u^{-1}, when selection is acting on the population. The histogram was determined from a numerical simulation of a finite population of individuals and is normalised, at each time, so the area under the histogram is unity. The parameter values used were $u = 10^{-2}, m = 0.2, \gamma = 0.5, s = 0.025, N = 10^4$ and the x values, at time $t = 0$, were drawn from a Gaussian distribution, with mean 2 and variance of 0.1.

tral case ($s = 0$) that was considered in Section 3.1 and illustrated in Fig. 4. As a consequence of the narrowness of the distribution, it is hard to see any evidence, at the extremes of the range of allelic value, of the long tails of the distribution of mutant effects, in contrast to what was seen in the neutral case, when the population was finite.

Let us analyse the properties of the equilibrium solution arising from Eq. (10). The equilibrium solution obeys

$$\hat{\varphi}(x) = u \frac{\int f(x - \gamma y) w(y) \hat{\varphi}(y) \, dy}{\bar{w} - (1 - u) w(x)},$$

$$\bar{w} = \int w(x) \hat{\varphi}(x) \, dx. \tag{11}$$

We can estimate the behaviour of the solution of Eq. (11), when mutation is weak compared with selection, as characterised by $u \Gamma(1/\alpha)/(\pi \alpha m^2) \ll s$. When this applies, the House of Cards approximation [17] yields $\hat{\varphi}(0) \propto \alpha m s/(u \Gamma(1/\alpha))$, i.e. in this case, $\hat{\varphi}(0)$ is very large, and the bulk of the normalisation of $\hat{\varphi}(x)$ lies very close to $x = 0$. Thus, in contrast to the neutral case, the shape of the equilibrium distribution of allelic effects, $\hat{\varphi}(x)$, in the presence of selection, does not gen-

erally take a similar profile to $f(x)$. The behaviour in the opposite limit, where selection is weak compared with mutation, is much harder to estimate, however, we can note that in this case the form of $\hat{\varphi}(x)$ will not be approximately Gaussian, in contrast to what is found when $f(\bullet)$ has a finite second moment [1,18].

Apart from the approximate local behaviour of $\hat{\varphi}(x)$, we can determine its leading asymptotic behaviour. To obtain this, note that when $|x| \gg 1/\sqrt{s}$, we can neglect $w(x)$ appearing in the denominator on the right-hand-side of Eq. (11). We shall assume $1/\sqrt{s} \gg m$, as is often typical in quantitative traits [15]. Then, when $|x| \gg 1/\sqrt{s}$, we can replace $f(x - \gamma y)$ by its asymptotic form, Eq. (3). It follows that $\hat{\varphi}(x) \simeq u A_\alpha \int |x - \gamma y|^{-\alpha-1} w(y) \hat{\varphi}(y) \, dy / \bar{w} \simeq u A_\alpha |x|^{-\alpha-1} \int w(y) \hat{\varphi}(y) \, dy / \bar{w}$, i.e.

$$\hat{\varphi}(x) \sim u A_\alpha |x|^{-\alpha-1} \equiv u \times \text{asymptotic form of } f(x).$$

$$\tag{12}$$

The general result of Eq. (12), for the tails of the distribution of $\hat{\varphi}(x)$, has two implications. (i) At large $|x|$, the distribution $\hat{\varphi}(x)$ approaches 0 with an identical power-law behaviour to that of $f(x)$, and since this large $|x|$ behaviour is responsible for the infinite variance of $f(x)$, it follows that $\hat{\varphi}(x)$ is, itself, a long-tailed distribution and, as an automatic consequence, has an infinite variance. (ii) The strength of the long tails, i.e. the coefficient, u, of $f(x)$ in Eq. (12), can be small. Typical mutation rates in asexual populations, where u represents the mutation rate of the trait, can be of order 10^{-2} (while in sexual populations, where u corresponds to the allelic mutation rate, u can be of order 10^{-5}) [15]. Thus, while the long tails are present in $\hat{\varphi}(x)$, they are present with low weight, since selection rapidly removes such extreme individuals. A value of x, that is randomly picked from the distribution $\hat{\varphi}(x)$, will lie in the region where Eq. (12) is applicable, with a probability that is substantially smaller than u. As a consequence, it requires a population of size much larger than u^{-1} in order to have an appreciable number of individuals in the tails of the distribution and corresponding to outliers of the distribution. This is the reason the result of the long time simulations illustrated in Fig. 4 do not exhibit any significant outlier effects associated with the infinite variance of the distribution.

Let us note that in the case where selection is weak, in the sense $s \ll 1$, we cannot generally make the

quadratic approximation to the fitness function of Eq. (9), that is commonly made in the literature, namely $w(x) \simeq 1 - sx^2$. The reason is that this approximation leads, in Eq. (11), to the asymptotic form, $\phi(x) \sim$ constant $\times |x|^{-\alpha-3}$, i.e. a result that corresponds to a *finite* second moment and is thus incompatible with the exact leading behaviour of Eq. (12), which yields an infinite second moment.

We end this Section by noting that census, in Eq. (10), is taken to occur after mutation has occurred, but before selection. If we censused immediately after selection, then the distribution describing the population is $w(x)\varphi(x)/\int w(x)\varphi(x)\, dx$ and, because of the factor $w(x)$, this is a short tailed distribution. Thus, every generation, the distribution that describes the population alternates between being long- and short-tailed and here we have concentrated on the anomalous, long-tailed aspect, of this.

5. Summary and discussion

Distributions of mutant effects with long tails have, as we have pointed out, infinite variances. We have indicated how a long-tailed distribution of mutant effects determines the asymptotic properties of the equilibrium distribution of allelic effects describing an infinite population, and how this distribution will also have long tails—and hence an infinite variance associated with it. In the case of a finite population, it has been seen that when stabilising selection is operating, extreme individuals are rapidly removed by section and the tails of the distribution will not be significantly populated, and hence not significantly contribute to variance, unless (population size)$\times u$ is not small compared with unity.

Apart from the influence that long-tailed distributions of mutant effects have on the equilibrium distribution of allelic effects, there may also be dynamical implications of such mutation distributions. In particular, it is interesting to ask whether long-tailed distributions of mutant effects have a substantial effect on the *rate of adaptation* of evolving populations. The nature of the fitness landscape is likely to affect the answer to this question—whether the landscape is rugged or smooth. Here, we shall only give the results of a very preliminary investigation, for the smooth, stabilising fitness function of Eq. (9).
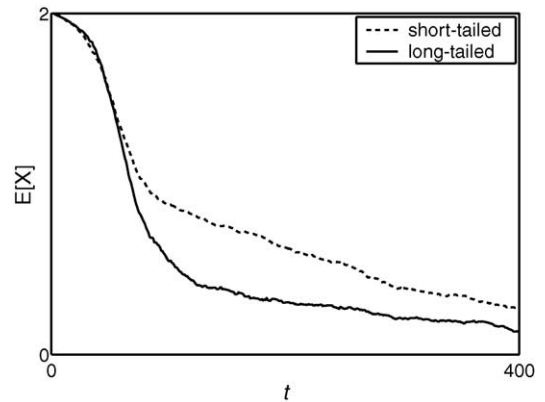

Fig. 5. A plot of the mean genotypic effect, $E[X]$, as a function of time, for two populations that were subject to stabilising selection of identical strength. Both populations had, initially, the same distribution of allelic effects. The populations differed in the distribution of mutant effects; one had a short tailed (Gaussian) distribution while the other had a long-tailed (Cauchy) distribution. The figure was obtained from individual based simulations of the two populations. The parameter values used were $u = 10^{-3}$, $m = 0.2$, $\gamma = 0.5$, $s = 0.025$ and $N = 10^4$. The different rates of adaptation exhibited in the figure are due to the different values of the shape parameter, α, of the two mutation distributions, and e.g. replacing m in just the short-tailed distribution by 0.8 or 1.2 m does not qualitatively affect the slower rate of adaptation associated with this distribution.

In Fig. 5, we plot the mean genotypic effect, $E[X]$, as a function of time, t, for simulations of two populations. Initially, both populations were identically distributed, but one has a short-tailed distribution of mutant effects, while the other has a long-tailed distribution. The populations have an initial mean genotypic effect that is some distance from the fitness optimum (which lies at $x = 0$) and are thus maladapted. The figure indicates that the range of the tails of the distribution of mutant effects *can* have a significant effect on the rate of evolutionary adaptation of a population.

We can therefore conclude that the length of tails of the distribution of mutant effects can affect both equilibrium and dynamical properties of a population, but the latter requires further investigation.

Appendix A

In this Appendix we give a number of ways of viewing probability distributions that possess so-called long tails.

Firstly, if we repeatedly draw random numbers (i.e. x's) from such a distribution, then while many of the numbers will be close to the mode of the distribution, some of the numbers will have substantial deviations from the mode. This is in sharp contrast to a Gaussian distribution, or any other distribution with a finite variance, where virtually none of the random numbers picked from this distribution deviate from its mode by more than a few standard deviations.

Another way of looking at long-tailed distributions is to consider the statistics of the sample mean, $S_n = \sum_{i=1}^{n} x_i/n$ of a finite number of x's that are independently drawn from such a distribution. If, for increasingly larger sample sizes, n, the distributions of the S_n, are compared, then long-tailed distributions result in anomalous behaviour, compared with short-tailed (i.e. finite variance) distributions. In particular, the sample means of a long-tailed distribution may not settle down (converge) to a definite value as the sample size, n, is increased. Similarly when the variances of the S_n

of random variables from a long-tailed distribution may have a distribution that approaches a limiting form [6].

Appendix B

In this Appendix we give four representative forms for the function $f(x)$ of Eq. (2) of the main text. The cases $\alpha = 1$ and 2 lead to straightforward integrations. We have found that writing $f(x)$ of Eq. (2) in the equivalent form $f(x) = \int_0^\infty \cos(kX) \exp(-k^\alpha)\, dk/(m\pi)$ with $X = x/m$, leads to integrals which for some values of α may be evaluated with the computer algebra package Maple®. These values include $\alpha = 1/2, 4/3, 3/2, 5/3, \dots$ although some of the resulting expressions are long. The values $\alpha = 1/2$ and $4/3$ are relatively simple and are given below. With $C(z)$ and $S(z)$ Fresnel integrals and $_2F_2(a, b; c, d; z)$ generalised hypergeometric functions [14] and with

$$X = |x|/m, \qquad F(X) = mf(x)$$

we have

$$F(X) = \begin{cases} \dfrac{4}{\sqrt{2\pi}}\left(\dfrac{1}{4X}\right)^{3/2}\cos\left(\dfrac{1}{4X}\right)\left[1 - 2C\left(\sqrt{\dfrac{2}{\pi}\dfrac{1}{4X}}\right)\right] + \dfrac{4}{\sqrt{2\pi}}\left(\dfrac{1}{4X}\right)^{3/2}\sin\left(\dfrac{1}{4X}\right)\left[1 - 2S\left(\sqrt{\dfrac{2}{\pi}\dfrac{1}{4X}}\right)\right], & \alpha = 1/2, \\[12pt] \dfrac{1}{\pi}\dfrac{1}{1+X^2}, & \alpha = 1, \\[12pt] \dfrac{3\Gamma(3/4)}{4}\,_2F_2\left(\dfrac{7}{12}, \dfrac{11}{12}; \dfrac{1}{2}, \dfrac{3}{4}; \dfrac{27X^4}{256}\right) - \dfrac{15}{128}\dfrac{\sqrt{2\pi}X^2}{\Gamma(3/4)}\,_2F_2\left(\dfrac{13}{12}, \dfrac{17}{12}; \dfrac{5}{4}, \dfrac{3}{2}; \dfrac{27X^4}{256}\right), & \alpha = 4/3, \\[12pt] \dfrac{1}{\sqrt{4\pi}}\exp\left[-\dfrac{X^2}{4}\right], & \alpha = 2. \end{cases}$$

are compared, for increasing values of n, they will not generally decrease with sample size n, as $1/n$, as they do for a short tailed distribution.

A third feature of long tailed distributions is that a sum of independently chosen random variables from such a distribution will not approach a normal distribution as the number of terms in the sum is increased. Thus, the Central Limit Theorem (i.e. the asymptotic approach of the distribution of a sum of random variables to a Gaussian) breaks down for such distributions. There are, however, extensions of the Central Limit Theorem which indicate the way suitably scaled sums

Appendix C

In this Appendix, we provide some theoretical results that are used in the main text.

We begin with Eq. (6) which we Fourier transform, by multiplying by e^{ikx} and integrating over all x. With $\psi(k) = \int e^{ikx}\varphi(x)\, dx$ the characteristic function associated with $\varphi(x)$, we find

$$\psi'(k) = (1 - u)\psi(k) + ug(k)\psi(\gamma k) \tag{13}$$

where $g(k) = \int e^{ikx} f(x)\, dx$.

Eq. (13) has a unique equilibrium solution, $\hat{\psi}(k)$, which satisfies $\hat{\psi}(k) = g(k)\hat{\psi}(\gamma k)$. It is natural to solve this equation by iteration. The first few iterates, starting from an arbitrary initial characteristic function, $\psi_0(k)$, are: $\psi_1(k) = g(k)\psi_0(\gamma k)$, $\psi_2(k) = g(k)\psi_1(\gamma k) = g(k)g(\gamma k)\psi_0(\gamma^2 k)$. Continuing in this way, we obtain $\psi_N(k) = [\prod_{n=0}^{N} g(\gamma^n k)]\psi_0(\gamma^N k)$. Since, for all $\gamma < 1$, $\lim_{N\to\infty} \psi_0(\gamma^N k) = \psi_0(0) = 1$ by virtue of $\psi_0(k)$ being a characteristic function, it follows that $\hat{\psi}(k) = \lim_{N\to\infty} \psi_N(k) = \prod_{n=0}^{\infty} g(\gamma^n k)$. Thus, the equilibrium distribution of genotypic effects, $\hat{\varphi}(x)$, is thus given by

$$\hat{\varphi}(x) = \int e^{ikx}\psi(k)\frac{dk}{2\pi} = \int e^{ikx}\left[\prod_{n=0}^{\infty} g(\gamma^n k)\right]\frac{dk}{2\pi}.$$
(14)

Using $g(k)$ of Eq. (2) in this equation quickly leads to Eq. (7) of the main text.

To establish convergence of Eq. (13) to the equilibrium solution, we approximate this equation by a continuous time equation: $\partial\psi(k,t)/\partial t = -u\psi(k,t) + ug(k)\psi(\gamma k,t)$. Setting $T = ut$, $\psi(k,t) = e^{-T}\chi(k,T)$ we find $\chi(k,T)$ obeys $\partial\chi(k,T)/\partial T = g(k)\chi(\gamma k,T)$ with solution $\chi(k,T) = \chi(k,0) + g(k)\int_0^T \chi(\gamma k,s_1)\,ds_1$. Iterating yields $\chi(k,T) = \chi(k,0) + g(k)T\chi(\gamma k,0) + \cdots$. Proceeding in this way, we infer that $\chi(k,T) = \sum_{j=0}^{\infty} C_j(k)(T^j/j!)\chi(\gamma^j k,0)$ where $C_0(k) = 1$ and for $j > 0$, $C_j(k) = \prod_{h=0}^{j-1} g(\gamma^h k)$. We can thus write the solution for $\psi(k,t)$ as $\psi(k,t) = e^{-ut}\sum_{j=0}^{\infty} C_j(k)[(ut)^j/j!]\psi(\gamma^j k,0)$. and this may be verified to solve Eq. (13). Since $\lim_{j\to\infty}\psi(\gamma^j k,0) = 1$ and since $\left|\prod_{h=0}^{\infty} g(\gamma^h k)\right| < 1$ we have a convergent series solution for $\psi(k,t)$. For large t the series is dominated by large j, in which case $\psi(k,t) \to C_\infty(k)e^{-ut}\sum_{j=0}^{\infty}(ut)^j/j! = C_\infty(k) \equiv \hat{\psi}(k)$ and we have convergence to the equilibrium solution.

References

[1] M. Kimura, A stochastic model concerning the maintenance of genetic variability in quantitative characters, Proc. Natl. Acad. Sci. U.S.A. 54 (1965) 731–736.
[2] J.F. Crow, M. Kimura, The theory of genetic loads, Proceedings of the XIth International Congress of Genetics, vol. 2, 1964, pp. 495–505.
[3] J. Welch, D. Waxman, Non-equivalent loci and the distribution of mutant effects, Genetics 161 (2002) 897–904.
[4] D. Waxman, J. Welch, Non-equivalent Loci and Mutation-Selection Balance, Theor. Popul. Biol. 63 (2003) 339–345.
[5] M.G. Bulmer, Maintenance of genetic variability by mutation-selection balance: a child's guide through the jungle, Genome 31 (1989) 761–767.
[6] B.V. Gnedenko, A.N. Kolmogoroff, Limit Distributions for Sums of Independent Random Variables, Addison-Wesley, Reading, MA, 1968.
[7] G. Rangarajan, M. Ding (Eds.), Processes with Long Range Correlations: Theory and Applications, Springer Lecture Notes in Physics 621, Springer Verlag, New York, 2003
[8] J. Beran, Statistics for long-memory processes, Chapman and Hall/CRC Press, London, 1994.
[9] J.F.C. Kingman, A simple model for the balance between selection and mutation, J. App. Prob. 15 (1978) 1–12.
[10] Z.B. Zeng, C.C. Cockerham, Mutation models and quantitative genetic variation, Genetics 133 (1993) 729–736.
[11] D. Waxman, J.R. Peck, The anomalous effects of biased mutation, Genetics 164 (2003) 1615–1626.
[12] D. Waxman, J.R. Peck, A one locus, biased mutation model and its equivalence to an unbiased model, BioSystems 78 (2004) 93–98.
[13] P. Lévy, Calcul des Probabilités, Gauthier Villars, Paris, 1925.
[14] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions, Dover, New York, 1965.
[15] M. Lynch, B. Walsh, Genetics and Analysis of Quantitative Traits, Sinauer, Sunderland, MA, 1998.
[16] J.B.S. Haldane, The measurement of natural selection, in: G. Montalenti, A. Chiarugi (Eds.), Proceedings of the 9th International Congress of Genetics, Part 1, Bellagio, Italy, August 1953, 1954, pp. 480–487.
[17] M. Turelli, Heritable genetic variation via mutation selection balance: Lerche's zeta meets the abdominal bristle, Theor. Popul. Biol. 25 (1984) 138–193.
[18] R. Lande, The maintenance of genetic variability by mutation in a polygenic character with linked loci, Genet. Res. 26 (1976) 221–235.