Opinion piece

On measuring selection in experimental evolution

Distributions of mutation fitness effects from evolution experiments are available in an increasing number of species, opening the way for a vast array of applications in evolutionary biology. However, comparison of estimated distributions among studies is hampered by inconsistencies in the definitions of fitness effects and selection coefficients. In particular, the use of ratios of Malthusian growth rates as ‘relative fitnesses’ leads to wrong inference of the strength of selection. Scaling Malthusian fitness by the generation time may help overcome this shortcoming, and allow accurate comparison of selection coefficients across species. For species reproducing by binary fission (neglecting cellular death), ln2 can be used as a correction factor, but in general, the growth rate and generation time of the wild-type should be provided in studies reporting distribution of mutation fitness effects. I also discuss how density and frequency dependence of population growth affect selection and its measurement in evolution experiments.

Keywords: mutation fitness effects; experimental evolution; population growth; density dependence; frequency dependence

1. INTRODUCTION

The aim of this opinion piece is to clarify the definition and measurement of selection coefficients in evolution experiments investigating fitness effects of mutations. Such studies, from the accumulation of deleterious mutations under relaxed selection over many generations [1–3] to the identification of spontaneous or induced mutations as they arise [4–7], have generated considerable interest recently (reviewed in [8,9]), notably because of their important bearings on many questions across evolutionary genetics. First, they allow testing theoretical predictions about the genetics of adaptation, such as the size of beneficial mutations that get fixed in a population, or the cost of complexity for adaptation (reviewed in [10]). Second, they uncover the intensity of deleterious mutation effects, enabling quantitative predictions to be made about, e.g. the evolution of recombination [11], of selfing [12], or the extinction of small populations [13]. Third, they allow us to address fitness interactions between genes (epistasis) and the ruggedness of fitness landscapes [14–16], which determines the very nature of the evolutionary process from a genetic standpoint. Here I argue that one of the most popular measurements of fitness in evolution experiments (the ratio of Malthusian parameters) does not have a clear evolutionary meaning, and should be replaced by a measurement relating more directly to the dynamics of frequency change of mutations.

2. FITNESS AND SELECTION

The primary goal of measuring fitness effects of mutations is to relate them to the evolutionary dynamics of alleles under selection. This is quantified by the selection coefficient \( s \). Consider an asexual (haploid) population consisting of two genotypes, a mutant and a wild-type, with population sizes (or density) \( M \) and \( N \), and frequencies \( p \) and \( (1 - p) \). With continuous growth and no age structure, the selection coefficient can be defined as

\[
s = \frac{d}{dt} \ln \left( \frac{p}{1 - p} \right)
\]  

[17], which has units of time\(^{-1}\). The frequency of the mutant increases if \( s > 0 \) and decreases if \( s < 0 \), at a speed determined by \( s \). Since the ratio of allelic frequencies is also the ratio of numbers (or densities) of each genotype, we may also write

\[
s = \frac{d \ln M}{dt} - \frac{d \ln N}{dt}.
\]  

In particular, if selection is density independent and there are no interactions between genotypes, then

\[
s = r_m - r_w,
\]  

where \( r \) is the Malthusian parameter [17] or intrinsic rate of increase of each genotype (\( m \), mutant; \( w \), wild-type).

In practice, \( r \) is estimated as the regression slope of log-population size against time in the exponential phase, that is, at low population density (assuming no Allee effects). Note that selection can be density independent even if population growth is density dependent, but this implies that the genotype which initially grows faster also has a higher carrying capacity (figure 1a, first row; electronic supplementary material, appendix; see also [18,19]). When this is valid, \( s \) is constant and from equation (2.1), the ratio of genotypic frequencies increases exponentially in time (figure 1a, second row), resulting in a logistic trajectory of the mutant frequency in time. Moreover, in this case, equation (2.3) holds whether each \( r \) is measured from cultures of each genotype in isolation, or in a competition experiment mixing both genotypes. However, when in competition, more statistical significance can be obtained by estimating \( s \) directly from equation (2.1) [20].

If selection is density dependent, the selection coefficient changes with population size (figure 1b), so that measuring the growth rate in the exponential phase is not sufficient to estimate \( s \). Similarly, with genotype-by-genotype interactions in fitness, \( s \) changes with the relative frequencies of each genotype (figure 1c).

Disentangling these two possible causes of changes in \( s \) (density dependence and frequency dependence) requires combining a competition experiment with cultures of each genotype in isolation (electronic supplementary material, appendix).

Most population genetic models are formulated in discrete non-overlapping generations rather than...
Figure 1. Selection and demography in isolation versus competition. The population sizes of two genotypes grown in isolation (first row: dark grey, wild-type; light grey, mutant) or in competition (second row: dark grey area, wild-type; light grey area, mutant) are shown together with the ratio of genotypic frequencies in competition (third row: note the logarithmic scale on the y-axis) for three demographic scenarios. Scenario (a) leads to frequency- and density-independent selection. Scenario (b) illustrates frequency-independent but density-dependent selection, in the particular case where both genotypes have the same carrying capacity $K$. In this case, $s$ tends to 0 as the population approaches the carrying capacity. In scenario (c), selection is density-independent but frequency-dependent. Recursions were made from the discrete-generation model in equation (A2) from the electronic supplementary material, appendix, with logistic population growth, using $\lambda_w = 1.08$, $\lambda_m = 1.09$, $K_w = 100\,000$ in all scenarios, and $K_m = K_w \ln(\lambda_m)/\ln(\lambda_w)$ and $t_w = t_m = 1$ in (a); $K_m = K_w$ and $t_w = t_m = 1$ in (b); $K_m = K_w \ln(\lambda_m/\lambda_w)$, $t_w = 1$ and $t_m = 0.98$ in (c).

In continuous ones as above, and in this case, an equivalent to equation (2.1) is

$$s_T = \ln \left( \frac{p'/\ln(1-p')}{p'/\ln(1-p)} \right) = \ln \left( \frac{M'/N'}{M/N} \right),$$

with primes denoting values in the next generation, and the subscript $T$ indicating that the evolutionary change is taken over a generation. In this case, if selection is density independent and there are no genotype-by-genotype interactions, then

$$s_T = \ln \left( \frac{\lambda_m}{\lambda_w} \right) \approx \frac{\lambda_m}{\lambda_w} - 1,$$

with $\lambda_m$ and $\lambda_w$ the density-independent components of the per-generation multiplicative growth rate of each genotype (electronic supplementary material, appendix). With overlapping (but discrete) generations, such that the population is structured by ages, or if life stages can be identified (such as larva, juvenile, adult), equation (2.5) is also valid for asexuals (and can be used as a weak-selection approximation for sexuals) if $\lambda$ is the leading eigenvalue of the projection matrix of transition rates (juvenile/adult survival and fecundity) [21,22, p. 61].

Hence, with discrete generations, the ratio $\lambda_m/\lambda_w$ of the fitness of the mutant to that of the wild-type (or ‘relative fitness’, unitless) determines the evolutionary dynamics. Note that $\lambda$ is necessarily positive (and the population size decreases if $\lambda < 1$), while $r$ can be negative for a decreasing population. Because a negative $r$ cannot be measured starting from a small population, some authors have pooled all mutants with (unmeasured) negative growth rates into the class $r = 0$ (sometimes labelled as ‘lethals’), thus artificially inflating this class [6,7,23].

Although equations (2.3) and (2.5) are classical results in population genetics [22,24], a persistent tradition in experimental evolution has been to measure relative fitness in continuously growing populations as $r_m/r_w$ that is, as a ratio of Malthusian parameters [6,7,23,25–28]. Accordingly, the selection coefficient of a mutation is often inferred from demographic parameters as $s_T = r_m/r_w - 1$, mirroring the definition in discrete-time models from equation (2.5) [7,9,29,30]. However this measure does not directly relate to the evolutionary dynamics of mutations, even for density-independent selection with no interactions between genotypes. Notably, since $s_T = s/\rho_m$, it would predict faster evolution per unit time in a population with the smaller growth rate $r_w$ of the wild-type, while the actual evolutionary dynamics is the same if $s$ is equal. This has important consequences for any application of these measures to questions in

evolutionary biology where the magnitude of fitness effects of mutations is critical [11–13,31].

3. COMPARING SELECTION COEFFICIENTS

The motivation for presenting results in terms of ‘relative Malthusian fitness’ in experimental evolution has been to obtain a dimensionless parameter [20], allowing for comparison between genotypes and species with different generation times. For instance, a selection coefficient measured in h⁻¹ (say, for *Escherichia coli*) is not comparable to one measured in days⁻¹ (say, for *Caenorhabditis elegans*). The important question for comparative purposes is how fast evolution occurs per generation. A dimensionless measurement of selection that allows comparison between studies carried out over different time scales is then

\[ s_T = (r_m - r_w)T, \]  

(3.1)

where \( T \) is the generation time [22, p. 178]. Electronic supplementary material, table S1, shows an overview of some of the relevant literature where fitness is measured in continuous time (that is, not from survival and fecundity per generation). In six out of nine articles, the selection coefficient per unit time \( s \) or per generation \( s_T \), is not provided and cannot be estimated, because only ratios of Malthusian fitness are given, not the fitness of the wild-type \( r_w \). In only one article is \( r_w \) given, so \( s_T \) can be estimated using the generation time reported in another study for the same organism (yeast). In two cases, the selection coefficient that is provided is directly \( s_T \) since it is a difference of Malthusian fitnesses measured as growth rates per generation, not per unit time. Overall, the available data make it difficult to compare selection coefficients across species and studies.

Generation times can also vary among mutants. This may even be the main cause of genetic variation in growth rates (per unit time) for some organisms and life cycles [32]. In this case, \( T \) should be replaced by \( T_{\text{w}} \), thus measuring how fast evolution occurs relative to the generation time of the wild-type (or ‘resident’) genotype [30]. In particular, selection coefficients are often measured for unicellular organisms that reproduce by binary fission. In this case, neglecting cellular deaths, the number of generations \( s_T \), or per generation \( s_T \), is not provided and cannot be estimated, because only ratios of Malthusian fitness are given, not the fitness of the wild-type \( r_w \). In only one article is \( r_w \) given, so \( s_{T,\text{bin}} \) can be estimated using the generation time reported in another study for the same organism (yeast). In two cases, the selection coefficient that is provided is directly \( s_{T,\text{bin}} \) since it is a difference of Malthusian fitnesses measured as growth rates per generation, not per unit time. Overall, the available data make it difficult to compare selection coefficients across species and studies.

Generation times can also vary among mutants. This may even be the main cause of genetic variation in growth rates (per unit time) for some organisms and life cycles [32]. In this case, \( T \) should be replaced by \( T_{\text{w}} \), thus measuring how fast evolution occurs relative to the generation time of the wild-type (or ‘resident’) genotype [30]. In particular, selection coefficients are often measured for unicellular organisms that reproduce by binary fission. In this case, neglecting cellular deaths, the number of generations (division events) of the wild-type per unit time is simply \( r_w/\ln 2 \), and the generation time of the wild-type is \( \ln 2/r_w \). A correct measure of the dynamics of selection per generation for species reproducing by binary fission (neglecting cellular death) is then

\[ s_{T,\text{bin}} = \frac{r_m - r_w}{r_w} \ln 2. \]  

(3.2)

Note that if the focus of the study is fixation probabilities rather than the (deterministic) evolutionary dynamics, then \( T_{\text{w}} \) should be used as a scaling factor instead of \( T_w \) thus measuring the expected lifetime reproductive output of the mutant. This should affect the shape of the distribution of \( s \), since a different scaling factor would be used for each mutant. None of the studies reported in electronic supplementary material, table S1, provided information about variation in generation times among mutants. Besides, comparison of expected fixation probabilities across species based solely on measures of selection coefficients from frequency changes is not to be recommended anyway, since for a given \( s \), fixation probabilities can differ widely depending on how stochasticity affects offspring numbers [33].

4. CONCLUSION

Although distributions of mutation fitness effects are being measured in a growing number of organisms, lack of consistency between estimates still prevents proper comparison of selection coefficients for different species, even when selection is density and frequency independent. In particular, the use of relative Malthusian fitness results in overestimating the per-generation strength of selection (in terms of evolutionary dynamics) by a factor \((\ln 2)^{-1} \approx 1.44\) for species that reproduce by binary fission (neglecting cellular deaths), and by an undetermined factor in other cases. It is worthwhile noting that this bias has no effect on the shape of the distribution and on the coefficient of variation of selection coefficients of mutations originating from a given genotype, an important parameter relating to the effective number of phenotypic traits under selection and to the ‘cost of complexity’ for adaptation [9,34]. Apart from this, the fitness effects of mutations across species should be re-examined in the light of the present argument, with particular attention to comparisons between species that reproduce by binary fission and others.

I thank Thomas Lenormand, Guillaume Martin, Arpat Ozgul, Lilia Perfeito, Gabriel Perron, Olivier Tenaillon and an anonymous reviewer for helpful discussions and comments on an earlier draft. This work was supported by a Newton International Fellowship by the Royal Society.

Luis-Miguel Chevin*

Division of Biology, Imperial College London, Silwood Park, Buckhurst Road, Ascot SL5 7PY, UK

*l.chevin@imperial.ac.uk


APPENDIX: A MODEL OF DENSITY- AND FREQUENCY-DEPENDENT POPULATION GROWTH AND NATURAL SELECTION

In this appendix, I use a haploid model (also valid for diploids in the absence of dominance and age or stage structure) to illustrate how genotype-specific demographic parameters relate to the selection coefficient. I describe demographic processes without explicitly modeling their ecological or physiological base (e.g., resource competition, resistance to toxins, etc). While such mechanistic models may account better for population dynamics in specific cases, the main argument exposed here applies in the more generic situation presented below, chosen for its simplicity. I indicate with a subscript ‘$w$’ variables for the wild-type genotype, and with a subscript ‘$m$’ variables for a mutant genotype.

With overlapping generations and continuous population growth, the dynamics of the number of mutant ($M$) and of the wild-type ($N$) genotypes can be written

$$\frac{d \ln M}{dt} = r_m \left(1 - \frac{f(M + N)}{f(K_m)}\right) + i_m \frac{N}{M + N} \quad (A1a)$$

$$\frac{d \ln N}{dt} = r_w \left(1 - \frac{f(M + N)}{f(K_w)}\right) + i_w \frac{M}{M + N}. \quad (A1b)$$

Setting $N = 0$ in eq. (A1a) and $M = 0$ in eq. (A1b) yields the dynamics when genotypes are grown separately. For each genotype $x$, its intrinsic rates of increase $r_x$ is the density-independent component of its per-capita rate of increase. Dependence of population growth on total population size (or density) is modeled by the function $f$, which depends on the details of population regulation (in particular $f(N) = N$)
corresponds to logistic population growth). Note that $f$ can also be non-monotonic if there are Allee effects. The carrying capacity $K_x$ is the equilibrium population size of genotype $x$ when grown in isolation. The last terms in eqs (A1a,b) represent interactions between genotypes, from competition ($i_m < 0$ and $i_w < 0$), to mutualism ($i_m > 0$ and $i_w > 0$), or exploitation ($i_m i_w < 0$). For each genotype, the probability of interacting is proportional to the frequency of the other genotype.

**Density- and frequency-independent selection**

Let us first consider the case where there are no interactions between genotypes ($i_m = i_w = 0$). From eq. (A1a,b), when $r_m / f(K_m) = r_w / f(K_w)$, the growth rate of the mutant and that of the wild-type genotypes are equally affected by density dependence of population growth. In this case, from eq. (2) in the main text, the selection coefficient of a mutation is $s = r_m - r_w$: it is constant in time, and is not affected by density regulation. It can thus be estimated by measuring the intrinsic rate of increase (growth rate in the exponential phase) of each individual genotype in isolation.

**Density-dependent selection**

Whenever the above condition is not met for equilibrium population sizes of genotypes grown in isolation, selection is density-dependent. In particular, if the carrying capacity is the same for both genotypes ($K_m = K_w = K$), then the selection coefficient of the mutation is $s = (r_m - r_w) [1 - f(M + N) / f(K)]$, which tends to 0 as the total population size (or density) approaches the carrying capacity (Figure 1B). More strikingly, if the carrying capacity is lower for the genotype that grows faster at low density – the classic trade-off between $r$ and $K$ strategies – then the sign of the selection coefficient changes with population density. Therefore, in order to assess whether the intrinsic rate of increase of each genotype measured in isolation can be
used to estimate a selection coefficient, a first test is to compare the equilibrium population sizes reached by each genotype in isolation. If it is larger for the fastest growing genotype, then $r_m - r_w$ should give a fairly accurate measure of selection. In other cases, the demographic dynamics should be measured again across a wider range of densities, to predict how the selection coefficient may change with population density, and whether this change can be neglected or not.

**Frequency- and density-dependent selection**

Finally, in the more general case where interaction(s) between genotypes cannot be neglected, the selection coefficient has an additional term $(i_mN_w - i_wN_m)/(N_w + N_m)$. Selection is thus frequency-dependent in this case, and no accurate measurement of selection coefficients can be performed without competition experiments where the two genotypes are placed in the same medium. This procedure is more complex, since it involves using a mechanism (genetic or phenotypic marker, antibiotic screening) to distinguish between the two genotypes. But even competition experiments may not produce definitive results in the presence of frequency-dependent selection, because the evolutionary dynamics depends on the initial frequency of both genotypes.

Moreover, changes in the evolutionary dynamics caused by frequency dependence may not be easily distinguished from those caused by density dependence, as mentioned above. Combining population dynamics measured for each genotype in isolation to a competition experiment should allow tearing apart the relative roles of these two processes, provided population dynamics is measured over a sufficiently wide range of population sizes.

Note also that some genotypes could affect the demographic parameters of others in an essentially frequency-independent manner. This occurs for instance when one genotype, regardless of its frequency, saturates the medium with a molecule that
is toxic to other genotypes. This limiting case is a special example of niche construction, that is, of environment partly determined by genes, in which different genotypes vary in how adapted they are to this shared environment. This too can be quantified by combining a competition experiment with demographic parameters measured separately for each genotype.

**Discrete generations**

The argument developed above extends to populations with discrete non-overlapping generations. Denoting with primes values in the next generation, the recursion for the number of individuals for each genotype can be written

\[
M' = \lambda_m \frac{1-f(M+N)}{f(K_m)} \frac{N}{t_{m+N}} M
\]

\[
N' = \lambda_w \frac{1-f(M+N)}{f(K_w)} \frac{M}{t_{w+N}} N
\]

where \(\lambda\) is the multiplicative fitness in the absence of density dependence, \(K\) is the carrying capacity, and \(t\) is an interaction term for each genotype. Selection is frequency- and density-independent if \(t_w = t_m = 1\) and \(\frac{1}{f(K_w)} = \frac{1}{f(K_m)}\).
This table is an online supplement to Chevin (Biography Letters) "On measuring selection in experimental evolution." For each parameter I report whether it is provided in each of the reported studies.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Can be measured from the paper</th>
<th>Mean value</th>
<th>Full distribution</th>
<th>Mean value and SD (2007)</th>
<th>Mean and SD (2007) (for a given line)</th>
<th>Mean and SD (2007)</th>
<th>Full distribution + mean value</th>
<th>Mean and SD (2007)</th>
<th>Full distribution + mean value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>No</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>