ELSEVIER

Experimental Gerontology 38 (2003) 373–386

Experimental Gerontology

www.elsevier.com/locate/expgero

Statistical tests of demographic heterogeneity theories

Laurence D. Mueller*, Mark D. Drapeau, Curtis S. Adams[1], Christopher W. Hammerle, Kristy M. Doyal, Ali J. Jazayeri, Tuan Ly, Suhail A. Beguwala, Avi R. Mamidi, Michael R. Rose

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697-2525, USA

Received 15 April 2002; received in revised form 31 October 2002; accepted 31 October 2002

Abstract

In this paper we develop predictions from models of life-long demographic heterogeneity. These predictions are then compared to observations of mortality in large laboratory populations of *Drosophila melanogaster*. We find that the demographic heterogeneity models either require levels of variation that far exceed what would be considered biologically plausible, or they predict a much larger number of very old individuals than we actually observe. We conclude that the demographic heterogeneity models are not reasonable explanations of demographic patterns and are weakly motivated biological models.

Keywords: Fruit fly; Medfly; Negative aging; Longevity; Frailty; Antagonistic pleiotropy; Mutation accumulation; Evolutionary; Stress; Survival; Ageing; Senescence

1. Introduction

1.1. Late-life mortality rate plateaus

It has been known for some time that human mortality rates plateau at advanced ages (e.g. Greenwood and Irwin, 1939; Gavrilov and Gavrilova, 1991). Gompertz himself suggested that his model of mortality would not apply to people aged 60 to 100 years (Gompertz, 1872). Recently, mortality plateaus have been found in a large number of other organisms (Carey et al., 1992; Curtsinger et al., 1992; Vaupel et al., 1998).

Olshansky and Carnes (1997) review the historical explanations for the departures from exponential mortality rate increases at advanced ages. One currently popular explanation for this phenomenon has been lifelong heterogeneity in robustness. Suppose that all individuals die at a rate that follows an exponentially rising probability with age, as given by the Gompertz equation

$$u(x) = A\exp(\alpha x), \quad (1)$$

where $u(x)$ is the mortality rate at age-x, A is the age-independent parameter, and α is the age-dependent parameter. Further, suppose that there is life-long heterogeneity in these mortality functions, such that more robust subgroups survive to later ages, slowing the rate of decline in average survival probabilities at late ages among large cohorts. At advanced ages the remaining individuals from a cohort are expected to be so robust that the mortality rate becomes a very shallow function of age, resembling a plateau. We describe the heterogeneity as lifelong because differences between individuals are in place when adulthood begins and after that time do not change (Vaupel et al., 1979). This is a special type of heterogeneity and should not be confused with simple genetic or environmental variation (Carnes and Olshansky, 2001). More recently demographic models have been proposed that permit mortality rates to stochastically vary continuously with age (Weitz and Fraser, 2001).

What has been lacking is an explicit analysis of lifelong heterogeneity theory. Recently Service (2000a) and Pletcher and Curtsinger (2000) have explored the behavior of several variants of the heterogeneity model. Variables that have been of little interest to demographers (e.g. Vaupel et al., 1979), such as variance in mortality rates, are treated explicitly by Service, Pletcher and Curtsinger, giving us our

* Corresponding author. Tel.: +1-949-824-4744; fax: +1-949-824-2181.
E-mail address: ldmuelle@uci.edu (L.D. Mueller).
[1] Present address: Department of Biology, University of California, Riverside, CA 92521, USA.

0531-5565/03/$ - see front matter © 2002 Published by Elsevier Science Inc.
PII: S0531-5565(02)00238-3

first opportunity to examine heterogeneity theory with care, at least as a theory considered in the abstract.

As a predictive, falsifiable theory, lifelong heterogeneity has already been evaluated elsewhere (e.g. Khazaeli et al., 1998; Drapeau et al., 2000; also see Mueller et al., 2000; Service, 2000b; Arking and Giroux, 2001; Drapeau, 2002). Khazaeli et al. found that populations of fruitflies from highly controlled larval environments that should produce adults with reduced life-long heterogeneity, showed plateaus as frequently as populations lacking these environmental controls. Drapeau et al. (2000) found that genetically robust populations did not experience increases in the age of onset of plateaus as predicted from basic theory. In this article we expand our previous theoretical and experimental work with a more detailed examination of the statistical properties of the heterogeneity models and some experimental tests that follow from these statistical properties.

1.2. Variance in mortality rates

Variances in estimated mortality rates are important for several reasons. We are not going to know the component terms of the heterogeneity model, that is the rates of ageing of individuals, so we have to study its properties indirectly, using variances inter alia. If particular patterns of variation are unique to heterogeneity models they could serve as a means of testing this theory.

Here we dissect the variability that arises in a standard cohort mortality assay. We do this because most experimental evidence testing heterogeneity models comes from large cohorts followed throughout adult life. This design is also used in computer simulations that generate mortality patterns under varied conditions.

For instance, if we start with a genetically variable population (Fig. 1), it is possible that mortality rates will vary due to genetic variation that affects the values of A and α of the Gompertz equation. Even if we select a single genotype and make many copies of it, mortality rates in this genetically homogeneous cohort will vary due to environmental-developmental factors that affect individuals (Fig. 1). We call these developmental factors, since it is presumed that their effects are in place at the time the organism starts to age. As with genetic variation, this environmental variation may presumably affect both A and α.

Proponents of heterogeneity theories suggest that this variation will be present even in controlled laboratory studies of inbred organisms (Fukui et al., 1996). Consequently, as a practical matter it would be impossible under

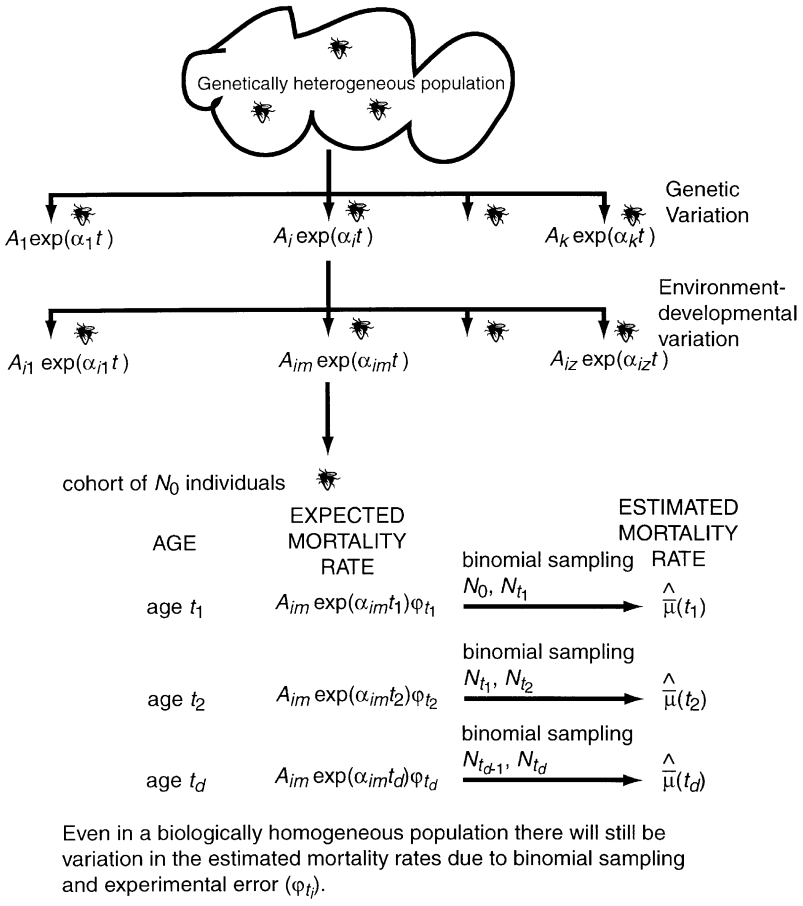

Fig. 1. Sources of variation in mortality rates. In a population where the underlying dynamics of mortality are governed by the Gompertz equation, variability in estimated mortality rates may be due to genetic variation, environmental/developmental factors, experimental error, and sampling variation.

this theory to have an entire population that aged according to a Gompertz model with one set of A and α parameters. But certainly on a computer we can create such a population to trace the other sources of variation in estimated mortality rates, as we have illustrated in Fig. 1.

Suppose we have N_0 individuals that age according to a Gompertz equation with one set of parameters. At various time intervals t_1, t_2, …, t_d, the number of survivors $(N_{t_1}, N_{t_2}, \ldots, N_{t_d})$ are recorded and this information is used to estimate the mean mortality rates $\hat{\boldsymbol{\mu}}(t_i)$ (Fig. 1). The observed number of deaths in any time interval in a real experiment will always be subject to some level of experimental error $(\varphi_{t_1}, \varphi_{t_2}, \ldots, \varphi_{t_d})$ that may increase or decrease the mortality rate from its expected Gompertz value. In *Drosophila* experiments, for instance, adults are transferred to fresh vials at regular intervals and it may be appropriate to consider the experimental errors independent identically-distributed random variables with a mean of 1. However, in experiments where flies are housed in a single cage over the course of the experiment the quality of that environment may degrade over time, due to accumulating wastes on the side of the cage or other factors that change with time. The experimental error terms in this case may depend on time, be autocorrelated, and may not have a mean of 1. Finally, since all experiments rely on finite samples there will be binomial sampling error at each time interval. The mortality estimate is the mean of N_{t_i} observations that are either 0 or 1 (1 = dead, 0 = alive) at each time interval.

Note that the binomial sampling will only hold in the biologically unrealistic case of no genetic or environmental variation. In all other (real) cases the population will be a mixture of binomial distributions corresponding to the different probabilities of death due to different genotypes and environmental-developmental types. The properties of these mixed distributions can depart substantially from the binomial. For instance consider a population with mixed binomial distribution (sometimes called a Lexian distribution) with a mean probability of success equal to p, and a variance of p values equal to θ^2. If we take a sample of size N, the number of successes is expected to be Np, as we would expect from the standard binomial, but the variance will equal $Np(1-p) + N(N-1)\theta^2$ (Johnson and Kotz, 1969, p. 78). Thus, the variance is substantially inflated over the binomial variance unless the variance of p-values, θ^2, is quite small. Thus, attempts to tease apart different components of variance in a genetic analysis of mortality rates should not assume binomial sampling (cf.Shaw et al., 1999).

1.3. Heterogeneity and variation in mortality rates

The best developed model of heterogeneity supposes that individuals vary in their age-independent mortality parameter (Vaupel et al., 1979). Individuals at the onset of adult life vary in frailty. We may suppose that frailty is measured by a random variable z. The probability of death of each individual is then governed by the Gompertz model with parameters α, and $A' = zA$. Of course as individuals age they die and the remaining population would have relatively fewer individuals with very large values of A'. Thus, we expect the distribution of z to change with age. Let $z(x)$ represent the random variable z in a population of individuals aged x time units. Vaupel et al. (1979) showed that if z has a gamma distribution with parameters λ and k and mortality follows the Gompertz equation then $z(x)$ has a gamma distribution with parameters $\lambda(x)$ and k where,

$$\lambda(x) = \lambda + A\alpha^{-1}[\exp(\alpha x) - 1]. \tag{2}$$

In the special case of z with mean 1 and variance σ^2 then the variance of $z(x)$ is given by

$$\mathrm{Var}[z(x)] = \frac{\sigma^2}{\{1 + \sigma^2 A\alpha^{-1}[\exp(\alpha x) - 1]\}^2}. \tag{3}$$

The variance of $u(x)$ is simply $u^2(x)\mathrm{Var}[z(x)]$. It is clear from Eq. (3) that as the population ages (x increases) the variance in frailty, $\mathrm{Var}[z(x)]$, decreases as one might expect. However, the variance of $u(x)$ has a factor of $u^2(x)$ in front of it, which is increasing exponentially with age. Consequently $\mathrm{Var}[u(x)]$ will generally increase with age.

Service (2000a), Fig. 3b) displays a graph showing the relationship between $\mathrm{Var}[\ln(u(x))]$ and age. These results were derived from simulations that included genetic and environmental variability in both A and α, together with sampling error. This curve shows a hump that Service attributes to environmental-developmental heterogeneity, suggesting that an interesting way to test the heterogeneity theory would be to look for these humped curves from replicate measurements of mortality. But, as we will show, the patterns displayed by Service are largely due to binomial (or mixed binomial) sampling variation (Fig. 1) and have little to do with the heterogeneity model.

1.4. Plan of this study

Past work on life-long heterogeneity models has focused on fitting heterogeneity models to demographic observations post hoc rather than developing testable predictions that would enable us to evaluate the validity of heterogeneity theory. In this paper we develop testable predictions based on life long heterogeneity theory and then compare these predictions with data obtained from large replicated cohorts of *Drosophila melanogaster*. We begin by considering the variance in mortality rates in aging cohorts, and find that binomial sampling variance dominates the age-specific pattern. We continue by deriving predictions concerning the number of deaths that will occur at extremely late ages, according to the lifelong heterogeneity model. Actual data from large *Drosophila* cohorts demonstrably fail to match these predictions.

2. Materials and methods

2.1. Computer simulations

We used computer simulations to study the variance in mortality rates due to several different random factors. In our computer simulations both A and α varied. The deaths of individuals in a cohort of 1250 individuals were simulated following Service (2000a). For each individual two random variables, z_1 and z_2, were chosen. z_1 had a lognormal distribution with a mean of -8.57 and variance of 0.766 while z_2 was sampled from a gamma distribution with mean 1 and variance of 0.0021. The mortality rates of that individual were then determined by a Gompertz equation with $A = \exp(z_1)$ and $\alpha = z_2\alpha_0$.

A random time of death for this individual was then generated by the inverse transform method (Fishman, 1996, p. 149) as $\ln(1 - \alpha\ln(1 - U)/A)\alpha$, where U is a uniform random number on (0,1). Gamma random variables were generated from the GKM1 algorithm of Fishman (1996) and the RGS algorithm of Best (1983). Lognormal random variables were generated from the Pascal version of the function *gasdev* (Press et al., 1992, p. 289). In simulations with no genetic variation, simulations for a given cohort size were repeated 100 times and the variance in $\ln(\bar{\mu}(x))$ estimated from these 100 values. In the simulation with genetic variation, we used the same conditions as Service (2000a), 24 cohorts of 1250 individuals each, repeated a total of 10 times.

2.2. Fly populations employed for experimental measurements

All stocks used in these experiments were ultimately derived from a sample of the Amherst, Massachusetts, Ives population that was collected in 1975 and cultured at moderate to large population sizes since (e.g. Ives, 1970). Individual populations have been subjected to a series of selection regimes. Each of the four stocks differs in their age of last reproduction and consists of five outbred replicate populations (Fig. 2). The four stocks are B_{1-5}, O_{1-5}, CO_{1-5}, and ACO_{1-5}. The ACO and B populations have an early age of last reproduction (9 and 14 days respectively), the CO populations have an intermediate last age of reproduction (28 days) and the O populations have a late last age of reproduction (70 days, see Fig. 2). These populations have each been maintained for more than 100 generations at effective population sizes > 1000. Together, these populations define a spectrum of selection on the age of reproduction, and thus a spectrum of patterns for the age-specific force of natural selection. It is probably the case that natural populations are more variable than lab populations and thus could in principle generate more life-long heterogeneity. However, since mortality plateaus are seen in genetically homogeneous lab populations, the populations and environments used here should be sufficient to test predictions of heterogeneity models.

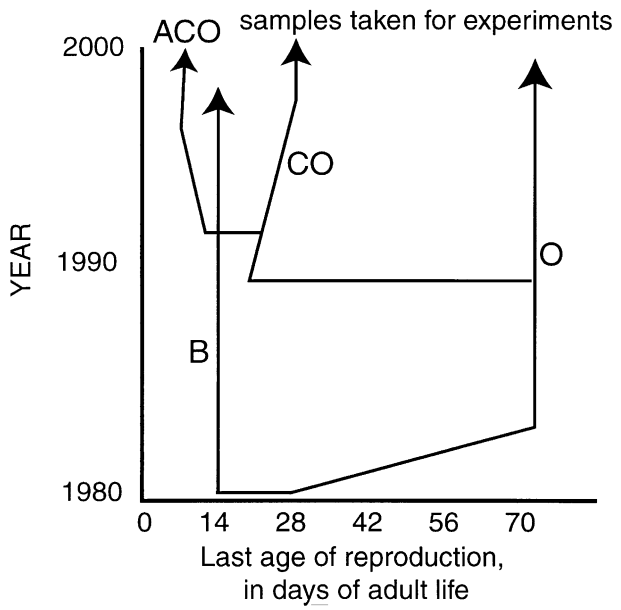

Fig. 2. The derivation of *Drosophila* stocks used in this experiment. Lines indicate the source population of each line and how the age of reproduction has changed over time.

We also measured mortality in all possible hybrid populations of the five B-populations. Every pair-wise combination of the cross $B_i \times B_j$ (both i and j varying from 1 to 5) was performed and non-hybrid cross progeny were handled in the same manner as hybrid cross progeny.

Rearing was parallel for two generations prior to assay. Densities were standardized among fly cultures. All cohorts were assayed for mortality by complete census of deaths every other day, similar to the mortality assays described above. Sample sizes were approximately 880 per cohort per sex, with a grand total of 43,937 flies.

2.3. Collection of fly mortality-rate data

The protocol used here closely matches that of our earlier study (Drapeau et al., 2000). For each replicate population, at least 80 8-dram glass banana-agar-corn syrup food vials each containing 60 ± 5 eggs were prepared. On the ninth and tenth days after the egg collection (but within a 24-hour period), all eclosed adult flies were sexed and placed into new food vials in groups of 24 (12:12 male:female) which would be used to start the mortality assay.

Adult survival was determined and living flies were transferred to new vials with fresh food every other day. Any flies that were living but irretrievably stuck to a part of the vial were scored as dead two days later. When necessary, flies within replicates were recombined among handling vials in order to maintain a density of approximately 20 flies/vial, to rule out any possible density effects on mortality rates. This procedure resulted in adult densities

varying between 12 and 24. For a culture that can support a carrying capacity of nearly 300 adults this is a very small range of adult densities, a range that is expected to have negligible effects on mortality rates (cf. Nusbaum et al., 1993; Carey et. al., 1993, 1995; Graves and Mueller, 1993, 1995; Curtsinger, 1995a,b; Khazaeli et al., 1995,1996). Survival assays were continued until every fly was dead. We used high cohort sizes (at least 1000 individuals per replicate) in order to reduce sampling variance in our estimations of mortality rates (Pletcher, 1999; Promislow et al., 1999) (see Table 3 for sample sizes).

More detail on the mortality patterns in these populations, and analysis with respect to evolutionary theories of late-life mortality (Mueller and Rose, 1996; Rose and Mueller, 2000), are presented elsewhere (Rose et al., 2002).

3. Results

3.1. Age-specific variance in B and O-populations

Previous work has documented large differences between the B and O populations for α, so we focus attention on those populations here. If there were zero deaths in any time interval we set the mortality rate to 1 over the number of survivors at that age class. We followed this procedure in the computer simulations in the next section also. Since the patterns of age-specific variance were very similar between sex and populations within a selection treatment (B vs. O) we have averaged the variance over sexes and similar populations. We also only estimated variance at ages where we had survivors across all populations. This reduced our ability to see the variance patterns at advanced ages. Since a whole population is used to estimate a single mortality rate the variances are among population means. This is important to keep in mind, since we will derive theoretical variances between individuals later.

In Fig. 3 we see that the B-populations show a peak in variance at very young ages. The O-populations, which have a much lower value of α than the B-populations, show a corresponding increase in the age of the first peak for mortality variance (Fig. 3). We also see a rise in variance late in life for the B-populations. Between these peaks, the variance is relatively constant. Likewise the late rise in variance in the O-populations occurs at a more advanced age, but the variance remains relatively flat at intermediate ages.

The general trend from both the B and O-populations is a peak in variance at an early age and at a very late age, with relatively constant variance at the intermediate ages. The difference between the B and O-populations is that the early and late peak start at younger ages in the B-populations compared to the O-populations.

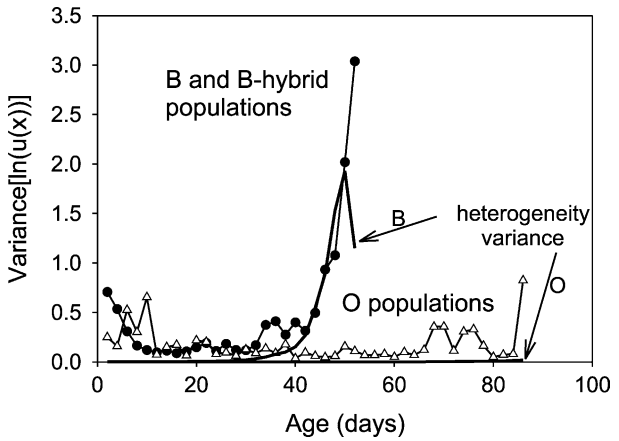

Fig. 3. The variance of the natural log of mortality as a function of age. Both sexes and 10 populations were used to estimate the variance in the O-populations. Both sexes and 25 populations were used to estimate the variance for the B and B-hybrid populations. The solid lines near the x-axis are the predicted levels of variation from the heterogeneity-in-A model. These were calculated from Eq. (4) and individual estimates of the model parameters given in Table 1.

3.2. Results of analytical and simulation work

We now develop some analytical results for the Vaupel et al. (1979) model that allows heterogeneity-in-A, in order to determine the expected patterns in age-specific variance for this model. It is fairly easy to show that $\mathrm{Var}[\ln(u(x))] = \mathrm{Var}[\ln(z(x))]$. This variance refers to the variance between individuals within a population. If we computed the variance between mean mortalities in populations each of size N, then this variance would be equal to, $\mathrm{Var}[\ln(u(x))]N^{-1}$. We now build on Vaupel et al.'s work to derive the density function of the log transformation of the random variable $z(x)$. If we simplify the notation, our problem is to determine the density function of the random variable $y = \ln(z)$. This can be accomplished by recalling that the relationship between the density function of y ($f_Y(y)$) and z ($f_Z(z)$) is,

$$f_Y(y) = \left|\frac{d\exp(y)}{dy}\right| f_Z(\exp(y))$$

(Mood et al., 1974, p. 200). In particular if z has a gamma distribution with parameters λ and k then the density function of $\ln(z)$ is,

$$\frac{\lambda^k(e^y)^k e^{-\lambda e^y}}{\Gamma(k)}. \tag{4}$$

We can use Eq. (4) to estimate $E[\ln(z(x))]$ and $\mathrm{Var}[\ln(z(x))]$-and hence the $\mathrm{Var}[\ln(u(x))]$-by letting

$$\lambda = \frac{1}{\sigma^2} + \frac{A[\exp(\alpha x) - 1]}{\alpha}$$

and $k = 1/\sigma^2$. If we assume mortality rates obey the Gompertz model with heterogeneity in A, we can use the observed patterns of mortality in the B and O populations to estimate the parameters of this model. We have used

estimates of A, α, and σ^2 for each B and O population (see Section 3.4 for details of the estimation procedure) to numerically estimate the variance in mean log mortalities due to variation in A alone. Before describing these results we give details of these numerical methods. Since the gamma function in Eq. (4) can get quite large, we first computed the log of the density function before placing that result in an exponential function to get Eq. (4). We used Romberg integration (the QROMB procedure described in Press et al., 1992, p. 140) to numerically integrate the density function. Polynomial interpolation with 10 ten points was used in this implementation. This increased the accuracy noticeably over five interpolating points. The allowable error estimate for this procedure was set to 10^{-6}.

The first result we obtained is that in an infinite population the variance between individuals for the natural log of mortality rates is constant with age (Fig. 4). It is clear from Eq. (3) that the Var[$z(x)$] goes to zero with increasing age. This is due to the scale of z, which must be a positive number. At advanced ages the only survivors have very small values of z and thus the numerical value of the variance gets very small. However, on a log scale there is no such change in variance as can be seen from the example in Fig. 4. This result appears to contradict Pletcher and Curtsinger (2000) who suggest that this variance will decrease with age (see their Fig. 1). Pletcher and Curtsinger did not derive the distribution or density function of their randomly varying mortality rates. Rather they relied on a Taylor series approximation of variance derived from the population mean mortality. Their differentiation of this equation was with respect to only one source of variation, the genetic variation, but not the environmental variation.

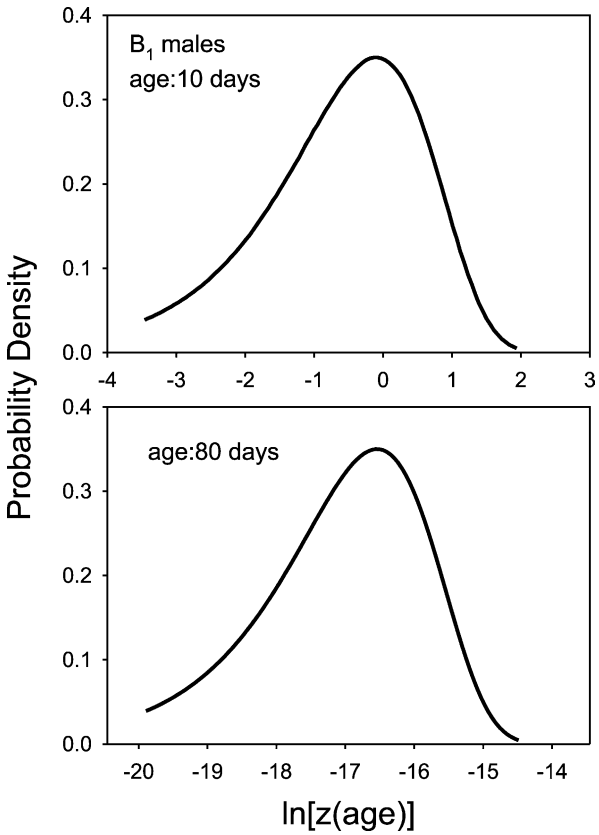

Fig. 4. The distribution of ln($z(x)$) at two different ages in the B1 male population. At ten days the mean values of z is about 1 while at 80 days it is 3.5×10^{-8}. At both ages the variance is the same, 1.86.

We took our numerical estimates of variance and divided by the total number of individuals alive at each age, in each population to get an estimate of the variance in mean mortality. We then compared these theoretical results to the observed variance in the B and O populations. In Fig. 3 we have illustrated the theoretical variance with solid lines. Heterogeneity in A will at best make a very small contribution to the total variation except possibly at the oldest ages where only a few individuals are still alive. In Fig. 3 the solid curve for the B-populations starts to rise at age 40 days. In the O-populations the variance is very small (on the x-axis) at essentially all ages shown. We next used computer simulations to study the variance of ln($\bar{\mu}(x)$) when the α parameter is also varied.

3.3. Magnitude of the different sources of variation

At very young ages, there are often no deaths between consecutive age classes using the values of A and α provided by Service. Service (2000a) followed the practice of setting these zero values to 1 divided by the number of adults at the start of the time interval. Examination of Fig. 3a in Service (2000a) reveals many values of log mortality at a boundary of -3.1, which corresponds to log(1/1250). The consequence of this type of truncated distribution is an apparent reduction in variance. Our first simulation (Fig. 5a) shows the variance in log mortality rates in a population with no variation in A and α. Thus, we are looking at sampling variance only. The variance peaks at about 10–15 days. The low variances before this peak correspond to the days when many populations appear not to vary since they all have zero observed deaths.

Our next two simulations show the effects of sampling variance plus environmental-developmental variation in A (Fig. 5b) and environmental variation in α (Fig. 5c). The most important conclusion is that the broad pattern is dominated by sampling variation. There is very little effect of within-population variation in A and α. This is in line with the results from the previous section. This conclusion is emphasized in Fig. 5d, which shows the results when three sources of variation are considered -sampling variation, variation in A, and variation in α- compared to just sampling variation (from Fig. 5a).

In each of the simulations shown in Fig. 5a–d there is a rapid decline in variance after the first peak and then a second peak at day 45–50. Let the estimated mortality rate at one of the young ages be p, then Np will have a mixed binomial distribution (assuming N individuals alive at the start of the time interval). Var[ln(p)] is approximately

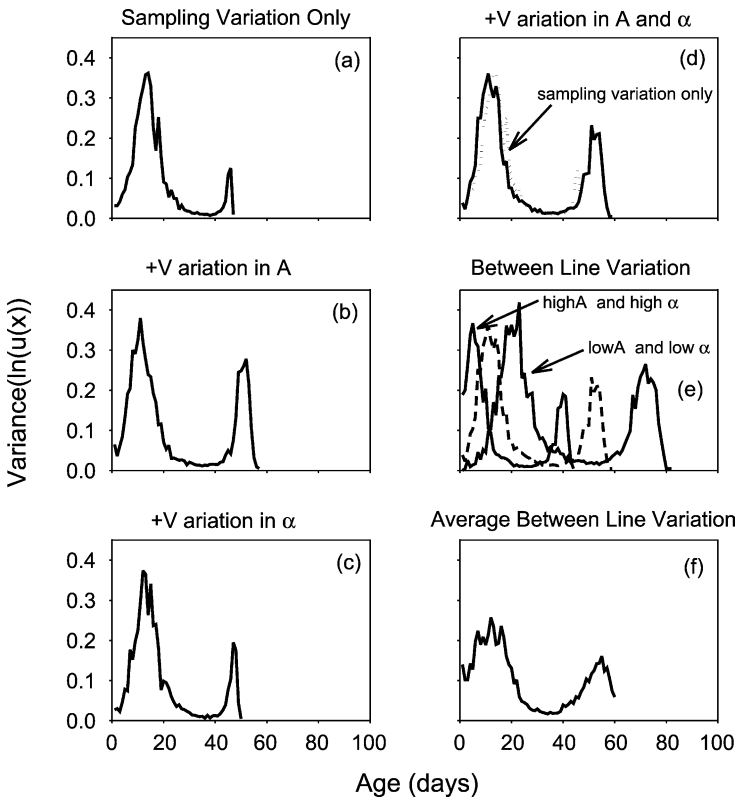

Fig. 5. The variance of ln[mortality rates] between replicate populations as a function of age. The values used in all these simulations were: $A = 0.00019$, $\mathrm{Var}[\ln(A)] = 0.766$, $\alpha = 0.193$, $\mathrm{Var}[\alpha] = 0.0021$. Mortality rates were determined from a Gompertz equation with either (a) only sampling variation, (b) sampling variation and variation in the A parameter of the Gompertz, (c) sampling variation and variation in α of the Gompertz, (d) sampling variation and variation in both A and α of the Gompertz, (e) the same as (d) except the high lines are one standard deviation greater than (d) and the low lines are one standard deviation smaller than (d), (e) average variation over 24 genetically different lines when within population variation is like (d).

$p^{-2}\mathrm{Var}(p) = (1-p)p^{-1}N^{-1} + (N-1)N^{-1}\theta^2 p^{-2} \cong d^{-1} + \theta^2 p^{-2}$, when $1 - p \cong 1$ and d is the number of deaths in the time interval. Since d increases rapidly from very young ages (5–15 days) into the middle ages (20–30 days), the variance declines rapidly in this age range. After this rapid decline in variance during ages 30–40, there are periods of relatively constant variance as we might expect from the analytic results derived earlier. The variance starts to increase again at later ages as the total number of survivors becomes low and thus the sampling variation increases, so when $p \cong 0.5$, $\mathrm{Var}[\ln(p)] \cong N^{-1}\theta^2 1/4$ which grows quickly at small values of N.

We next consider the effects of genetic variation. Service (2000a) allowed the mean values of A and α to vary between genetically different lines. In Fig. 5e we show the results from Fig. 5d (dashed line) along with the results after making the mean value of A and α one standard deviation smaller and one standard deviation larger than the means used in the previous simulations. Between-line genetic variation shifts the peaks to the left or right but maintains the same general pattern that is due to sampling variation. In fact it is the late life peak that is shifted most by the between-line genetic variation. In Fig. 5f we have reproduced the simulations of Service by sampling 24 computer-generated genetically variable lines. The results prior to day 40 closely resemble Service's results. Service did not see the variance peaks late in life due to additional constraints on his simulations. Service (2000a) required that at least 10 individuals be alive at the start of each age class to estimate mortality. Additionally, Service computed the between-line variation only at ages where all 24 lines had survivors, which effectively eliminates the older age classes from his analysis. Our analysis suggests that variation between individuals within populations for either A or α does not lead to substantial changes in the variance of ln(mortality) with age at least for the sample sizes and parameters values used here. The observed patterns of Service (2000a) are largely a consequence of sampling variation in estimated mortality rates or the effects of artificially truncated distributions on the estimates of variance.

We conclude from these analyses that, except for unrealistically large populations, observed laboratory patterns of age-specific variance in mortality rates will be dominated by sampling variance. There is little hope of detecting patterns that arise from different demographic heterogeneity models that may be present in infinite populations (Pletcher and Curtsinger, 2000), given the degree of replication in most laboratory experiments.

These predictions are consistent with observed patterns in the B and O-populations (Fig. 3). Both populations show the variance peaks at early and late life predicted from our

simulations. Since the value of α in the O-populations is much less than in the B-populations, our simulations predict that the peaks in the O-populations should be displaced towards older ages, which they are.

3.4. Tests of the life-long heterogeneity models using survival patterns

3.4.1. The heterogenity-in-A model

When there is variability in the A parameter of the Gompertz equation, Vaupel et al. (1979) have derived the average mortality at age-x as,

$$\bar{\mu}(x) = \frac{A\exp(\alpha x)}{1 + \sigma^2 A\alpha^{-1}[\exp(\alpha x) - 1]}. \quad (5)$$

This model is sometimes called the logistic model (Promislow et al., 1996). Two observations point to environmentally generated variation in A as a reasonable form of the heterogeneity model. First, populations that are highly inbred, and therefore likely to have little genetic variation, show plateaus (Fukui et al., 1993). Secondly, several environmental perturbations that have demonstrable effects on longevity only seem to affect the age-independent parameter of the Gompertz (Nusbaum et al., 1996; Joshi et al., 1996).

As a test of the lifelong heterogeneity theory, we estimated the parameters of Eq. (5) in ten populations. Mortality data from these ten populations were collected at the same time using the same experimental protocol. Accordingly we would expect levels of environmental variation to be identical in each population. The ten populations tested consist of five populations called B's and five called O's described above. All ten populations were originally derived from the same source population but have been subject to different regimes of age-specific selection (Rose, 1984). There is now substantial genetic differentiation between these populations. Nusbaum et al. (1996) showed that there were no significant differences between the B and O populations for A of the Gompertz equation but that there were significant differences for α. We estimated the parameters of the logistic model, Eq. (5), A, α, and σ^2 using maximum likelihood techniques described in Mueller et al. (1995), treating Eq. (5) as a density function (Table 1). Of course fitting this model to our data doesn't show that the model is valid. It merely provides us with an estimate of the model parameters under the *assumption* that the model is valid. Fitting the mortality data from large cohorts reveals significant differences in σ^2 between the B and O populations (two-way ANOVA, $p = 0.00084$, see Table 2).

These results seem difficult to reconcile with a theory based on environmental heterogeneity in A. One might argue that there is also genetic variability in A and that there is substantially more variation in the B populations than in the O's. This argument is difficult to accept since the selection that has taken place has apparently not affected the mean value of A and the effective population sizes in the two populations is roughly equal. Thus, we don't expect to have lost more variation in the O populations due to direct selection on A or by loss from genetic drift. If anything, the B populations have undergone a larger number of generations since splitting from the O's and might be expected to have less variation that the O's.

Table 1
The estimated parameters of the logistic model for five B and five O populations. The estimates were obtained by the maximum likelihood technique without approximation as described in Mueller et al. (1995)

Population	Sex	A	α	σ^2
B1	Males	0.00183	0.268	1.09
B2		0.00531	0.185	0.524
B3		0.00178	0.298	1.06
B4		0.00140	0.272	1.11
B5		0.000759	0.245	1.10
O1		0.000660	0.0984	0.813
O2		0.000887	0.0924	0.469
O3		0.00105	0.0841	0.669
O4		0.00124	0.0674	0.335
O5		0.000833	0.07997	0.607
B1	Females	0.00134	0.325	2.02
B2		0.00569	0.172	0.905
B3		0.00195	0.290	1.82
B4		0.00673	0.166	0.972
B5		0.00243	0.244	1.50
O1		0.00141	0.0822	0.586
O2		0.00293	0.0686	0.299
O3		0.00277	0.0692	0.364
O4		0.00296	0.0657	0.497
O5		0.00181	0.0750	0.532

In studies with *Drosophila,* several environmental variables are known to affect longevity and presumably mortality rates. These include mating status (Maynard Smith, 1958), nutritional status (Chapman and Partridge, 1996; Chippindale et al., 1993), and presence of urea in the adult food (Joshi et al., 1996). Estimates of Gompertz parameters for the experiments varying levels of food and urea have shown that only the A parameter of the Gompertz is affected and this presumably accounts for the differences in mortality between cohorts receiving different food levels (Nusbaum et al, 1996).

From the estimates of σ^2 in Table 1 we can ask if the levels of variation in $A'(= Az)$ required to fit the datasets are reasonable. If we focus on the B-males, as an example,

Table 2
The analysis of variance for σ^2 obtained from the heterogeneity-in-A model (Table 1)

Source	*d.f.*	Sum of squares	Mean square	F	p
Population	1	1.58	1.58	16.8	0.00084
Sex	1	0.031	0.031	0.33	0.57
Sex × population	1	0.163	0.163	1.73	0.21
Error	16	1.51	0.0941		

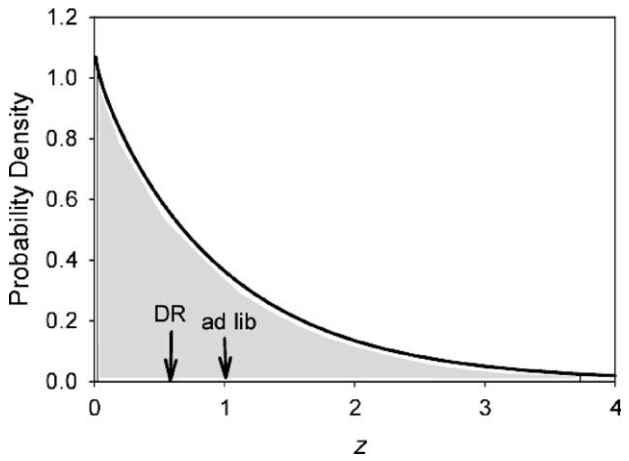

Fig. 6. The shaded portion shows the range that 95% of the z values would fall in for the average B-male population. This also represents the relative range for the different A'A'-values. The value of A' for B-males fed ad lib food is arbitrarily put at one then the relative reduction in A' due to dietary restriction (DR) is shown on the figure.

the average variance of z (σ^2) is 0.976 (Table 1). Since z has a gamma distribution with a mean of 1 and a variance of 0.976, 95% of the values of z would be between 0.023 and 3.73 (Fig. 6, shaded area). Thus, the largest and the smallest value of A' differ by a factor of 162. We need to recall that this large level of variation is mostly environmental and arises in experimental populations where all efforts have been made to make the environment as constant as possible. This type of variation also results in very different age-specific mortality rates motivating our earlier claim that the observed number of deaths has a Lexian distribution. For instance, B_1 males produce a 95% confidence interval on z of (0.017,3.48). This means that the 95% confidence interval on probabilities of mortality between day 10 and day 11 is (0.00052,0.101).

When food levels are purposely changed from ad lib to low levels (DR) in these same populations of B-males, the mean value of the A-parameter changes by only a small fraction of the variation that is supposedly present in these populations (Fig. 6). The heterogeneity in A theory requires that subtle environmental forces, that can't be controlled even in the laboratory, produce variation in the age-independent Gompertz parameter that is 80 times greater than the mean variation that is produced by brute force intervention in diet.

3.4.2. The heterogeneity-in-α model

We call the second heterogeneity model 'heterogeneity-in-α'. In this model the age-dependent parameter, α, is a random variable equal to $\zeta\tilde{\alpha}$. The random variable, ζ, has a gamma distribution with a mean of one and variance equal to k^{-1}. The mean instantaneous mortality rate for individuals aged-x under the heterogeneity-in-α model is (Pletcher and Curtsinger, 2000),

$$\bar{u}(x) = \int_0^\infty \frac{Az^{k-1}\exp(\tilde{\alpha}zx - \phi(x,z))\mathrm{d}z}{\int_0^\infty z^{k-1}\exp(-\phi(x,z))\mathrm{d}z} \tag{6}$$

where $\phi(x,z) = kz + A(\tilde{\alpha}z)^{-1}[\exp(\tilde{\alpha}zx - 1)]$. The expression for $\phi(x,z)$ given here corrects a typographical error in Pletcher and Curtsinger (2000).

Using Eq. (6) we can get estimates of the parameters of this model, A, $\tilde{\alpha}$, and k, using maximum likelihood techniques, as we did for the heterogeneity-in-A model. These estimates reveal much smaller variances for the gamma random variable, ζ. For the O-males the average variance that is observed is 0.037. The interval 0.66 to 1.42 includes 95% of the values of ζ, roughly a two fold difference between the smallest and largest. Natural selection for late-life reproduction can lead to genetic changes in populations that affect the age-specific parameter of ageing. Thus, the value of $\tilde{\alpha}$ in the B-males is roughly twice as large as O-males (0.21 vs. 0.081) and females also show a two-fold difference (0.19 vs. 0.067).

This heterogeneity-in-α model assumes that a small portion of the population will have very small values of α and will be very long lived. One possible sign that this model does not work well is that it requires very long lived individuals due to the assumption of heterogeneity in α. An indication of possible problems is revealed in Service (2000a). In his simulations, when α was varied he generated populations with average longevities of 50 days, which is reasonable for *Drosophila*, but simulated maximum life-spans of 365 days, which are unknown for this species.

To study this further, we have used the maximum likelihood estimates of A, $\tilde{\alpha}$, and k to simulate deaths in large cohorts for the heterogeneity in α model. We have averaged the results of these simulations across the five replicate populations within a selection treatment for each sex (solid lines in Fig. 7). Along with these predictions the actual probabilities of living beyond a specific age are shown in Fig. 7 (circles). While the model does a good job at young ages we see that at advanced ages, especially in the longer lived lines, the model predicts more survivors than we typically see.

We have tested this late-survivor prediction of the heterogeneity models with observations from the B, O, CO, and ACO populations. Since male and female data are analyzed separately there are a total 40 populations analyzed (2 sexes×4 selection regimes×5 replicates). With the estimated values of A, $\tilde{\alpha}$, and k for each of the 40 populations we simulated deaths of 100,000 individuals following the procedures used in Fig. 5. For each population the frequency of individuals that lived longer than some critical age were determined.

We also determined from our survival data the observed numbers surviving beyond these critical days. The results from the five replicates were pooled and the observed

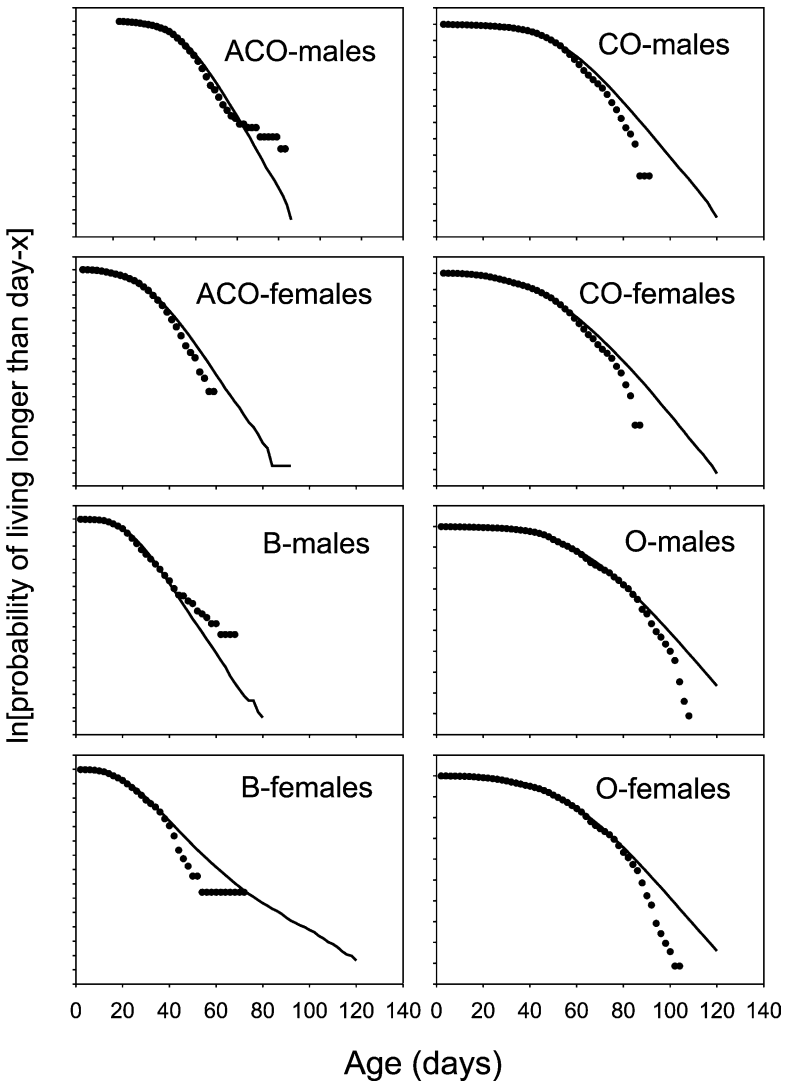

Fig. 7. The probability of surviving beyond a certain age in four different populations and both sexes. The probability scales are not identical for all figures. The solid circles are the observed probabilities in each group while the solid lines are the values predicted by the heterogeneity-in-α model. Since the model predictions are based on simulated, finite numbers of deaths, they do not always produce smooth curves especially at the advanced ages.

numbers were compared to the expected number from the heterogeneity-in-α theory. To determine if the expected frequencies were too large, we determined the 5% and 1% limits to our observed numbers from binomial sampling theory (Table 3). While we stated earlier that real populations are likely to be mixed binomial samples we believe the simple binomial distribution will be adequate here. First of all there are few surviving adults in the real population so the number of different p-values is likely to be small. Secondly, these extremely long lived individuals are all more likely to have similar probabilities of survival because deaths have reduced the variation of p (at least on a linear scale), making θ^2 very small.

These results show that for five out of eight comparisons the heterogeneity-in-α model predicts an excessive number of very old individuals. This suggests that some fundamental feature of the heterogeneity-in-α model is flawed.

4. Discussion

In biology, one of our most important organizing principles is adaptive evolution by natural selection. Natural selection rests on genetic variation, differential survival or reproduction, and transmission of genetic information from parent to offspring. Many profound ‘why’ questions in biology can be answered with simple principles of evolutionary biology, including questions in the fields of aging and demography. Unlike evolutionary theories of demography, the lifelong heterogeneity theories do not rest upon such well-established principles of biology. The lack of a mechanistic basis for these heterogeneity theories also makes it difficult to design critical experiments to test heterogeneity theories. The only support for these theories comes from their ability to mimic the post hoc patterns of mortality seen in biological populations. This is a weak form

Table 3
Number of flies with longevities greater than the critical age. The observed numbers for each population are polled from the five replicates. The confidence levels are based on the observed numbers and the expectations from binomial sampling theory. The expected numbers for the heterogeneity-in-α model are based on the separate maximum likelihood estimates for each sex and population

Population	Critical age (days)	Sex	Sample size	Observed number > critical age	Expected number
ACO	60	Males	12,444	8	9
CO	60	Males	11,987	1346	1613**
O	100	Males	8854	23	53.1**
B	40	Males	4867	49	41.9
ACO	60	Females	14,084	1	3.8
CO	60	Females	12,361	739	804*
O	100	Females	10,037	2	28.6**
B	40	Females	5,143	99	138**

*Probability of observed numbers this large or greater <0.05; **probability of observed numbers this large or greater <0.01.

of support for models in biology since there are often many post hoc models with these properties (Mueller and Joshi, 2000, Chapter 1).

4.1. The heterogeneity models

Nevertheless, in this paper we have shown that the heterogeneity theories are demonstrably flawed. Most importantly, the limited predictions that can be derived from them are not corroborated, as summarized in Table 4. The heterogeneity-in-α theory predicts too many long-lived individuals. This shortcoming is even more telling when we recall that populations that have too few long-lived individuals are the same ones used to estimate the parameters of these post hoc heterogeneity-in-α models. The heterogeneity in A models require extraordinarily high levels of between-individual variation. The known biological mechanisms of generating differences between individuals for the A parameter only account for a small fraction of the variation needed by this model.

An advantage of models that have no mechanistic basis for their equations is that a failure to fit a particular body of data is only an invitation to add higher-order terms, or new mathematical functions, to the theory. For example, the failure of the heterogeneity theory to predict a deficiency in extreme late-life deaths could be met by the use of an age-dependent mortality function that produces a steeper acceleration in the mortality rate at very late ages. In the history of science, such stratagems are well known (Popper, 1968), one use of them being the Ptolemaic epicycles added to save the geocentric model for the astronomical universe. Heterogeneity theory may continue to survive, but its practitioners will probably resort to progressively more elaborate model tinkering over the course of repeated empirical failures.

4.2. Evolutionary models

We have proposed an evolutionary model to explain mortality plateaus (Mueller and Rose, 1959), hereafter referred to as the Mueller–Rose theory. The Mueller–Rose theory relies on the joint forces of natural selection and random genetic drift. It is postulated that throughout most of the adult life span natural selection molds the patterns of adult mortality. Since the force of natural selection declines with age (Charlesworth, 1980) mortality rates increase

Table 4
A summary of experimental tests of models of heterogeneity and natural selection. Results that are inconsistent with these predictions are called 'negative' and those consistent with the predictions are 'positive'

Theory	Prediction	Result	Reference
Heterogeneity	Reduction of early life variability should reduce appearance of plateaus	Negative	Khazaeli et al., 1998
Heterogeneity	Selection to resist stress should affect characteristics of plateau	Drapeau et al., 2000 Negative	Drapeau et al., 2000; Mueller et al., 2000
Heterogeneity	Number of survivors in extreme age-classes	Negative	This paper
Heterogeneity	Similar variance in different populations raised in the same environment	Negative	This paper
Evolution	Decrease in age of onset of plateau with early life selection	Positive	Rose, et al., 2002
Evolution	Increase in age of onset of plateau with late life selection	Positive	Rose, et al., 2002

exponentially with age due to genetic mechanisms like mutation accumulation and antagonistic pleiotropy. The theory further supposes that at some advanced age selection becomes so weak that random genetic drift becomes the more important evolutionary force. At that age and at later ages mortality remains high but shows no trend, since all ages are equivalent from the perspective of selection and drift. Quantitative support for this theory was derived from computer simulations.

Several authors have criticized the Mueller–Rose theory. In addition, there are now several alternative evolutionary models that have been used to analyze mortality plateaus. We provide a brief review of these theories and critiques below.

4.2.1. The Mueller–Rose theory fails to identify the true stationary states

Wachter (1999) has criticized the Mueller–Rose theory suggesting that the plateaus observed in our computer simulations were simply transient states of a stochastic process. The true stationary states, according to Wachter, show exponential increases in mortality with age. It should be noted that the transient plateaus that observed by Mueller–Rose exist for the equivalent of millions of generations. It is also not clear from Wachter's work if the stationary states he identifies could be reached in a biologically meaningful period of time. In fact his only justification for interest in the stationary states rather than long lived transient states is that a theory that relies on transient states is 'unappealing'.

There is no reason to suppose that any environment would remain constant for millions of generations. As environments change, gains made in age-specific survival would be lost and the process of adaptation would begin anew, preventing them from reaching a stationary state. Many scientists have noted the importance of environmental heterogeneity in short and long term evolution (for instance see Gillespie, 1991, for a detailed treatment of this type of theory in molecular evolution).

Wachter's analysis only considers the special case of mutations that affect one age-class at a time, presumably because more realistic models were mathematically intractable. The Mueller–Rose theory concentrated on mutations that affect broad windows of contiguous age-classes. Charlesworth (2001) has recently shown that mutation accumulation models can lead to mortality plateaus, consistent with our findings. However, Charlesworth points out that pleiotropy is needed to generate these plateaus. In particular he assumed that mutations have both an age-specific effect and a pleiotropic non-age-specific effect. The non-age-specific effect is important for generating these plateaus. Charlesworth's conclusions suggest that Wachter's results may not hold for Mueller–Rose models with broad pleiotropic effects.

4.2.2. Alternative models use entropy as a measure of fitness rather than r

Demetrius (2001) developed an evolutionary model to explain the existence of mortality plateaus that uses a demographic statistic called entropy to predict the outcome of evolution, as opposed to the intrinsic rate of increase used in the Mueller–Rose and Charlesworth theories. Demetrius shows that the partial derivative of entropy with respect to age-specific survival is positive late in life but negative in mid-life, if population size is regulated. He suggests, without proof, that the age-specific pattern of entropy-sensitivity predicts that late life mortality should plateau. However, his results are also consistent with actual declines in late life mortality. The Mueller–Rose theory does not predict declines in late-life mortality. Thus, distinguishing between the utility of these contrasting evolutionary theories rests on empirical patterns of late-life mortality.

One can find occasional reports in the literature of unreplicated populations that show declines in mortality late in life (e.g. Carey et al., 1992). However, the broad pattern that has emerged by considering many species and replicates of similar populations is that mortality rates plateau at late life; they do not decline (Rose et al., 2002). Thus, existing empirical evidence is not consistent with the theory developed by Demetrius.

4.2.3. The Mueller–Rose theory rests on special assumptions or artifacts

Pletcher and Curtsinger (1998) suggest that our method of generating mutants artifactually leads to mortality plateaus. They assert that without these artifacts, late-life mortality would go to 100%. They support their arguments with computer simulations. These simulations assume that there is no reproduction at the late ages where mortality rises to 100%. None of the examples used to develop the Mueller–Rose theory make this assumption. Contrary to their assertions the alternative methods for generating mutants described by Pletcher and Curtsinger do result in mortality plateaus. Finally, Pletcher and Curtsinger derive theoretical results that suggest that the Mueller–Rose models will always equilibrate mortalities at intermediate levels. However, as pointed out by Wachter (1999) and Rose and Mueller (2000) these derivations rely on several incorrect assumptions.

4.2.4. Variants of the heterogeneity theories can explain plateaus

Weitz and Fraser (2001) develop a model that assumes individual mortality rates are random variables. Thus, an initially homogeneous population may develop heterogeneity since every individual experiences a different combination of random variation that alters their age-specific survival. The model assumes there is a linear increase in instantaneous mortality rates with age, so it assumes, rather than explains, Gompertz-like kinetics, unlike the evolutionary models (Mueller and Rose, 1996; Charlesworth, 2001).

Their model also assumes that mortality rates are subject to perturbations from uncorrelated Gaussian noise with a mean of zero. The functional form of this stochastic model has no deep biological motivation. Thus, the model's main merits are its ability to mimic the behavior of real populations. More work with this model will be required to determine if it is an advance over previous heterogeneity models.

As an alternative to heterogeneity models, demographers and gerontologists could consider using the evolutionary theory of late life (Rose and Mueller, 2000). This theory explains late-life mortality rate plateaus using equations articulated by Hamilton (1966) and Charlesworth (1980), especially the force of natural selection. There are simple qualitative predictions that can be derived from this theory (Mueller and Rose, 1996; Rose and Mueller, 2000; Charlesworth, 2001), and tested using practicable experiments (cf. Rose et al., 2002). The strongest support for the evolutionary models comes from their ability to make predictions that have been corroborated (Rose et al., 2002). Few experimental tests of the evolutionary theory of late life have been published to date, and this theory is undoubtedly in need of improved mathematical definition and analysis. Nonetheless, at a minimum, it deserves more of the attention now given to lifelong heterogeneity theories of late life.

Acknowledgements

We thank the students in the Rose laboratory for help with the collection of experimental results, several anonymous referees for helpful comments, and the National Institutes of Health and the estate of Robert Tyler for financial support.

References

Arking, R., Giroux, C., 2001. Antioxidant genes, hormesis, and demographic longevity. Journal of Anti-Aging Medicine 4, 125–136.

Best, D.J., 1983. A note on gamma variate generators with shape parameter less than unity. Computing 30, 185–188.

Carey, J.R., Liedo, P., Orozco, D., Vaupel, J.W., 1992. Slowing of mortality rates at older ages in large medfly cohorts. Science 258, 457–461.

Carey, J.R., Curtsinger, J.W., Vaupel, J.W., 1993. Fruit-fly aging and mortality-response. Science 260, 1567–1569.

Carey, J.R., Liedo, P., Vaupel, J.W., 1995. Mortality dynamics of density in the Mediterranean fruit fly. Experimental Gerontology 30, 605–629.

Carnes, B.A., Olshansky, S.J., 2001. Heterogeneity and its biodemographic implications for longevity and mortality. Experimental Gerontology 36, 419–430.

Chapman, T., Partridge, L., 1996. Female fitness in Drosophila melanogaster: An interaction between the effect of nutrition and of encounter rate with males. Proceedings of the Royal Society of London Series B—Biological Sciences 263, 755–759.

Charlesworth, B., 1980. Evolution in Age-Structured Populations, Cambridge University Press, New York, USA.

Charlesworth, B., 2001. Patterns of age-specific means and genetic variances of mortality rates predicted by the mutation-accumulation theory of ageing. Journal of Theoretical Biology 210, 47–65.

Chippindale, A.K., Leroi, A.M., Kim, S.B., Rose, M.R., 1993. Phenotypic plasticity and selection in Drosophila life-history evolution. 1. Nutrition and the cost of reproduction. Journal of Evolutionary Biology 6, 171–193.

Curtsinger, J.W., 1995a. Density and age-specific mortality. Genetica 96, 179–182.

Curtsinger, J.W., 1995b. Density, mortality, and the narrow view. Genetica 96, 187–189.

Curtsinger, J.W., Fukui, H.H., Townsend, D.R., Vaupel, J.W., 1992. Demography of genotypes: failure of the limited life-span paradigm in *Drosophila melanogaster*. Science 258, 461–463.

Demetrius, L., 2001. Mortality plateaus and directionality theory. Proceedings of the Royal Society of London B 268, 1–9.

Drapeau, M.D., 2002. Empirically testing heterogeneity theories of mortality. Journal of Anti-Aging Medicine 5, 235–236.

Drapeau, M.D., Gass, E.K., Simison, M.D., Mueller, L.D., Rose, M.R., 2000. Testing the heterogeneity theory of late-life mortality plateaus by using cohorts of *Drosophila melanogaster*. Experimental Gerontology 35, 71–84.

Fishman, G.S., 1996. Monte Carlo Concepts, Algorithms, and Applications, Springer, New York.

Fukui, H.H., Ackert, L., Curtsinger, J.W., 1996. Deceleration of age-specific mortality rates in chromosomal homozygotes and heterozygotes of *Drosophila melanogaster*. Experimental Gerontology 31, 517–531.

Fukui, H.H., Xiu, L., Curtsinger, J.W., 1993. Slowing of age-specific mortality rates in *Drosophila melanogaster*. Experimental Gerontology 28, 585–599.

Gillespie, J.H., 1991. The cause of Molecular Evolution, Oxford, New York.

Gompertz, B., 1872. On one uniform law of mortality from birth to extreme old age, and on the law of sickness. Journal of the Institute of Actuaries 16, 329–344.

Graves, J.L. Jr, Mueller, L.D., 1993. Population density effects on longevity. Genetica 91, 99–109.

Graves, J.L. Jr, Mueller, L.D., 1995. Population density effects on longevity revisited. Genetica 96, 183–186.

Gavrilov, L.A., Gavrilova, N.S., 1991. The Biology of Lifespan: A Quantitative Approach, Harwood, New York.

Greenwood, M., Irwin, J.O., 1939. The biostatistics of senility. Human Biology 11, 1–23.

Hamilton, W.D., 1966. Moulding of senescence by natural selection. Journal of Theoretical Biology 12, 12–45.

Ives, P.T., 1970. Further genetic studies of south Amherst population of *Drosophila melanogaster*. Evolution 24, 507–518.

Johnson, N.L., Kotz, S., 1969. Distributions in Statistics: Discrete Distributions, Wiley, New York.

Joshi, A., Shiotsugu, J., Mueller, L.D., 1996. Phenotypic enhancement of longevity by dietary urea in *Drosophila melanogaster*. Experimental Gerontology 31, 533–544.

Khazaeli, A.A., Xiu, L., Curtsinger, J.W., 1995. Effect of adult Cohort density on age-specific mortality in *Drosophila melanogaster*. Journal of Gerontology 50A, B262–B269.

Khazaeli, A.A., Xiu, L., Curtsinger, J.W., 1996. Effect of density on age-specific mortality in *Drosophila*: a density supplementation experiment. Genetica 98, 21–31.

Khazaeli, A.A., Pletcher, S.D., Curtsinger, J.W., 1998. The fractionation experiment: reducing heterogeneity to investigate age-specific mortality in Drosophila. Mechanisms of Ageing and Development 105, 301–317.

Maynard Smith, J., 1958. The effects of temperature and of egg-laying on the longevity of Drosophila subobscura. Journal of Experimental Biology 35, 832–842.

Mood, A.M., Graybill, F.A., Boes, D.C., 1974. Introduction to the Theory of Statistics, third ed., McGraw-Hill Book Company, New York, USA.

Mueller, L.D., Joshi, A., 2000. Stability in Model Populations, Princeton University Press, Princeton, NJ.

Mueller, L.D., Rose, M.R., 1996. Evolutionary theory predicts late-life mortality plateaus. Proceedings of the National Academy of Sciences, USA 93, 15249–15253.

Mueller, L.D., Drapeau, M.D., Rose, M.R., 2000. Stress resistance, heterogeneity, and mortality plateaus: response by the authors. Experimental Gerontology 35, 1089–1091.

Mueller, L.D., Nusbaum, T.J., Rose, M.R., 1995. The Gompertz equation as a predictive tool in demography. Experimental Gerontology 30, 553–569.

Nusbaum, T.J., Graves, J.L., Mueller, L.D., Rose, M.R., 1993. Fruit-fly aging and mortality. Science 260, 1567.

Nusbaum, T.J., Mueller, L.D., Rose, M.R., 1996. Evolutionary patterns among measures of aging. Experimental Gerontology 31, 507–516.

Olshansky, S.J., Carnes, B.A., 1997. Ever since Gompertz. Demography 34, 1–15.

Pletcher, S.D., 1999. Model fitting and hypothesis testing for age-specific mortality data. Journal of Evolutionary Biology 12, 430–439.

Pletcher, S.D., Curtsinger, J.W., 1998. Mortality plateaus and the evolution of senescence: why are old-age mortality rates so low? Evolution 52, 464–545.

Pletcher, S.D., Curtsinger, J.W., 2000. The influence of environmentally-induced heterogeneity on age-specific genetic variance for mortality rates. Genetical Research 75, 321–329.

Popper, K.R., 1959. The Logic of Scientific Discovery, Harper and Row, New York, USA, printing.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in C, second ed., Cambridge, UK.

Promislow, D.E.L., Tatar, M., Pletcher, S., Carey, J.R., 1999. Below-threshold mortality: implications for studies in evolution, ecology, and demography. Journal of Evolutionary Biology 12, 314–328.

Promislow, D.E., Tatar, M., Khazaeli, A.A., Curtsinger, J.W., 1996. Age-specific patterns of genetic variance in *Drosophila melanogaster*. I. Mortality. Genetics 143, 839–848.

Rose, M.R., 1984. Laboratory evolution of postponed senescence in *Drosophila melanogaster*. Evolution 38, 1004–1010.

Rose, M.R., Drapeau, M.D., Yazdi, P.G., Shah, K.H., Moise, D.B., Thakar, R.R., Rauser, C.S., Mueller, L.D., 2003. Evolution of late-life mortality in *Drosophila melanogaster*. Evolution 56, 1982–1991.

Rose, M.R., Mueller, L.D., 2002. Ageing and immortality. Philosophical Transactions of the Royal Society of London Series B—Biological Sciences 355, 1657–1662.

Service, P.M., 2000a. Heterogeneity in individual mortality risk and its importance for evolutionary studies of senescence. American Naturalist 156, 1–13.

Service, P.M., et al., 2000b. Stress resistance, heterogeneity, and mortality plateaus: a comment on Drapeau. Experimental Gerontology 35, 1085–1087.

Shaw, F.H., Promislow, D.E.L., Tatar, M., Hughes, K.A., Geyer, C.J., 1999. Toward reconciling inferences concerning genetic variation in senescence in *Drosophila melanogaster*. Genetics 152, 553–566.

Vaupel, J.W., Manton, K.G., Stallard, E., 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16, 439–454.

Vaupel, J.W., Carey, J.R., Christensen, K., Johnson, T.E., Yashin, A.I., Holm, N.V., Iachine, I.A., Kannisto, V., Khazaeli, A.A., Liedo, P., Longo, V.D., Zeng, Y., Manton, K.G., Curtsinger, J.W., 1998. Biodemographic trajectories of longevity. Science 280, 855–860.

Wachter, K.W., 1999. Evolutionary demographic models for mortality plateaus. Proceedings of the National Academy of Sciences, USA 96, 10544–10547.

Weitz, J.S., Fraser, H.B., 2001. Explaining mortality rate plateaus. Proceedings of the National Academy of Sciences USA 98, 15383–15386.