Title

Evolutionary shaping of demographic schedules

Journal

Proceedings of the National Academy of Sciences of the United States of America, 111(supplement_3)

ISSN

0027-8424

Authors

Wachter, Kenneth W
Steinsaltz, David
Evans, Steven N

Publication Date

2014-07-22

DOI

10.1073/pnas.1400841111

Peer reviewed

Evolutionary shaping of demographic schedules

Kenneth W. Wachter[a,b,1], David Steinsaltz[c], and Steven N. Evans[b,d]

Departments of [a]Demography, [b]Statistics, and [d]Mathematics, University of California, Berkeley, CA 94720; and [c]Department of Statistics, Oxford University, Oxford OX1 3TG, England

Edited by Brian Skyrms, University of California, Irvine, CA, and approved May 1, 2014 (received for review February 12, 2014)

Evolutionary processes of natural selection may be expected to leave their mark on age patterns of survival and reproduction. Demographic theory includes three main strands—mutation accumulation, stochastic vitality, and optimal life histories. This paper reviews the three strands and, concentrating on mutation accumulation, extends a mathematical result with broad implications concerning the effect of interactions between small age-specific effects of deleterious mutant alleles. Empirical data from genomic sequencing along with prospects for combining strands of theory hold hope for future progress.

senescence | hazard functions | biodemography | evolutionary demography | genetic load

Darwinian natural selection is a story about demographic success. Creatures pass on their genes thanks to the survival and fertility they achieve as they age across the life course. It makes sense to try to understand the age-specific patterns realized in demographic schedules from species to species in the light of evolution. Three main lines of inquiry are being actively pursued by demographers, mutation accumulation, stochastic vitality, and optimal life histories, described below. Of these, the first, mutation accumulation, draws the most specific connections between genomes and demographic outcomes. The last few years have seen the consolidation in refs. 1 and 2 of mathematical theory for the demographic consequences of this evolutionary process. This paper situates mutation accumulation within the context of the other demographic approaches, extends a mathematical result with demographic implications, and considers emerging empirical and theoretical opportunities.

Mutation accumulation is an idea of Sir Peter Medawar (3). It posits large numbers of deleterious alleles, each with small age-specific effects on survival, imposing genetic load on the population. Natural selection weeds out more slowly bad alleles that only or mainly affect an organism when its days for procreating, parenting, and grandparenting are running short. More late-acting alleles will be found in any equilibrium state where inflow of new mutations balances outflow in "mutation–selection balance." Basic theory is found in refs. 4 and 5 extended to demographic settings in refs. 6 and 7. It has gained relevance from discoveries of large numbers of rarely occurring single-nucleotide variants (SNVs) in sequenced genomes by the authors of ref. 8, many apparently of the kind posited for mutation accumulation.

Stochastic vitality, a second approach, aims to represent population heterogeneity as it affects survival across the life course. Differing genetic endowments and vagaries of development create heterogeneity in physiology and susceptibility to environmental shocks. Unobserved heterogeneity fixed across life is called frailty, modeled by ref. 9 within a demographic framework of proportional hazards. Stochastic vitality generalizes fixed frailty; "vitality" changes across life in a stochastic, usually Markovian process.

Large-scale models have been developed over many years by Kenneth Manton, Anatoli Yashin, and many collaborators (e.g., ref. 10) in the form of "stochastic risk factor models" that are Markov processes with high-dimensional state spaces. States of a system represent large suites of physiological indicators, with transitions estimated from data in longitudinal surveys. Small-scale, stylized models have also proved useful in identifying generic properties and demographic implications of stochastic vitality. Exemplars are refs. 11–14. In one example easy to picture, vitality is represented by a unidimensional Brownian motion and death by hitting a lower barrier or by remaining below a lower barrier for some random waiting time. Probability models developed to study bankruptcy of firms are enriching the mathematical tools for demographic analysis.

Optimal life history theory, a third approach, is familiar from a long tradition in biology studying organisms making adaptive trade-offs over the life course to maximize reproductive success. Trade-offs can be viewed as being programmed into the genome or implemented as dynamic behavioral responses. They could be manifest in genetic variation along the lines of "antagonistic pleiotropy" or they could be found in norms of reaction established by alleles long gone to fixation. The "disposable soma" approach of Thomas Kirkwood (15) has inspired work within this framework emphasizing investments in growth, maintenance, reproduction, and repair.

The demographic side of the enterprise, represented for instance by refs. 16–20, emphasizes roles for intergenerational and intergroup transfers in social species. Analysis of returns to investments in different background environments give insight into differences among species and taxa, especially into a distinction between "fast" and "slow" life histories. The optimization at issue is optimization under constraints. In practice, in formal models for demographic schedules, constraints tend to have to be invented to produce desired shapes, and the mathematics on its own does not add much predictive power. However, the descriptive pictures provided by these models give a wide range of qualitative insights.

It might well be feared that the baffling diversity of body plans and environmental challenges among living things might make the linkages between natural selection and demographic schedules too complicated for words. Reassurance comes, however, from discoveries of commonalities in the shapes of survival curves in populations of organisms as diverse as fruit flies, primates, and worms. These discoveries helped define an agenda for a new field of "biodemography" seeded by remarkable findings in refs. 21–24 and consolidated in the 1997 volume *Between Zeus and the Salmon* (25). As a practical goal, biodemographers seek to give an account of how our human evolutionary heritage allows our extended survival and whether and how it gives scope for yet further enhancements.

Chief among the commonalities highlighted by biodemographic research are exponential increases in mortality rates with age at moderately old ages ("Gompertz hazards") and

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "In the Light of Evolution VIII: Darwinian Thinking in the Social Sciences," held January 10–11, 2014, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS website at www.nasonline.org/ILE-Darwinian-Thinking.

Author contributions: K.W.W., D.S., and S.N.E. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: wachter@demog.berkeley.edu.

abating increases or "plateaus" at extreme ages. Alongside these patterns are a host of allometric scaling relationships connecting life span parameters with physical size and metabolic rate.

Processes of mutation accumulation, stochastic vitality, and life history optimization are not alternatives, but complements. It is to be hoped that the future will bring models that show these processes working in combination. So far, though, these lines of research have barely intersected.

Each strand has its advantages. Mutation accumulation has the advantage of focusing on features of demographic schedules that are fairly ubiquitous across the tree of life. It has strong connections to genomic observations through its explicit modeling of alleles subject to mutation, selection, and recombination. The highly specific content of models for natural selection endows mathematical results with predictive power and potential for surprises.

Stochastic vitality has the advantage of being well-suited to accommodating physiological realism. High-dimensional models can draw on extensive survey data and assimilate medical knowledge. Low-dimensional models bring out mathematical relationships clearly.

The subject of optimal life histories has the advantage of a vast array of field observations and studies of natural history and potential for explicit modeling of environments. It permits an emphasis on explaining demographic differences, from niche to niche, population to population, and species to species. It finds application to many timescales from evolutionary time to short-term adaptive strategies. It helps provide understanding of "negative senescence" in populations from some species—some turtles are examples—where mortality rates fall rather than rising with age (26). Optimal life history theory has proved well-suited for taking account of benefits of cooperation and sharing in social species. On the mathematical side, its flexibility reduces predictive force, but it directs attention firmly toward the conceptualization and investigation of constraints.

In this paper, our primary focus is mutation accumulation, the age-specific manifestation of genetic load. Agrawal and Whitlock (27) give a recent review of the broad subject of genetic load. They list six topics calling for elaboration: epistasis, asexual reproduction, nonrandom mating, spatial effects, temporal effects, and genetic drift. Epistasis is the one under study here. The spotlight is being directed at a particular source of epistasis grounded in the logic of age-specific demographic schedules. Among geneticists, the word "demography" typically comes up with regard to histories of effective population size. Demography here is much more. It is the warp and woof of fitness. Fitness is to be calculated from rates of fertility and mortality depending systematically on age across the life course, reflecting batches of alleles carried by individual members of a population.

Three features of our modeling of epistatic genetic load stand out. First, we are not concerned with mutant alleles at specific sites but rather with sets or "teams" of alleles. Alleles are assigned to teams on the basis of their age-specific profiles of demographic action. Alleles belonging to one team may act through very different physiological pathways, but the resulting impact on age-specific rates is shared. Widely scattered sites contribute members to each team. Whereas genetic drift plays its usual role in driving the highly random trajectories of counts of derived alleles in the population for each site, we are looking at aggregate counts for teams of alleles, for which mean values are paramount.

Second, in this paper, we are concerned only with one piece of the whole load of deleterious mutant alleles. Although the model of ref. 1 also applies to fertility and to infant and child mortality, here we are restricting attention to adult mortality. This restriction has a benefit. It gives a handle on what is often, as Agrawal and Whitlock point out, an elusive abstraction, the "load-free genotype." Adult survival across ages of reproduction and nurturing cannot be better than 100% and cannot plausibly be better than curves implied by reasonable levels of age-independent, extrinsic mortality. As described under *Discussion*, anthropologists' life tables for contemporary hunter-gatherers bring a load-free baseline and bounds on fitness costs from this particular piece of genetic load within reach of empirical assessment.

Third, for this piece of genetic load, an adjustment for epistasis can be directly calculated when a rate for new mutations is to be compared with a total fitness deficit. Haldane's principle, which equates the two in the absence of epistasis, has a closed-form generalization proved in ref. 2.

We adhere to a demographers' point of view in expecting size and density-dependent homeostatic feedbacks to keep net population growth near zero on a short-term timescale from one generation to the next across prehistory. Levels of fertility and offspring survival are taken to be compensating for what would otherwise be a loss in population viability. Optimal life history theory should shed light on processes of compensation and on pathways through which genetically determined biochemical differences work themselves out in impacts on age-specific demographic schedules.

We now turn to a specific result in mutation accumulation theory and its interpretation. The result is to be understood in the light of an idea of Brian Charlesworth (28). Charlesworth assumed (*i*) a linear approximate model for mutation accumulation; (*ii*) mutant alleles with effects restricted to each and every range of later ages; (*iii*) heavy constant background exogenous mortality. He showed that these assumptions taken together would arrange for mutant alleles to imprint a Gompertz-like exponential pattern on mortality rates by age. Numerous proposals have been made over the decades for explaining Gompertz mortality schedules. However, many manage to get exponential patterns out by putting exponential patterns in. In Charlesworth's story, the exponential pattern arises spontaneously from conditions that are highly generic and fairly ubiquitous.

Charlesworth's model is a linear one derived via approximations that hold when overall effects of genetic load are low. It equates the selective cost of a batch of deleterious mutant alleles to the sum of their separate selection costs. However, nonlinearity is an essential feature of mutation accumulation. The loss of net reproduction from one mutant allele reduces what is left to be lost by other mutant alleles. In any full model, the selective cost of a batch of mutant alleles has to be less than the sum of the costs of each of them separately. Thus, in a full model, natural selection has less force for weeding out mutant alleles over generations to preserve mutation–selection balance.

In ref. 2, we proved that under a full nonlinear model, within the setting investigated by Charlesworth, taking nonlinear interactions into account destroys the equilibrium and drives survival to zero at all reproductive ages. That proof required somewhat stringent conditions on the age-specific profiles of demographic effects restricted to later ages. We conjectured that those restrictive conditions were unnecessary. We now go on to establish that conjecture and prove the result in greater generality.

Results

We preview our nonlinear model for mutation accumulation with some brief heuristics. We are interested in a description with a resolution along the time axis of, say, fifties of generations. On this timescale, recombination events, more than one per chromosome per generation, break up most linkage between sites numbering in, say, the hundreds, widely scattered across the genome. Thus, it is reasonable to expect something like statistical independence when adding up, across different sites, alleles belonging to the same team. At each separate site, Wright–Fisher resampling may be taken to govern the random fates of derived mutant alleles, whereas team-by-team totals of load for population members are sensibly described in terms of distributions of sums around mean values and by a deterministic model governing mean values over time.

In this setting, recombination is rapid compared with mutation and selection, and both are rapid compared with aggregate effects of genetic drift. Effective population sizes on the order of 10,000 or more are contemplated. The theory applies to alleles that are mildly deleterious, intermediate between those that are strongly selected and those for which selection is negligible compared with drift. At any one site, when effective population size is not very small and counts of derived alleles are not very large, each round of (suitably scaled) Wright–Fisher resampling from the population as it selects the next group of carriers should not differ too much from low-probability Bernoulli trials. Adding up across sites, loads for individuals should not differ too much from Poisson-distributed independent variates. At this level, the argument remains informal, but formal results relating recombination to Poissonization are given in chapters 4–6 of ref. 1, where assumptions about the rapidity of the recombination process are made precise.

In short, demographic applications call for a model in which the genetic load for a randomly selected member of the population is represented by some form of Poisson process on some space indexing teams of alleles. Poisson processes are uniquely determined by intensity measures or intensity functions, so some such object should play a central role. When aggregate effects of drift are small, an infinite population model is indicated. When relevant intervals of time encompass many generations, a continuous-time model is indicated; hence, the model of ref. 1.

Our formal model for mutation accumulation with demographic selective costs from ref. 1 is presented in this section with supporting details in *Materials and Methods*, where the proof of our theorem is also to be found. As we have mentioned, the model can accommodate effects on fertility and juvenile mortality, but in this paper we restrict attention to adult mortality schedules. The model has four essential constituents. First is a space $\mathcal{M}$ indexing teams that are each composed of mutant alleles sharing a profile of demographic action. This index space may be any complete, separable metric space and is sometimes just a finite set. Here, it is the positive real axis. Second are the action profiles themselves. For each m in $\mathcal{M}$, we define an age-specific profile $\theta(m,x)$ specifying the effect of m at age x. Formally, θ is a bounded measurable function from $\mathcal{M}\times\mathbb{R}^+$ to $\mathbb{R}^+$ nondecreasing in age x for every m. For flexibility in comparing effects of different sizes, a positive scaling parameter η is taken to multiply all of the profiles $\theta(m,x)$. Often (as in our theorem) η is set to 1. Our third constituent is a mutation rate $q(m)$. It is a nonnegative measurable real function on $\mathcal{M}$. Our fourth constitutent is a family of intensity functions. For each time $t\geq 0$, $r_t(m)$ is a measurable function representing the predicted mean density of alleles from team m at time t in the population.

The model is an infinite population model in continuous time with weak selection and random mating. At sites contributing to any team, alleles are either wild type or semidominant deleterious mutants with no backmutation. This framework builds on the well-known work of Kimura and Maruyama (29), but recombination leads to different and simpler formulas. For demographic calculations, the genotype of an individual is identified with the collection of elements m carried by the individual, each repeated as many times as there are alleles from the team indexed by m. A genotype is specified by an element g taken from the space $\mathcal{G}$ of integer-valued Borel measures on $\mathcal{M}$. An individual with the "null" genotype $g\equiv 0$ has wild-type alleles at every contributing site.

Demographic structure enters the model through the profiles θ. They determine the survival function through the cumulative hazard. The survival function $\ell_x(g)$ is the probability of an individual with genotype g surviving to age x. The logarithm of its reciprocal is the cumulative hazard, whose slope, when it exists, is the hazard function itself (see, e.g., chapter 8 of ref. 30). In the model, the cumulative hazard implied by g is formed from adding up a baseline cumulative hazard plus an increment $\theta(m,x)$ for each m carried in the genotype g. The selective cost of mortality due to these alleles is measured by the reduction in a net reproduction ratio (*NRR*) calculated from the product of $\ell_x(g)$ times a baseline fertility schedule f_x. The product (the "net maternity function") is assumed to be nonnegative, integrable, and bounded with support $[\alpha,\beta]$ with $0\leq\alpha<\beta\leq\infty$]. Its integral is the *NRR*. The model does not distinguish individuals by sex and selective cost is calculated for individuals rather than couples, but inheritance is diploid. (See ref. 2, p. 10,141.) For any intensity function r, an age-specific force of natural selection F_r can be calculated from formulas given in *Materials and Methods*. A fixed-point argument in chapter 2 of ref. 1 shows that a family of intensity functions $r_t(m)$ varying over time can be defined as the unique solution (given r_0) to a dynamical equation as follows:

$$\frac{dr_t(m)}{dt}=q(m)-r_t(m)F_{r_t}(m). \qquad [1]$$

The outcomes under scrutiny in our theorem occur when each mutant allele is associated with an age of onset. Effects on the cumulative hazard are restricted to ages beyond the age of onset. We identify the index m with the associated age of onset and require that the mutation rate $q(m)$ be bounded below by some constant $q_0>0$ for m within the support (possibly infinite) of the net maternity function. We let q vanish outside the support.

In linear approximate models that ignore interactions, these kinds of profiles lead to well-behaved equilibrium limits for $r_t(m)$ as $t\to\infty$ starting from the null genotype. In noteworthy cases inspired by W. D. Hamilton and featured by Charlesworth (28), the profiles are taken to be step functions with a single step of some size η at m, so called "point-mass increments." For contrasts in qualitative behavior between linear and nonlinear predictions, shapes of profiles beyond their ages of onset turn out not to matter. However, until now, it has been an open question as to whether some minimal step in profiles at age of onset is behind the disappearance of equilibria.

In ref. 2, we proved that a finite limit for $r_t(m)$ as $t\to\infty$ fails to exist for the setup just described under the special condition that there exists $\eta_0>0$ such that $\theta(m,x)\geq\eta_0$ for all $x>m$ and all m. In other words, the proof depends on assuming some minimal step at the age of onset. This condition is undesirable. A step added onto a cumulative hazard function is like a Dirac delta function added onto the hazard function itself. Such abstractions are unrealistic for actual effects on mortality. In ref. 2, we conjectured that the condition could be removed, but the specter that it might turn out to be required or that some complicated alternative might be necessary has hung over the subject. The following theorem, proved in *Materials and Methods*, confirms that the condition can indeed be relaxed.

Theorem. *Let the mutation rate $q(m)$ be bounded below by a constant $q_0>0$ beyond an age of maturity $\alpha\geq 0$ and vanish for $m<\alpha$. With effect sizes $\eta_0\equiv 1$, let the action profiles θ satisfy*

$$\theta(m,x)=0 \quad for \quad x\leq m,$$

$$\theta(m,x)>0 \quad for \quad x>m.$$

Then, starting from $r_0\equiv 0$, the intensity function $r_t(m)$ diverges monotonically to infinity as $t\to\infty$ for all $m>\alpha$.

The eventually unbounded increase of the intensity function described in the theorem can be pictured in terms of changes in the predicted hazard function over time. At all times, each age is subject to alleles arriving in the population with the capacity to drive up hazards at and beyond that age. At older ages, there is never much selective pressure removing alleles that only affect

such older ages, and the frequencies of these alleles increase linearly with time and head toward infinity. At younger ages, early on, there is plenty of remaining reproductive potential to be lost by death and selective pressure removes many of the incoming alleles, temporarily keeping hazard rates low. In cross-section, there is a stretch of low young-age hazards, a stretch of rising hazards where remaining net fertility is dropping off, and a stretch of high old-age hazards steadily trending upward. As time goes by, remaining reproductive potential at middling ages is gradually driven down. Under the assumptions of the theorem, no selective pressure is arising from ages before the age of onset for the effect of each allele. Thus, for alleles with onsets at middling ages, it comes to be the case that selective pressure can no longer keep outflow close to inflow. The stretch of rising hazards shifts toward ever younger ages, until hazards at all ages are marching toward infinity.

Discussion

The theorem delineates an outcome which nature has to avoid. It reveals hidden dangers lurking in a kind of adaptive change that would have been expected, on the face of it, to be advantageous. Alterations in coding or in the regulation of gene expression that postpone processes of deterioration to later ages would seem to offer an efficient path to higher fitness. The theorem shows, to the contrary, that there can be too much of a good thing. If early-age effects of many deleterious alleles are altogether erased and onsets turn up widely scattered late in the span of reproductive or nurturing ages, then control over mutation accumulation can be utterly lost.

One way to forestall the destruction of equilibria is easy to envision but has far-reaching implications. Instead of erasing early-age effects, processes of postponement can merely reduce them, leaving some traces in hazard functions at early ages of pathways of debilitation or vulnerability that come to be largely concentrated at later ages. In other words, there is an important difference between effects restricted to older ages and effects only concentrated at older ages. Conditions that impose lower bounds on selective costs including effects at early ages are proved to be sufficient to guarantee the existence of mutation–selection equilibria in corollary 3.17 of ref. 1. Small fixed costs independent of age would be indistinguishable from lower bounds on early-age effects.

Another plausible way to keep mutation accumulation from running amok is to have mutation rates drop off for very late-acting alleles. On its own, this alternative would not be efficient. It only suffices if mutation rates drop all of the way to zero, and then only under certain conditions. According to the theorem, even the tiniest lower bound on $q(m)$ spells trouble. However, in combination with early-age effects for late-acting alleles, diminishing mutation rates for very late-acting alleles offer a realistic antidote.

There is something particularly interesting about these ways of avoiding equilibrium destruction. Under ordinary conditions, they imply plateaus in hazard rates at extreme ages. The age-specific force of natural selection trends downward as the age profile of effects is concentrated at later and later ages, but even when there is next to no selective cost to the late-age effects, the costs of trace effects, if bounded below, hold the representation of the alleles in check at equilibrium. Thus, mutation accumulation with demographic cost functions inherently promotes plateaus as a late-age concomitant of early accelerating increases in hazard functions.

Other processes featured in the other approaches described at the beginning of this paper also give good reasons for expecting plateaus. Plateaus in hazard functions are a generic feature of a wide class of Markovian stochastic vitality models, thanks to the property of quasistationarity leading to powerful results by refs. 12, 14, and 31. Heterogeneity in fixed frailty models promotes plateaus, and there are ways of arranging for plateaus in optimal life history models. Mutation accumulation must be regarded as only one of several contributors to the occurrence of plateaus. However, it is especially interesting as an enabling feature, a kind of prerequisite. Each of the other processes remains vulnerable to disruption by deleterious mutations and the plateaus that they foster could be destroyed, if mutation accumulation were not being held in check.

A contrary view is expressed by Danko et al. (32), who claim that "mutation accumulation may be a minor force in shaping life history traits." However, these authors impose an arbitrary upper bound of 10 on the load of mutant alleles, excluding most of the effect. Under their model with the bound removed, mutations do accumulate in large numbers (see theorem 5.1 in ref. 33), supporting the opposite conclusion.

Evolutionary demography is at a stage where new infusions of empirical findings are needed. It is not to be hoped that age-specific profiles for the action of the kinds of alleles involved in mutation accumulation can be estimated from data, because these are alleles with small effects, likely below any attainable threshold of detectability. Estimation of age-specific action is rather uncertain even for the few alleles so far identified with measurable effects on longevity, like polymorphisms in the human forkhead box gene *FOX03A*. However, statistical findings from whole-genome sequencing efforts hold out promise for calibrating the parameters of mutation accumulation models and for confronting predictions with data.

Mutation accumulation is a story about small effects. Small effects have long clearance times. In hazard functions, we view the outcome of interaction between genetic vulnerabilities and environmental challenges, challenges progressively reduced at least over the last thousands of years and impressively over the last few hundreds of years. Such reductions slow clearance further. One implication is that deleterious alleles found today in human genomes may often have ancient origins. Their frequencies may reflect long periods of culling by natural selection in older, more arduous environments. We expect their age-specific impacts to reflect selective forces from times when much more severe mortality impinged on ages of reproduction and nurturing than is true today.

How might results from genome sequencing be brought to bear? Our infinite-population model with epistasis does not incorporate genetic drift and does not apply directly to single sites. Trajectories for derived alleles from a mutation at any single site should obey a standard Wright–Fisher model with selection. The heuristic arguments under *Results* suggest that the random site trajectories should be compatible at appropriately coarse timescales with the mean totals for teams predicted by our model.

The model presupposes a picture in which senescence is shaped by individual endowment with a modest sample from a very large stockpile of reasonably rare deleterious alleles. The first test is whether such a stockpile exists. Only recently has it become possible to answer this question clearly in the affirmative. Relevant results from ref. 8 are discussed below. More delicate empirical questions, such as how many of these alleles have effects that vary with age and what the nature and distribution of age specificity might be, have hardly been addressed.

Predictions within our model concern a special class of deleterious alleles—dubbed MA alleles for this discussion—with age-specific effects on adult survival chances held at equilibrium in mutation–selection balance. This special subset of mildly deleterious alleles, although only one part of the broader phenomenon, is perhaps more amenable than others to comparisons with genome studies.

Some empirical findings bear primarily on stocks, others on flows. We first consider stocks. An accounting of SNVs based on full sequencing of coding regions for 2,044 individuals is given in ref. 8. Some of these SNVs occur with sufficiently high frequency to qualify as SNPs. Figure 4*C* of ref. 8 shows more than

75% of functional SNVs per individual with frequencies above 5% in their European population. However, among all SNVs, functional or not, 86% are found with minor allele frequencies less than 0.005. The 23 authors of ref. 8 observe that individuals carry on average 13,959 SNVs out of a pool of 503,481 SNVs detected in their sample. In the pool, 58% of SNVs are nonsynonymous, entailing differences in protein products. The authors apply seven criteria for separating out functional SNVs. Although 47% of SNVs are categorized as deleterious by at least one criterion, overlap is small, and the authors arrive at figures of 318 and 580 for average numbers carried per individual by combined criteria of varying strictness.

By this reckoning, as expected, most SNVs are either altogether neutral or nearly neutral. Deleterious SNVs constitute a small minority. Mildly deleterious alleles are surely more common than any highly deleterious alleles still to be found in an individual. Among mildly deleterious alleles, MA alleles affecting adult mortality are presumably themselves a minority.

Comparisons with mutation accumulation models depend on inferences about genetic load for populations that could still have been in mutation–selection equilibrium. Reductions in mortality within environments altered by processes of civilization leave us far from equilibrium today. An important question thus concerns the number and character of SNVs detected today that derive from mutations occurring long enough in the past to have been shaped by natural selection when a mutation–selection equilibrium could have been in force. The authors of ref. 8 do not report estimated distributions for the antiquity of deleterious alleles, but a variety of methods, reviewed in ref. 34, could be marshalled for the purpose.

The authors of ref. 8, in line with others, attribute the large aggregate numbers of neutral and nearly neutral SNVs to population growth. The demographic model that they fit to their data exhibits an *NRR* for their European population of 1.0038 up to 5,115 y ago and an *NRR* of 1.0195 thereafter, corresponding to a population growth rate a little less than 1 per thousand per year. The later *NRR* for their African population is 1.0166. For the European population, the end-state effective population size quoted at 512,000 implies an effective population size around 10,000 at the estimated time of transition to higher growth, conveniently close to the date of the ancient Egyptian palette of Narmer and the dawn of recorded history.

Demographic scenarios are undergoing rapid refinement as datasets expand (see, e.g., ref. 35). The timing and pace of population growth certainly affect numbers of neutral and nearly neutral alleles. However, the authors of ref. 36, p. 969, observe that "while population growth dramatically increases the number of deleterious segregating sites in the population, it only mildly increases the number carried by each individual." Similarly, Simons et al. (37) find that the "deleterious mutation load is insensitive to recent population history," using simulations that, unlike our model, incorporate back mutation. Overall, on our reading of the literature, the last several hundred generations of population growth do not appear to be a dominant influence on counts per person of the kinds of alleles featured in our mutation accumulation models.

The counts in the hundreds from ref. 8 are for deleterious SNVs per individual detected within the exome, within protein-coding regions of the genome. The counts need to be increased by an as-yet-uncertain allowance for mildly deleterious alleles per individual from other parts of the genome. They need to be decreased by a fraction reflecting the share of MA alleles among all mildly deleterious alleles. These adjustments may be expected to offset each other to some extent. Progress may soon bring greater clarity. In the meantime, working forward from ref. 8, a guess in the hundreds for the stock of MA alleles held in mutation–selection equilibrium seems compatible with what is known so far.

From stocks, we turn to consideration of flows. Estimation of total rates of new mutations in humans is a topic of active research reviewed in ref. 38. See also ref. 39. Rates on the order of 70 mutations per individual per generation have been proposed (ref. 38, p. 299). Most new mutations are altogether neutral or nearly neutral. A small fraction is expected to be mildly deleterious. A portion of those are the MA alleles on which we are focusing. Possible distributions of fitness effects can be studied through simulations as in ref. 40.

When MA alleles on their own are under consideration, as we have mentioned, the demographic structure in our model has the useful feature that a rough estimate of total mutation rate for the class can be ventured. The estimate relies on comparisons of survivorship schedules for a presumed baseline and a possibly observable outcome. It rests on a generalization of the principle of J. B. S. Haldane (41). Haldane's principle applies to linear approximate models, that is to say, it applies in the absence of epistasis, where it dictates that the total mutation rate for relevant alleles should equal the total loss in fitness at equilibrium. This equality breaks down in the face of interactions, but a closed-form generalization turns out to hold in the full nonlinear model, encapsulated in theorem 3 of ref. 2.

An illustrative calculation is described in ref. 42 with parameters specified under *Materials and Methods*. The calculation is informed by the near-contemporary hunter-gatherer life tables reported in ref. 43 and enhanced by an allowance for fitness effects from nurturance on the part of surviving parents and grandparents. It generates an estimate on the order of 13% for loss in fitness from a load of MA alleles, which the generalized Haldane's principle transforms into a rate of 0.20 per individual per generation for new MA mutations. Extreme cases from ref. 43 generate rates for MA mutations no higher than 0.50. These rates are subject to downward revision, the greater the extent to which adaptive life history processes are shaping baseline adult survival and contributing to long-term signatures of senescent mortality. However, pegging MA mutation rates very much lower would be hard to square with the numbers of alleles turning up in genome sequences.

Taken together, an estimate of flow along with an estimate of stock for the special class of MA alleles yield an estimate of average selective cost. For the right-hand side of Eq. **1** to vanish, the average value of the age-specific force of natural selection $F_r(m)$ at equilibrium, averaged over earlier-acting and later-acting alleles, has to equal the quotient of the total mutation rate for MA alleles per individual divided by the average count of MA alleles per individual. For example, 0.20/300 would be 1/1,500, a selective cost qualifying as mildly deleterious when effective population sizes equal or exceed 10,000. Ages of derived alleles at single sites spreading out around as much as 1,500 generations would be plausible.

As we have said, mortality reduction accompanying the march of civilization has transformed present-day manifestations of our genetic legacy. Mutation accumulation models describe equilibria shaped by natural selection when loss of life during ages of reproduction and nurturing was much more severe. With small effect sizes and extended clearance times, frequencies of alleles will be long in adjusting to recently beneficent environments, if beneficent environments are indeed sustained. All strands of biodemography face the need for an account of how patterns shaped by evolution are to be translated into patterns in the demographic schedules of nowadays.

Formally speaking, our mutation accumulation model allows for simple adjustments in level preserving features of shape, by retaining the equilibrium intensity $r_t(m)$ and profiles $\theta(m,x)$ while reducing a scaling parameter η, which governs sizes of effects. Exogenous mortality reflected in the baseline schedule can also be removed. However, this simple expedient is not empirically adequate. Regularities in hazard functions—exponential increases and incipient plateaus—are now seen at ages that are too far

beyond the reach of ages primarily subject to selective pressure. Contributions by grandmothers to their descendants' fitness, along with contributions from parenting and grandparenting and extended male ages of fertility are all now well appreciated, but they do not reach far enough into extreme old age to account for the regularities we continue to see.

A puzzle remains: Why are mortality rates today in human populations at extreme ages not irregular functions of age? All strands of evolutionary demography face this challenge.

Hints of an answer to this question may perhaps be found in what demographers are observing in recent patterns of mortality decline. A thoughtful overview is given in ref. 26. As James Vaupel points out there, it appears that the pace of increase of mortality with age, the senescent component reflected in the slope parameter for Gompertz curves, has been relatively steady over a number of decades. Steady increases in life expectancy have been associated with steady decreases in level. With an exponential hazard curve, sinking the level of the curve is equivalent to sliding the curve along the age axis toward higher ages. Popular sayings like "sixty is the new fifty" actually capture the formal structure of hazard rate changes. The extreme ages above 100 at which plateaus in hazard rates in contemporary humans appear to form are also compatible with the idea that there are processes that have been sliding schedules outward in age.

In such an account, regularities shaped by natural selection at younger, reproductive and nurturing ages would now be seen at older ages, translated into postreproductive life. The study of formal models that might give this idea substance offers a productive direction for future research. Mutation accumulation depends upon a fund of potential mutations, which impinge upon realized demographic schedules and so are shaped by natural selection. These demographic influences must be mediated by influence on biochemical pathways, whose abstract structure could be modeled by effects on transition rates in stochastic vitality models. Reductions in environmental challenges might be represented in terms of a lowered threshold of lethality as well as in terms of exogenous changes in transition rates governing vitality trajectories.

Whether or not this particular approach turns out to be productive, we believe that future progress demands models that combine mutation accumulation with stochastic vitality and ultimately with optimal life history theory. The three approaches described at the beginning of this article, largely pursued along separate lines, offer in concert better hope for understanding the evolutionary shaping of demographic schedules.

Materials and Methods

The ingredients of the model and conditions on the profiles $\theta(m,x)$ and baseline survivorship ℓ_x and fertility f_x are stated under *Results*. The same model is described at more length in ref. 2 and in full detail in ref. 1, where our specification of demographic selective costs is treated in sections 1.4 and 3.9 and existence and uniqueness of a solution $r_t(m)$ satisfying Eq. **1** are established in theorems 2.9 and 2.10 and remark 2.13. This solution is jointly measurable in t and m and continuously differentiable in t for all m. The solution is determined by the choice of formula for the age-specific force of natural selection $F_r(m)$, in our case given by the following:

$$F_r(m) = \int_\alpha^\infty \left(1 - e^{-\eta\theta(m,x)}\right) f_x \, \mathcal{L}_x^{(r)} \, dx. \qquad [2]$$

In other settings, the scaling parameter η can depend on m, but for present purposes we take η fixed, typically at unity. The function $\mathcal{L}_x^{(r)}$ is the aggregate population survivorship function given by the following:

$$\mathcal{L}_x^{(r)} := \mathbb{E}_r[\ell_x(G)] = \ell_x(0)\exp\left(-\int_{\mathcal{M}} \left(1 - e^{-\eta\theta(m',x)}\right) r(m')\, dm'\right). \qquad [3]$$

These formulas are derived from measuring the selective cost of carrying a batch of mutant alleles by the associated decrease in the *NRR* and from taking the genotype of a randomly selected member of the population to be the realization of a Poisson process with intensity $r_t(m)$. (See ref. 2, p. 10,146.) Selective costs are based on the *NRR* rather than on the more familiar parameter Lotka's r, because alleles are not invading the population but are being maintained in an equilibrium.

Useful properties of $r_t(m)$, $\mathcal{L}_x^{(r)}$, and $F_r(m)$ are collected in the following lemma proved in ref. 2, supplementary information, p. 1.

Lemma. *The solution $r_t(m)$ to* Eq. **1** *starting from $r_0 \equiv 0$ is a nondecreasing function of t for every m. The aggregate population survivorship function $\mathcal{L}_x^{r_t}$ is a nonincreasing function of t for every x. The age-specific force of natural selection $F_{r_t}(m)$ is a nonincreasing function of t for every m.*

We now proceed to the proof of the theorem.

Proof: The lemma implies that the intensity r_t increases pointwise on $\mathcal{M}$ to an extended real-valued limit r_*. Similarly, $\mathcal{L}_x^{r_t}$ is nonnegative and nonincreasing in t and converges pointwise in x to some (always finite) limit $\mathcal{L}_x^*$. The limit is nonincreasing as a function of x and, by monotone convergence, may be written as $\mathcal{L}_x^* = \exp(-H_x)\ell_x(0)$, where the genetic contribution H_x to the aggregate cumulative hazard at equilibrium is given by the equation

$$H_x := \int_{\mathcal{M}} \left(1 - e^{-\theta(m',x)}\right) r_*(m')\, dm'.$$

We define the maximum age of survival ω by the expression

$$\omega := \inf\{x : \mathcal{L}_x^{r_t} \downarrow 0\}, \qquad [4]$$

with the usual convention that $\inf\emptyset = \infty$. It follows from the dynamical equation (Eq. **1**) via monotone convergence arguments spelled out in detail in the proof of theorem 1 of ref. 2 that $r_*(m) < \infty$ implies $m \in (\alpha,\omega)$ and on any such interval $q(m) = r_*(m)\, F_{r_*}(m)$. Thus, it suffices to prove $\omega = \alpha$.

We divide our proof by contradiction into two cases, one starting with an assumption that $\alpha < \omega < \infty$ and the other starting with $\alpha < \omega = \infty$. Consider first the case where ω is assumed finite.

We begin by showing that $\lim_{x\uparrow\omega} e^{-H_x} = 0$, which implies $\lim_{x\uparrow\omega} \mathcal{L}_x^* = 0$. The latter limit is obvious if $\theta(m,x)$ and $\ell_x(0)$ are continuous functions of x for each m, because by two applications of monotone convergence we have the equalities

$$\lim_{x\uparrow\omega} \mathcal{L}_x^* = \mathcal{L}_\omega^* = \lim_{x\downarrow\omega} \mathcal{L}_x^* = 0.$$

However, we are not excluding profiles with jumps that might bunch up at ω and raise the threat of a discontinuity at ω in $\mathcal{L}_x^*$.

We rule out such a discontinuity with another set of monotone convergence arguments. At an equilibrium, as we have said, we have $q(m) = r_*(m)\, F_{r_*}(m)$ for all m in (α,ω). Because $(1 - e^{-\theta(m,x)})$ vanishes for x in (α,m), we can write the equilibrium condition in terms of an integral over x in (m, ω):

$$q_0 \le q(m) = \int_m^\omega r_*(m)\left(1 - e^{-\theta(m,x)}\right) f_x \, \mathcal{L}_x^* \, dx. \qquad [5]$$

Thanks to the upper bound on the baseline net maternity function and the inequality $\mathcal{L}_x^* \le \ell_x(0)$, we have the inequality

$$q_0 \le \sup_y(f_y\ell_y(0)) \int_m^\omega r_*(m)\left(1 - e^{-\theta(m,x)}\right) dx. \qquad [6]$$

For each m, the monotone nondecreasing function $(1 - e^{-\theta(m,x)})$ has a limit $u(m)$ defined by the condition

$$u(m) := \lim_{x\uparrow\omega}\left(1 - e^{-\theta(m,x)}\right).$$

Integrating over x and using the upper bound on the integrand provided by u, we have the result

$$\int_m^\omega r_*(m)\left(1 - e^{-\theta(m,x)}\right) dx \le (\omega - m) r_*(m) u(m). \qquad [7]$$

Hence, we obtain the relationship

$$\frac{q_0}{\sup(f_x\ell_x(0))} \frac{1}{\omega - m} \le r_*(m) u(m).$$

For any δ in $(0,\omega-\alpha)$, we integrate both sides over values of m from α to $\omega-\delta$:

$$\frac{q_0}{\sup(f_x\ell_x(0))}\log\left(\frac{\omega-\alpha}{\delta}\right)\le\int_\alpha^{\omega-\delta} r(m)u(m)dm.$$

Letting δ go to zero, we have shown that the integral of $r_*(m)u(m)$ from α to ω is infinite.

By another application of monotone convergence for the equilibrium aggregate cumulative hazard H_x, we find:

$$\lim_{x\uparrow\omega} H_x=\int_\alpha^\omega r_*(m)u(m)dm=\infty.$$

Thus, as $x\uparrow\omega$, we have $e^{-H_x}\downarrow 0$, implying along the way that $\mathcal{L}_x^*$ is continuous at $x=\omega$.

We now suppose that ω is strictly greater than α and derive a contradiction. Thanks to the limit at $x=\omega$ which has just been established, for any $\epsilon>0$ we can find $\delta>0$ such that $x>\omega-2\delta$ implies $e^{-H_x}<\epsilon$, and δ can be chosen so that $\alpha<\omega-\delta$.

For any Borel subset $B\subseteq[\alpha,\omega]$ and each age x, define

$$Y_B(x)=\int_B r_*(m)\left(1-e^{-\theta(m,x)}\right)dm. \quad [8]$$

The function Y is measurable, thanks to the joint measurability of θ, and Y is nondecreasing in x for every B.

When, as in our case, ω is finite, set $z=\omega-2\delta$ and partition the whole of $[\alpha,\omega)$ into three intervals $I=[\alpha,z)$ and $J=[z,z+\delta)$ and $K=[z+\delta,\omega)$. Then $H_x=Y_I(x)+Y_J(x)+Y_K(x)$. When $x\le z$, we have $Y_J(x)=Y_K(x)=0$ so that $H_x=Y_I(x)$.

At an equilibrium, as before, we have $q(m)=r_*(m)\,F_{r_*}(m)$ for all m in (α,ω), which means that $q(m)$ equals the integral

$$\int_\alpha^\omega r_*(m)\left(1-e^{-\theta(m,x)}\right)e^{-Y_I(x)}e^{-Y_J(x)}e^{-Y_K(x)}\,\ell_x(0)f_x\,dx. \quad [9]$$

We concentrate on the middle interval J of length δ, integrate both sides over m in J, and use the nonnegativity of the integrands to justify reversing the order of integrations over x and m, obtaining a factor of $Y_J(x)$ inside the integral with respect to x:

$$\int_J q(m)dm=\int_\alpha^\omega Y_J(x)e^{-Y_I(x)}e^{-Y_J(x)}e^{-Y_K(x)}\,\ell_x(0)f_x\,dx. \quad [10]$$

On the right, the lower limit of integration α can be replaced by z, because $Y_J(x)$ vanishes for smaller values of x. We have chosen δ so that the factor $e^{-Y_I(x)}=e^{-H_x}$ is smaller than ε over this range of integration. The product $Y_J(x)\,e^{-Y_J(x)}$ is bounded above by $1/e=\sup(y\exp(-y))$. The factor $e^{-Y_K(x)}$ is always less than 1. The range of integration (z,ω) has length $2\,\delta$. The right-hand side of Eq. **10** is therefore bounded above by $(\epsilon)(1/e)(\sup(f_x\ell_x(0))(2\,\delta)$. The left-hand side is bounded below by $q_0\,\delta$. Dividing out δ, we find that

$$q_0<\epsilon(2\sup(f_x\ell_x(0))/e).$$

We can therefore choose ε small enough to make the right-hand side of Eq. **10** strictly smaller than the left-hand side, contradicting the equilibrium condition. Thus, we conclude that $\omega<\infty$ implies $\omega=\alpha$.

We now turn to the case with $\omega=\infty$. For this case, we choose any fixed $\delta>0$, and for any $\epsilon>0$ we use the integrability of $\ell_x(0)f_x$ to choose z large enough so that $\int_z^\infty \ell_x(0)f_x\,dx<\epsilon$. We bound $\exp(-Y_I(x))$ by unity. Eq. **10** then forces $q_0\,\delta<\epsilon/e$, and for fixed δ and small enough ε, we again have a contradiction. We conclude that ω must be finite and so, as before, equal to the minimum age of reproduction α as claimed.

Q.E.D.

The generalized form of Haldane's principle from theorem 3 of ref. 2 may be applied using a formula given there in supplementary equation S8:

$$\int q(m)dm=\int H_x e^{-H_x}\,f_x\ell_x(0)dx. \quad [11]$$

The illustrative calculation mentioned under *Discussion* is based on the estimated average hunter-gatherer life table in ref. 43. The function H_x is set equal to $\int_\alpha^x(0.013\zeta+0.000147\exp(0.086y))dy$ and $\ell_x(0)=\exp(-(1-\zeta)(0.013)(x-\alpha))$. Here, $\zeta=0.50$ is the share of age-independent mortality attributed to genetic load and $\alpha=13$. The f_x schedule is proportional to the density for a gamma distributed random variable with shape parameter 2 and scale parameter 9 shifted to start at age α and normalized so that $\int f_x\,\mathcal{L}_x^*\,dx=1$. This schedule gives good fit at reproductive ages to a standard Coale–Trussell schedule for natural fertility (ref. 30, section 6.8) along with a tapering contribution for benefits to children and grandchildren from surviving parents and grandparents. Motivation and discussion may be found in ref. 42; future research should refine the picture.

ACKNOWLEDGMENTS. We thank Iain Mathieson for illuminating discussions, the referees for helpful suggestions, and Richard Suzman for long nurturing of biodemography. The research of K.W.W. has been supported by the University of California Berkeley Center for the Economics and Demography of Aging under Grant 2P30AG012839-18 from the National Institute on Aging. The research of S.N.E. has been supported in part by Grant DMS-0907630 from the National Science Foundation and Grant 1R01GM109454-01 from the National Institutes of Health.

1. Evans SN, Steinsaltz DR, Wachter KW (2013) *A Mutation-Selection Model for General Genotypes with Recombination*, Memoir 222 (American Mathematical Society, Providence, RI).
2. Wachter KW, Evans SN, Steinsaltz DR (2013) The age-specific force of natural selection and biodemographic walls of death. *Proc Natl Acad Sci USA* 110(25):10141–10146.
3. Medawar P (1952) *An Unsolved Problem in Biology: An Inaugural Lecture Delivered at University College, London, 6 December, 1951* (H. K. Lewis, London).
4. Finch C (1990) *Longevity, Senescence, and the Genome* (Univ of Chicago Press, Chicago).
5. Bürger R (2000) *The Mathematical Theory of Selection, Recombination, and Mutation* (Wiley, Chichester, UK).
6. Charlesworth B (1994) *Evolution in Age-Structured Populations* (Cambridge Univ Press, Cambridge, UK).
7. Charlesworth B (2000) Fisher, Medawar, Hamilton and the evolution of aging. *Genetics* 156(3):927–931.
8. Tennessen JA, et al.; Broad GO; Seattle GO; NHLBI Exome Sequencing Project (2012) Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 337(6090):64–69.
9. Vaupel JW, Manton KG, Stallard E (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16(3):439–454.
10. Yashin AI, Manton KG (1997) Effects of unobserved and partially observed covariate processes on system failure: A review of models and estimation strategies. *Statistical Science* 12(1):20–34.
11. Le Bras H (1976) Lois de mortalité et âge limite. *Population (Paris)* 31(3):655–691.
12. Aalen OO, Gjessing HK (2001) Understanding the shape of the hazard rate: A process point of view. *Stat Sci* 16(1):1–22.
13. Weitz JS, Fraser HB (2001) Explaining mortality rate plateaus. *Proc Natl Acad Sci USA* 98(26):15383–15386.
14. Steinsaltz D, Evans SN (2004) Markov mortality models: Implications of quasistationarity and varying initial distributions. *Theor Popul Biol* 65(4):319–337.
15. Kirkwood T (1977) Evolution of ageing. *Nature* 270(5635):301–304.
16. Kaplan H (1997) The evolution of the human life course. *Between Zeus and the Salmon: The Biodemography of Longevity*, eds Wachter KW, Finch CE (National Academy Press, Washington, DC), pp 175–211.
17. Kaplan HS, Robson AJ (2002) The emergence of humans: The coevolution of intelligence and longevity with intergenerational transfers. *Proc Natl Acad Sci USA* 99(15):10221–10226.
18. Robson AJ, Kaplan HS (2003) The evolution of human life expectancy and intelligence in hunter-gatherer economies. *Am Econ Rev* 93(1):150–169.
19. Baudisch A (2008) *Inevitable Senescence? Contributions to Evolutionary-Demographic Theory* (Springer, Berlin).
20. Chu CY, Lee RD (2006) The co-evolution of intergenerational transfers and longevity: An optimal life history approach. *Theor Popul Biol* 69(2):193–201.
21. Carey JR, Liedo P, Orozco D, Vaupel JW (1992) Slowing of mortality rates at older ages in large medfly cohorts. *Science* 258(5081):457–461.
22. Curtsinger JW, Fukui HH, Townsend DR, Vaupel JW (1992) Demography of genotypes: Failure of the limited life-span paradigm in *Drosophila melanogaster*. *Science* 258(5081):461–463.
23. Vaupel JW, Carey JR (1993) Compositional interpretations of medfly mortality. *Science* 260(5114):1666–1667.
24. Vaupel JW, et al. (1998) Biodemographic trajectories of longevity. *Science* 280(5365): 855–860.
25. Wachter KW, Finch CE, eds (1997) *Between Zeus and the Salmon: The Biodemography of Longevity* (National Academy Press, Washington, DC).
26. Vaupel JW (2010) Biodemography of human ageing. *Nature* 464(7288):536–542.
27. Agrawal AF, Whitlock MC (2012) Mutation load: The fitness of individuals in populations where deleterious alleles are abundant. *Annu Rev Ecol Evol Syst* 43:115–135.
28. Charlesworth B (2001) Patterns of age-specific means and genetic variances of mortality rates predicted by the mutation-accumulation theory of ageing. *J Theor Biol* 210(1):47–65.
29. Kimura M, Maruyama T (1966) The mutational load with epistatic gene interactions in fitness. *Genetics* 54(6):1337–1351.

30. Wachter KW (2014) *Essential Demographic Methods* (Harvard Univ Press, Cambridge, MA).
31. Gjessing HK, Aalen OO, Hjort NL (2003) Frailty models based on Lévy processes. *Adv Appl Probab* 35(2):532–550.
32. Danko MJ, Kozlowski J, Vaupel JW, Baudisch A (2012) Mutation accumulation may be a minor force in shaping life history traits. *PLoS One* 7:e34146.
33. Steinsaltz D, Evans SN, Wachter KW (2005) A generalized model of mutation-selection balance with applications to aging. *Adv Appl Math* 35(1):16–33.
34. Slatkin M, Rannala B (2000) Estimating allele age. *Annu Rev Genomics Hum Genet* 1: 225–249.
35. Gazave E, et al. (2014) Neutral genomic regions refine models of recent rapid human population growth. *Proc Natl Acad Sci USA* 111(2):757–762.
36. Gazave E, Chang D, Clark AG, Keinan A (2013) Population growth inflates the per-individual number of deleterious mutations and reduces their mean effect. *Genetics* 195(3):969–978.
37. Simons YB, Turchin MC, Pritchard JK, Sella G (2014) The deleterious mutation load is insensitive to recent population history. *Nat Genet* 46(3):220–224.
38. Keightley PD (2012) Rates and fitness consequences of new mutations in humans. *Genetics* 190(2):295–304.
39. Kong A, et al. (2012) Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* 488(7412):471–475.
40. Kousathanas A, Keightley PD (2013) A comparison of models to infer the distribution of fitness effects of new mutations. *Genetics* 193(4):1197–1208.
41. Haldane JBS (1937) The effect of variation on fitness. *Am Nat* 71(735):337–349.
42. Wachter KW (2014) Alleles, mortality rates, and the evolutionary theory of senescence. *Sociality, Hierarchy, Health: Comparative Biodemography*, ed Weinstein M (National Academy Press, Washington, DC).
43. Gurven M, Kaplan H (2007) Longevity among hunter-gatherers: A cross-cultural examination. *Popul Dev Rev* 33(2):321–365.