Cellular senescence in the Penna model of aging

Avikar Periwal*
Montgomery Blair High School, 51 University Boulevard East, Silver Spring, Maryland 20901, USA
(Received 24 January 2013; revised manuscript received 10 July 2013; published 5 November 2013)

Cellular senescence is thought to play a major role in age-related diseases, which cause nearly 67% of all human deaths worldwide. Recent research in mice showed that exercising mice had higher levels of telomerase, an enzyme that helps maintain telomere length, than nonexercising mice. A commonly used model for biological aging was proposed by Penna. I propose a modification of the Penna model that incorporates cellular senescence and find an analytical steady-state solution following Coe, Mao, and Cates [Phys. Rev. Lett. **89**, 288103 (2002)]. I find that models corresponding to delayed cellular senescence have younger populations that live longer. I fit the model to the United Kingdom's death distribution, which the original Penna model cannot do.

DOI: 10.1103/PhysRevE.88.052702 PACS number(s): 87.23.−n

I. INTRODUCTION

From early alchemists looking for the elixir of life, to modern-day researchers, humans have always wanted to understand aging. As people age, their cells go through replication cycles. Each replication reduces the length of the telomeres in the cells. If a cell's telomeres are too short, it may not be able to replicate [1]. Cells that can no longer replicate are termed senescent. Cellular senescence is thought to play a major role in age-related diseases, which cause nearly 67% of all human deaths worldwide [2]. Recent research in mice showed that exercising mice had higher levels of telomerase, an enzyme that helps maintain telomere length, than nonexercising mice [3]. In humans, runners had longer telomeres than nonrunners [3]. Since shortened telomeres are thought to be related to death, this research would seem to indicate that people who exercise live longer lives. Population studies do indeed show this; however, most studies show that only the mean lifespan increases in exercising populations, not maximum lifespan [4,5].

A commonly used model for biological aging was proposed by Penna in 1995 [6,7]. The model looks at death from a mutation accumulation standpoint. In the past 10 years, papers have been published describing methods for finding the age distributions created by the Penna model without actually simulating the model [8–10]. However, the research on cellular senescence indicates that age-related death is not caused by mutation accumulation but rather by an inability to reproduce after a number of cycles . I propose a modification of the Penna model that uses this mechanism instead of mutation accumulation and show that a mean lifespan can increase without affecting the maximum lifespan, which cannot be done by changing parameters in the original Penna model. Furthermore, I show that this modification allows the Penna model to fit actual data.

II. MODEL

The Penna model assigns a bit string to each individual in the population. Each bit corresponds to a timestep of the simulation. A 1 in the string represents a mutation, and a 0 means no mutation. If an individual has gone through T 1s, then it dies. Each individual can have offspring, with probability b. The child's bit string is derived from the bit string of the parent, where each 0 has probability $(1 - e^{-\beta})$ of becoming a 1. The Penna model ignores positive mutations, because they are rare. The length of the bit string provides a hard limit for lifespan.

In the proposed modification, which we will call the senescent Penna (SP) model, each individual can only get one disease, which is essentially the beginning of aging. After the individual starts to age, it has probability p of staying alive to reproduce at each time step. The maximum number of replication cycles is M, so an individual can stay alive to reproduce for up to M future cycles. People who exercise generally have longer telomeres, so once cellular senescence starts, they have a smaller probability of dying. Higher values of p could represent exercising populations.

The first 1 in an individual's bit string is the age, l, at which cellular senescence begins. The number of people alive at time j, with age x, is $n_j(x,l,m) = pn_{j-1}(x-1,l,m-1)$ if $x > l$, where p is the probability of living after cellular senescence, and m is the time since the inception of cellular senescence. If $x \leqslant l$, then $n_j(x,l) = n_{j-1}(x-1,l)$. If b is the probability of birth, and $e^{-\beta}$ is the probability of an individual going unmutated, then the number of children born in the next time step with disease acquisition time l, $n_{j+1}(0,l,0)$ is given by

$$n_{j+1}(0,l,0) = be^{-\beta l}\sum_x\sum_m n_j(x,l,m) + (1-e^{-\beta})be^{-\beta l}\sum_{l'>l}\sum_x\sum_m n_j(x,l',m). \quad (1)$$

Since m is just the maximum of 0 and $x - l$,

$$\sum_x\sum_m n_j(x,l,m) = \sum_{x=0}^{l-1} n_j(x,l) + \sum_{x=l}^{M+l-1} n_j(x,l), \quad (2)$$

where M is the maximum number of replication cycles allowed after disease. Assuming a steady state, $n_{j+1}(x,l) = n_j(x,l) \equiv n(x,l)$. Therefore,

$$\sum_{x=0}^{l-1} n(x,l,0) = l \cdot n(0,l,0), \quad (3)$$

*avperiwa@mbhs.edu

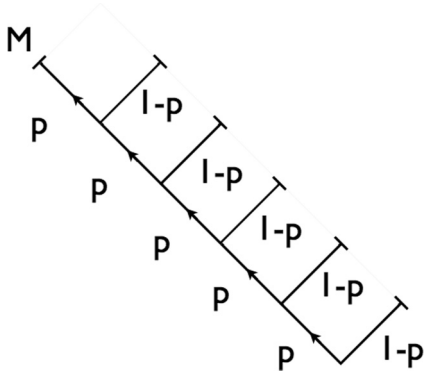

FIG. 1. The possible paths until death, after reaching l, in the SP.

since each x has the same number of people. Since $n_j(x,l) = pn_{j-1}(x-1,l)$ for $x > l$, and since $n_{j+1}(x,l) = n_j(x,l)$ in a steady state,

$$\sum_{x=l}^{M+l-1} n(x,l) = n(0,l)\sum_{x=1}^{M} p^x = p\frac{1-p^M}{1-p}n(0,l). \quad (4)$$

Defining q_l as

$$q_l = l + p\frac{1-p^M}{1-p}, \quad (5)$$

and $n(l) = n(0,l)$, Eq. (1) can be simplified to

$$0 = be^{-\beta l}n(l) - \frac{n(l)}{q_l} + (1-e^{-\beta})be^{-\beta l}\sum_{l'>l+1} n(l'). \quad (6)$$

Writing the same equation for $l+1$, some algebra leads to

$$n(l+1) = n(l)\frac{be^{-\beta l} - 1/q_l}{be^{-\beta(l+1)} - e^{\beta}/q_{l+1}}. \quad (7)$$

This equation leads to some limiting cases, in order to maintain a steady state. Neither the numerator nor the denominator should vanish in Eq. (7).

$$q_{\max} < \frac{1}{1-e^{-\beta}}, \quad (8)$$

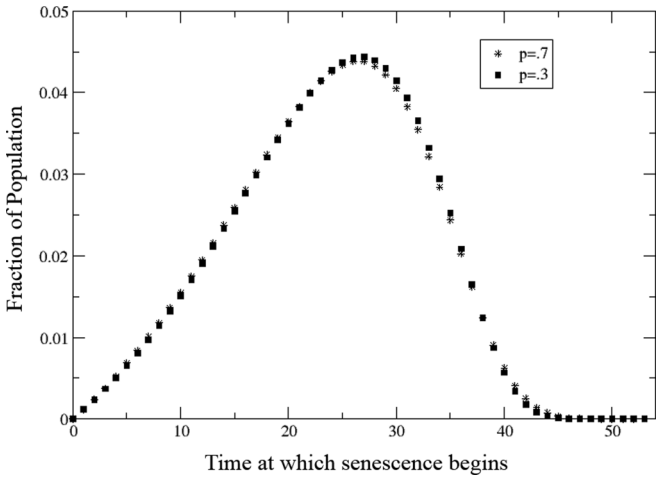

FIG. 2. Higher values of p result in earlier times of cellular senescence in the SP. In this picture, $M = 5$. Age in constant birthrate units (CBRU).

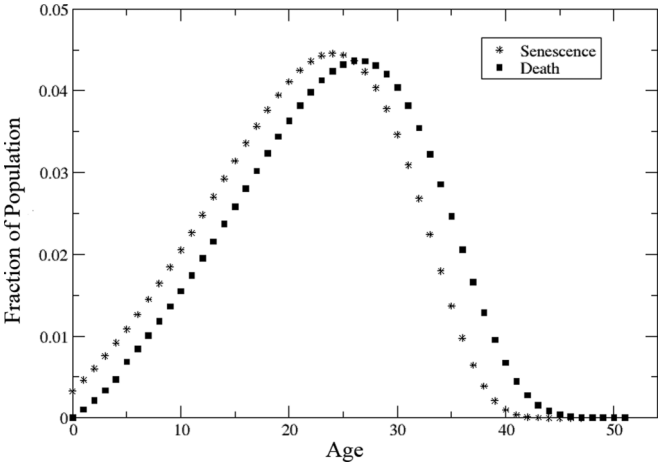

FIG. 3. The death curve in the SP is a shift of the cellular senescence curve, with the amount shifted varying on p. Here, $M = 5$ and $p = 0.8$. Age in CBRU.

which means that the latest age a person can get a mutation, $l_{\max}$, is given by

$$l_{\max} = q_{\max} - p\frac{1-p^M}{1-p}. \quad (9)$$

Note that $l_{\max}$ is not the maximum age an individual can live to but rather the latest age of senescence. The maximum lifespan is given by $l_{\max} + M$. This also sets the birthrate,

$$b = \frac{1}{q_{\max}e^{-\beta l}}. \quad (10)$$

However, Eq. (7) only gives the time of cellular senescence, not the lifespan. In the original Penna model, at l the individual dies, but here the individual has only a $(1-p)$ probability of dying. The number of individuals who die at age t is just

$$D(t) = p^M n(t-M) + \sum_{x=t-M+1}^{t} (1-p)p^{t-x}n(x), \quad (11)$$

assuming $t \geqslant M$. The first term ensures that when $t = l + M$, all the remaining people alive with $n(l)$ die (Fig. 1).

III. RESULTS

A. Model

Interestingly, in the SP model higher values of p show a "younger population" (Fig. 2). With lower p values, an individual with low l will not have as many opportunities to reproduce, since chances are it will die out soon. This provides an evolutionary pressure for higher values of l. However, if p is high, then individuals with low l can still reproduce.

While cellular senescence begins later for populations with lower p, death comes earlier once cellular senescence is reached (Fig. 3). The average death time for $n(l)$ is $l + (1-p) + 2p(1-p) + 3p^2(1-p) + \cdots + (M+1)p^M$, where the last term has no $1-p$ factor, since everybody left alive has to die. Higher values of p result in later deaths, as would be expected. The later deaths are more apparent after $l_{\max}$, when any living members of the population are undergoing cellular senescence. Figures 4 and 5 show the

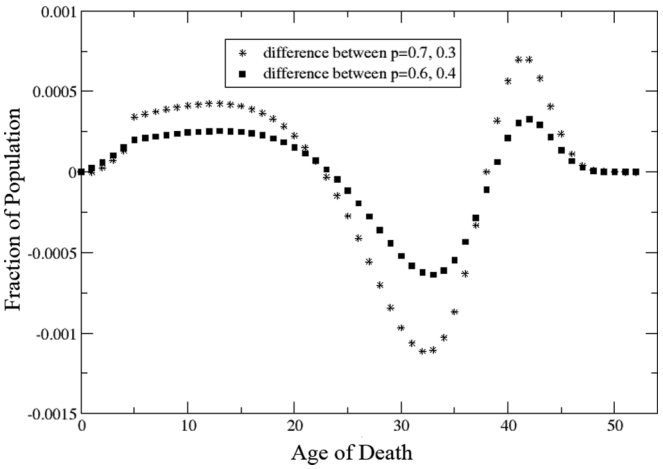

FIG. 4. Higher values of p have a higher proportion of their population reach larger ages, but that proportion difference is countered in the middle of the death distribution. This graph shows the differences in the death distributions at each time unit for different values of p and $M = 5$. Age in CBRU.

differences in percentages, which have to sum to 0. The increased proportion of older individuals for higher values of p has to be balanced by a reduced proportion of the population at lower ages. The jump in the differences of death proportions early in Figs. 4 and 5 is caused by M. For larger values of p, the $1 - p$ term in Eq. (10) is small enough to make the p^{t-x} term negligible. However, once M is reached, the final term has no $(1 - p)$ factor, causing the jump.

The SP model does not show an increase in the mean for higher values of p, but it does show that the probability of living to a higher age is greater. Even though the probability of living to a high age is greater, the maximum age for both populations with higher and lower p is still $l_{\max} + M$. Higher values of M push the time of cellular senescence further forward, since there is less evolutionary incentive for an individual to have a higher l.

However, just like p, higher values of M also afford a longer time until death, balancing out the earlier cellular senescence

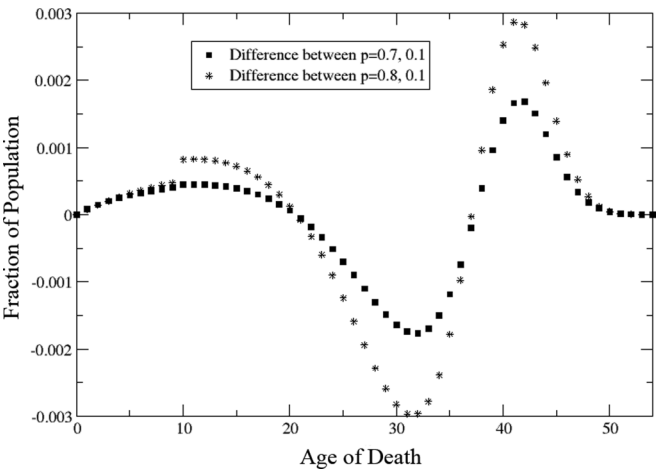

FIG. 5. The jump in differences of the SP death distributions with different values of p in Fig. 4 is also observed here, except $M = 10$. The jump always occurs at position M. Age in CBRU.

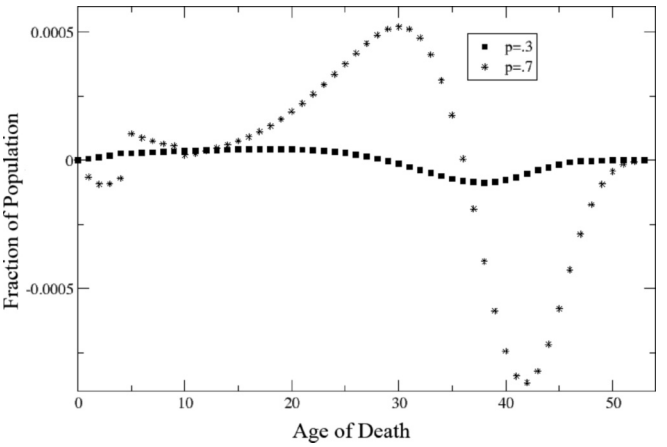

FIG. 6. The death distributions for different values of p look approximately the same, but there is a slightly higher proportion of people alive at later times for higher values of p. Age in CBRU.

times (Fig. 6). Since higher values of M allow for a longer life, the proportion of people alive at a higher age is greater for higher values of M. The breaks in Fig. 7 are caused by the same mechanism as the breaks in Figs. 4 and 5, except instead of the deaths coming at one specific M, they come at the values chosen for the simulations: $M = 5$ and $M = 10$.

B. Fitting the model

I will compare the SP model and the Penna model to data from the United Kingdom, averaged from 1981 to 2011 [11]. Since the time step in the Penna model is completely unspecified, the model should fit any constant birthrate data. However, humans do not have a constant birth rate (see Fig. 8). The number of deaths per unit time (Fig. 9), which in this case is deaths per year, needs to be transformed into deaths per equal birth rate unit. Years are an astronomical time unit that have no significance in actual human life. We transform the death distribution in years to a distribution in equal birthrate units. To find an equal birthrate unit, the cumulative distribution function of the birthrate should be split into equal intervals

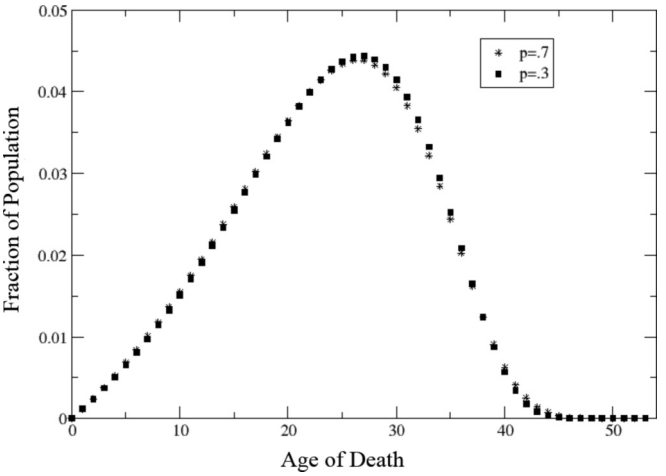

FIG. 7. The differences in the death distribution curves of the SP of $M = 5$ and $M = 10$, for different values of p. Age in CBRU.

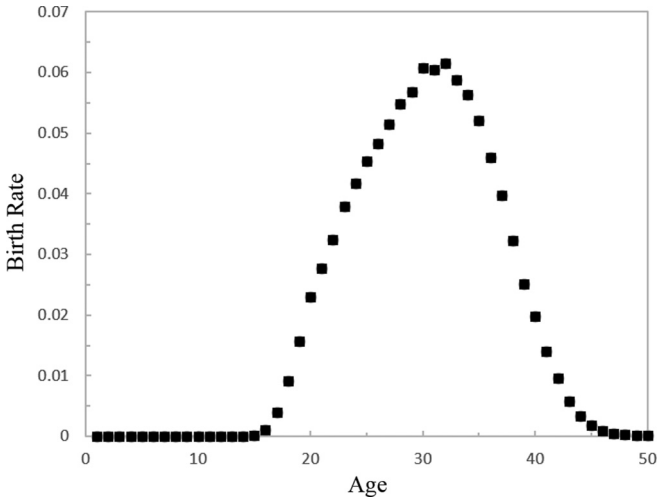

FIG. 8. The normalized birthrate (sums to 1) of the United Kingdom averaged from 1981 to 2011. Age in years.

along the birthrate axis. This creates bins with an equal number of births. Finding the birthrate adjusted death distribution is then just a matter of integrating the nonadjusted death distribution between the end points of the equal birth bins. The number of bins is limited by the number of data points.

Notice that in Fig. 10, the number of deaths stays very close to 0 before dramatically rising at the last point. This is because the majority of deaths occur after the age of 50, while nearly all births occur before 50. Since there are no births after the age of 50, the majority of deaths are always placed in the last birth bin, regardless of the number of bins. My simulations confirmed this. The original Penna model cannot fit this sudden jump.

For the first T time units, the original Penna model has no deaths. The Penna model assumes a bit string, so individuals cannot pick up two mutations at once. It therefore takes at least T time units before an individual can die. The recursive form for computing the death distribution of the Penna model

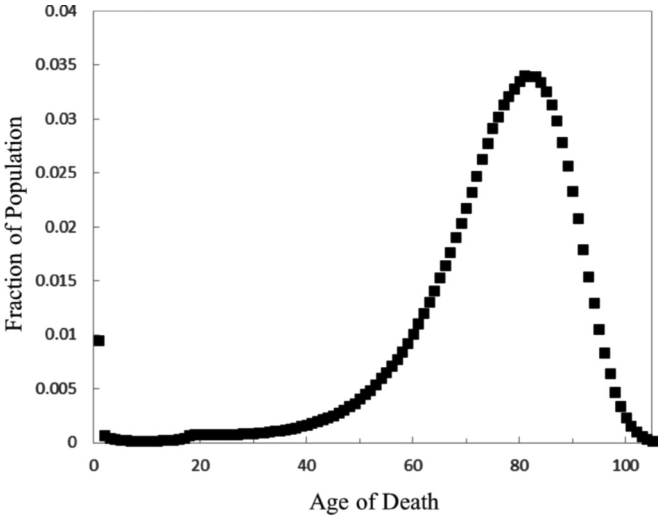

FIG. 9. The fraction of the population that died at a given age in the United Kingdom, averaged from 1981 to 2011. Age in years.

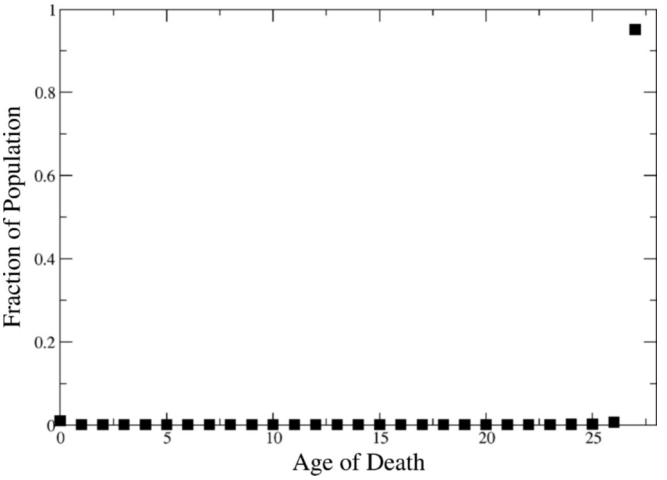

FIG. 10. The United Kingdom's death distribution, in 27 CBRU.

is given by

$$n(l+1) = n(l)\frac{(l+1)(e^{\beta(l-T+1)} - bl)}{l(e^{\beta(l-T+1)} - b(l+1)(1-T+Te^{-\beta})} \quad (12)$$

in Ref. [10].

The best fits to the data using the original Penna model come when $T = l_{\max}$. This means that the best fit from the Penna model predicts that every member of the population dies at $l_{\max}$. This is a trivial solution. There are no population dynamics if everybody dies at the same time. If $T < l_{\max}$, then the Penna model fails to even come close to the actual data.

The SP model is able to fit the observed data well. p ensures that there are nonzero deaths for the first $T-1$ time points, while M ensures that the number of deaths jumps dramatically at time T. The solutions that minimized the sum of the mean-squared residuals had high values of p, and $M = l_{\max} - 1$ (Fig. 11).

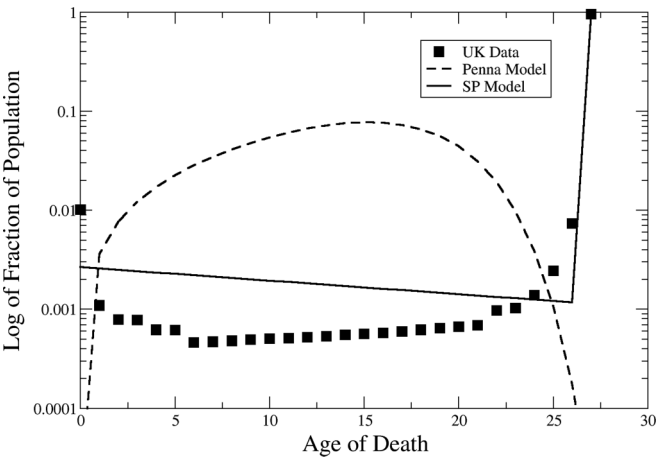

FIG. 11. A comparison of the SP and Penna models with data from the United Kingdom. Parameters were chosen that minimized the sum of the squared residuals. The sum for the SP model was 0.00014, and 0.95994 for the Penna model. Age in CBRU.

IV. CONCLUSION

In this paper, I showed that simple modifications to the Penna model allow for shifts in the lifespan distribution without changing the maximum lifespan. Higher values of p in the SP model result in younger populations, but they die later.

Thinking about the original context for this modification, higher values of p can represent exercising populations. The maximum age of two populations with different p and the same M is the same, but the exercising population has a higher chance of reaching later ages. It is intriguing that this model shows that the exercising populations have a longer period of senescence, if higher p corresponds to exercising populations. The average time of cellular senescence for high p is lower than low p, yet the probability of reaching a high age is greater.

The Penna model is a tool used to help us understand population dynamics. My modifications of the Penna model take into account cellular senescence of a population, which is a critical part of the aging process, and help to explain the changes in lifespan observed in exercising and nonexercising populations. By adjusting $e^{-\beta}, p$, and M, the SP model fits observed data significantly better than the Penna model. Further improvements to the models could take into account recent research in autism which suggests that $e^{-\beta}$ is actually a function of time [12], and also looking at positive mutations. The birthrate reparametrization could be applied to exercising and nonexercising populations, provided that birthrate data was available.

[1] U. Herbig, W. A. Jobling, B. P. Chen, D. J. Chen, and J. M. Sedivy, Mol. Cell **14**, 501 (2004).
[2] A. D. N. J. de Grey, Studies Ethics, Law, Technol. **1**, 1941(2007).
[3] C. Werner, T. Fuerster, T. Widmann, J. Pöss, C. Roggia, M. Hanhoun, J. Scharhag, N. Buechner, T. Meyer, W. Kindermann, J. Haendeler, M. Boehm, and U. Laufs, Circulation **120**, 2438 (2009).
[4] J. Holloszy, J. Appl. Physiol. **82**, 399 (1997).
[5] S. Sarna, T. Sahi, M. Koskenvuo, and J. Kaprio, Med. Sci. Sports Exercise **25**, 237 (1993).
[6] T. Penna, J. Stat. Phys. **78**, 1629 (1995).
[7] D. Stauffer, Bioinformat. Biol. Insights **1**, 91 (2007).
[8] J. B. Coe and Y. Mao, Phys. Rev. E **67**, 061909 (2003).
[9] J. B. Coe and Y. Mao, Phys. Rev. E **69**, 041907 (2004).
[10] J. B. Coe, Y. Mao, and M. E. Cates, Phys. Rev. Lett. **89**, 288103 (2002).
[11] Office for National Statistics, *Population Estimates for England and Wales, Mid-2011 (2011 Census-based)*, Tech. Rep. (Government of the United Kingdom, 2012).
[12] A. Kong, M. L. Friggeand, G. Masson, S. Besenbacher, P. Sulem, G. Magnusson, S. A. Gudjonsson, A. Sigurdsson, A. Jonasdottir, A. Jonasdottir, W. S. W. Wong, G. Sigurdsson, G. B. Walters, S. Steinberg, H. Helgason, G. Thorleifsson, D. F. Gudbjartsson, A. Helgason, O. T. Magnusson, U. Thorsteinsdottir, and K. Stefansson, Nature (London) **488**, 471 (2012).