Cite this article as: Tarkhov A.E., Denisov K.A., & Fedichev P.O. (2024). Aging Clocks, Entropy, and the Challenge of Age Reversal, *AgingBio*, *2*, e20240031. doi: 10.59368/agingbio.20240031.

Research Paper

Aging Clocks, Entropy, and the Challenge of Age Reversal

Andrei E. Tarkhov,[1,2,*] Kirill A. Denisov,[1] and Peter O. Fedichev[1,*]
[1]Gero PTE, Paya Lebar Square, Singapore
[2]Present address: Retro Biosciences Inc., Redwood City, CA, USA
[*]Corresponding authors: at@gero.ai; pf@gero.ai
https://doi.org/10.59368/agingbio.20240031
Received: 2/28/2023, Revised: 6/3/2024, Accepted: 6/12/2024, Published: 8/12/2024

The ready availability of large longitudinal datasets, such as the UK Biobank, enables analyses of complex aging traits at previously unattainable levels. We analyze the aging signatures of DNA methylation and longitudinal electronic medical records from the UK Biobank and demonstrate that their dynamics can be recapitulated by rare and independent stochastic transitions among numerous metastable states, so that the accumulated effect of aging changes can be captured by a single stochastic variable, termed thermodynamic biological age (tBA), in agreement with other aging omics. In the proposed theoretical model, tBA increases linearly with age, tracks the entropy produced (and hence information lost) during the aging process, and causes an irreversible drift in physiological state variables, reduced resilience, and an exponential acceleration of the incidence of chronic diseases and mortality risks. The entropic nature of aging drift may constrain the possibility of complete age reversal and highlights important distinctions between aging in humans and mice, thus necessitating a re-examination of strategies for engineering negligible human senescence.

Introduction

Aging is a complex process manifesting itself across different organismal levels (see hallmarks of aging[1]) and leading to the exponential acceleration of the incidence of chronic diseases[2] and mortality[3]. It is both practically and intellectually appealing to reduce the effects of the multitude of phenotypic changes to a few, or, even better, a single actionable indicator, mostly referred to as "biological age" (BA). BA models can be trained to predict the chronological age or mortality risks of an individual from different sources of biomedical data, ranging from DNA methylation (DNAm)[4–14] to physical activity records from wearable devices[15,16]. Excessive BA (or BA acceleration) is associated with all-cause mortality as well as the prevalence, future incidence, and severity of chronic[10,17,18] and transient diseases, such as COVID-19[16,19–21]. No wonder, BA predictors have increasingly gained traction in clinical trials[22–24].

The dynamic properties of BA and the exact relation between BA variation and aging are not entirely understood. For example, DNAm age may increase without an appreciable increase in all-cause mortality in negligible senescent species[25,26]. Moreover, even in the most healthy individuals, BA levels can transiently change throughout the day following circadian rhythms[27] or in response to stress factors and lifestyle choices such as smoking[18,28]. The characteristic time required for an organism state to relax to homeostatic equilibrium and the range of BA fluctuations progressively increase as a function of age[18]. The number of individuals exhibiting slow recovery increases exponentially and doubles approximately every 8 y, which is close to the mortality doubling time in humans[16]. Further applications of BA models in aging research and medicine require a better understanding of the dynamics and causal relation between, on the one hand, underlying biological and physiological variations of the organism state captured by various BA indicators and, on the other hand, mortality, prevalence and severity of diseases, and the effects of medical interventions.

To address these fundamental questions, we reviewed the universal features of aging signatures in biomedical data. We performed a principal component analysis (PCA) in a large cross-sectional white-blood-cell DNAm dataset[29] and the longitudinal electronic medical records (EMRs) from the UK Biobank[30]. In both cases, aging dynamics can be recapitulated by rare and independent stochastic transitions among numerous metastable states of the methylation of individual CpG sites (5'-C-phosphate-G-3' sequence of nucleotides) or the incidence of specific diseases during life. At the same time, most of the variance in the data could be explained by a single factor linearly increasing with age and demonstrating the strongest correlation with Horvath's DNAm age or the number of chronic diseases in the DNAm and EMR datasets, respectively.

To explain the dynamics behind the universally observed aging signatures, we put forward a semiquantitative model of aging in a complex regulatory network. We assumed that living systems are collections of a vast number of interacting functional units (FUs) that are initialized to a metastable state at the end of development. Aging then results from relaxation of the organism state toward equilibrium through a sequence of stochastic transitions, representing microscopic state changes in all FUs.

Dynamically accessible states are countless, as well as the number of stochastic transitions among them. Hence, the accumulated

effect of random transitions on individual biological processes can be quantified by a stochastic variable with a linearly increasing mean and variance. The quantity progressively increases over time in a sufficiently large regulatory network and hence may emerge as a natural aging clock—the thermodynamic BA (tBA). We argue that tBA is thus the fundamental aging variable. It is best associated with the dominant principal component (PC) score in biomedical data and Horvath's methylation clock. It is proportional to configuration entropy and quantifies information lost during aging.

Materials and Methods

PCA of the DNAm data

We took the white-blood-cell methylation data from the GSE87571 dataset[29]. It contains 729 samples (more than 440k features each), collected from patients of both genders (341 males and 388 females), covering the age range between 14 and 94 y.

To focus the analysis on aging, we filtered out patients younger than 20 y old (620 samples remaining). We filtered out CpG sites according to Pearson's correlation between the DNAm levels and the chronological age at the level of $p < 0.005/\mathrm{N}$ (where N is the total number of the reported features), thus obtaining 96,536 sites. We performed and reported the results of the PCA on the resulting data.

We computed Horvarth's methylation age[4], a few CpG sites (cg17099569, cg00431549, cg11025793, and cg14409958) were not present in the data, and hence we had to exclude them from the calculation.

DNAm-PC3 increased with age at a rate faster than linear. We collected all the pairs of the DNAm-PC3 scores and the chronological age for every patient n in the dataset and used the available age range to produce a fit of the data to average from Equation (8):

$$\mathrm{DNAm-PC3}^n \approx \frac{a}{\mathrm{t_{max}} - \mathrm{t}^n} + \mathrm{b}\cdot \mathrm{t}^n + c \qquad (1)$$

with the uniform Gaussian error and $\mathrm{t_{max}}$, a, b, and c being the parameters of fit. The calculation returned $\mathrm{t_{max}} = 129.9$ y. We also performed the linear fit of the inverse variance of DNAm-PC3 and obtained 90% CI [114.5,122.2] for $\mathrm{t_{max}}$.

Gene set enrichment analysis

We collected the CpG sites best associated with DNAm-PC1 and DNAm-PC3 according to the values of the respective vector components. We retrieved the gene IDs from Illumina's 450k methylation arrays documentation. Finally, we performed gene ontology and disease ontology enrichment with the help of the R "clusterProfiler4.0" package[31].

Preprocessing of EMRs from the UK Biobank

To avoid using the disease labels corresponding to the transient diseases, we selected 111 chronic diseases diagnoses using Chronic Condition Indicators for ICD-10[32]. Overall, 389,494 patients are included in the EMR dataset, mostly of Caucasian origin (366,715 or 94%), of both sexes (179,032 males and 210,462 females) in the age range of 38–74y.

Entropy/entropy production rate determination

In the DNAm dataset, we computed the configuration entropy as the mean Shannon entropy over all individual's CpG sites indexed by i in a sample (patient) indexed by n as follows:

$$S^n = -\,<\sigma^n{}_i \log_2(\sigma^n{}_i) + (1-\sigma^n{}_i)\log_2(1-\sigma^n{}_i)>_i \qquad (2)$$

where σ_i is the probability of finding the CpG state i in the polarized state and $<\ldots>_i$ represents averaging over all CpG sites in an individual. The dots in **Figure 4** were obtained by averaging the mean entropy values of all patients in the subsequent 5-y-old age bins.

In the EMR dataset, the probabilities shown in Equation (2) corresponded to the incidence of the diseases in the subsequent 5-y-old age bins.

Theory: Aging in a complex regulatory network

We propose to model the effect of the interactions among FUs with the help of the auxiliary variables—the effective "regulatory fields" h_i evolving over time according to

$$\dot{\mathrm{h}}_i = \sum_j \mathrm{k}_{ij}\sigma_j + \sum_{jk} \mathrm{g}_{ijk}\sigma_j\sigma_k + \mathrm{J}_i{}^0 + f_i \qquad (3)$$

where k_{ij} and g_{ijk} describe the first linear and first-order nonlinear interaction between the individual units, respectively. The force terms $\mathrm{J}_i{}^0$ and f_i represent the effects of constant (such as smoking or diets) and stochastic (social status and deleteriousness of the environment[33]) factors, respectively. For simplicity, we assume that the noise factors have zero mean and are not correlated over time. The states of individual FUs i can be observed depending on the regulatory field h_i according to the Boltzmann distribution: $\bar{\sigma}_i = (1+\exp(-\mathrm{h}_i/\mathrm{T})^{-1}$, where T is the effective temperature.

We start from Equation (3) and observe that the regulatory fields change over time in response to the deterministic (the direct linear and the higher-order nonlinear interactions between units) and stochastic forces f_i. We naturally assume that the stochastic force terms are not correlated over long time intervals: $<f(\mathrm{t})f(\mathrm{t}')> = \mathrm{B}\delta(\mathrm{t}-\mathrm{t}')$, where B is the power of stochastic noise, $<\ldots>$ represents the averaging along the individual trajectory and over all specimen, and $\delta(\mathrm{t})$ is the Dirac delta function.

In spite of apparent simplicity, Equation (3) is nonlinear and may have highly nontrivial solutions leading to applications in condensed matter physics[34] and neurophysiology[35]. In our discussion, it is important that the stochastic noise drives the system toward equilibrium at an effective temperature controlled by the power of the noise $\mathrm{T} \sim \mathrm{B}$.

The data suggest that there is a large "bulk" of units characterized by excessive lifetimes. Mechanistically, this may be explained by operating within a vicinity of a metastable state with a very high activation energy $\mathrm{U_{act}}$ relative to the effective temperature, $\mathrm{U_{act}} \ll \mathrm{T}$ (**Fig. 3A**).

We will assume that the effects of aging are small on the scale of $\mathrm{U_{act}}$ and hence the depolarization rates $\mathrm{R}_i \sim \mathrm{U_{act}}/\mathrm{T}$ are not only very small but also do not considerably depend on age. Accordingly, the depolarization is on average a linear function of age and the total number of configuration transitions Z_t: $<\Delta\sigma_i> = \mathrm{R}_i\mathrm{t} \propto Z_\mathrm{t}$ and $<\Delta\sigma_i^2> \propto \mathrm{R}_i\mathrm{t} \propto Z_\mathrm{t}$.

Let us think that the aging drift in the form of simultaneously occurring configuration transitions progresses slowly compared

with fast functional responses in the organism. We linearize the equations for the regulatory fields next to the youthful state $\bar{h}_i$:

$$\dot{\delta h}_i = \sum_j K_{ij}\delta h_j + \sum_j K_{ij}<\Delta\sigma_j> + \sum_{j,k=j} g_{ijk}<\Delta\sigma_j{}^2> + \sum_{j,k} g_{ijk}\frac{d\sigma_j}{dh_j}\delta h_j<\Delta\sigma_k> + J_i{}^0 + f_i \quad (4)$$

where $\delta h_i = h_i - \bar{h}_i$ and $\Delta\sigma_j$ describe the deviations of the fields and depolarization of the units, respectively, whereas the averages $<\ldots>$ involve the averaging over the "bulk" uncorrelated states only.

The solutions of the linearized Equation (4) can be best understood with the help of a linear decomposition: $\delta h_i \approx \sum_A z_A b_i{}^A$, where z_A is the pathway activations and $b_i{}^A$ is the right eigenvectors of the interaction matrix corresponding to the smallest eigenvalues r_A (the matrix K is nonsymmetric and hence its complete eigensystem must include the left eigenvectors $a^A K = -r_A a^A$, and the right eigenvectors $Kb^A = -r_A b^A$). The components of the vector $b_i{}^A$ characterize the participation of the FU i in the pathway A.

Substituting the solution into the equation and multiplying both sides by the corresponding left eigenvector, we find that

$$\dot{z}_A = -(r_A - r'_A Z_t)z_A + \beta_A Z_t + J_A + f_A \quad (5)$$

where $J_A = a^A J$ and $f_A = a^A f$. The effect of aging comes through the mean field on the pathway activation $\beta_A Z_t = a^A K <\Delta\sigma> + O(g)$ and the nonlinear correction to the eigenvalue $r_A \approx r_A - r'_A Z_t$.

It is important to understand that all the relevant vectors and constants cannot be derived and could only be measured experimentally. By virtue of the central limit theorem, the large number of configuration transitions ensures that the effect of the mean field is exactly linear in Z_t.

Qualitatively, the net effect of the rare transitions and the associated mean field Z_t together produce a persistent pathway activation, on average, slowly increasing with age. This is often referred to as an enslavement principle: stochastic depolarization transitions produce a slowly evolving mean field Z_t that disturbs pathways characterized by fast relaxation times having thus enough time to adjust to its current level.

Results

Aging signatures in cross-sectional DNAm data

We start by analyzing a dataset of DNAm in aging white blood cells[29]. Each of the reported DNAm levels $\bar{\sigma}_i$ is the average of a binary single-cell signal over a bulk tissue sample comprising many cells. In other words, $\bar{\sigma}_i$ is the probability of finding a CpG site in a methylated state.

To avoid complications due to the crossover between development and aging, we only analyzed donors older than 25 y. Furthermore, to counter the "curse of dimensionality"[36] due to the shallow nature of the dataset (450k CpGs measured in less than 800 patients), we focused our analysis only on CpGs significantly correlated with age (after the Bonferroni correction for multiple testing, $p = 0.005/450\text{k}$). Of approximately 100k CpGs significantly correlated with age (almost 25% of all reported), most were either initially hypermethylated ($\bar{\sigma}_i > 0.9$, 26%) or hypomethylated ($\bar{\sigma}_i < 0.1$, 28%).

To normalize the distribution of the DNAm signal confined in the interval [0,1], we converted DNAm levels to log odds ratios $h_i \sim \ln(\bar{\sigma}_i/(1-\bar{\sigma}_i))$. We refer to h_i as "regulatory fields" by analogy with condensed matter physics (see the Materials and Methods section).

The PCA of regulatory fields reveals a few PCs associated with age (DNAm-PC1 and DNAm-PC3 explained 41.3% and 3.2% of variance in the data but changed with age, respectively, and DNAm-PC2 explained 3.8% of variance but did not change with age and was omitted in the analysis of aging changes; see **Figs. S1 and S2**). The dominant PC (DNAm-PC1) evolves approximately linearly as a function of age (**Fig. 1A**; Pearson's $r = 0.68$, $p = 3 \cdot 10^{-98}$). The variance of DNAm-PC1 also increases linearly with age (**Fig. 1B**), which is a feature of a stochastic process (random walk).

Aside from DNAm-PC1, the best correlation with chronological age was produced by the third PC, DNAm-PC3 (Pearson's $r = 0.56$, $p = 3 \cdot 10^{-62}$). DNAm-PC3 increased faster in the subsequent age-adjusted bins than at a linear pace as a function of age (**Fig. 1C**). The variance of DNAm-PC3 also increased faster than linearly, so that the inverse variance decreased approximately linearly in the patients older than 40 y old (**Fig. 1D**). By extrapolation, the inverse variance of DNAm-PC3 would approach zero (and hence the variance would diverge) at some age within the age range of 120–150 y. This behavior hints at a nonlinear coupling of DNAm-PC3 to (and hence the dependence on) DNAm-PC1.

The loading vectors corresponding to DNAm-PC1 and DNAm-PC3 describe two distinct methylation profile changes with age. The distribution of the PC1 loading vector's components is non-Gaussian and bimodal. Hence, the dominant aging signature in DNAm data involves two large groups of CpG sites (**Fig. 1E**) changing their methylation ("polarization") with age in opposite directions. The first PC score is then proportional to the total number of polarization transitions.

In contrast, the distribution of the loading vector's components from DNAm-PC3 has a single peak and clear leading contributions from non-Gaussian tails (**Fig. 1E**). The gene set enrichment analysis of methylation regions associated with the PC3 variation reveals pathways involved in innate immunity and cancer (**Fig. 1G,H**).

The age-associated PC scores demonstrate the best correlation with Horvarth's DNAm age[4] (**Fig. 1F**). The corresponding Pearson's correlation coefficients were $r = 0.75$ ($p = 2 \cdot 10^{-131}$) and $r = 0.52$ ($p = 10^{-52}$) for DNAm-PC1 and DNAm-PC3, respectively (see also **Figs. S1 and S2**, for a summary of other PC scores' correlation with age and Horvath's DNAm age).

Next, we checked which characteristic features of dynamics associated with aging in DNAm can be observed in other forms of biomedical data. To confirm the stochastic character of the dominant aging signature in humans, one would need to analyze a large longitudinal dataset. We did not have access to a high-quality set of longitudinal DNAm measurements. Instead, we turned to an extensive EMRs collection from the UK Biobank. Irrespective of the age at the first assessment, the EMRs provided information on the prevalence of chronic diseases from inception until the end of the follow-up (slightly more than 10 y after enrollment, on average). We represent each patient by a vector of binary variables indicating the presence or absence of a disease (see the Materials and Methods section).

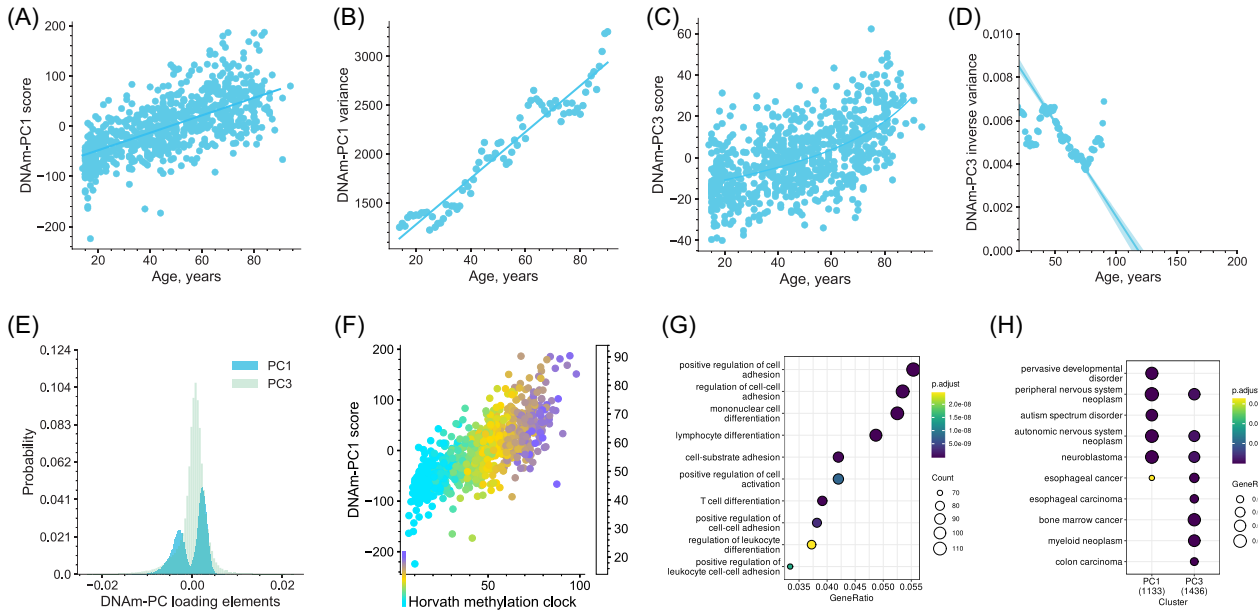

Figure 1. PCA of age-dependent methylation profiles from white-blood-cell samples in GSE87571 dataset. (A) DNAm-PC1 and **(B)** variance of DNAm-PC1 increase, on average, linearly as functions of age. **(C)** DNAm-PC3 increases faster than linearly with age. **(D)** Inverse variance of DNAm-PC3, the extrapolation for the range of 40–75 y vanishes at approximately 120 y old. **(E)** Distribution of PC vector components. **(F)** DNAm-PC1 scores correlate with Horvath's DNAm age. The color bar represents patients' chronological age. **(G)** Gene set enrichment analysis: CpG sites comprising DNAm-PC3 are associated with the regulation of innate immune response. **(H)** Methylation profiles driven by DNAm-PC1 and DNAm-PC3 are associated with developmental and mental diseases and internal organs' diseases.

Most of UK Biobank's subjects are healthy early in life. Hence, the states representing the presence of diseases are initially polarized ($\sigma_i = 0$). Most chronic diseases are relatively infrequent: the most prevalent diseases are metabolic disorders (with the prevalence of 15%), joint disorders (14%), and arthrosis (12%). Hence, most of the states stay polarized for life, with only a small fraction of patients exhibiting depolarization transitions leading to the incidence of specific diseases.

The PCA of binary-valued vectors representing a health state for the EMRs of UK Biobank's subjects at the time of the first assessment look like the PCA results from the white-blood-cell DNAm study above. This time, we observed only two PCs significantly associated with age (see the blue and green lines and the respective ranges corresponding to the mean levels and one standard deviation in **Fig. 2A**).

The dominant aging signature, the first PC in the UK Biobank's EMR data (EMR-PC1), evolves approximately linearly as a function of age and is linearly associated with the total number of diagnosed diseases (**Fig. 2B**). Hence, in line with the results of our DNAm analysis above, the first PC correlates with the total number of depolarization transitions (this time being equal to the disease burden at the time of measurement).

As expected, the variance of EMR-PC1 increases linearly with age (**Fig. 2C**), which is a feature of a stochastic process. This time, however, due to the longitudinal nature of the EMR dataset, we can make a stronger claim by computing the autocorrelation function of EMR-PC1. We observe that the autocorrelator increases linearly as a function of the time lag between the observations, which is typical for a result of a stochastic process with a drift (**Fig. 2D**).

Aging in a complex regulatory network

To explain the key features of dynamics of aging signatures, let us consider an organism as a network of interacting FUs. Each of the units can be observed in multiple states of varying physiological capacity. We have already presented examples of such microscopic states corresponding to the different methylation levels of CpGs or disease states. However, the language may be used to describe other situations involving, e.g., mutations or conformation changes in biomolecules.

For any given FU *i*, we will focus on the two most-occupied microscopic states (**Fig. 3A**) corresponding to two adjacent potential wells in the free energy landscape shaped by regulatory interactions. We encode the pair of states by a binary variable σ_i taking values of $\sigma_i = 0$ and $\sigma_i = 1$, respectively. At the end of development, most FUs are polarized, so that most of the subjects occupy one of the selected states.

According to the model, the initial states set during development are metastable states. Hence, over time, the organism state relaxes toward thermal equilibrium via a series of configuration transitions between microscopic states driven by fluctuations. Both the DNAm and EMR data suggest that, in most cases, the transitions are infrequent. On average for each FU, we observe fewer than a single transition between the states over the lifetime $\bar{t}$ of an organism. In other words, the corresponding transition rates R_i are slow ($R_i\bar{t} \ll 1$, see **Fig. 3B**). Slow transition rates depend on the activation energies U^{act}_i, and the effective temperature T exponentially, $R_i \sim \exp(-U_i^{act}/T)$[37]. Therefore, we expect that $U_i^{act} \gg T$ are barely affected by the effects of aging.

The effective temperature T characterizes the statistical properties of regulatory noise, which may depend on the fidelity of regulatory interactions and the deleteriousness of the environment[33]. The effective temperature shall not be confused with (although maybe related to) the body or environmental temperature (see the Materials and Methods section).

Quantitatively, stochastic fluctuations leading to configuration transitions change the average polarization of every FU linearly over time t:

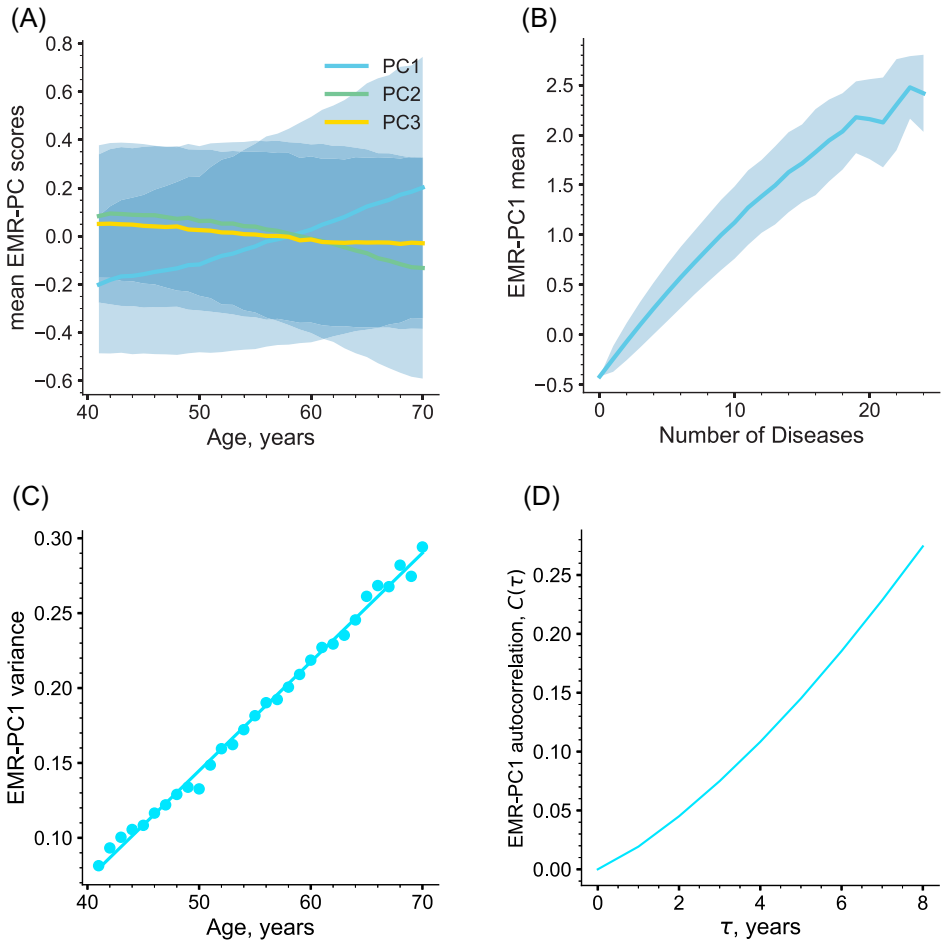

Figure 2. PCA of disease-state vectors from the electronic medical records (EMRs) from the UK Biobank. (A) Age dependence of the first three PC scores. **(B)** PC1 correlates with the total number of chronic diseases. The solid lines and the ranges represent the means and the standard deviations in subsequent age-matched cohorts. **(C)** Variance of PC1 increases linearly with age. **(D)** Autocorrelation function C(τ) of PC1 increases linearly as a function of the time lag τ.

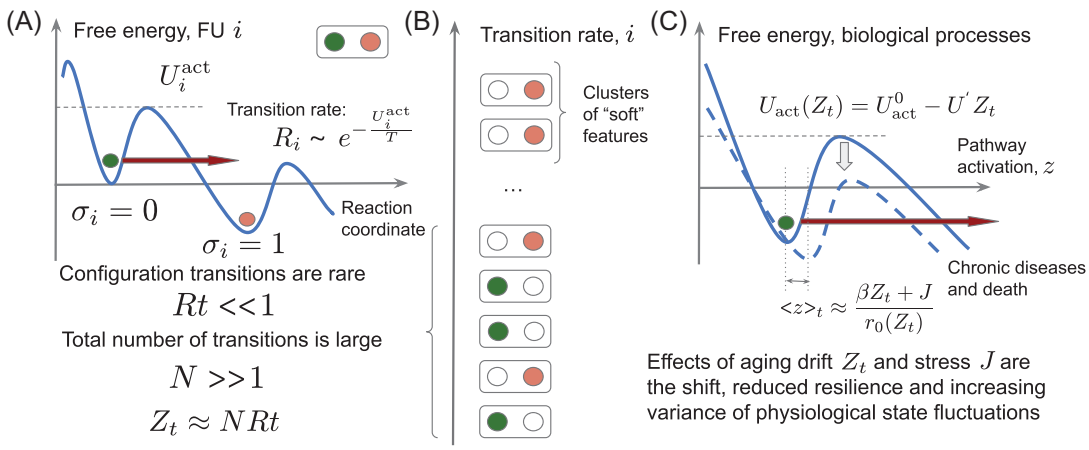

Figure 3. Basic features of the proposed aging model. (A) Schematic representation of relaxation dynamics of a functional unit (FU) *i* residing in a potential well (the blue curve). The subsequent metastable states are labeled by the "polarization" $\sigma_i = 0$ and $\sigma_i = 1$ (the red arrows indicate thermally activated configuration transition between microscopic states). The initial "polarized" state is protected by the activation barrier characterized by the activation energy U_i^{act}. The configuration transition rates R_i are presumed small and depend exponentially on the effective temperature T. **(B)** Human organisms consist of a macroscopically large number N of FUs. We classify them according to the mean activation rates R_i. Most configuration transitions are very rare ($R_i\bar{t} \ll 1$). **(C)** Dynamics of the "soft" FUs with low activation barriers $R_i\bar{t} \sim 1$. Aging drift causes the reduced resilience and diverging variance of physiological state fluctuations, all proportional to the overall number of the configuration transitions to date $Z_t \approx NRt$.

$$<\sigma_i>_t = <\sigma_i>_{t_0} + \epsilon_i R_i (t - t_0) \quad (6)$$

Here, the averaging $<\ldots>_t$ occurs over the samples produced at age t and $\epsilon_i = \pm 1$ is the direction of a depolarization transition.

The PCA of a dataset modeled in Equation (6) would produce the first PC, which is directly proportional to the total number of configuration transitions $Z_t \approx NRt$, where R is the average depolarization rate and N is the total number of FUs. Only a tiny fraction of all FUs is practically observable in any given experiment. However, we may expect that the total number of the depolarization transitions in any sufficiently large subset of the data (such as DNAm or EMRs) is proportional to Z_t.

The depolarization rate for each FU may be slow, $R_i\bar{t} \ll 1$. However, the total number of FUs available for the configuration transitions is practically infinite $N \gg 1$, and their compound effect is not necessarily small. If an organism is sufficiently long lived, the number of configuration transitions Z_t is substantial, and the aging signature described in Equation (6) would dominate the variance in real-life biomedical data. This is consistent with our observations in the PCA of the DNAm (**Fig. 1A**) and EMR data (**Fig. 2A**) above. The distribution of the first PC vector's components would be bimodal for DNAm (bidirectional gain and loss of methylation, **Fig. 1E**) and unimodal for EMRs (the number of diagnoses would only grow with age).

Given that Z_t is a cumulative number of random transitions, its dynamics would obey a stochastic Langevin equation with a drift (diffusion). Hence, the variance of Z_t in age-adjusted bins should increase linearly with age. This prediction is consistent with the observed behavior of the first PC in the DNAm (**Fig. 1B**) and EMR (**Fig. 2C**) data.

More evidence in favor of the stochastic character of Z_t could be produced by the investigation of the autocorrelation function $C(\tau) = <(Z_{t+\tau} - Z_t)^2>$, where $<\ldots>$ represents the averaging, first, along the individual trajectory and, then over all patients. The autocorrelation function of the leading PC in the EMR data increased linearly as a function of the time lag in the range between 2 and 10 y (**Fig. 2D**). The diffusion coefficient's estimates from the variance and autocorrelation increase turned out to be close: 0.012 and 0.009 per year, respectively, thus confirming the association of the leading PC score with the increasing number of configuration transitions Z_t.

Assuming that we start from highly polarized states, $<\sigma_i>_t \approx 1 - R_i t$, we show that the configuration entropy is equal to the number of depolarization transitions Z_t

$$S(t) = N < R_i t \log(1/R_i t) > \sim Z_t \quad (7)$$

up to a proportionality coefficient, thus highlighting the stochastic character of aging drift in the model. Hence, the model explains the dominant aging DNAm changes (DNAm-PC1) and the incidence of chronic diseases (EMR-PC1) as the growth of configuration entropy. Notably, the entropy production rates turned out to be very close: the regression line slopes were 0.0122 ($p = 0.001$) and 0.0375 ($p = 0.025$) bits per FU per year in the DNAm and the EMR datasets (**Fig. 4A,B**), respectively.

Due to the inverse exponential dependence between the transition rates and the activation barrier, the lowest activation barriers would lead to the highest depolarization rates (the top FUs in **Fig. 3B**). Below, we show that the interactions between such FUs can no longer be neglected and thus the FUs should form coregulated clusters (see the Materials and Methods section). We expect that the joint activation of FUs forming a cluster (or a pathway) labeled by A affects all other FUs *i* in the cluster via a shift of regulatory fields according to $\delta h_i \approx z_A b_i^A$, where z_A and b_i^A are the pathway activation strength and the participation vectors' components, respectively.

In **Figure 3C**, the solid blue line represents the cross-sectional view of the free energy as a function of the pathway activation variable z_A experiencing stochastic fluctuations in response to stress factors. The dynamics of the pathway activation depend on the power of stochastic noise (proportional to the effective temperature T) and persistent stress factors J_A. The effects of the regulatory interactions can be described by the recovery rate, r_A, which is directly related to the curvature of the basin of attraction for z_A. The recovery rate is the inverse recovery time and characterizes the pathway's ability to respond to stress and relax toward the equilibrium position after a shock.

The model suggests that the depolarization transitions occur independently from pathway activation, but their cumulative effect slowly reshapes the free energy landscape of regulatory fields for other FUs (see the dashed blue line in **Fig. 3C**). In other words, stochastically accumulated changes are tiny and not critical in the short term, but their accumulated effect slowly gnaws

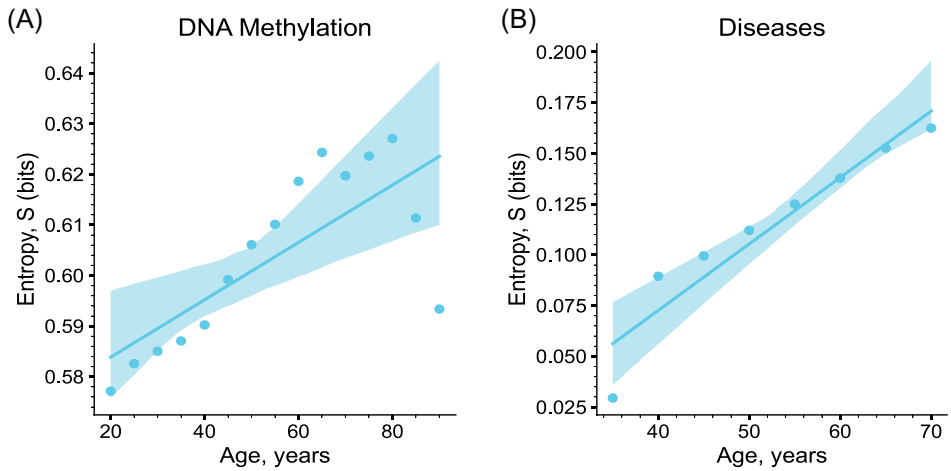

Figure 4. Computed configuration entropy. (A) For the distribution of the DNA methylation and **(B)** for the EMR features in subsequent age-matched cohorts increasing linearly as a function of age after 40 y.

away the organism's resilience and stability. Because the number of transitions is enormous, the central limit theorem comes into play[38] and ensures that the net effect of configuration changes on any physiological process must be proportional to the total number of depolarization transitions Z_t.

Over longer time scales, well exceeding pathway equilibration times $\sim r_A^{-1}$, the stochastic component of z_A fluctuations averages out. The mean pathway activation and variance are then given by (see the Materials and Methods section)

$$<z_A>_t \approx \frac{\beta_A Z_t + J_A}{r_A(Z_t)}$$

$$\sigma_z^2(t) \sim \frac{T}{r_A(Z_t)} \quad (8)$$

Here, $r_A(Z_t) = r_A^0 - r'_A Z_t$ is the age-adjusted recovery rate, whereas β_A, and r'_A are the small and pathway-specific quantities characterizing the weak mode-coupling effects leading to the compound (and hence proportional to Z_t) effect of depolarization processes on pathway activation and resilience, respectively.

Accordingly, the fluctuations of organism state variables other than stochastic variables described in Equation (6) can be attributed to a few clusters of coregulated features participating in pathways characterized by the slowest recovery rates (vanishing denominators in Eq. 8). In this case, the participation vectors $b^A i$ and the pathway activation variables z_A should approximately coincide with the leading PC loading vectors and scores, respectively.

According to Equation (8), an increasing number of depolarization transitions Z_t causes progressive shifts in pathway activation. Notably, this effect is indistinguishable, albeit smaller than the effects of constant stress modeled by J_A. More subtly, aging in the form of progressive depolarization of an organism state measured by Z_t also affects the recovery rates in the denominator of Equation (8). The two effects combine and cause the mean pathway activations and therefore the leading PC scores in the data depend on age in a nonlinear—hyperbolic fashion (see the dynamics of DNAm-PC3 in **Fig. 1C** and EMR-PC2 in **Fig. 2A**).

The nonlinear coupling of organism-state fluctuations with depolarization transitions may reduce one of the smallest recovery rates to zero: $r(Z_t) = r^0 - r' Z_t = r^0(1 - t/t_{max})$ at some point late in life at age $t_{max} = r^0/(r' dZ_t/dt)$. The situation corresponds to the critical point corresponding to the complete loss of resilience, that is, the inability of the system to retain its homeostasis equilibrium and hence it is incompatible with survival[18].

There is no way to measure the recovery rate in cross-sectional data. However, according to Equation (8), the vanishing recovery rate should lead to the simultaneous divergence of one of the leading PC scores and its variance at a certain advanced age. In our analysis, DNAm-PC3 increases faster than linearly as a function of chronological age. The fit of the DNAm-PC3 scores to the hyperbolic solution for the average z_A from Equation (8) gives $t_{max} \approx$ 130 y (see the solid line in **Fig. 1C** and the Materials and Methods section, for the details of the calculations).

In agreement with Equation (8), the extrapolation shows that the inverse variance of DNAm-PC3 hits zero and hence the variance of DNAm-PC3 diverges at approximately 120 y (see the solid line in **Fig. 1D**). The estimations of the limiting age from the behavior of DNAm-PC3 mean and its variance are comfortably close. Hence, our calculations support the existence of a critical point in the age range of 100 – 150 y.

In reality, the disintegration of the organism state happens well before reaching the criticality at the limiting age t_{max}. Stress factors and the depolarization of the organism state do not merely shift the mean pathway activation levels. Both factors may also decrease the activation energy separating the organism state from disintegration and death (**Fig. 3C**). In the linear regime, the activation energy linearly depends on the mean field, $U_{act}(Z_t) = U^0{}_{act} - U' Z_t$, where $U' = dU/dZ$.

Mortality in the model is nothing else but the probability of barrier crossing per unit time: $M \sim \exp(-U^0{}_{act}/T) \exp(U' Z_t/T)$. Therefore, the aging drift in the form of the linearly increasing number of the configuration transitions registered by the progressively increasing $Z_t \sim t$ may drive the exponential acceleration of all-cause mortality with age: $M \sim \exp(\Gamma t)$. The mortality doubling rate, $\Gamma = T^{-1} U' dZ_t/dt$, in the model depends on the details of the regulatory interactions (through U'), the rate of the aging drift dZ_t/dt, and the effective temperature T.

Discussion

We put forward a semiquantitative model of aging in a complex regulatory network and applied it to the analysis of human aging signatures in a cross-sectional white-blood-cell DNAm dataset[29] and the extensive collection of longitudinal EMRs from the UK Biobank[30]. The model explains the dynamics of aging signatures in both signals by the cumulative effect of numerous stochastic configuration changes accompanied by increasing entropy.

The data suggest that the rates of transitions among microscopic states of methylation or the incidence of specific diseases are slow. In most cases, fewer than a single transition occurs throughout lifetime for each FU. Even though the transition rates are slow, the number of transitions is vast: more than 25% of all CpG sites exhibited age-related dynamics. Hence, the compound effects of transitions accumulate and dominate the dynamics of the physiological state in the long term. We observed that in the leading aging signature explaining most of the variance in the data (the first PC score), which increased linearly with age in the DNAm and EMR data. The first PC score was proportional to the number of configuration transitions Z_t (the number of DNAm level changes or chronic diseases). Simultaneously, the first PC variance grew linearly with age in both datasets, as is expected for a stochastic quantity representing accumulation of a large number of independent random transitions. Accordingly, we propose using Z_t as a quantitative measure of the net effect of entropic changes on the aging organism—tBA (tBA $\sim Z_t$). The first PC score is then a good proxy of tBA for a specific signal well correlated with age and Horvath's DNAm age. Due to the central limit theorem[38] for a large number of random transitions, tBA would increase linearly with age with a high accuracy. That explains why it is almost always possible to build an accurate predictor of chronological age for different biomedical signals[4,5,8].

Configuration transitions do not only work as a natural clock in aging organisms but also define the thermodynamic arrow of time. Our model suggests that tBA is proportional to entropy produced (information lost) during aging. Particular depolarization patterns may differ in various cells of one organism or among different organisms of the same age. However, the number of transitions would be comparable and hence quantify the overall aging state of an organism. In other words, the older an organism gets

the more "rust" it accumulates, which is captured by tBA. We expect that the present model can generalize to other FUs and readouts changing over time, such as gene expression changes[39], conformation or chemical modifications of macromolecules, DNA damage, etc. Because configuration changes occur simultaneously in every part of an organism and increase its tBA, observation of a change in any single data modality cannot be interpreted as the cause of aging drift in other modalities. The inherent stochasticity of aging epigenome may explain why genetically identical pairs of twins diverge with age[40,41] and may drive transcriptomic dysregulation in cancer[42].

The aging drift manifests itself as a "mean field" causing the shift of physiological indices that is proportional to tBA. The mean-field theory is a powerful approximation for understanding the behavior of interacting systems first developed in physical sciences[43] and since then applied in statistical inference[44] (see, e.g., protein structure prediction[45,46]). Here, we modeled the overwhelming complexity of FU interactions by a simpler approximation, where each FU or large FU clusters operate independently and substitute cumulative effects of all other FUs by their mean field represented by tBA. Other than the dominant aging PC1 ~ tBA in biological signals, other leading PCs with the longest recovery times and strongest fluctuations would also change with age. The aging drift affects those modules or pathways in a way similar to other stresses (such as smoking or diet). Because tBA increases linearly with age, we expect that, in the first approximation, all pathways "follow" the aging process by increasing (or decreasing) activation barriers linearly with age. In a higher-order approximation, the nonlinearity of regulatory interactions produces significant deviations from a simple linear age dependence of nondominant PC scores in biomedical data (DNAm-PC3 and EMR-PC2). Nonlinear regulatory interactions let the configuration changes (but not stresses) affect the resilience defined as the ability of an interacting FU cluster to respond to a perturbation and relax to equilibrium afterward. If the recovery rate is slow, this may lead to a divergence of organism state fluctuations at a critical age, where the recovery rate would vanish completely. For example, this happens for the dynamics of DNAm-PC3 because the extrapolation of the mean and variance of DNAm-PC3 diverges at an age close to $t_{max} \approx 130$ y. Recently, we also demonstrated that linear log-mortality predictors built from complete blood counts and physical activity[18] exhibited similar diverging fluctuations and a vanishing recovery rate at about the same limiting age $t_{max} \approx 130$ y.

Hence, the prediction of mortality (or the remaining lifespan) in humans requires an estimate for tBA ~ Z_t and for a few most crucial pathway activations (also, on average, depending on Z_t). Hence, no single BA measure fully describes longevity in humans. We expect that the BA models trained to predict chronological age should yield better estimates of tBA. On the other hand, the models trained to predict the remaining lifespan (such as PhenoAge[47], GrimAge[10], DOSI[18], etc.) should return a combination of pathway activations associated with the prevalence of diseases and accelerated mortality[13] and hence be better suited for the detection of reversible effects of diseases, lifestyles, and medical interventions[15].

The PCA of human biomedical data is peculiar because it produces more than a single age-dependent feature, which is not the case in simpler animals such as worms[48], flies[49], or mice[50], where aging could be explained by a dynamic instability leading to the exponential disintegration of an organism state[49,51]. We expect that the entropic contribution to aging is not dominant in those cases, and the BA is a dynamic (not entropic) factor, and the aging effects may be reversible in those animal models[50]. The loss of stability signs in the DNAm-PC3 hint at a two-stage aging process in humans. Our model explains how the entropic changes reduce resilience and the recovery rates of protective mechanisms (hallmarks of aging). The stochastic accumulation of regulatory noise or deleteriousness of the environment lead to an exponential destabilization of the organism state. Therefore, we conclude that the cumulative effect of entropic changes captured by tBA explains the exponential mortality and disease incidence acceleration—a characteristic feature of human aging.

The present model along with other direct dynamics stability analyses of organism state fluctuations in longitudinal biomedical data[16,18] support the idea that the human organism state (and potentially other long-lived mammals, such as naked mole rats) stays metastable until very late in life, and slowly loses its stability and resilience with age due to accumulation of entropic changes. According to the model, human aging has a significant entropic component, which not only dominates the variance in biomedical data but also causes increasing stress on adaptive subsystems (stress responses). Our approach is in line with Hayflick's proposal[52] that distinguishes the genetic determinism of longevity from the stochasticity of aging. If the proposition is accurate, we must expect that although the hallmarks of aging (features or activations of specific adaptive pathways leading to mortality and morbidity acceleration[1]) can, in principle, be reverted, the expected effects of such interventions on lifespan may be transient and limited.

We expect that attempts to reduce the dominant aging signature tBA would require availability and timely application of an immense number of precise interventions. This is, to say the least, technologically challenging. Accordingly, we predict that aging in humans can be reversed only partially. The fact that there is a strong entropic contribution to aging does not necessarily mean that one cannot reset some of the organismal subsystems closer to a younger state. The entropic character of aging implies that age reversal would be limited to a specific organismal subsystem without a full rejuvenation of the whole organism. For example, recent epigenetic reprogramming experiments[53–55] led to the reversal of epigenetic clock readouts.

However, our model suggests that achieving strong and lasting rejuvenation effects in humans may remain a remote perspective. However, there may be a more practical way to intercept aging by dramatically slowing the rate of aging. The rates of entropic transitions between any two states depend exponentially on the effective temperature. Hence, even minor alterations of the effective temperature may cause a dramatic drop in the rate of aging. In condensed matter physics, this situation is known as glass transition, where the viscosity and relaxation times may grow by 10–15 orders of magnitude in a relatively narrow temperature range[56–63]. We note that living organisms are nonequilibrium open systems, and hence the effective temperature is not the same as body or environment temperatures. Rather, the effective temperature is a measure of how deleterious the environment is for an organism[33].

We speculate that the evolution of long-lived mammals may have provided an example of tuning the effective temperature. Naked mole rats are known for their exceptional stress resistance, DNA repair efficacy[64–67], and translational fidelity[68,69]. Those factors should reduce noise in regulatory circuits and lower the effective temperature of the system. One example of such tuning may be used to explain the recent studies indicating that naked

mole rats breeders age slower than their nonbreeding peers, at least according to the DNAm clock[25]. Social status and mental health also impact the aging rate measured by DNAm and other clocks in humans[31,70], possibly via neuroendocrine system. Higher socioeconomic status, somewhat counter-intuitively, significantly increases the mortality doubling rate and simultaneously reduces the age-independent mortality in such a way that mortality in the highest and lowest income groups converge at an age close to our t_{max} estimates[71]. Such a behavior of mortality is consistent with a reduction in the effective temperature in the higher-income cohorts in our model.

Future studies should help establish the best ways to "cool down" the organism state and reduce the rate of aging in humans. The simple linear PCA exemplified here may only help gain a qualitative understanding of underlying processes. We expect that increasing availability of high-quality longitudinal biomedical data will lead to a better understanding of the most critical factors behind the kinetics of aging and diseases, including those controlling entropy production in the course of aging. This should lead to a discovery of actionable targets influencing the rate of aging, help slow down aging, and thus produce a dramatic extension of human healthspan.

P.S. Since our first publication of a preprint of this article, it inspired a number of follow-up works by other groups. For example, three papers investigated the stochastic nature of aging by extensive simulations and analyses of gene expression and DNAm data and supported our original proposal that the dominant component of aging changes was stochastic: on a single-cell level[72], for DNAm-based clocks[73], and gene expression[74]. The studies of dynamical properties of stochastic changes in DNAm[75] and other longitudinal signals in mice[50] in the course of aging and in response to antiaging interventions also corroborated our analyses and predictions regarding the distinction between entropic/irreversible and dynamic/reversible components of aging signatures.

Acknowledgments

We would like to thank A. Velikanova, D. Kriukov, K. Avchaciov, and T. Pyrkov for insightful discussions and help with data preparation and M. Kholin and A. Kadet for stimulating discussions and comments on the article. The work was funded by Gero LLC (Singapore).

Author Contributions

A.E.T. and P.O.F. proposed, formulated, and developed the theory. A.E.T, K.A.D., and P.O.F. conducted data analyses and validated the model assumptions against the data. A.E.T. and P.O.F. wrote and prepared this article.

Conflicts of Interests

P.O.F. is a shareholder of Gero PTE. A.E.T., K.A.D., and P.O.F. were employed by Gero PTE during the work on the article. The study was funded by Gero PTE.

Supplementary Materials

Supplemental information can be found here: Supplementary.

References

1. López-Otín C., Blasco M.A., Partridge L., Serrano M., & Kroemer G. (2013). The hallmarks of aging. *Cell* **153**(6), 1194–1217. PMID: 23746838; doi: 10.1016/j.cell.2013.05.039.
2. Zenin A., Tsepilov Y., Sharapov S., Getmantsev E., Menshikov L.I., Fedichev P.O., & Aulchenko Y. (2019). Identification of 12 genetic loci associated with human healthspan. *Commun. Biol.* **2**(1), 41. PMID: 30729179; doi: 10.1038/s42003-019-0290-0.
3. Gompertz B. (1825). XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. *Philos. Trans. R. Soc. Lond.* 513–583. doi: 10.1098/rstl.1825.0026.
4. Horvath S. (2013). DNA methylation age of human tissues and cell types. *Genome Biol.* **14**(10), R115. PMID: 24138928; doi: 10.1186/gb-2013-14-10-r115.
5. Hannum G., Guinney J., Zhao L., Zhang L., Hughes G., Sadda S., … Zhang K. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**(2), 359–367. PMID: 23177740; doi: 10.1016/j.molcel.2012.10.016.
6. Levine M.E. (2013). Modeling the rate of senescence: Can estimated biological age predict mortality more accurately than chronological age? *J. Gerontol. Ser. Biomed. Sci. Med. Sci.* **68**, 667–674. PMID: 23213031; doi: 10.1093/gerona/gls233.
7. Pal S., & Tyler J.K. (2016). Epigenetics and aging. *Sci. Adv*. **2**(7), e1600584. PMID: 27482540; doi: 10.1126/sciadv.1600584.
8. Petkovich D.A., Podolskiy D.I., Lobanov A.V., Lee S.G., Miller R.A., & Gladyshev V.N. (2017). Using DNA methylation profiling to evaluate biological age and longevity interventions. *Cell Metab*. **25**(4), 954–960.e6. PMID: 28380383; doi: 10.1016/j.cmet.2017.03.016.
9. Levine M.E., Lu A.T., Quach A., Chen B.H., Assimes T.L., Bandinelli S., … Horvath S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**(4), 573. PMID: 29676998; doi: 10.18632/aging.101414.
10. Lu A.T., Quach A., Wilson J.G., Reiner A.P., Aviv A., Raj K., … Horvath S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**(2), 303. PMID: 30669119; doi: 10.18632/aging.101684.
11. Levine M.E. (2020). Assessment of epigenetic clocks as biomarkers of aging in basic and population research. *J. Gerontol. A* **75**, 463–465. PMID: 31995162; doi: 10.1093/gerona/glaa021.
12. Levine M., McDevitt R.A., Meer M., Perdue K., Di Francesco A., Meade T., … Ferrucci L. (2020). A rat epigenetic clock recapitulates phenotypic aging and co-localizes with heterochromatin. *Elife* **9**, e59201. PMID: 33179594; doi: 10.7554/eLife.59201.
13. Porter H.L., Brown C.A., Roopnarinesingh X., Giles C.B., Georgescu C., Freeman W.M., & Wren J.D. (2021). Many chronological aging clocks can be found throughout the epigenome: Implications for quantifying biological aging. *Aging Cell* **20**(11), e13492. PMID: 34655509; doi: 10.1111/acel.13492.
14. Trapp A., Kerepesi C., & Gladyshev V.N . (2021). Profiling epigenetic age in single cells. *Nat. Aging* **1**(12), 1189–1201. PMID: 36211119; doi: 10.1038/s43587-021-00134-3.
15. Pyrkov T.V., Getmantsev E., Zhurov B., Avchaciov K., Pyatnitskiy M., Menshikov L., … Fedichev P.O. (2018). Quantitative characterization of biological age and frailty based on locomotor activity records. *Aging* **10**(10), 2973. PMID: 30362959; doi: 10.18632/aging.101603.
16. Pyrkov T.V., Sokolov I.S., & Fedichev P.O. (2021). Deep longitudinal phenotyping of wearable sensor data reveals independent markers of longevity, stress, and resilience. *Aging* **13**(6), 7900. PMID: 33735108; doi: 10.18632/aging.202816.
17. Chen B.H., Marioni R.E., Colicino E., Peters M.J., Ward-Caviness C.K., Tsai P.C., … Horvath S. (2016). DNA methylation-based measures of biological age: Meta-analysis predicting time to death. *Aging* **8**, 1844. PMID: 27690265; doi: 10.18632/aging.101020
18. Pyrkov T.V., Avchaciov K., Tarkhov A.E., Menshikov L.I., Gudkov A.V., & Fedichev P.O. (2021). Longitudinal analysis of blood markers reveals

progressive loss of resilience and predicts human lifespan limit. *Nat. Commun.* **12**(1), 2765. PMID: 34035236; doi: 10.1038/s41467-021-23014-1.

19. Santesmasses D., Castro J.P., Zenin A.A., Shindyapina A.V., Gerashchenko M.V., Zhang B., … Gladyshev V.N. (2020). COVID-19 is an emergent disease of aging. *Aging Cell* **19**(10), e13230. PMID: 33006233; doi: 10.1111/acel.13230.
20. Kuo C.-L., Pilling L.C., Atkins J.L., Masoli J.A., Delgado J., Tignanelli C., … Levine M.E. (2020). COVID-19 severity is predicted by earlier evidence of accelerated aging. *Medrxiv*. PMID: 32676624; doi: 10.1101/2020.07.10.20147777.
21. Belsky D.W., Caspi A., Corcoran D.L., Sugden K., Poulton R., Arseneault L., … Moffitt T.E. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife* **11**, e73420. PMID: 35029144; doi: 10.7554/eLife.73420.
22. Fahy G.M., Brooke R.T., Watson J.P., Good Z, Vasanawala S.S., Maecker H., … Horvath S. (2019). Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell* **18**(6), e13028. PMID: 31496122; doi: 10.1111/acel.13028.
23. Rejuvant™ Safety and Biomarker Study. https://ClinicalTrials.gov/show/NCT04821401.
24. Nielsen J., Bakula D., & Scheibye-Knudsen M. (2022). Clinical trials targeting aging. *Front Aging* **3**, 820215. PMID: 35821843; doi: 10.3389/fragi.2022.820215.
25. Horvath S., Haghani A., Macoretta N., Ablaeva J., Zoller J.A., Li C.Z., … Gorbunova V. (2021). DNA methylation clocks tick in naked mole rats but queens age more slowly than nonbreeders. *Nat. Aging* **2**(1), 46–59. PMID: 35368774; doi: 10.1038/s43587-021-00152-1.
26. Kerepesi C., Meer M.V., Ablaeva J., Amoroso V.G., Lee S.G., Zhang B., … Gladyshev V.N. (2022). Epigenetic aging of the demographically non-aging naked mole-rat. *Nat. Commun.* **13**(1), 355. PMID: 35039495; doi: 10.1038/s41467-022-27959-9.
27. Oh G., Ebrahimi S., Carlucci M., Zhang A., Nair A., Groot D.E., … Petronis A. (2018). Cytosine modifications exhibit circadian oscillations that are involved in epigenetic diversity and aging. *Nat. Commun.* **9**, 1–11. PMID: 29440637; doi: 10.1038/s41467-018-03073-7.
28. Bergstedt J., Azzou S.A.K., Tsuo K., Jaquaniello A., Urrutia A., Rotival M., … Milieu Intérieur Consortium (2022). The immune factors driving DNA methylation variation in human blood. *Nat. Commun.* **13**(1), 5895. PMID: 36202838; doi: 10.1038/s41467-022-33511-6.
29. Johansson Å., Enroth S., & Gyllensten U. (2013). Continuous aging of the human DNA methylome throughout the human lifespan. *PLoS ONE* **8**(6), e67378. PMID: 23826282; doi: 10.1371/journal.pone.0067378.
30. Bycroft C., Freeman C., Petkova D., Band G., Elliott L.T., Sharp K., … Marchini J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**(7726), 203–209. PMID: 30305743; doi: 10.1038/s41586-018-0579-z.
31. Protsenko E., Yang R., Nier B., Reus V., Hammamieh R., Rampersaud R., … Wolkowitz O.M. (2021). "GrimAge," an epigenetic predictor of mortality, is accelerated in major depressive disorder. *Transl. Psychiatry* **11**(1), 1–9. PMID: 33820909; doi: 10.1038/s41398-021-01302-0.
32. Healthcare and Utilization Project (HCUP). (2023) Chronic Condition Indicator Refined (CCIR) for ICD-10-CM Diagnoses, v2024.1. Agency for Healthcare Research and Quality (AHRQ). Retreived from: https://hcup-us.ahrq.gov/toolssoftware/chronic_icd10/CCIR-User-Guide-v2024-1.pdf.
33. Strehler B.L., & Mildvan A.S. (1960). General theory of mortality and aging: A stochastic model relates observations on aging, physiologic decline, mortality, and radiation. *Science* **132**(3418), 14–21. PMID: 13835176; doi: 10.1126/science.132.3418.14.
34. Amit D.J., Gutfreund H., & Sompolinsky H. (1985). Spin-glass models of neural networks. *Phys. Rev. A* **32**(2), 1007. doi: 10.1103/PhysRevA.32.1007.
35. La Camera G. (2021). The mean field approach for populations of spiking neurons. *ArXiv Prepr. ArXiv210901279*.
36. Teschendorff A.E. (2019). Avoiding common pitfalls in machine learning omic data science. *Nat. Mater.* **18**(5), 422–427. PMID: 30478452; doi: 10.1038/s41563-018-0241-z.
37. Pitaevskii L.P., & Lifshitz E. (2012). *Physical Kinetics*. vol. 10, Butterworth-Heinemann.
38. Fischer H. (2010). *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer Science & Business Media.
39. Perez-Gomez A., Buxbaum J.N., & Petrascheck M. (2020). The aging transcriptome: Read between the lines. *Curr. Opin. Neurobiol.* **63**, 170–175. PMID: 32563038; doi: 10.1016/j.conb.2020.05.001.
40. Czyz W., Morahan J.M., Ebers G.C., & Ramagopalan S.V. (2012). Genetic, environmental and stochastic factors in monozygotic twin discordance with a focus on epigenetic differences. *BMC Med.* **10**(1), 1–12. PMID: 22898292; doi: 10.1186/1741-7015-10-93.
41. Panzeri I., & Pospisilik J.A. (2018). Epigenetic control of variation and stochasticity in metabolic disease. *Mol. Metab.* **14**, 26–38. PMID: 29909200; doi: 10.1016/j.molmet.2018.05.010.
42. Hogg S.J., Beavis P.A., Dawson M.A., & Johnstone R.W . (2020). Targeting the epigenetic regulation of antitumour immunity. *Nat. Rev. Drug Discov.* **19**(11), 776–800. PMID: 32929243; doi: 10.1038/s41573-020-0077-5.
43. Kadanoff L.P. (2009). More is the same; phase transitions and mean field theories. *J. Stat. Phys.* **137**(5-6), 777–797. doi: 10.1007/s10955-009-9814-1.
44. Opper M., & Saad D. (2001) *Advanced Mean Field Methods: Theory and Practice*. MIT Press.
45. Koehl P., & Delarue M. (1994). Application of a self-consistent mean field theory to predict protein side-chains conformation and estimate their conformational entropy. *J. Mol. Biol.* **239**(2), 249–275. PMID: 8196057; doi: 10.1006/jmbi.1994.1366.
46. Morcos F., Pagnani A., Lunt B., Bertolino A., Marks D.S., Sander C., … Weigt M. (2011). Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc. Natl. Acad. Sci.* **108**(49), E1293–E1301. PMID: 22106262; doi: 10.1073/pnas.1111471108.
47. Levine M.E., Lu A.T., Quach A., Chen B.H., Assimes T.L., Bandinelli S., … Horvath S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**(4), 573–591. PMID: 29676998; doi: 10.18632/aging.101414.
48. Tarkhov A.E., Alla R., Ayyadevara S., Pyatnitskiy M., Menshikov L.I., Shmookler Reis R.J., & Fedichev P.O. (2019). A universal transcriptomic signature of age reveals the temporal scaling of Caenorhabditis elegans aging trajectories. *Sci. Rep.* **9**(1), 1–18. PMID: 31089188; doi: 10.1038/s41598-019-43075-z.
49. Kogan V., Molodtsov I., Menshikov L.I., Reis R.J.S., & Fedichev P. (2015). Stability analysis of a model gene network links aging, stress resistance and negligible senescence. *Sci. Rep.* **5**(1), 1–12. PMID: 26316217; doi: 10.1038/srep13589.
50. Avchaciov K., Antoch M.P., Andrianova E.L., Tarkhov A.E., Menshikov L.I., Burmistrova O., … Fedichev P.O . (2022). Unsupervised learning of aging principles from longitudinal data. *Nat. Commun.* **13**(1), 6529. PMID: 36319638; doi: 10.1038/s41467-022-34051-9.
51. Podolskiy D., Molodtcov I., Zenin A., Kogan V., Menshikov L.I., Gladyshev V.N., … Fedichev P.O. (2015). Critical dynamics of gene networks is a mechanism behind ageing and Gompertz law. *ArXiv Prepr. ArXiv150204307*.
52. Hayflick L. (2007). Entropy explains aging, genetic determinism explains longevity, and undefined terminology explains misunderstanding both. *PLoS Genet.* **3**(12), e220. PMID: 18085826; doi: 10.1371/journal.pgen.0030220.
53. Olova N., Simpson D.J., Marioni R.E., & Chandra T. (2019). Partial reprogramming induces a steady decline in epigenetic age before loss of somatic identity. *Aging Cell* **18**(1), e12877. PMID: 30450724; doi: 10.1111/acel.12877.
54. Lu Y., Brommer B., Tian X., Krishnan A., Meer M., Wang C., … Sinclair D.A. (2020). Reprogramming to recover youthful epigenetic information and restore vision. *Nature* **588**(7836), 124–129. PMID: 33268865; doi: 10.1038/s41586-020-2975-4.
55. Gill D., Parry A., Santos F., Okkenhaug H., Todd C.D., Hernando-Herraez I., … Reik W. (2022). Multi-omic rejuvenation of human cells by

maturation phase transient reprogramming. *eLife* **11**, e71624. PMID: 35390271; doi: 10.7554/eLife.71624.

56. Kob W., & Barrat J.-L. (1997). Aging effects in a Lennard-Jones glass. *Phys. Rev. Lett.* **78**(24), 4581. doi: 10.1103/PhysRevLett.78.4581.
57. Yue Y., Jensen S., & Christiansen J.deC. (2002). Physical aging in a hyperquenched glass. *Appl. Phys. Lett.* **81**(16), 2983–2985. doi: 10.1063/1.1514386.
58. Lubchenko V., & Wolynes P.G. (2004). Theory of aging in structural glasses. *J. Chem. Phys.* **121**(7), 2852–2865. PMID: 15291595; doi: 10.1063/1.1771633.
59. Sellitto M., & Kurchan J. (2005). Shear-thickening and entropy-driven reentrance. *Phys. Rev. Lett.* **95**(23), 236001. PMID: 16384320; doi: 10.1103/PhysRevLett.95.236001.
60. Lunkenheimer P., Wehn R., Schneider U., & Loidl A. (2005). Glassy aging dynamics. *Phys. Rev. Lett.* **95**(5), 055702. PMID: 16090889; doi: 10.1103/PhysRevLett.95.055702.
61. Rittigstein P., & Torkelson J.M. (2006). Polymer–nanoparticle interfacial interactions in polymer nanocomposites: Confinement effects on glass transition temperature and suppression of physical aging. *J. Polym. Sci. Part B Polym. Phys.* **44**(20), 2935–2943. doi: 10.1002/polb.20925.
62. Greinert N., Wood T., & Bartlett P. (2006). Measurement of effective temperatures in an aging colloidal glass. *Phys. Rev. Lett.* **97**(26), 265702. PMID: 17280429; doi: 10.1103/PhysRevLett.97.265702.
63. Jabraoui H., Ouaskit S., Richard J., & Garden J.-L. (2020). Determination of the entropy production during glass transition: Theory and experiment. *J. Non-Cryst. Solids* **533**, 119907. doi: 10.1016/j.jnoncrysol.2020.119907.
64. Delaney M., Nagy L., Kinsel M., & Treuting P. (2013). Spontaneous histologic lesions of the adult naked mole rat (Heterocephalus glaber): A retrospective survey of lesions in a zoo population. *Vet. Pathol.* **50**(4), 607–621. PMID: 23355517; doi: 10.1177/0300985812471543.
65. Tian X., Nagy L., Kinsel M.J., & Treuting P.M. (2013). High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature* **499**(7458), 346–349. PMID: 23783513; doi: 10.1038/nature12234.
66. Grimes K.M., Reddy A.K., Lindsey M.L., & Buffenstein R. (2014). And the beat goes on: maintained cardiovascular function during aging in the longest-lived rodent, the naked mole-rat. *Am. J. Physiol.-Heart Circ. Physiol.* **307**(3), H284–H291. PMID: 24906918; doi: 10.1152/ajpheart.00305.2014.
67. Evdokimov A., Kutuzov M., Petruseva I., Lukjanchikova N., Kashina E., Kolova E., … Lavrik O. (2018). Naked mole rat cells display more efficient excision repair than mouse cells. *Aging* **10**(6), 1454. PMID: 29930219; doi: 10.18632/aging.101482.
68. Buffenstein R., Amoroso V., Andziak B., Avdieiev S., Azpurua J., Barker A.J., … Smith E.S.J. (2022). The naked truth: A comprehensive clarification and classification of current 'myths' in naked mole-rat biology. *Biol. Rev.* **97**(1), 115–140. PMID: 34476892; doi: 10.1111/brv.12791.
69. Azpurua J., Ke Z., Chen I.X., Zhang Q., Ermolenko D.N., Zhang Z.D., … Seluanov A. (2013). Naked mole-rat has increased translational fidelity compared with the mouse, as well as a unique 28S ribosomal RNA cleavage. *Proc. Natl. Acad. Sci.* **110**(43), 17350–17355. PMID: 24082110; doi: 10.1073/pnas.1313473110.
70. Ng T.K.S., Matchar D.B., Pyrkov T.V., Fedichev P.O., Chan A.W., & Kennedy B. (2021). Association between housing type and accelerated biological aging in different sexes: Moderating effects of health behaviors. *Aging* **13**(16), 20029. PMID: 34456185; doi: 10.18632/aging.203447.
71. Chetty R., Stepner M., Abraham S., Lin S., Scuderi B., Turner N., … Cutler D. (2016). The association between income and life expectancy in the United States, 2001-2014. *Jama* **315**(16), 1750–1766. PMID: 27063997; doi: 10.1001/jama.2016.4226.
72. Tarkhov A.E., Lindstrom-Vautrin T., Zhang S., Ying K., Moqri M., Zhang B., … Gladyshev V.N. (2024). Nature of epigenetic aging from a single-cell perspective. *Nat. Aging* **4**(6), 854–870. PMID: 38724733; doi: 10.1038/s43587-024-00616-0.
73. Tong H., Dwaraka V.B., Chen Q., Luo Q., Lasky-Su J.A., Smith R., & Teschendorff A.E. (2024). Quantifying the stochastic component of epigenetic aging. *Nat. Aging* **4**(6), 886–901. PMID: 38724732; doi: 10.1038/s43587-024-00600-8.
74. Meyer D.H., & Schumacher B. (2024). Aging clocks based on accumulating stochastic variation. *Nat. Aging* **4**(6), 871–885. PMID: 38724736; doi: 10.1038/s43587-024-00619-x.
75. Perevoshchikova K., & Fedichev P.O. (2024). Differential responses of dynamic and entropic aging factors to longevity interventions. *bioRxiv 2024–02*.