Evolution and change: a principle of least selection

Michael Yarus
Department of Molecular, Cellular and Developmental Biology
University of Colorado Boulder
Boulder, CO 80309-0347

Keywords: RNA world, dispersion, variance, standard deviation, Normal distribution
Mobile: (303) 618-8017
Office: (303) 492-8376
Email: yarus@colorado.edu

Abstract

Favored characteristics evolve at two roughly-defined rates: selection of fitness < truncation ≈ truncation with fitness. The fastest route is also the most probable path; the path of least selection. When possible, truncations, requiring $x \geq$ threshold x_t, implement least selection for x. Least selection is also speeded by maximal σ (standard deviation) and maximal threshold x_t (consistent with some survivors). Least selection applies to both chemical and biological evolution. Non-genetic truncation is free of the Haldane cost of natural selection; and can yield very rapid non-Darwinian evolution. Least selection also unites prior examples of brisk evolution. For example: evolution via 'starting bloc selection' (in which initial action is crucial) and 'chance utility' (opportune fluctuation) and 'near-ideal reactions' (where undependable reactants recurrently yield a product) occur via transitory paths of least selection. Least selection's optimized path may still be clear even for transitions whose origins are lost. 'Bayesian convergence' (which defines a most probable path) is guaranteed by least selection's extremum. 'Distribution fitness' (specific outcomes from diffuse concurrent variation) is least selection from rare highly functional minorities. Inheritance of encoded chemical capabilities emerges because least selection exploits coexistence of templating and catalysis in a single molecular complex. The 'crescendo' of codes leading to the Standard Genetic Code is a specific device that makes a least selection. Thus, plausible evolutionary paths reach biological goals via least selections. Extinctions and radiations are events with coercive thresholds, ensuring that Earth's history reflects truncation/least selection.

Introduction

Evolution readily takes a subtle route. A natural population may include an opportune variant accidentally satisfying a proto-biological need. Thus, less abundant structures, less related to previous materials, can still provide the most likely evolutionary path. Further, a conceptual barrier to study of evolution is that even the simplest replication, metabolism and expression are complex molecular events – thus, there will be no obvious primordial route to them.

Instead, research must identify the *least improbable* routes to biological complexities. Moreover, forms known from laboratory experiments are unlikely to completely sample possible protobiological conformations or chemistry: we are unsure of initial conditions.

Given these conceptual and experimental barriers, there is an acute need for well-founded predispositions. Is there a general way to anticipate selection's rate? Is there a coherent way to compare hypotheses to project a more probable evolutionary transition? This text concerns these problems.

Following sections are meant to be semi-independently readable: first analysis (**immediately below**), then numerical illustrations (start at **Selection on a Normal distribution**), then previously-analyzed evolution reconsidered as least selections (start at **A principle of least selection**).

Selection from a varied population. Putting implications in explicit form begins with selection for x, denoted:

$$\varepsilon = a + bx \qquad (1)$$

where ε measures the number of descendants of individuals possessing x; this is usually called fitness and ε is called fitness below. Within Eqn 1, a is increase independent of molecule x. Increase dependent on the advantageous quality is bx, with fitness proportional to x availability.

The probability of x is often describable by a normalized ***p***robability ***d***istribution ***f***unction, *pdf*:

$$\int_{-\infty}^{\infty} pdf(x)dx = 1 \qquad (2)$$

To follow a population's fate, average x, μ , is followed. If $pdf(x)$ is the probability of x:

$$\mu = \int_{-\infty}^{\infty} x\, pdf(x)dx \qquad (3)$$

Selection changes $pdf(x)$:

$$pdf(x) \rightarrow \varepsilon\, pdf(x) = (a + bx)\, pdf(x) \qquad (4)$$

Selection for fitness. Consider selected increase of a favored quality x; such selected increase in mean x is described by the Price equation (Price 1970). The Price equation mathematically partitions evolutionary change into that due to selection and that due to all other frequency changes for qualities of interest (Frank 2012). We are particularly interested in selection:

$$\mu_{fit} = \mu + cov(\varepsilon, x)/\bar{\varepsilon}$$

Where *cov* is covariation and $\bar{\varepsilon}$ is population mean fitness:

$$\bar{\varepsilon} = \int_{-\infty}^{\infty} \varepsilon\, pdf(x)\, dt = a \int_{-\infty}^{\infty} pdf(x)dx + b \int_{-\infty}^{\infty} x\, pdf(x)\, dt = a + b\mu$$

$$cov(\varepsilon, x) = \int_{-\infty}^{\infty} (\varepsilon - \bar{\varepsilon})(x - \mu)\, pdf(x)\, dx = b\sigma^2$$

Thus $$\mu_{fit} = \mu + ({}^{b}/_{a + b\mu})\, \sigma^2 \qquad (5)$$

Fractional selected change in mean x is:

$$\Delta\mu_{fit} = (\mu_{fit} - \mu)/\mu = ({}^{b}/_{a + b\mu})\ {}^{\sigma^2}/_{\mu} \qquad (6)$$

Defined this way, $\Delta\mu_{fit}$ reflects selection's improvement. $\Delta\mu_{fit} = 0$ for no progress, and, scaled to the initial mean, $\Delta\mu$ increases as evolutionary change increases. $\Delta\mu = 1$ is selected doubling.

Fitness selection entirely dependent on x is particularly simple.

$$\Delta\mu_{fit} = {}^{\sigma^2}/_{\mu^2} \qquad (7)$$

This plausible situation is emphasized in Fig. 2 below: selection requiring x and $\propto x$ $(a = 0)$ increases average x precisely by dispersion squared, $(\sigma/\mu)^2$, independent of b. $\Delta\mu_{fit}$ is wholly defined by the initial distribution's shape. As intuition suggests, complete dependence on x $(a = 0)$ yields maximal fitness selection; larger by the factor $(b\mu + a)/b\mu$, thus necessarily the most fit condition. The Price equation is a mathematical identity (Frank 2012) and therefore always valid. And, no specific system property is assumed in calculations. Thus Eqn, 6, 7 are true of any integrable $pdf(x)$. $\Delta\mu_{fit}$ describes any system selected for fitness; genetic, non-genetic, prebiotic or solely chemical. Blue (before) and dotted red lines (after selection) in Fig. 1 illustrate fitness effects.

Truncation selection. Evolution may also truncate a distribution, propagating a favored greater subset (Fig. 1). That is, more than a threshold x_t may be required to survive and reproduce.

Z_t provides a generalized way to express truncation. Discussion of truncation is facilitated by Z co-ordinates. Z values are distances from a mean in σ units:

$$x_t = \mu + Z_t\sigma \qquad (11)$$

Z and Z_t can also be negative, measuring distance below the mean (Fig. 1). Use of Z for positions eases comparisons because standardized Normal distributions all have the same probability at the same Z (Methods). $\Phi(x_t)$ and $\Phi(Z_t)$ reference the same x and are unaltered by switching to Z (as are ϕ and ${}^{\phi}/_{\Phi}$), but x_t and Z_t are distinct (Eqn 11).

The Normal distribution as an example. To quantitate the fate of exceptional x_t or Z_t, x frequencies must be known. The Normal, or Gaussian, distribution (Methods) has a special status documented by the Central Limit Theorem (Feller, William 1970). The theorem shows that any property which is the sum of random variables, which may individually be distributed in any way, nevertheless approaches a Normal distribution as the number of variables increase. Normal and near-Normal behavior is accordingly widespread in real-world systems, making it a cogent example. Moreover, because of the broad usage of the Normal distribution, its integrals are evaluated in standard software. For example,

$$\int_{x_t}^{\infty} pdf(x)dx = \Phi(x_t) \qquad (8)$$

where $\Phi(x_t)$ from spreadsheet or mathematical software is the population fraction exceeding a threshold: e.g., for x between x_t and ∞ (Methods).

Mean x for selection above a threshold, μ_{trunc} measures selection's success, and can be evaluated by rearrangement:

$$\mu_{trunc} = \int_{x_t}^{\infty} x\, pdf(x)dx = \Phi(x_t)\mu_{trunc} = \int_{x_t}^{\infty} \mu\, pdf(x)dx + \sigma^2 \int_{x_t}^{\infty} {(x-\mu)}/{\sigma^2}\; pdf(x)\; dx$$

where $\Phi(x_t)$ is defined by Eqn 8, the integral-normalizing factor for μ_{trunc}. The rearrangement on the right can be integrated, because the second rightward term integrates the negative derivative of the Normal $pdf(x)$:

$$\Phi(x_t)\mu_{trunc} = \mu\, \Phi(x_t) - \sigma^2\, [pdf(\infty) - pdf(x_t)]$$

$$\Phi(x_t)\mu_{trunc} = \Phi(x_t)\mu + \sigma^2\, pdf(x_t)$$

$$\mu_{trunc} = \mu + \sigma^2\, pdf(x_t)/\Phi(x_t) \qquad (9)$$

Later we will use the fractional truncation-selected change in mean, in units of initial μ:

$$\Delta\mu_{trunc} = (\mu_{trunc} - \mu)/\mu = \left({\sigma^2}/{\mu}\right) \mathrm{pdf}(x_t)/\Phi(x_t) = \left({\sigma}/{\mu}\right)\phi(Z_t)/\Phi(Z_t) \qquad (10)$$

where ϕ is a sometimes-convenient form related to pdf (Methods).

$\Delta\mu_{trunc}$ assumes a Normal distribution, but utilizes no underlying system properties. Thus, Eqn 10's truncation-selected change, like Eqn 6 above, is general. Eqn 10 is valid in any Normal system: genetic, pre-genetic, biological or chemical. Progress, $\Delta\mu_{trunc}$, will be ∝ to dispersion, ${\sigma}/{\mu}$, and to ${\phi}/{\Phi}$. The hatched blue region rightward in Fig. 1 illustrates a truncation selection.

Truncation is somewhat enhanced by concurrent fitness. Threshold selection can also be somewhat more effective. Suppose that a characteristic being selected undergoes a truncation. For example, toleration of elevated temperature can aid survival, but a sudden temperature spike still truncates those beneath a new maximum. In Fig. 1, this selects the rightward blue hatched area, but now including individuals below the red dotted line. Below, such truncation on a selected background is quantitated.

We have already written the pdf for a fitness selection in Eqn 4, and normalized it so that its range includes all valid x (Eqn 5).

$$pdf_{fit}(x) = (a + bx/a + b\mu)\; pdf(x) \qquad (12)$$

If we instead truncate the fitness distribution:

$$\mu_{fit,\ trunc} = \left(\frac{1}{a\Phi(x_t)+b\mu_{trunc}}\right) \int_{x_t}^{\infty} x\,(a+bx)\, pdf(x)dx = \left(\frac{1}{a\Phi(x_t)+b\mu_{trunc}}\right)(a\mu_{trunc} + b\overline{x^2}_{trunc}) \qquad (13)$$

where $\overline{x^2_{trunc}}$ is the mean square x above the truncation. And $\overline{x^2_{trunc}} = \mu^2_{trunc} + \sigma^2_{trunc}$ where, importantly, all quantities are for the initial Normal, and all are known. μ_{trunc} is Eqn 9, and

$$\sigma^2_{trunc} = \sigma^2[1 + Z_t \frac{\phi(Z_t)}{\Phi(Z_t)} - \left(\frac{\phi(Z_t)}{\Phi(Z_t)}\right)^2]$$

comes from standard sources (Greene, William H. 2003). Combining all above as in Eqn 13:

$$\mu_{fit,\ trunc} = \frac{1}{a\Phi(Z_t)+b\mu_{trunc}}\left(a\left[\mu + \sigma^2\frac{\phi(Z_t)}{\Phi(Z_t)}\right] + b\left[\sigma^2 + \mu^2 + (Z_t\sigma + 2\mu)\frac{\sigma\phi(Z_t)}{\Phi(Z_t)}\right]\right) \qquad (14)$$

All quantities refer to the initial Normal, so are available via standard software. Moreover, combined selection and truncation in Eqn 14, like separated fitness (Eqn 5) and truncation (Eqn 9) above, assume only properties of the distribution itself. The result, Eqn 14, is therefore again valid for any Normal system: protobiological, chemical or biological, independent of genetic or chemical detail.

Selection on a Normal distribution. Fig. 1 shows unselected and selected distributions for x. The logarithmic vertical axis points up lesser change near the mean (including no change <u>at</u> μ), and resolves the increase in upper tail x. The initial Normal distribution (blue) is symmetrical around mean μ = 10. After selection for *x*, the selected, normalized population (Fig. 1, red dotted) has depressed *x* frequencies on the left, and elevated *x* frequencies above the mean on the right (required for selected frequencies to normalize to 1.0). In the example, fitness selection increases μ_{fit} by ≈ 4% ($\Delta\mu_{fit} \cong 0.04$; Eqn 7; dashed red arrow, Fig. 1).

Truncation and fitness selection compared. Fig. 1 also previews a more general conclusion; threshold selection is usually more effective; in fact, $\Delta\mu_{trunc}/\Delta\mu_{fit} > 12$ for this example distribution, comparing downward red and blue dashed arrows for selected μ in Fig. 1.

In greater detail: Fig. 2A shows that here, a small ε effect quickly reaches a maximum $\Delta\mu_{fit} = 0.04$ (blue data, $\sigma = 2$) as *b* (and emphasis on x) increases (Eqn 7). Non-linear improvement of $\Delta\mu_{fit}$ with σ^2 (Eqn 6) is also shown by results for varied σ values (dotted and solid lines are explained below).

In Fig. 2B, truncation behaves differently: as Z_t approaches the mean, $\Delta\mu_{trunc}$ becomes increasingly significant, and ultimately increases steeply and linearly. Notably, with assumed values, a threshold $Z_t \geq 4.8$ (Fig. 2B) more than doubles initial μ ($\Delta\mu_{trunc} \geq 1$). Substantial evolutionary change occurs via one truncation, in contrast to fitness selection (Fig. 2A). The superiority of truncation selection is retained even if x truncation is gradual rather than sharp (Kimura and Crow 1978; Crow and Kimura 1979). Accordingly, threshold selection has a special role below.

As the apposed straight dotted line in Fig. 2B indicates, truncation selections ($\Delta\mu_{trunc}$ and $\Delta\mu_{fit,trunc}$ below) improve ≈linearly above $Z_t = 1$ (≈ 0.3-fold selected increase).

Effect of σ on evolutionary change. Fig. 3A shows how initial variability σ alters selected improvement, $\Delta\mu$. Different selections exploit initial variation σ with varied efficiency.

Selected $\Delta\mu_{trunc}$ increases precisely linearly with σ (at constant Z_t) with no selection possible if there is no variability (Eqn 10; Fig. 3A).

Calculation of $\Delta\mu_{fit,trunc}$ shows that allowing fitness with truncation (Fig. 1) adds a few percent to the superior effect of truncation on selection. $\Delta\mu_{fit}$ is generally less effective than either truncation, even given its benefit from larger initial σ (Eqn 6, Fig. 2A). Most effective selection is via truncation or truncation+fitness, perhaps slightly better for the latter (Fig. 3B).

Fitness with increased variability creates truncations. $\Delta\mu_{fit}$ responds to σ^2 (Eqn 6), decreasing the gap between fitness and truncation selections as σ increases (Fig. 3A). But, increased σ does not yield highly effective fitness selection. Large σ broadens distributions, so that individuals with $x = 0$ will exist. For usual qualities of evolutionary interest, $x < 0$ is not meaingful. Thus, enhancing $\Delta\mu_{fit}$ by increasing σ creates truncations at $x = 0$ with better $\Delta\mu_{fit,\ trunc}$. This is shown in Fig. 2A, where dotted curves include the effect of truncation due to broad distribution (Eqn 14). For this reason, and given a tendency to smaller fitness selections (Fig. 2A, 3A), over a large range of conditions, truncations yield faster evolution.

Effect of Z_t on evolutionary change. The above conclusion is reinforced by the effect of varied thresholds (Fig. 3B). Here fractional selected mean increase, $\Delta\mu$, is presented versus varied threshold Z_t (constant standard deviation, $\sigma = 2$). $\Delta\mu_{trunc}$ is small at $Z_t < -2$, but then rapidly becomes significant, decisively increasing above the mean.

With simultaneous fitness and truncation, $\Delta\mu_{fit,trunc}$ is significantly increased even at $Z_t = -5$ because it includes an intrinsic small but complete fitness selection (Eqn 6). However, such fitness is a substantial benefit only far below μ; truncation becomes the major event at Z just below the mean (Fig. 3B).

Because of the rapid decline in Normal *pdf(x)* at higher Z_t (Fig. 1), linear increase in $\Delta\mu_{trunc}$ and $\Delta\mu_{fit,trunc}$ at high Z_t (Fig. 3B) are understandable. $\Delta\mu_{trunc}$ approaches Z_t as Z_t increases because selected individuals above Z_t become increasingly close to Z_t, assuring similar behavior for Z_t and $\Delta\mu$. Further, comparing the strict limit for $\Delta\mu_{fit}$ in Fig. 2B with this unlimited linear increase in $\Delta\mu_{trunc}$ and $\Delta\mu_{fit,trunc}$ (Fig. 3B), truncation superiority is reinforced.

The cost of truncation. However, truncation at greater Z_t is strongly constrained by the rarity of exceptional x: population size required to expect one improved individual above Z_t is shown on the righthand axis in Fig. 3B. Normal tails decline increasingly steeply with Z (Fig. 1), and the population needed to recover one improved x individual increases steeply and non-linearly with $\Delta\mu$. In Fig. 4, the population expected to contain one selected individual, $1/\Phi(Z_t)$, is compared to selected improvement $\Delta\mu_{trunc}$ for different initial variabilities, σ.

Substantial improvements (e.g., 50% of μ, $\Delta\mu_{trunc} = 0.5$) can be selected in populations with varied standard deviations. But such selection is critically dependent on initial variation - decreasing initial σ 3-fold increases the population required more than 5 orders of magnitude (Fig. 4). Thus, selection of an upper tail for moderate improvements is undemanding. But, if exceptional increases are vital – much larger populations, unusual properties, low survivals, and/or elevated variabilities (σ or σ/μ) are required. This constraint can be severe, with near-exponential population increases required for small improvements. Accordingly, finely divided populations with small individuals, unusually varied populations or both were surely crucial to rapid chemical evolution. Tiny, varied entities likely executed any emergent biology (Fig. 4).

Chemical selection is not slowed by the cost of natural selection. The cost discussion above has a passing resemblance to the classical discussion of the cost of natural selection, due to JBS Haldane (Haldane 1957). However, biological and chemical cases differ greatly. Haldane's cost was due to replacement of genes in a modern, genetic population: possessors of an old gene configuration must perish to be replaced by a new genetic constellation. Too much replacement yields too much perishing, and thus extinction. Haldane's cost may be mitigated by recombination (Hickey and Golding 2019), so may help account for the prevalence of sexual reproduction. Still, the rate of natural selection is cost-limited. However, in a non-genetic population, any fraction of a population may be replaced in a short time as long as a small selected portion persists. Thus, chemical least selection may be sporadic, but much more rapid than natural selection. As previously emphasized, inheritance itself might emerge in a few days (Yarus 2017).

A principle of least selection. A particularly useful idea arises from time-to-select. Maximal $\Delta\mu$ minimizes selections and thus minimizes time-to-evolve a particular x magnitude, thereby also indicating the most likely evolutionary route. Minimal time and maximal change can accommodate multiple evolutionary effects, incorporating both threshold and selection, evoking the faintly parallel physical principle of least action (Feynman, RP, Leighton, RB, Sands, M 2010). Finding this extremum identifies the most likely route to an improbable outcome. Call minimization of selections or time-to-select "least selection". In Discussion, the value of least selection as an interpretive principle is illustrated.

Discussion

Outline. Fluctuation frequently is the crux of evolutionary progress (Yarus 2013). This makes some evolutionary progress predictable (Yarus 2021a). Here this idea is quantitated, in that functions of standard deviation σ, dispersion σ/μ, threshold Z_t , and selection parameters *a* and *b*, explicitly determine selection's outcome with varied fitness and/or truncation. The results, viewed broadly, are simple. For Normally-distributed x, truncating by requiring minimal x yields maximal change. Truncation-selected change is usually greater than selection of fitness, and is accurately proportional to standard deviation σ or dispersion σ/μ. Improvement in a favored property after truncation also increases proportionately to a threshold for acceptable x, Z_t, at

241 larger Z_t. However, threshold selection is effective even at $Z_t = 0$ because a threshold at the
242 mean implies summing over increased $x > \mu$. Populations with highest σ and truncation
243 threshold Z_t require least selection, and will evolve most quickly. However, exploiting upper-
244 tail x for greater change can require large populations (Fig. 2B, 4). These properties allow
245 estimation of selection's relative effects, or discussion of selection adequate for specific
246 change.

247 **Example: Starting bloc selection exploits an inevitable threshold.** The starting bloc is an
248 example of effective evolution clarified by parallels to least selection.

249 The starting bloc (Yarus 2018) was encountered while calculating whether known
250 ribonucleotide reaction rates suffice for evolution of encoded, and therefore heritable,
251 chemical capacities (Yarus 2017). In fact, emergence of inheritable chemical reactions appeared
252 plausible.

253 Heritability's emergence depends on two well-known RNA capabilities. The first is existence of
254 simple catalytically active RNAs, homologues of the adenylate-dinucleotide-containing enzyme
255 cofactors (White 1976; Yarus 2011). The second is templated synthesis of similar molecules via
256 an ultimately simple, geochemically available (Ferris et al. 1996) RNA "gene", poly-U
257 (Puthenvedu et al. 2015). Selection of such template-directed catalytic synthesis of cofactors (in
258 preference to untemplated chemical synthesis) is uniquely effective early, when only some of
259 potentially capable sites have actually begun synthesis (called the "starting bloc", (Yarus 2018)).
260 Increases of templated synthesis selected in such a population are nearly proportionate to
261 dispersion (σ/μ) in the chemically active product over a large range of product concentrations
262 (Yarus 2018), as also anticipated for least selection.

263 **Starting bloc selection employs distribution tails.** The original starting bloc (Yarus 2018) also
264 produced roughly proportionately faster change when carried out with lower survival after
265 selection, a condition that was called "stringency". This parallels present calculations: increased
266 Z_t and decreased survival in Fig. 2B, 3B produces faster selection. Moreover, combining
267 variability (as in dispersion σ/μ or standard deviation σ) with survival in a single index best
268 predicted starting bloc success (Yarus 2018).

269 **Starting bloc selection exploits an early threshold.** The difference between functioning pools,
270 which make an active molecule early, and inactive ones with no product is particularly readily
271 selected. Starting bloc success via some, versus no, early activity exploits early truncation,
272 paralleling least selection in Fig. 3B.

273 Accordingly, starting bloc selection of a genetically encoded chemical capacity (Yarus 2017,
274 2018) follows, with almost formulaic precision, the variability, survival and threshold
275 dependence intrinsic to least selection (Fig. 3A, 3B, 4).

276 Starting blocs may be of fundamental interest: such nascent biological capabilities are
277 intrinsically selectable. In particular, encoded (genetic) expression is a crucial prebiotic

innovation. While replication is often cited as the foundational requirement for biology (Joyce 2002), there seems small advantage in replicating unexpressed sequence information. Thus, effective expression likely predates replication (Yarus 2011). Accordingly, the starting bloc's least selection potentially fixes expression, speeding expression united in a genome, and thus initiating vertical inheritance and advent of the Darwinian era.

Example: Chance utility captures favorable variability. Prebiotic systems must evolve toward biological complexity before genomes, or biology would be inaccessible. Selected temporary, or especially, permanent change in chemical systems before the advent of genetic inheritance has been called 'chance utility'. For example, variation in a 100-fold more concentrated inhibitor frequently provides an opportunity to capture a useful, but apparently overwhelmed minor reaction (Yarus 2016). This accidental contravariation of a substrate and its competitor relies on means and favorable variations for these reactants; that is, on chance resemblance to a favorable result, an opportunity for least selection.

Example: Random reactant supplies necessarily support near-ideal reactions. Even with random, uncontrolled nucleotide supplies, a hypothetical small RNA replication recurs (Yarus 2013). Notably, most of the replication in such a randomly-fed reactor comes from a recurrent subset of near-ideal substrate sequences. So, a randomly-fed pool can produce recurrent specific results from utterly unregulated inputs, an opportunity for least selection.

Least selection opportunities can flicker. The starting bloc, the chance utility example, and the mechanism of near-ideal reactions illustrate a common theme: selection can exploit kinetics. Occasions for least selection can flicker: transiently appearing, though usually absent. Truncation exploits favorable transients; relatively rapid selection in an evolving world that varies more slowly greatly extends truncation, and consequent least selection.

Example: Stable optima also imply evolutionary paths. Just as flickering opportunities for truncation predict evolutionary transitions, stable optimal selections also offer probable paths. An example is advent of initial, simplified wobble coding. In a forming genetic code, there appears an optimal time for wobble advent (Yarus 2021c). Such a time of maximal SGC resemblance offers a least selection, marking that era as plausible for initiation of wobble.

Example: Bayesian convergence follows from least selection's extremum. Bayes' theorem says that a correct origin hypothesis explains all aspects of a phenomenon. However, an explanation is uniquely effective when independent data are explained, and the hypothesis' probability is elevated by the product of independent confirmations (Yarus et al. 2005). For example, a fusion route to the Standard Genetic Code (Yarus 2021b) explains how metabolic, chemical and wobble order coexist without mutual interference in a near-universal coding table. Bayesian convergence is robustly supported by least selection, which maintains that the historical route to the genetic code was distinct, perhaps unique, because it traverses an extremum in relevant factors. This in turn overlaps Orgel's dictum that evolution exhibits Continuity (Orgel 1968); that is, later evolutionary states are necessarily linked to their precursors by continuous change.

Example: Distribution fitness is least selection, using different language. The fitness of a superior group within a distribution is a key to apparently improbable outcomes. For example, a small minority of nascent coding tables has unlikely coding properties (Table2; Yarus 2021a). These become ordered quickly, having made accurate Standard Genetic Code (SGC) assignments (Yarus 2021a), leaving some unassigned triplets that will later be critical during final assembly of a complete code (Yarus 2021b). They appear quickly because they require few slow events, like assignment decays (Yarus 2021a). This useful minority short-circuits the outward problem of evolving a complex pattern by accidently evolving it simply. Apparently implausible behavior; for example, showing what seems to be foreknowledge of late coding early in code history, is possible because they are a rare minority selected late. Their properties ultimately seem near-ideal, but they pre-exist without purpose. When selected as most functional (Yarus 2022), they immortalize their improbable histories, presenting a seeming paradox. This peculiar quality is routine in truncation selections, which capture relatively improbable states (Fig. 1, 4).

Thus, distribution fitness anticipates least selection - success from rare precursors with close resemblance(s) to a complex biological achievement.

More recent work on coding (Yarus 2022) clarifies the search for excellent minorities. Fusion of early partial codes strongly enforces code uniformity, because fused unlike codes assign ambiguous codon meanings, and are less fit. Fusion therefore generates a long-lived series of codes that increasingly resemble an underlying consensus, such as the Standard Genetic Code. Given a progressive increase in code uniformity (called the 'crescendo') that presents varying SGC-like encodings, code fusion proficiently implements SGC advent by least selection.

Example: Least selection appears as similarity to archaic forms. Least selection can be literal; vast gaps in time can link two similar things. Ribosomal peptidyl transferase is an example. CCdApPuro (Welch et al. 1995) is a rationally-designed oligonucleotide inhibitor that binds both halves of the ribosomal active site, contacting A- and P-sites in a configuration that allows peptide transfer between them. So, CCdApPuro emulates the transition state for peptide synthesis (Pauling 1948). When an RNA was selected from randomized sequences to bind CCdApPuro with affinity similar to ribosomes (Welch et al. 1997), the two most probable binding sites had a 17 nucleotide homology with the peptidyl transferase loop of rRNA. This homology included an exact 8-mer match with among the most conserved RNA sequences known, the ribosomal peptide transferase active site AUAACAGG. CCdApPuro also bound this RNA octamer in the large ribosomal subunit (Ban et al. 2000), in an orientation similar to selected RNA sequences (Welch et al. 1997; Yarus and Welch 2000). Selected ribonucleotides even include a likely catalytic feature, an A with a shifted pKa, though catalysis was not used in selection (Yarus and Welch 2000). These results strongly support initial peptide synthesis in an RNA cradle, and then least selection of a ribonucleoprotein descendant, still recognizably related to its RNA ancestor billions of years later.

Truncation is a major evolutionary event: extinctions. We know evolution best from recent Earth history, chiefly from durable fossils of relatively modern, multicellular creatures. Such evolution is sporadic, with long periods of relatively slower change interrupted by shorter eras containing mass extinctions (Jablonski 1986, 2002). There have been ≥ 7 relatively recent, major extinctions. Hodgskiss et al (Hodgskiss et al. 2019) present isotopic evidence for one such ending the Great Oxidation Event 2 gigayears ago. For the last 540 million years, see Muscente et al (Muscente et al. 2018) for such alternation of stability and catastrophe in the "big five" extinctions (Benton 1995). Ceballos (Ceballos et al. 2020) provide evidence of a current, human-induced catastrophe. More evident quick species losses are followed by re-diversification of surviving organisms (Hull 2015). Thus major extinctions are also radical truncations, making least selection unmistakably significant to life's history on Earth.

Truncation is a major evolutionary event: radiations. The mostly sessile oceanic life of the Edicaran era (588 - 539 Mya) had made major progress: it included metazoans who were bilateral, for example (Droser and Gehling 2015). However, these benthic creatures were almost completely replaced in a 'Cambrian explosion' (539 – 488 Mya) of novelty (Conway-Morris 2003), in which most modern phyla were established (Briggs 2015; Budd and Jensen 2017). Exceptionally-preserved worldwide Cambrian fossils in Canadian Burgess shale (Conway Morris 1999) and at Qingjiang in Southern China (Fu et al. 2019) validate the Cambrian radiation at the consensus origin of modern biota. Radiations, however radical, do not re-originate life. Thus major radiations are also radical truncations, making least selection unmistakably significant to life's history on Earth.

Plausible biological paths: least selections and biological intermediates. A small, rather than a grand, claim is made here; pursuit of least selection will not suddenly manifest biotic history. Neutral and near-neutral chemical change (Kimura and Ohta 1971) have not been considered. Even more significantly, plausible precursors may not yet be identified, so least selection cannot yet be defined.

Specifically, our present molecular portfolio is deeply partial. For example, prebiotic nucleotide activation has been an indispensable, but elusive missing link to both ancient translation and replication. Imidazole activates AMP for templated oligoRNA synthesis (Weimann et al. 1968). Existence of a possible prebiotic, UV-activated imidazole activation of AMP (Mariani et al. 2018) has guided related chemistry with 2-amino-imidazole and ensuing templated RNA synthesis from activated CMP and GMP (Zhang et al. 2020). Thus, a vital nucleotide activation gap may be filled.

Least selection complements the previous suggestion that the most plausible pathway yields a known intermediate with greatest probability (Yarus 2021a). Assembling least selections to plausibly intersect a conserved biological result creates a credible evolutionary path. When that path can be reduced to selection for a selected character, one will likely see regions with slower change (fitness selections, Fig. 2A, 3A) interspersed with larger steps an order-or-more steeper (truncations, Fig. 3A).

Acknowledgement. An earlier, less complete form of this discussion is available (Yarus 2021d).

Methods.

Numerical illustrations. Calculations were executed on a Dell XPS computer with an Intel Core i9-8950HK CPU @ 2.90GHz and 32 GB of RAM, running 64-bit Windows 10 Enterprise v. 20H2. Figures were prepared with Microsoft Excel 2016 16.0.5161.1002, making particular use of the Excel norm.dist function for Normal pdf, ϕ, and cumulative distributions, Φ.

As an example, in Eqn 8, $\Phi(x_t)$ is an integral not expressible in closed form using simple functions. But $\Phi(x_t)$ is readily accessible via the widely-available cumulative Normal distribution function $C(\mu, \sigma, x_t)$ (which yields the fraction of the distribution between -∞ and x_t): $\Phi(x_t) = 1 - C(\mu, \sigma, x_t)$.

The Normal probability distribution function. The Normal probability distribution function used here (e.g., in Eqn 8) appears in multiple forms:

$$pdf(x) = e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \Big/ \sqrt{2\pi}\sigma = e^{-\frac{Z^2}{2}} \Big/ \sqrt{2\pi}\sigma = \phi(Z) \big/ \sigma$$

using text notation. The denominator's standard deviation σ at right is significant because it differs from what is sometimes called the Standard Normal distribution ϕ, often seen in mathematical references (Greene, William H. 2003). The Standard Normal has $\mu = 0$, $\sigma = 1$, thus is not well suited to investigation of variations. For example, the above denominator σ cancels an external factor of σ, clarifying $\Delta\mu_{\ trunc}$ variation as $\propto \sigma$, not σ^2 (Eqn 10, Fig. 3A). In the text, pdf and ϕ are interspersed where each was thought clearest. A useful introduction to truncated Normal distributions can be found in (Greene, William H. 2003).

Competing interests. The author is aware of no competing interests.

Figure legends

Figure 1. **Selection and threshold selection on a Normal distribution**. The Normal probability of x, $pdf(x)$, is plotted as a solid blue line against a logarithmic ordinate: mean $\mu = 10$, standard deviation σ = 2. This distribution is selected, multiplying by $\varepsilon = a + bx$ (dotted red line); $a = 0.1, b = 0.5$. A labeled red dashed arrow marks the resulting selected mean, $\mu_{fit} = 10.39$, and black arrows illustrate the relation between the distributed quantity x and the generalized Z co-ordinate for the initial Normal. The dashed blue vertical line marks a selection threshold at $x_t = 14, Z_t = 2$, and its selected mean ($\mu_{trunc} = 14.746$) is right-adjacent, and dashed blue. Visible threshold-selected area is hatched with light blue diagonal dash-and-dotted lines.

Figure 2A. Fitness increase in mean $\boldsymbol{x}$, $\Delta\boldsymbol{\mu}_{\boldsymbol{fit}}$, with varied dependence on $\boldsymbol{x}$. $\epsilon = a + bx$, with a = 0.1 and b varying. Normal distribution parameters are those in Fig. 1. Solid lines are calculated for full distributions (Eqn 6), dotted lines for truncation of broad fitness distributions at $x = 0$ (Eqn 14). The lines for $\sigma = 2$ are too similar to be resolved.

Figure 2B. Selected truncation increase, $\Delta\boldsymbol{\mu}_{\boldsymbol{trunc}}$, across a Normal distribution as the acceptable threshold $\boldsymbol{Z_t}$ varies. Z_t ranges from below the mean ($Z_t = -3$) to far above ($Z_t = 5$). An arbitrary linear red dotted line has been superposed on Eqn 10 to illustrate linearity at $Z_t \geq 1$. Normal parameters are those in Fig. 1.

Figure 3A. Selected increase in mean $\mathbf{x}$, , $\Delta\boldsymbol{\mu}_{\boldsymbol{fit}}$, $\Delta\boldsymbol{\mu}_{\boldsymbol{trunc}}$, $\Delta\boldsymbol{\mu}_{\boldsymbol{fit,\ trunc}}$ as a function of increasing initial standard deviation $\boldsymbol{\sigma}$. $Z_t = 1$ throughout. Other Normal parameters are those in Fig. 1.

Figure 3B. Selected increase in threshold-selected means, $\Delta\boldsymbol{\mu}_{\boldsymbol{trunc}}$ and $\Delta\boldsymbol{\mu}_{\boldsymbol{fit,\ trunc}}$ as a function of threshold $\mathbf{Z_t}$, holding standard deviation $\sigma = 2$. The righthand ordinate is the population size required to find one improved individual above threshold Z_t, using truncation alone. Other Normal parameters are those in Fig. 1.

Figure 4. Population sizes to recover one superior individual as a function of selected improvement, $\Delta\boldsymbol{\mu}_{\boldsymbol{trunc}}$. Population sizes , $1/\Phi(Z_t)$, are plotted on against a logarithmic ordinate, assuming different starting variabilities, σ (Eqn 10). Normal parameters are those in Fig. 1.

References.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* **289**: 905–20.

Benton MJ. 1995. Diversification and extinction in the history of life. *Science* **268**: 52–58.

Briggs DEG. 2015. Extraordinary fossils reveal the nature of Cambrian life: a commentary on Whittington (1975) 'The enigmatic animal Opabinia regalis, Middle Cambrian, Burgess Shale, British Columbia.' *Philos Trans R Soc B Biol Sci* **370**: 20140313.

Budd GE, Jensen S. 2017. The origin of the animals and a "Savannah" hypothesis for early bilaterian evolution. *Biol Rev Camb Philos Soc* **92**: 446–473.

Ceballos G, Ehrlich PR, Raven PH. 2020. Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *Proc Natl Acad Sci U S A* **117**: 13596–13602.

Conway-Morris S. 2003. The Cambrian "explosion" of metazoans and molecular biology: would Darwin be satisfied? *Int J Dev Biol* **47**: 505–515.

Crow JF, Kimura M. 1979. Efficiency of truncation selection. *Proc Natl Acad Sci U S A* **76**: 396–399.

Droser ML, Gehling JG. 2015. The advent of animals: The view from the Ediacaran. *Proc Natl Acad Sci* **112**: 4865–4870.

Feller, William. 1970. An introduction to probability theory and its applications, v II, 2nd edition, p. 258-274.

Ferris JP, Hill AR, Liu R, Orgel LE. 1996. Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* **381**: 59–61.

Feynman, RP, Leighton, RB, Sands, M. 2010. *The Feynman Lectures on Physics*. Millenium Edition. California Institute of Technology.

Frank SA. 2012. Natural selection. IV. The Price equation. *J Evol Biol* **25**: 1002–1019.

Fu D, Tong G, Dai T, Liu W, Yang Y, Zhang Y, Cui L, Li L, Yun H, Wu Y, et al. 2019. The Qingjiang biota-A Burgess Shale-type fossil Lagerstätte from the early Cambrian of South China. *Science* **363**: 1338–1342.

Greene, William H. 2003. *Econometric Analysis; Section 22.2, Truncated DIstributions*. 5th ed. Prentice Hall.

Haldane JBS. 1957. The cost of natural selection. *J Genet* **55**: 511.

Hickey DA, Golding GB. 2019. Sex solves Haldane's dilemma. *Genome*. https://cdnsciencepub.com/doi/abs/10.1139/gen-2019-0051 (Accessed June 16, 2021).

Hodgskiss MSW, Crockford PW, Peng Y, Wing BA, Horner TJ. 2019. A productivity collapse to end Earth's Great Oxidation. *Proc Natl Acad Sci U S A* **116**: 17207–17212.

Hull P. 2015. Life in the Aftermath of Mass Extinctions. *Curr Biol CB* **25**: R941-952.

Jablonski D. 1986. Background and mass extinctions: the alternation of macroevolutionary regimes. *Science* **231**: 129–133.

Jablonski D. 2002. Survival without recovery after mass extinctions. *Proc Natl Acad Sci U S A* **99**: 8139–8144.

Joyce GF. 2002. The antiquity of RNA-based evolution. *Nature* **418**: 214–221.

Kimura M, Crow JF. 1978. Effect of overall phenotypic selection on genetic change at individual loci. *Proc Natl Acad Sci U S A* **75**: 6168–6171.

Kimura M, Ohta T. 1971. On the rate of molecular evolution. *J Mol Evol* **1**: 1–17.

Mariani A, Russell DA, Javelle T, Sutherland JD. 2018. A Light-Releasable Potentially Prebiotic Nucleotide Activating Agent. *J Am Chem Soc* **140**: 8657–8661.

Muscente AD, Prabhu A, Zhong H, Eleish A, Meyer MB, Fox P, Hazen RM, Knoll AH. 2018. Quantifying ecological impacts of mass extinctions with network analysis of fossil communities. *Proc Natl Acad Sci* **115**: 5217–5222.

Orgel LE. 1968. Evolution of the Genetic Apparatus. *J Mol Biol* **38**: 381–393.

Pauling L. 1948. Nature of Forces between Large Molecules of Biological Interest*. *Nature* **161**: 707–709.

Price GR. 1970. Selection and covariance. *Nature* **227**: 520–521.

Puthenvedu D, Janas T, Majerfeld I, Illangasekare M, Yarus M. 2015. Poly(U) RNA-templated synthesis of AppA. *RNA* **21**: 1818–25.

S. Conway Morris. 1999. *The crucible of creation*. Oxford University Press http://archive.org/details/crucibleofcreati00conw (Accessed July 26, 2022).

Weimann BJ, Lohrmann R, Orgel LE, Schneider-Bernloehr H, Sulston JE. 1968. Template-directed synthesis with adenosine-5'-phosphorimidazolide. *Science* **161**: 387.

Welch M, Chastang J, Yarus M. 1995. An inhibitor of ribosomal peptidyl transferase using transition-state analogy. *Biochemistry* **34**: 385–90.

Welch M, Majerfeld I, Yarus M. 1997. 23S rRNA Similarity from Selection for Peptidyl Transferase Mimicry. *Biochemistry* **36**: 6614–6623.

White HB. 1976. Coenzymes as Fossils of an Earlier Metabolic State. *J Mol Evol* **7**: 101–104.

Yarus M. 2011. Getting past the RNA world: the initial Darwinian ancestor. *Cold Spring Harb Perspect Biol* **3**: doi: 10.1101/cshperspect.a003590. PMID: 20719875; PMCID: PMC3062219.

508 Yarus M. 2013. A ribonucleotide origin for life - fluctuation and near-ideal reactions. *Orig Life Evol Biosph*
509 **43**: 19–30.

510 Yarus M. 2016. Biochemical Refinement Before Genetics: Chance Utility. *J Mol Evol* **83**: 89–92.

511 Yarus M. 2017. Efficient Heritable Gene Expression Readily Evolves in RNA Pools. *J Mol Evol* **84**: 236–252.

512 Yarus M. 2018. Eighty routes to a ribonucleotide world; dispersion and stringency in the decisive
513 selection. *RNA N Y N* **24**: 1041–1055.

514 Yarus M. 2021a. Evolution of the Standard Genetic Code. *J Mol Evol* **89**: 19–44.

515 Yarus M. 2021b. Fitting the standard genetic code into its triplet table. *Proc Natl Acad Sci U S A* **118**:
516 e2021103118.

517 Yarus M. 2021c. Optimal Evolution of the Standard Genetic Code. *J Mol Evol* **89**: 45–49.

518 Yarus M. 2021d. PreDarwinian evolution, prebiotic change: a principle of least selection.
519 2021.06.27.450095. https://www.biorxiv.org/content/10.1101/2021.06.27.450095v2 (Accessed
520 September 9, 2022)

521 Yarus M. 2022. A crescendo of competent coding (c3) contains the Standard Genetic Code. *RNA N Y N*
522 rna.079275.122.

523 Yarus M, Caporaso JG, Knight R. 2005. Origins of the genetic code: the escaped triplet theory. *Annu Rev*
524 *Biochem* **74**: 179–198.

525 Yarus M, Welch M. 2000. Peptidyl transferase: ancient and exiguous. *Chem Biol* **7**: R187-190.

526 Zhang SJ, Duzdevich D, Szostak JW. 2020. Potentially Prebiotic Activation Chemistry Compatible with
527 Nonenzymatic RNA Copying. *J Am Chem Soc* **142**: 14810–14813.

528

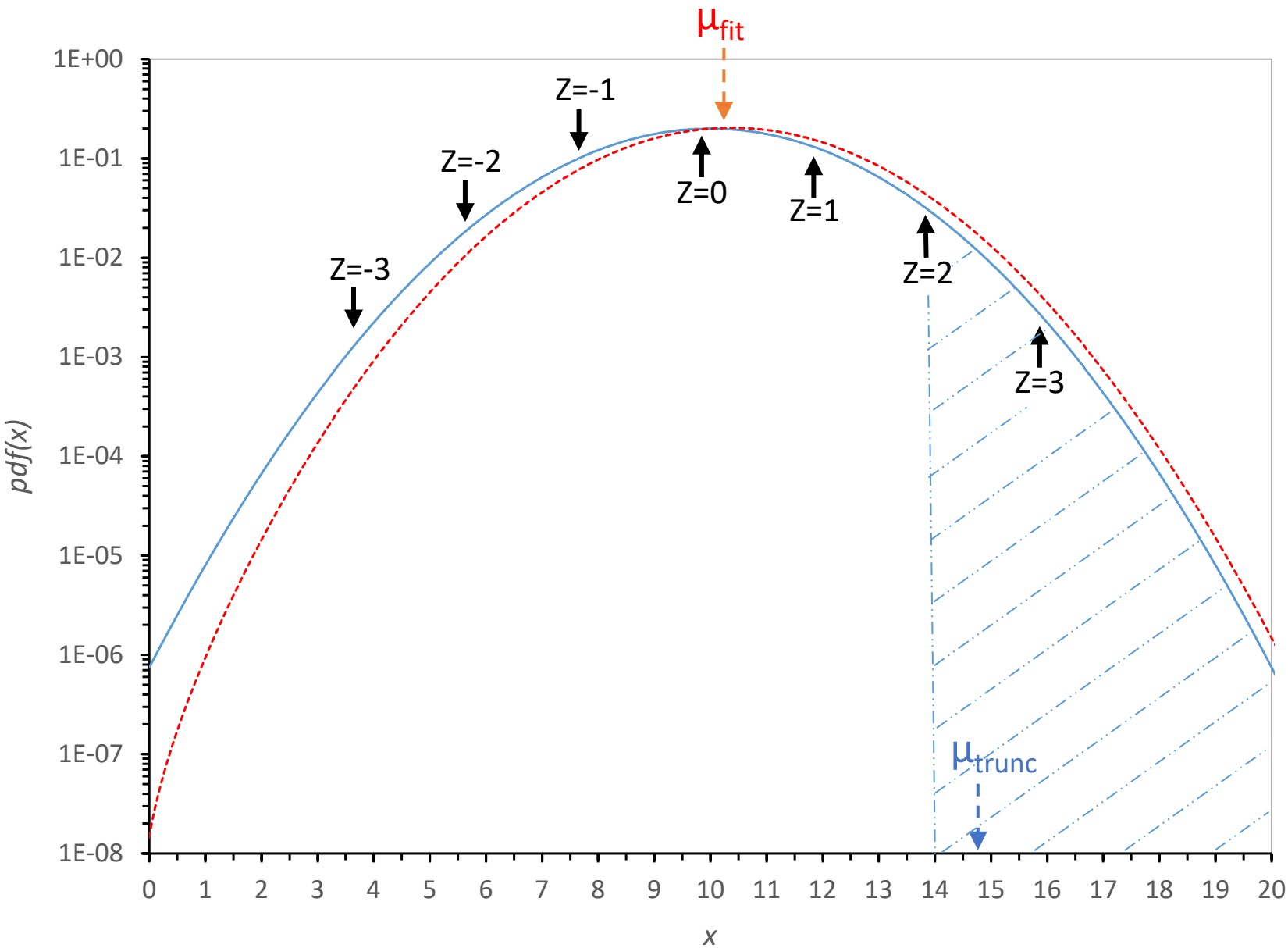
μfit
Z=-1
Z=-2
Z=-3
Z=0
Z=1
Z=2
Z=3
μtrunc
1E+00
1E-01
1E-02
1E-03
1E-04
1E-05
1E-06
1E-07
1E-08
pdf(x)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
x

Fig. 1

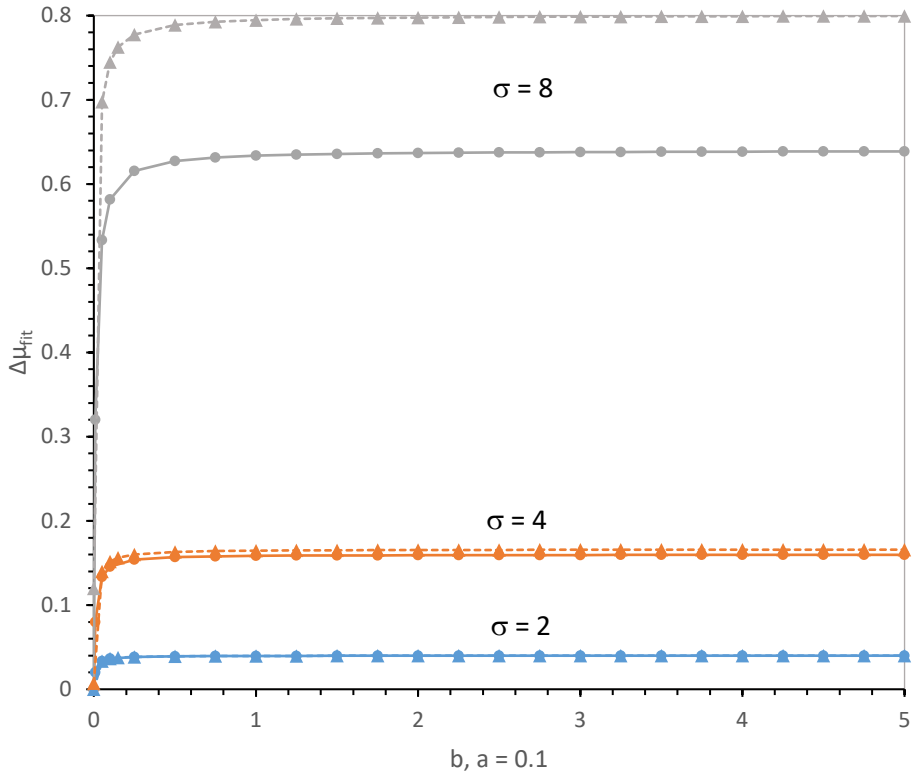

Fig. 2A

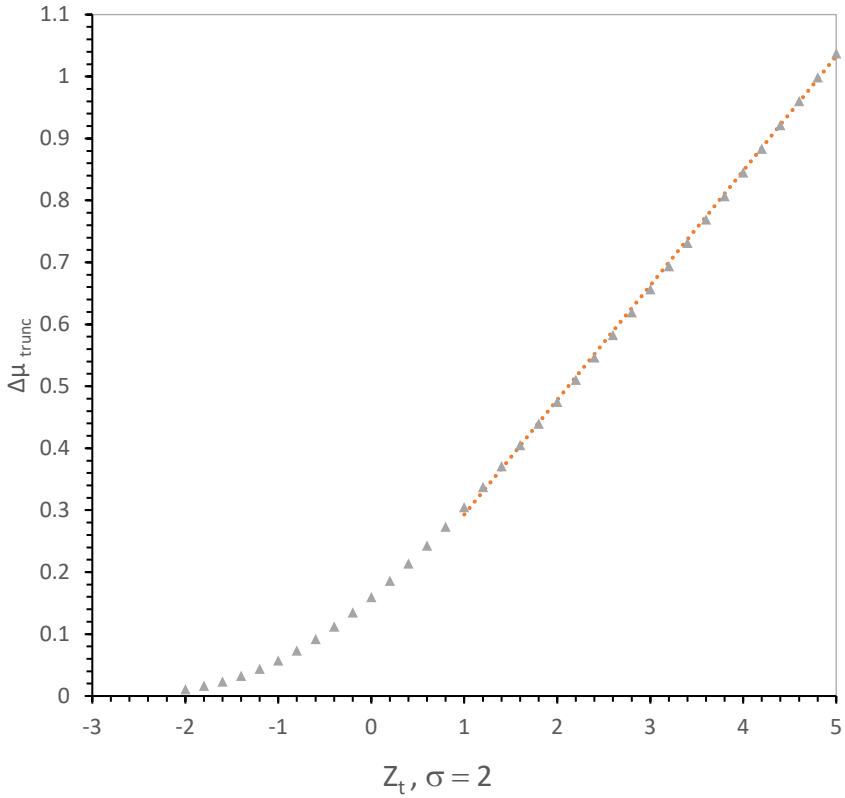

Fig. 2B

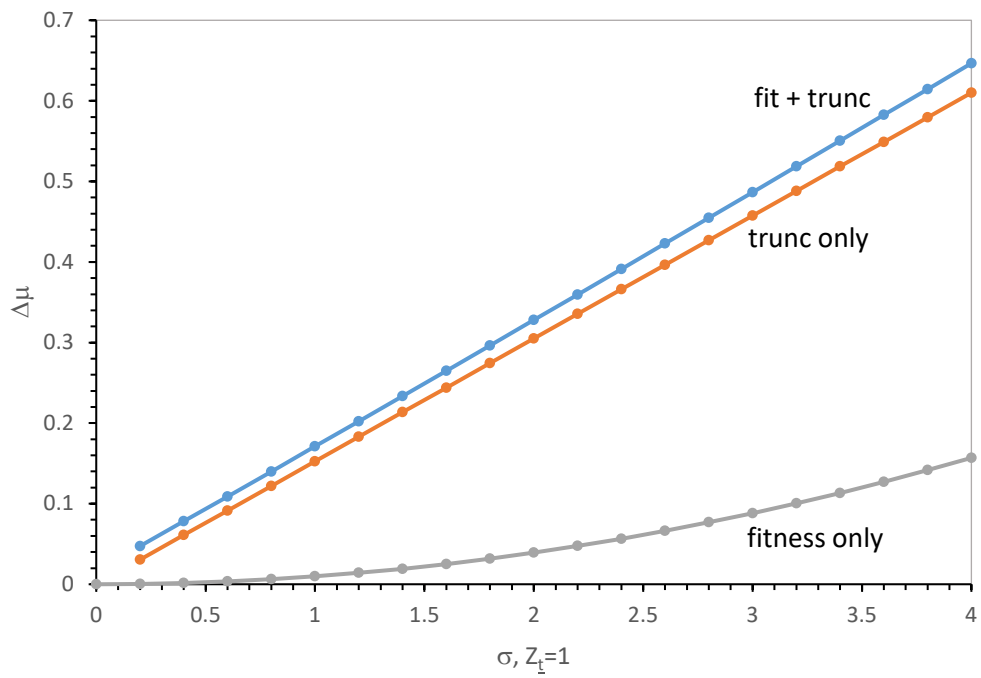
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Δμ
fit + trunc
trunc only
fitness only
0
0.5
1
1.5
2
2.5
3
3.5
4
σ, Z_t=1

Fig. 3A

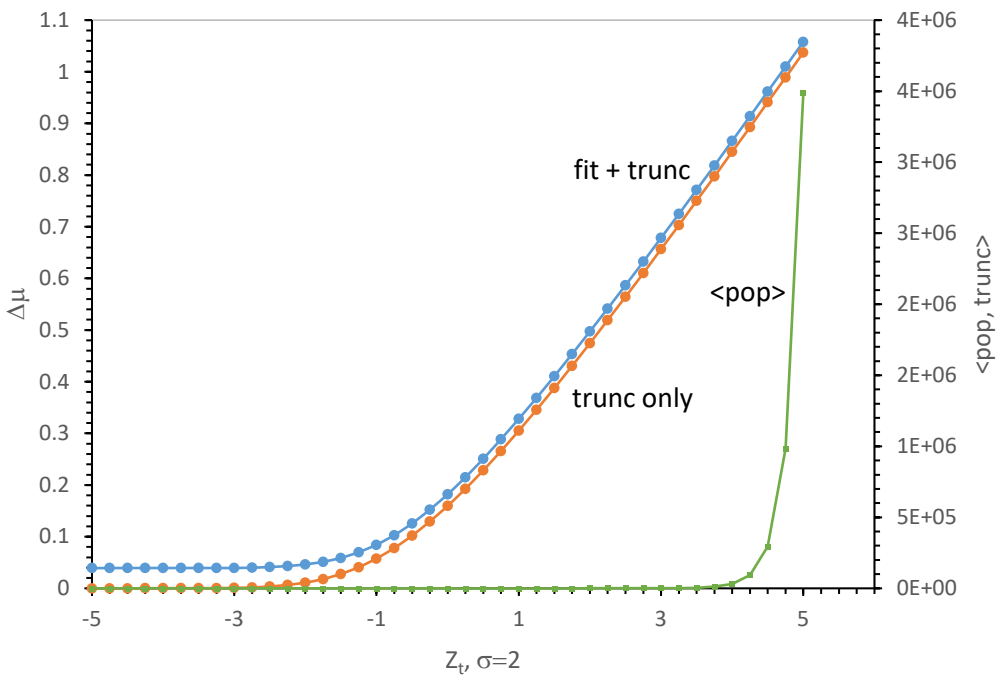
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Δμ
4E+06
4E+06
3E+06
3E+06
2E+06
2E+06
1E+06
5E+05
0E+00
<pop, trunc>
fit + trunc
<pop>
trunc only
-5
-3
-1
1
3
5
Z_t, σ=2

Fig. 3B

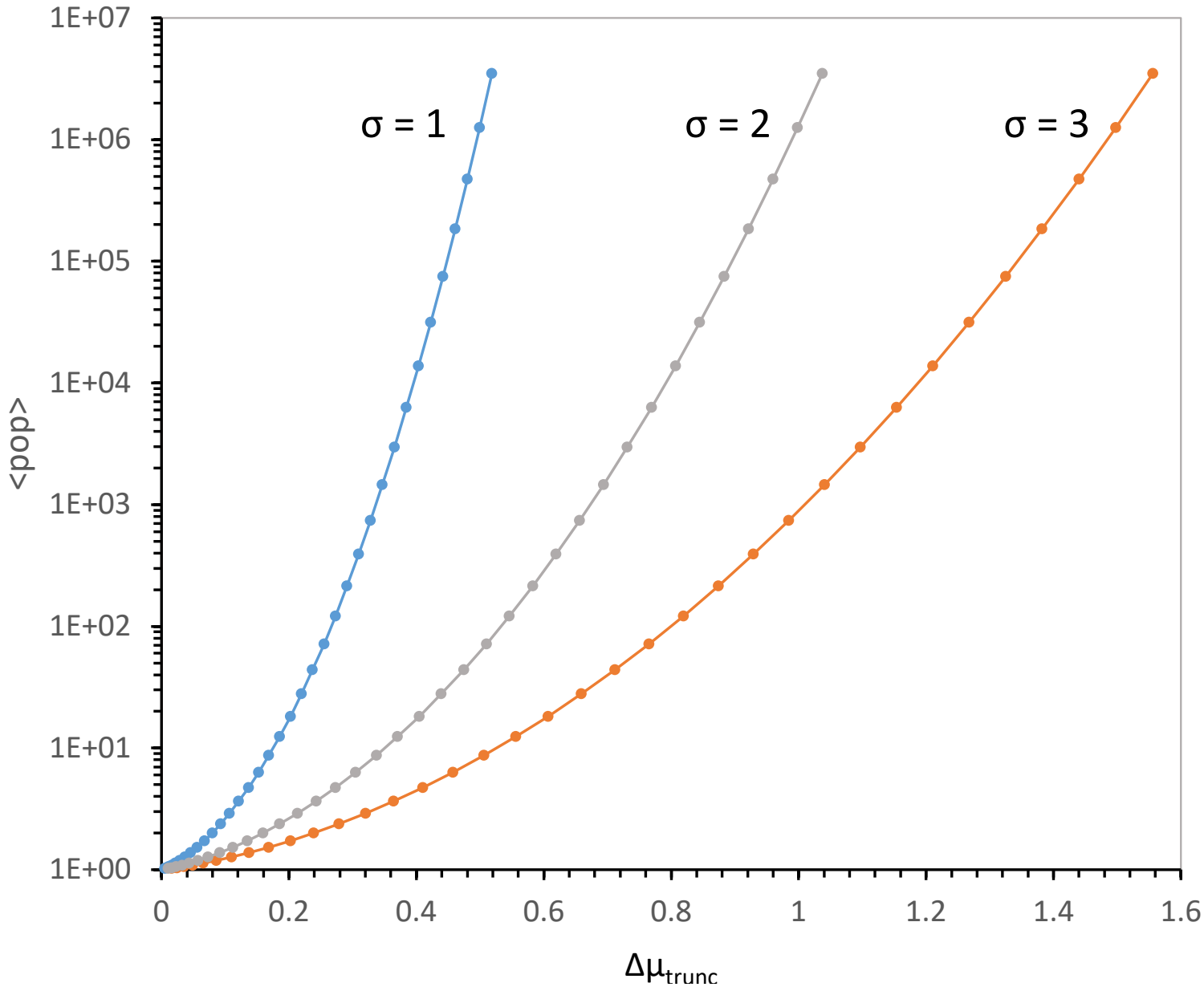
1E+07
1E+06
1E+05
1E+04
1E+03
1E+02
1E+01
1E+00
<pop>
σ = 1
σ = 2
σ = 3
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
Δμtrunc

Fig. 4