On possible indicators of negative selection in germinal centers

Bertrand Ottino-Loffler[1] and Gabriel D. Victora[2]

[1]Center for Studies in Physics and Biology, Rockefeller University, 1230 York Avenue, New York, 10065, New York, USA.
[2]Laboratory of Lymphocyte Dynamics, Rockefeller University, 1230 York Avenue, New York, 10065, New York, USA.

Contributing authors: bottino@rockefeller.edu; victora@rockefeller.edu;

Abstract

A central feature of vertebrate immune response is affinity maturation, wherein antibody-producing B cells undergo evolutionary selection in microanatomical structures called germinal centers, which form in secondary lymphoid organs upon antigen exposure. While it has been shown that the median B cell affinity dependably increases over the course of maturation, the exact logic behind this evolution remains vague. Three potential selection methods include encouraging the reproduction of high affinity cells ("birth/positive selection"), encouraging cell death in low affinity cells ("death/negative selection"), and adjusting the mutation rate based on cell affinity ("mutational selection"). While all three forms of selection would lead to a net increase in affinity, different selection methods may lead to distinct statistical dynamics. We present a tractable model of selection, and analyze proposed signatures of negative selection. Given the simplicity of the model, such signatures should be stronger here than in real systems. However, we find a number of intuitively appealing metrics – such as preferential ancestry ratios, terminal node counts, and mutation count skewness – require nuance to properly interpret.

Keywords: Evolutionary Dynamics, Dynamical Systems, Population Dynamics, Germinal Centers, Negative Selection

1 Introduction

In mathematical biology, the term "evolution" has two common associations. The first is the natural selection of living organisms over time [1] – the second is the algorithmic optimization of some fitness variable [2–6]. At times, these two meanings are conflated. In both, a system contains species of diverse fitnesses, with fitter genotypes persisting longer in time, and mutation generating novel genotypes. In the case of the computational algorithm, "fitness" is a well defined attribute which directs the dynamics to a concrete objective. However, real living systems have no universal definition of fitness, so it is hard to frame it as an optimization process. However, there is one place where natural and algorithmic evolution coexist – affinity maturation.

A vertebrate's body is exposed to an endless onslaught of pathogens, to which it responds by producing a large variety of tailored antibodies that bind to and ultimately neutralize these threats. Antibodies are initially generated by a genetic reshuffling process known as V(D)J recombination [7]. While the efficacy of these initial antibodies is poor, during infection the body starts generating higher and higher quality antibodies [8, 9], thanks to process known as *affinity maturation*. Affinity maturation is a process which occurs in germinal centers (GCs), microanatomical structures that form within secondary lymphoid organs (e.g. lymph nodes, spleen) upon exposure to antigen. Here, a diverse population of B cells (each producing their own antibody) is subjected to evolutionary pressure and high levels of mutation [10]. The objective is to find an antibody with a high binding affinity to the target antigen [11, 12]. We should note there is also interest in the topic of broadly neutralizing antibodies [13–15], but for this paper we will focus on narrow affinity maximization

In essence, the immune system is running an evolutionary optimization algorithm, with fitness corresponding to antigen-antibody binding affinity. But what is unknown is the *actual* algorithm. Historically, some GC models used birth-limited selection (also known as positive selection), where high fitness cells have accelerated division rates, whereas others used death-limited selection (also known as negative selection), where high fitness cells have diminished death rates [16–19]. Presently, there is empirical evidence for birth-selection, but a lack of evidence of affinity-based death selection [11, 20, 21].

A naive way to model fitness is to just treat it as the difference of birth and death rates, making it a 1D variable [22]. However, there are many contexts in which it is important to make fitness multidimensional. For example, in the world of network Moran models, there is a split between birth-selective and death-selective models [23, 24]. Despite what the 1D perspective would imply, these seemingly equivalent Moran models can have different outcomes. This is starkly true in fractional takeover times (the time for a single strain to take over x% of the population), where swapping the selection method can cause drastic distributional changes [25].

Unfortunately, results for network Moran model takeover times are not amiable to GCs. Not only do GCs lack a convenient graph structure, but Moran takeover times are usually calculated in a low-mutation limit, whereas GCs feature Somatic Hyper-Mutation (SHM). Thanks to SHM, a B cell's antibody-encoding gene experience on average one mutation per 10^3 base pairs per division, a million times the baseline rate [10, 11, 26].

Interestingly, there is some evidence that this mutation rate is not constant. So-called "clonal bursts" may occur in germinal centers, where a single B cell divides over and over again with apparently low rates of mutation [11, 27]. This suggests the possibility of there being *mutational selection*. Imagine two otherwise identical cell lines with different mutation rates: the genotype with the higher rate will persist less over time, simply by virtue of its of its offspring constantly turning into new strains. Therefore, having a differential mutation rate between high and low affinity B cells would introduce a third form of selection, independent of birth or death selection.

In this paper, we will build an analytically tractable model for selection that incorporates birth, death, and mutational selection as independent parameters, rendering a multidimensional fitness. For the sake of having our results be generalizable to other, non-GC evolutionary systems, we will avoid incorporating detailed germinal center mechanics, such as cyclic reentry, interactions with follicular dendritic cells or helper T cells, and receptor-antigen molecular dynamics [11, 28–31]. While we will continue to reference GCs as an core example, our results should apply to a wide array of asexual evolutionary systems.

By building such a simple model, we avoid messy confounding factors, so all signals of negative selection should appear with maximum fidelity. Therefore, we can pre-evaluate the usefulness of metrics, and see if they can successfully and qualitatively discriminate between positive and negative selection. This is an important consideration for experimentalists, since each additional observed metric may represent a substantial monetary cost, so we demand to know if it can or can not accurately measure negative selection.

When it comes to experimental design, we ideally wish for this measure of negative selection to be a static quantity (to prevent the need for multiple measurements across time), and to have a strong qualitative signal (to prevent the need to account for microscopic model details and calculations).

Despite the simplicity of our model, we unintuitively find that many seemingly promising signatures of negative selection confer weak information. That is to say, many candidate measures of negative selection can not, on their own, be used to qualitatively prove or disprove the presence of negative selection. However, there are cases where quantitative bounds can be put on the strength of negative selection, but that may require system-specific knowledge.

In this paper, we will start by describing our model, as well as its steady state. Then, we examine patterns in how different genotypes emerge, and analyze phylogenetic tree statistics. Finally, we examine the statistics and dynamics of the mutation count distribution. We conclude that, in many cases, more sophisticated methods will required to accurately assess if negative selection is present in a system or not.

2 Model of Generic Selection

We will develop a reduced model of n cells evolving in a well-mixed environment. For the sake of simplicity, we will assume there to be only two distinct affinity phenotypes, high affinity (H) and low affinity (L). This is not an unprecedented restriction [32], especially since there are cases where a single mutation can increase affinity by a factor

of ten [33], essentially rendering the rest of the genome into a high-dimensional neutral space.

We use a discrete time formulation, where exactly one event (a division or a death) occurs every timestep. We say that the germinal center has a carrying capacity of N, and the event is a birth with probability proportional to n, and a death with probability proportional to n^2/N (so the population is stable at $n = N$, with logistic-style growth [34, 35].) The odds that a cell dies during a death step is inversely proportional to its *death fitness*, which is 1 for low affinity cells and $r_D \geq 1$ for high affinity cells – so a higher level of death fitness corresponds to longer-living strains. Similarly, the odds that a cell divides during a birth step is proportional to its *birth fitness*, which is 1 for low affinity cells and $r_B \geq 1$ for high affinity cells. That is, r_B controls the level of positive selection, and r_D controls negative selection. We will take r_B and r_D to be finite in the main body of the text, and cover the infinite cases in appendix B in the supplemental information. If $r_D = 1$, then there is no negative selection whatsoever, and $r_B = 1$ corresponds to no positive selection.

Since we may have mutational selection, the mutation rates for one affinity line may be larger than the other. On division, H cells will have a mutation rate of ρ_H, with a η_H chance of the mutant being low affinity and $1-\eta_H$ chance of remaining high affinity (thus being a neutral mutation). Similarly, L cells have a mutation rate of ρ_L, with a η_L chance of the mutant being high affinity, and a $1 - \eta_L$ chance of remaining low affinity. It is often more useful to use the net transfer rates, so

$$\alpha_H = \rho_H(1 - \eta_H), \quad \beta_H = \rho_H\eta_H,$$
$$\alpha_L = \rho_L(1 - \eta_L), \quad \beta_L = \rho_L\eta_L.$$

α is the rate of neutral mutations, that is, H→H and L→ L mutations. Meanwhile β is the rate of mutating into a different affinity H→L and L→H. Note $\alpha_H + \beta_H = \rho_H$ and $\alpha_L + \beta_L = \rho_L$.

On a technical note, cellular division is typically symmetric, so one would expect both daughter cells to have equal odds of being a mutant ρ_X. However, somatic hypermutation in the germinal center is powered by the action of AID on a single strand of DNA at a time. So it is possible for one daughter cell to have a mutation rate that is a million times greater than its sister cell [36, 37]. With that in mind, and for the sake of notational cleanness, we will only allow one daughter cell of the two to be a potential mutant, with the other being a copy of the mother cell. In appendix A, we cover the generic case of arbitrary asymmmetric mutation rates, allowing for both daughters to be mutants.

3 Net Selection

The most immediate quantity to calculate is the overall level of selection – that is, what is the final percentage of high affinity cells, h = number of H cells/N?

Per timestep, let P_B be the probability of a birth/division, and P_D the probability of death. So,

$$P_B = \frac{1}{1 + n/N},$$

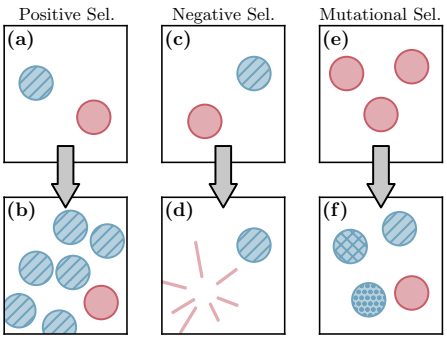

Fig. 1: Schematic of different forms of selection. Hatched blue circles represent high affinity cells, solid red circles represent low-affinity cells. (a) and (b) illustrate positive/birth selection – as time passes, high affinity cells divide more rapidly than low affinity cells, crowding them out. (c) and (d) illustrate negative/death selection – as time passes, the low affinity cells die faster, leaving the high affinity cells to survive. (e) and (f) illustrate mutational selection – as time passes, the transition gradient causes the low affinity cells to create a diversity of mutants, with a net increase of affinity.

$$P_D = \frac{n/N}{1 + n/N}.$$

To find the probability of a certain strain dividing, we let $\ell = 1 - h$ and use

$$\begin{aligned} Z_B &= \frac{P_B}{r_B h + \ell}, \\ P_{BH} &= r_B h Z_B, \\ P_{BL} &= \ell Z_B, \end{aligned}$$

where Z_B is just a normalization factor, and the individual probabilities are proportional to the strain's population and birth fitness.

Similarly, to find the probability of a certain strain having a death, we use

$$\begin{aligned} Z_D &= \frac{P_D}{h/r_D + \ell}, \\ P_{DH} &= (h/r_D) Z_D, \\ P_{DL} &= \ell Z_D, \end{aligned}$$

where Z_D is a normalization factor, and the individual probabilities are proportional to the strain's population and inversely proportional to its death fitness.

We will now use these probabilities to estimate the average change in $h = \#$ H cells$/n$ over time. First, the probability that the H population goes down by one is simply the probability that an H cell dies, so $P_{-1} = P_{DH}$. To find the probability of a new H cell appearing, we need to consider three sources: an H cell divides (w/o

mutation), an H cell divides (with an H→ H mutation), and an L cell divides (with an L → H mutation). Putting these together, we get:

$$\begin{aligned} P_{+1} &= P_{BH}(1-\rho_H) + P_{BH}\alpha_H + P_{BL}\beta_L \\ &= r_B h Z_B(1-\beta_H) + \ell Z_B \beta_L. \end{aligned}$$

The average population of high affinity cells changes by $P_{+1} - P_{-1}$ every time step, so for large N we can approximate by

$$\partial_t h = r_B h Z_B(1-\beta_H) + \ell Z_B \beta_L - (h/r_D) Z_D. \tag{1}$$

We want to find the steady-state value of h, so let's assume we already hit $n \equiv N$, and therefore $P_D = P_B = 1/2$. So we want to solve

$$0 = r_B \frac{1-\beta_H}{h(r_B-1)+1} + \frac{\beta_L(1-h)}{h(r_b+1} - \frac{h/r_D}{h(1/r_D-1)+1}.$$

For reasons which will be obvious later, let's assume $h = 1/(1+r_B g)$. Substituting in and rearranging, we get

$$0 = \; r_B r_D \beta_L g^2 + (r_B r_D(1-\beta_H) + \beta_L - 1)\, g - \beta_H. \tag{2}$$

Solving for g (and taking the $g \geq 0$ solution) gives

$$g = \frac{1}{2\beta_L r_B r_D}\left(1 - \beta_L + r_B r_D(\beta_H - 1) + \sqrt{\hat{g}(r_B r_D)}\right),$$

where

$$\begin{aligned} \hat{g}(x) = 1 &+ \beta_L^2 + x^2(1+\beta_H^2) + 2x\beta_H(1-x) \\ &+ 2\beta_L(-1+x) - 2x(1-\beta_H\beta_L). \end{aligned}$$

Therefore,

$$h(r_B, r_D) = \frac{1}{1 + r_B g(r_B r_D)}, \tag{3}$$

with the special case at $r_B = r_D = 1$ being

$$h(1,1) = \frac{\beta_L}{\beta_H + \beta_L}. \tag{4}$$

Notably, while g is symmetric in r_B and r_D, the net level of selection h is asymmetric in birth and death selectivity, as seen in figure 2. Consider the case where $r_B \gg 1$ and $r_D = 1$, so high affinity cells are dividing at a much higher rate than low affinity cells. However, because the transition rate β_H is nonzero, many of those new cells will

be low affinity as well, and will stick around for a while because the negative selection pressure is nonexistent. In the limit of large r_B, then

$$\lim_{r_B \to \infty} h(r_B, r_D) = \left(1 + \frac{\beta_H}{r_D(1-\beta_H)}\right)^{-1} < 1.$$

This means that there is an effective *ceiling of selection* for positive selection, wherein you will always have a mixture of high and low affinity, not matter how selective the system is. Meanwhile, with negative selection,

$$\lim_{r_D \to \infty} h(r_B, r_D) = 1.$$

So negative selection has no ceiling whatsoever.

In particular, if a real-world evolutionary system is observed to have

$$h_{\text{measured}} > h(\infty, 1) = 1 - \beta_H, \tag{5}$$

then said system **must have negative selection** $r_D > 0$, since the force of positive selection alone is insufficient to attain that level of pressure. This boundary is indicated by the dashed contour line in figure 2.

Conversely, because increasing r_D monotonically increases h until it reaches 1, if we know h and the mutation rates β_H and β_L, we can establish an upper bound on r_D. If we rearrange equation 2 relating $r_B r_D$ and g, we find

$$r_B r_D = \frac{\beta_H/g + 1 - \beta_L}{\beta_L g + 1 - \beta_H}.$$

Using $h = 1/(1 + r_B g)$ simplifies this to

$$r_D = \frac{h}{1-h} \frac{\beta_H r_B h + (1-\beta_L)(1-h)}{(1-\beta_H) r_B h + \beta_L (1-h)}. \tag{6}$$

This is monotone increasing or decreasing in r_B, depending on the strength of mutation. In the case of weak mutation ($\beta_H + \beta_L \leq 1$), we can bound the strength of negative selection via

$$r_D \leq \frac{h}{1-h} \frac{\beta_H h + (1-\beta_L)(1-h)}{(1-\beta_H) h + \beta_L (1-h)}. \tag{7}$$

Meanwhile, if mutation is high ($\beta_H + \beta_L > 1$), then

$$r_D \leq \frac{h}{1-h} \frac{\beta_H}{1-\beta_H}. \tag{8}$$

Returning to the specific case of germinal centers, it is well established that a diversity of affinity levels persists. In experiments, even after three weeks, cells can be

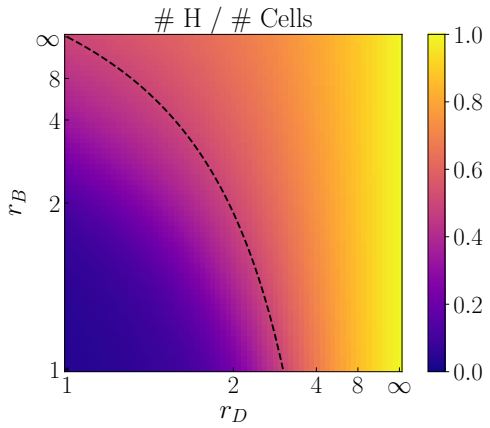

Fig. 2: Heat map of h via equation (3). Here, $\alpha_H = 0.196$, $\alpha_L = 0.679$, $\beta_H = 0.504$, and $\beta_L = 0.021$. Axes are plotted according to $1 - 1/r_X$ to span from one to infinity. The dashed contour indicates $h = 1 - \beta_H$.

observed with a *lower* affinity than the initial seed population [38]. With this logic, if negative selection is present in germinal centers it should not be extremely strong, though this may also be an artifact of the gap between division and selection being long enough for us to observe cells slated for death.

As a quick and dirty, back-of-the-envelope example, we can use the standardly quoted mutation rate of 10^-3/per base pair to get $\rho \approx 0.7$ [10]. We can use Araki's estimates and affinity binning to get transition rates $\eta_H \approx 0.25$ and $\eta_L \approx 0.024$, as well as a single GC measurement in Araki's figure 2.30 of $h \approx 68/79$ [38]. Naively putting these numbers into our bounding equation (7) to get an upper bound on the level of negative selection $r_D \lesssim 2.5$. However, since $h \approx 0.86$ and $1-\beta_H \approx 0.825$, we we satisfy the condition in equation (5), which would indicate the presence of negative selection. So just based on this one limited dataset, with this one specific affinity binning, we would say $1 < r_D \lesssim 2.5$.

Again, a larger, more dedicated dataset (as well as more GC-centric model) would be required to put confidence in these numbers. We are not using this paper to specifically claim proof of negative selection in germinal centers. What we are demonstrating here is that such bounds are simple to calculate for a given evolutionary system, and the parameters required for this calculation already can be measured in reasonable expirements. Therefore, methods such as these should be considered in future expiremental designs.

4 The Ancestry Hypothesis

While this model may only two affinity levels, we can still have a diversity of genotypic strains via neutral mutations (e.g., high → high affinity mutation). In principle, we can keep track of where these strains come from, therefore giving each strain an ancestor strain. We assume that the genotype space is large compared to the population of

cells, so every new mutation produces a novel strain. That is to say, if strain A has a mutation during division, it will produce strain B, and if strain A has another mutation later on, that second mutation will be distinct from strain B.

Looking at such phylogenetic trees, one might hypothesize that there may be a hidden signature of negative selection. One candidate signature is *preferential ancestry*. That is, we count how many high affinity cells have a high affinity strain as their ancestor, denoted by F_H. Similarly, we let F_L be the fraction of low affinity cells with a high affinity strain as an ancestor.

There are three distinct sources of high affinity cells. Per timestep, we attain an average of $P_{BH}(1-\rho_H)$ high affinity cells from nonmutating high affinity divisions, $P_{BH}\alpha_H$ from mutating high affinity divisions, and $P_{BL}\beta_L$ from low to high affinity divisions. Similarly, there are two sources of H cells with H ancestors: on average we get $F_H P_{BH}(1-\rho_H)$ such cells from non-mutating divisions, and $P_{BH}\alpha_H$ from H cells mutating into other H cells.

Taking a steady-state, this means

$$F_H = \frac{F_H P_{BH}(1-\rho_H) + P_{BH}\alpha_H}{P_{BH}(1-\rho_H) + P_{BH}\alpha_H + P_{BL}\beta_L}.$$

Keeping in mind $P_{BL}/P_{BH} = (1/h - 1)/r_B = g(r_B r_D)$, then we have

$$F_H = \frac{\alpha_H}{\alpha_H + \beta_L g(r_B r_D)}. \tag{9}$$

By a similar accounting, we find that

$$F_L = \frac{\beta_H}{\beta_H + \alpha_L g(r_B r_D)}. \tag{10}$$

Notably, both F_H and F_L only depend on positive and negative selection through their product $r_B r_D$. As seen by figures 3a and 3b, these preferential ancestry ratios are perfectly symmetric in these two fitness parameters. Therefore, in this model, such metrics are ill suited for qualitatively identifying the presence or absence of negative selection, since it makes no distinction between r_B and r_D.

It should be noted that we can combine these results with equation for h in (3). By doing so, we can solve for positive fitnesses, with

$$r_B = \frac{\beta_L}{\alpha_H}\frac{F_H}{1-F_H}\frac{1-h}{h}, \tag{11}$$

and

$$r_B = \frac{\alpha_L}{\beta_H}\frac{F_L}{1-F_L}\frac{1-h}{h}. \tag{12}$$

Therefore, we can solve for r_D exactly using equation (6). While in principle we can use these equations to estimate r_B and r_D for real-world expiremental systems, caution must be taken. The fact remains that F_H shows no *qualitative* transition, so it

makes for a poor heuristic. The equations above rely on specific, nonuniversal model details, so their application to real data must be done toughtfully.

As an example of thoughtless application, we can again use the transition rates from [10] and the phylogenetic tree in figure 2.30 of [38]. While there is ambiguity as to how to bin low versus high affinity cells, we can eyeball $F_H = 1$ and $F_L = 8/11$. Using equations (11) and (12), we get $r_B = \infty$ and $r_B \approx 1.7$. The reason for the extreme range is the small sample size, combined with the $F/(1-F)$ terms leading to some nasty behavior around $F = 1$. These values lead to $r_D \approx 1.3$ and $r_D \approx 2.0$, which are unsuprisingly within the $1 < r_D < 2.5$ bounds we calculated before.

While not conclusive, this "smell test" seems to indicate that F_H and F_L are underpowered for estimating fitnesses, at least compared to the easier to measure h.

Terminal Nodes

One related metric worth discussing here is the concept of "terminal nodes." As the system of cells continues to divide and reproduce, strains will produce novel mutant descendants, creating a phylogenetic tree of ancestries. At any point of time, this tree will have terminal nodes, which are strains with no extant descendants. It can be conjectured that if most low affinity cells are in terminal nodes, then there ought to be negative selection at work.

A proper analytic exploration of this metric is somewhat out of scope of this paper, but we can generate examples numerically. In figure 3c, we show the number of L cells in terminal nodes divided by the total number of L cells. Notice that this figure appears symmetric in r_B and r_D. Intuitively, this makes sense: if L cells only appear as terminal nodes, then either they die before they divide (high negative selection), or they don't get a chance to divide in the first place (high positive selection).

Similarly, the total fraction of cells which are in terminal nodes shows poor signal in r_D in figure 3d. While it might be tempting to tease out a trend, it should be noted in real-world datasets, phylogenetic trees have to be attained via statistical reconstruction. The simulations here create the tree with perfect knowledge, whereas in practice there will be ambiguity over cell terminality, which would likely drown out whatever slight signal may be present.

In addition, most other ratios involving number of terminal cells can be derived from the prior two metrics, f_H, F_L, and the total selection level h. While some terminal node metrics may have signals of negative selection, they would simply be from the signal present in the much easier to measure h. Moreover, pilot simulations for counting number of genotypes (instead of number of cells) suggest the same results.

As such, it is hard to recommend any terminal node metric as a qualitative signal of negative selection.

5 Mutation Count Distribution

Even though this model has only two phenotypes (low and high affinity), because of the large neutral space, the number of potential genotypes is extremely high. As before, we will assert that no two mutation events will ever be identical. That is, it is unlikely two separate mutation events will lead to the exact same base pair sequence.

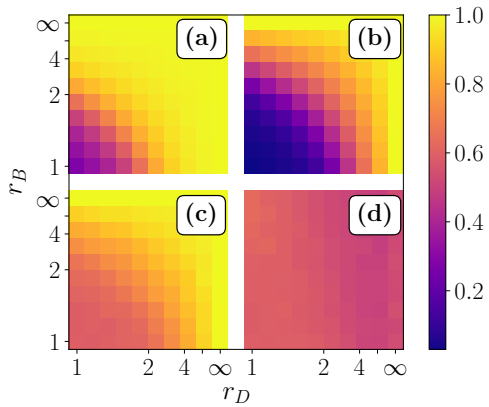

Fig. 3: Heat maps of various tree metrics. (a) Shows F_H via equation (9). (b) Shows F_L via equation (10). (c) Shows the final fraction of L cells which are in terminal nodes. (d) Shows the final fraction of cells which are in terminal nodes. For all, $\alpha_H = 0.196$, $\alpha_L = 0.679$, $\beta_H = 0.504$, and $\beta_L = 0.021$. Axes scale as $1 - 1/r_X$ to span from one to infinity. Simulations for (c) and (d) were ran over $N = 1e4$ and $T = 6e5$ steps, for a total of 60 generations.

If we initialize the germinal center with a population of genetically identical cells, we should be able to count how many mutations are accumulated in each cell line.

If a species grows at a rate r and has a mutation rate ρ, then you'd expect the population to accrue mutations at an average rate of $r\rho$. In our model, we have two separate growth rates and mutation rates, r_B & ρ_H and 1 & ρ_L, but the actual average accumulation rate is somewhat more complicated due to the L$\rightarrow$ H and H $\rightarrow$ L mutations coupling the two mutation distributions.

More to the point, dynamically measuring a population's mutational distribution is not always a good option. For many biological systems, doing a genotypic survey can be expensive and/or destructive (e.g., needing to destroy a germinal center to analyze its B cells), so measuring dynamical properties is usually unappealing.

However, if the distribution of mutation counts is different between low and high affinity cells, then intuitively the overall mutation count distribution ought to be asymmetric. That is, we can hypothesize that the mutation count distribution ought to have some level of *skew* over long times, as in figure 4. Moreover, a back of the envelope calculations suggests that the skew would approach a constant value over time. Skew is calculated via the central moments C_a, and a naive examination would suggest $C_a \propto t^a$. Therefore, the skew would look like

$$\text{Skew} = \frac{C_3}{(C_2)^{3/2}} \approx \frac{t^3}{(t^2)^{3/2}} = \text{constant}.$$

Because this is a prediction of the long-term shape of the distribution, and not its dynamics, it can be measured and approximated by a single snapshot genotypic survey. The experimental advantage of taking one measurement over a longitudinal study

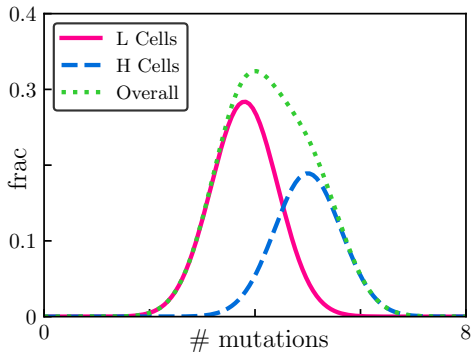

Fig. 4: Schematic showing the intuition behind the skew hypothesis. Here, different levels of mutational activity causes a split between the high fitness and low fitness B cells, producing skewness in the overall distribution.

need not be elaborated upon. Moreover, this final skew value should be a function of the different fitness parameters. Therefore, one can hypothesize that measuring mutational skew may be a good indicator of negative selection.

What follows is an explanation as to why this hypothesis is false.

To investigate the mutation count distribution, let's assume that the overall affinity of the population has hit a steady state. Next, let us define h_m to be the number of high affinity cells with m mutations, divided by the total number of cells n. Similarly, we let ℓ_m be the fraction of cells which are low affinity and have m mutations, and $n_m = h_m + \ell_m$. Note that $\sum_m h_m = h$, and $\sum_m \ell_m = \ell = 1 - h$.

To construct the dynamics for h_m, we do some basic accounting. Sources for H cells with m mutations are H and L cells with $m-1$ mutations, as well as H cells with m mutations. In the single-mutant dynamics, the only sink is the natural death rate. Therefore, the dynamics are given by

$$\begin{aligned}\partial_t h_m &= (1-\rho_H)(h_m r_B Z_B) + \alpha_H (h_{m-1} r_B Z_B) + \beta_L (\ell_{m-1} Z_B) - (h_m / r_D) Z_D \\ \partial_t \ell_m &= (1-\rho_L)(\ell_m Z_B) + \alpha_L (\ell_{m-1} Z_B) + \beta_H (h_{m-1} r_B Z_B) - \ell_m Z_D.\end{aligned}$$

To get the central moments C, we first need the regular moments. To find the moments of the mutational distributions, we define the moments $\mathcal{H}_k := \sum_m m^k h_m$ and $\mathcal{L}_k := \sum_m m^k \ell_m$ for the H and L populations respectively. Taking ∂_t of both sides, we get

$$\partial_t \mathcal{H}_k = \left[(1-\beta_H) r_B Z_B - \frac{Z_D}{r_D}\right] \mathcal{H}_k + \beta_L Z_B \mathcal{L}_k + \sum_{w=0}^{k-1} \binom{k}{w} (r_B \alpha_H Z_B \mathcal{H}_w + \beta_L Z_B \mathcal{L}_w) \tag{13}$$

$$\partial_t \mathcal{L}_k = \left[(1-\beta_L)\, Z_B - Z_D\right] \mathcal{L}_k + r_B \beta_H Z_B \mathcal{H}_k + \sum_{w=0}^{k-1} \binom{k}{w} (r_B \beta_H Z_B \mathcal{H}_w + \alpha_L Z_B \mathcal{L}_w), \tag{14}$$

where we handled h_{m-1} and ℓ_{m-1} by changing the sum index and using the binomial theorem, and taking $h_{-1} = \ell_{-1} = 0$.

While unpleasant to look at, this is fundamentally a linear system of equations. Moreover, since each k'th moment only depends on moments k to 0, this gives it a block lower-triangular structure, which should be amenable to analysis. In principle, a closed-form solution involving matrix exponentials should be possible. However, in appendix C, we show that this approach is both numerically and theoretically fraught, due to the dynamical matrix becoming poorly conditioned over time. Therefore, a slightly different approach is needed.

To make analysis easier, we will instead use the following change of variables, and assume at least one of r_B and r_D are larger than 1:

$$\mathcal{M}_k := \mathcal{H}_k + \mathcal{L}_k, \tag{15}$$

$$\mathcal{S}_k := (r_B Z_B \beta_H / \ell) \mathcal{H}_k + (Z_B \beta_L / h) \mathcal{L}_k. \tag{16}$$

$\mathcal{M}_k$ represents the k'th moment of the mutation count distribution for the overall population, including both H and L cells. Meanwhile, the definition of $\mathcal{S}_k$ was just chosen to make the dynamics cleaner, as given by

$$\partial_t \mathcal{M}_k = -\tau \mathcal{M}_k + \mathcal{S}_k + M_M \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{M}_w + M_S \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{S}_w, \tag{17}$$

$$\partial_t \mathcal{S}_k = S_M \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{M}_w + S_S \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{S}_w. \tag{18}$$

The prefactors are given by

$$M_M = Z_B \frac{\rho_L \beta_H - r_B \rho_H \beta_L g}{\beta_H - \beta_L g},$$

$$M_S = \frac{r_B \rho_H - \rho_L}{r_B \beta_H / \ell - \beta_L / h},$$

and

$$S_M = \frac{r_B Z_B^2 \beta_H \beta_L}{r_B \beta_H h - \beta_L \ell} \left(\beta_H / g - r_B \beta_L g + \alpha_L - r_B \alpha_H\right),$$

$$S_S = Z_B \frac{r_B \beta_H \alpha_H + \beta_H \beta_L (r_B g - 1) - \alpha_L \beta_L g}{\beta_H - g \beta_L}.$$

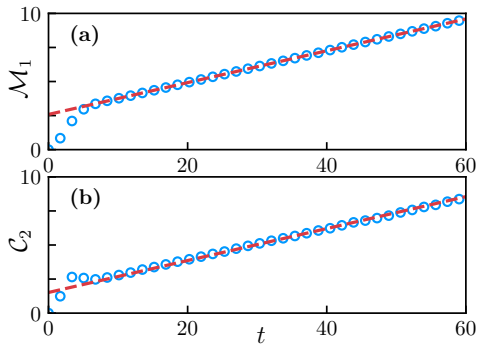

Fig. 5: Plot of the mutation count's (a) average $\mathcal{M}_1$ and (b) variance C_2. he circles are values attained from simulation, using $N = 2.5e5$, $r_B = 10$, and $r_D = 5$, with $\alpha_H = 0.196$, $\alpha_L = 0.679$, $\beta_H = 0.504$, and $\beta_L = 0.021$. The dashed line indicates the slope as predicted by (19) in (a), and by (24) in (b).

τ is given by

$$\tau = Z_B \left(\frac{\beta_H}{g} + \beta_L r_B g \right).$$

Notice that $\tau > 0$ always, so it functions as a proper timescale for the system.

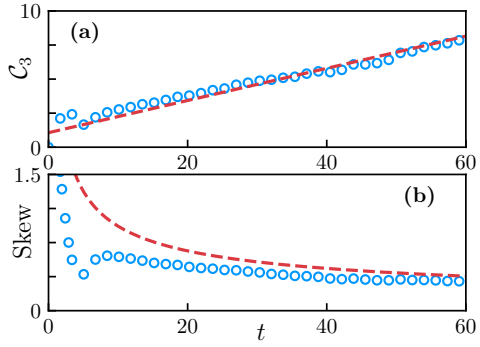

Fig. 6: Plot of the mutation count's (a) third central moment C_3 and (b) skew. The circles are values attained from simulation, using $N = 2.5e5$, $r_B = 10$, and $r_D = 5$, with $\alpha_H = 0.196$, $\alpha_L = 0.679$, $\beta_H = 0.504$, and $\beta_L = 0.021$. The dashed line in (a) indicates the predicted slope of C_3 (via appendix D), and in (b) a $1/\sqrt{t}$ curve to illustrate the gradual decay, as predicted by equation (25).

Also note that $\mathcal{S}_k$ only depends on lower order terms, meaning that it is possible to get exact solutions by iteratively solving. Both $\mathcal{S}_k$ and $\mathcal{M}_k$ will be polynomials in

time t with a maximum degree k, plus some other terms that decay exponentially with rate τ.

Looking at the definitions of $\mathcal{M}$ and $\mathcal{S}$ in (15) and (16), it is easy to see $\mathcal{M}_0 = 1$ and $\mathcal{S}_0 = \tau$. By solving the differential equations for $k = 1$, we get

$$\begin{aligned}\mathcal{S}_1(t) &= \tau\mu t + \theta_1^S, \\ \mathcal{M}_1(t) &= \mu t + \left(M_M + \tau M_S + \theta_1^S\right)/\tau + \left(\theta_1^M - (M_M + \tau M_S + \theta_1^S)/\tau\right) e^{-\tau t},\end{aligned}$$

where the θ terms are just initial conditions, and $\mu := S_M/\tau + S_S$. That is to say, the growth rate of the average number of mutations over time is μ, which can be rewritten as

$$\mu = r_B Z_B \frac{\alpha_H \beta_H / g + \alpha_L \beta_L g + 2\beta_H \beta_L}{\beta_H / g + r_B \beta_L g}. \tag{19}$$

Omitting the details of the recursion (which can be found in appendix D), we get that the leading order behavior for the k'th moment is given by

$$\mathcal{M}_k[t^k] = \mu^k. \tag{20}$$

As expected, the k'th moment grows like t^k, and the leading rate follows a simple form. The expression for the first correction (which is needed to calculate the distribution's skew) is somewhat more unwelcoming, with

$$\mathcal{M}_k[t^{k-1}] = \frac{k}{\tau}\mu^{k-1}\theta_1^S + k\mu^{k-2}\left[\frac{(k-1)}{2}\mu + \left(\frac{S_M}{\tau}(k-1) + \mu\right)\left(\frac{M_M}{\tau} + M_S - \frac{\mu}{\tau}\right)\right]. \tag{21}$$

In order to calculate the mutation distribution's skew and variance, we will need the central moments. The a'th central moment is defined by

$$\begin{aligned}C_a &= \langle (m - \mathcal{M}_1)^a \rangle \\ &= \sum_{k=0}^{a} \binom{a}{k} \mathcal{M}_k {\mathcal{M}_1}^{a-k} (-1)^{a-k}.\end{aligned}$$

Just as with the regular moments, these should be polynomials of maxium degree a and decaying exponential corrections.

We don't need the full solution for the central moment, just the leading growth term. But by some careful manipulation (also found in appendix D), we find that the supposed "leading" term is always zero,

$$C_a[t^a] = 0. \tag{22}$$

When we try the next leading order term, we similarly find for $a \neq 2$ that

$$C_a[t^{a-1}] = 0. \tag{23}$$

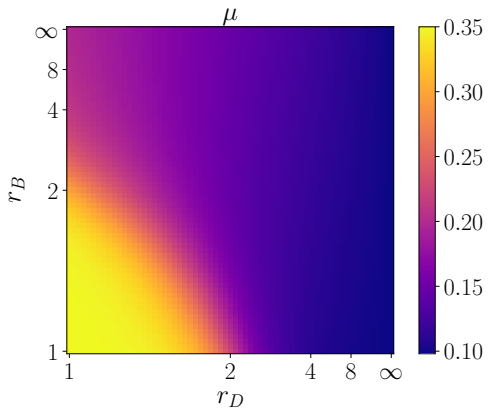

Fig. 7: Heat diagram of the growth rate μ of the average mutation count $\mathcal{M}_1$, via equation (19). Here, $\alpha_H = 0.196$, $\alpha_L = 0.679$, $\beta_H = 0.504$, and $\beta_L = 0.021$. Axes are plotted according to $1 - 1/r_X$ to span from one to infinity.

However, for the special case of $a = 2$, we get

$$C_2[t^1] = \mu + \frac{2S_M}{\tau}\left(\frac{M_M}{\tau} + M_S - \frac{\mu}{\tau}\right). \tag{24}$$

While C_2 has a t^1 term, we find C_3 has no t^2 term, C_4 has no t^3 term, and so on.

In particular, we now know that the variance (given by C_2) and the third central moment (C_3) must grow linearly in time. This is corroborated by simulation in figures 5a,5b, and 6a.

With this knowledge, we can finally estimate the skew of the mutant distribution. Plugging into the definition of skew, we find

$$\text{Skew} = \frac{C_3}{{C_2}^{3/2}} = \frac{\mathbb{O}(t)}{\mathbb{O}(t^{3/2})} \to 0. \tag{25}$$

This is supported by simulation results in 6b, where the skew indeed decays as $1/\sqrt{t}$. Therefore, there will be no long-term skew in the mutation count histogram, regardless of the choice of fitnesses. And so, the hypothesis outlined at the start of the section is false.

6 Discussion

Uncovering the mechanisms behind an evolutionary process is key – not only for advancing understanding and recreating natural systems [39–41], but also for improving our own optimizing algorithms [2, 42, 43], and providing avenues for better medical interventions [26].

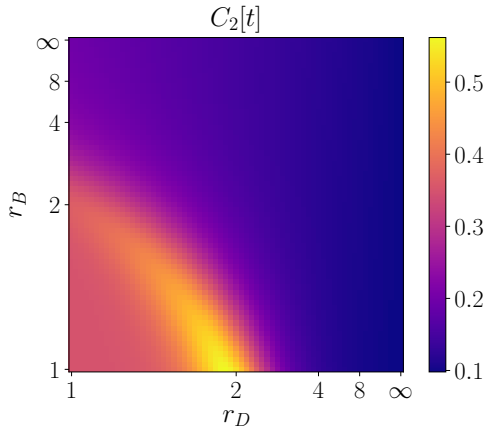

Fig. 8: Heat diagram of the growth rate $C_2[t]$ of the mutation count's variance, via evaluating equation (24) at $a = 2$. Here, $\alpha_H = 0.196$, $\alpha_L = 0.679$, $\beta_H = 0.504$, and $\beta_L = 0.021$. Axes are plotted according to $1 - 1/r_X$ to span from one to infinity.

Based on our work, certain proposed metrics for measuring such mechanisms are somewhat underwhelming, since we lack clear qualitative signals present in other evolutionary models [25]. For example, preferential ancestry ratios and mutational count skewness are symmetric in r_B and r_D, so we don't get an easy rule of thumb.

The net selectivity h, however, is asymmetric in positive and negative fitness, and can be used to set bounds on r_D. That being said, doing so may require detailed knowledge of mutation rates. Since an evolutionary system may employ time-variable mutational adaptation (e.g., a cell line mutating rapidly while exposed to stress), rates measured in vitro may differ in vivo. So while we have demonstrated that it is easy to create these inequalities, care must be taken before using them in a real-world system.

It should be noted that many other metrics we could propose here are simply redundant with h. For example, if we were to take the average number of mutations in L cells and divide by the average among H cells, the ratio would converge to $r_B g$. While this is certainly asymmetric, it gives no new information compared to the more accessible metric h.

This is not to say that no signature of mechanism exist – it is just that snapshot-style metrics have difficulty capturing dynamical differences. For example, the growth rate of average mutation count expresses mild asymmetry in figure 7, which only grows more pronounced when looking at the variance in figure 8. However, genotypic surveys, especially for microscopic cells, tend to be highly invasive, so the experimental burden of measuring any of these quantities may be too great to recommend.

However, perfect symmetry is rare in the real world. In an actual examination, one might expect that a quantity such as F_H would have some asymmetry, or that the mutational skew would take on some long-term value, simply by the vagaries of nature. While the germinal center may have many black boxes, there are many in silica and in vitro experiments one can do which have highly controlled evolutionary mechanisms, some of which may contradict such perfect results. However, the model

presented here is highly generic, especially considering the results from appendix A. If such a discrepancy were to be acknowledged, then it would cast doubt on many classes of mathematical evolutionary models.

One final complication specific to GCs may lie in the details of cyclic reentry. The essence of this scheme is that B cells move back and forth between two zones within the GC – a "dark zone" where division takes place, and a "light zone" where B cells are evaluated for affinity [11, 28, 29, 31, 44]. Many evolutionary models treat reproduction and fitness measurement as simultaneous. Here, there is ostensibly a period of "fitness blindness," where cells live and divide unaware of their own affinity/fitness, introducing noise in the fitness landscape [45, 46]. Moreover, depending on how the mutation-proffering AID is used, the actual mutation rates of cell lines may wildly fluctuate, leading to hypothetical scenarios where the number of mutations is unproportional to the number of divisions. If that is the case, we may need to use a multistage, seasonal growth model [47].

Acknowledgments

This research was supported by the Center for Studies in Physics and Biology at Rockefeller University (B.O.-L.). We thank Arup Chakraborty, Kevin O'Keffee, Daniel Abrams, Jörn Dunkel, and Juhee Pae for the helpful comments.

Author Contributions

Bertrand Ottino-Loffler provided the main text, mathematical analysis, and numerics. Gabriel D. Victora provided the biological background and motivation.

Conflict of interest

The authors declare there is no conflict of interest.

Data Availability

There is no data associated with this paper.

References

[1] Darwin, C.: On the origin of species. Routledge London (1859)

[2] Vrugt, J.A., Robinson, B.A.: Improved evolutionary optimization from genetically adaptive multimethod search. Proceedings of the National Academy of Sciences **104**(3), 708–711 (2007)

[3] Bäck, T., Schwefel, H.-P.: An overview of evolutionary algorithms for parameter optimization. Evolutionary computation **1**(1), 1–23 (1993)

[4] Bartz-Beielstein, T., Branke, J., Mehnen, J., Mersmann, O.: Evolutionary algorithms. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **4**(3), 178–195 (2014)

[5] Yu, X., Gen, M.: Introduction to Evolutionary Algorithms. Springer, New York (2010)

[6] Whitley, D., Rana, S., Dzubera, J., Mathias, K.E.: Evaluating evolutionary algorithms. Artificial intelligence **85**(1-2), 245–276 (1996)

[7] Tonegawa, S.: Somatic generation of antibody diversity. Nature **302**(5909), 575–581 (1983)

[8] Eisen, H.N., Siskind, G.W.: Variations in affinities of antibodies during the immune response. Biochemistry **3**(7), 996–1008 (1964)

[9] Ada, G.L., Nossal, S.G.: The clonal-selection theory. Scientific American **257**(2), 62–69 (1987)

[10] McKean, D., Huppi, K., Bell, M., Staudt, L., Gerhard, W., Weigert, M.: Generation of antibody diversity in the immune response of balb/c mice to influenza virus hemagglutinin. Proceedings of the National Academy of Sciences **81**(10), 3180–3184 (1984)

[11] Victora, G.D., Nussenzweig, M.C.: Germinal centers. Annual review of immunology **30**, 429–457 (2012)

[12] Berek, C., Berger, A., Apel, M.: Maturation of the immune response in germinal centers. Cell **67**(6), 1121–1129 (1991)

[13] Sok, D., Laserson, U., Laserson, J., Liu, Y., Vigneault, F., Julien, J.-P., Briney, B., Ramos, A., Saye, K.F., Le, K., *et al.*: The effects of somatic hypermutation on neutralization and binding in the pgt121 family of broadly neutralizing hiv antibodies. PLoS pathogens **9**(11), 1003754 (2013)

[14] Doria-Rose, N.A., Joyce, M.G.: Strategies to guide the antibody affinity maturation process. Current opinion in virology **11**, 137–147 (2015)

[15] Ferretti, F., Kardar, M.: Universal characterization of epitope immunodominance from a multi-scale model of clonal competition in germinal centers. arXiv preprint arXiv:2310.10966 (2023)

[16] Amitai, A., Mesin, L., Victora, G.D., Kardar, M., Chakraborty, A.K.: A population dynamics model for clonal diversity in a germinal center. Frontiers in microbiology **8**, 1693 (2017)

[17] Gitlin, A.D., Shulman, Z., Nussenzweig, M.C.: Clonal selection in the germinal centre by regulated proliferation and hypermutation. Nature **509**(7502), 637–640

(2014)

[18] Liu, Y., Joshua, D., Williams, G., Smith, C., Gordon, J., MacLennan, I.: Mechanism of antigen-driven selection in germinal centres. Nature **342**(6252), 929–931 (1989)

[19] Meyer-Hermann, M.: A molecular theory of germinal center b cell selection and division. Cell Reports **36**(8) (2021)

[20] Mayer, C.T., Gazumyan, A., Kara, E.E., Gitlin, A.D., Golijanin, J., Viant, C., Pai, J., Oliveira, T.Y., Wang, Q., Escolano, A., *et al.*: The microanatomic segregation of selection by apoptosis in the germinal center. Science **358**(6360), 2602 (2017)

[21] Stewart, I., Radtke, D., Phillips, B., McGowan, S.J., Bannard, O.: Germinal center b cells replace their antigen receptors in dark zones and fail light zone entry when immunoglobulin gene mutations are damaging. Immunity **49**(3), 477–489 (2018)

[22] Murray, J.D.: Mathematical Biology. Springer, New York (1993)

[23] Kaveh, K., Komarova, N.L., Kohandel, M.: The duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem. Royal Society open science **2**(4), 140465 (2015)

[24] Yagoobi, S., Sharma, N., Traulsen, A.: Categorizing update mechanisms for graph-structured metapopulations. Journal of the Royal Society Interface **20**(200), 20220769 (2023)

[25] Ottino-Loffler, B., Scott, J.G., Strogatz, S.H.: Evolutionary dynamics of incubation periods. ELife **6**, 30212 (2017)

[26] Young, C., Brink, R.: The unique biology of germinal center b cells. Immunity **54**(8), 1652–1664 (2021)

[27] Tas, J.M., Mesin, L., Pasqual, G., Targ, S., Jacobsen, J.T., Mano, Y.M., Chen, C.S., Weill, J.-C., Reynaud, C.-A., Browne, E.P., *et al.*: Visualizing antibody affinity maturation in germinal centers. Science **351**(6277), 1048–1054 (2016)

[28] Oprea, M., Van Nimwegen, E., Perelson, A.S.: Dynamics of one-pass germinal center models: implications for affinity maturation. Bulletin of mathematical biology **62**, 121–153 (2000)

[29] Yaari, G., Benichou, J.I., Vander Heiden, J.A., Kleinstein, S.H., Louzoun, Y.: The mutation patterns in b-cell immunoglobulin receptors reflect the influence of selection acting at multiple time-scales. Philosophical Transactions of the Royal Society B: Biological Sciences **370**(1676), 20140242 (2015)

[30] Victora, G.D., Dominguez-Sola, D., Holmes, A.B., Deroubaix, S., Dalla-Favera,

R., Nussenzweig, M.C.: Identification of human germinal center light and dark zone cells and their relationship to human b-cell lymphomas. Blood, The Journal of the American Society of Hematology **120**(11), 2240–2248 (2012)

[31] Bannard, O., Horton, R.M., Allen, C.D., An, J., Nagasawa, T., Cyster, J.G.: Germinal center centroblasts transition to a centrocyte phenotype according to a timed program and depend on the dark zone for effective selection. Immunity **39**(5), 912–924 (2013)

[32] Oprea, M., Perelson, A.S.: Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. Journal of immunology (Baltimore, Md.: 1950) **158**(11), 5155–5162 (1997)

[33] Allen, D., Simon, T., Sablitzky, F., Rajewsky, K., Cumano, A.: Antibody engineering for the analysis of affinity maturation of an anti-hapten response. The EMBO journal **7**(7), 1995–2001 (1988)

[34] Verhulst, P.-F.: Notice sur la loi que la population poursuit dans son accroissement. Correspondance Mathématique et Physique **10**, 113–121 (1838)

[35] Cramer, J.S.: The origins of logistic regression. Report, Tinbergen Institute (2002)

[36] Petersen-Mahrt, S.K., Harris, R.S., Neuberger, M.S.: Aid mutates e. coli suggesting a dna deamination mechanism for antibody diversification. Nature **418**(6893), 99–104 (2002)

[37] Muramatsu, M., Sankaranand, V., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., Honjo, T.: Specific expression of activation-induced cytidine deaminase (aid), a novel member of the rna-editing deaminase family in germinal center b cells. Journal of Biological Chemistry **274**(26), 18470–18476 (1999)

[38] Araki, T.: Replaying life's tape with intraclonal germinal center evolution. Phd thesis, Rockefeller University (2023). Available at https://digitalcommons.rockefeller.edu/student_theses_and_dissertations/720/

[39] Sepkoski, D.: "replaying life's tape": Simulations, metaphors, and historicity in stephen jay gould's view of life. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences **58**, 73–81 (2016)

[40] Gould, S.J.: Wonderful Life: the Burgess Shale and the Nature of History. WW Norton & Company, New York (1989)

[41] Hilfinger, A., Paulsson, J.: Separating intrinsic from extrinsic fluctuations in dynamic biological systems. Proceedings of the National Academy of Sciences **108**(29), 12167–12172 (2011)

[42] Sharma, N., Traulsen, A.: Suppressors of fixation can increase average fitness beyond amplifiers of selection. Proceedings of the National Academy of Sciences **119**(37), 2205424119 (2022)

[43] Tkadlec, J., Pavlogiannis, A., Chatterjee, K., Nowak, M.A.: Limits on amplifiers of natural selection under death-birth updating. PLoS computational biology **16**(1), 1007494 (2020)

[44] Victora, G.D., Schwickert, T.A., Fooksman, D.R., Kamphorst, A.O., Meyer-Hermann, M., Dustin, M.L., Nussenzweig, M.C.: Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. Cell **143**(4), 592–605 (2010)

[45] Trubenová, B., Krejca, M.S., Lehre, P.K., Kötzing, T.: Surfing on the seascape: Adaptation in a changing environment. Evolution **73**(7), 1356–1374 (2019)

[46] Merrell, D.J.: The Adaptive Seascape: the Mechanism of Evolution. U of Minnesota Press, Minneapolis (1994)

[47] Swartz, D.W., Ottino-Löffler, B., Kardar, M.: Seascape origin of richards growth. Physical Review E **105**(1), 014417 (2022)

Appendix A Asymmetric Mutation Rates

In this section, we will show that having non-identical mutation rates between the daughter cells produces the same results.

In particular, if an H cell divides, its first daughter has a α_{H1} chance of mutating into another H cell and a β_{H1} chance of becoming a L cell, and the second daughter has α_{H2} and β_{H2}. Similarly, the individual daughters of an L cell have an L→L transition rate of α_{L1} and α_{L2}, and an L→H transition rate of β_{L1} and β_{L2} respectively. We define the total transition rates as $\bar{\beta}_X = \beta_{X1} + \beta_{X2}$, $\bar{\alpha}_X = \alpha_{X1} + \alpha_{X2}$, $\rho_{Xn} = \beta_{Xn} + \alpha_{Xn}$, and $\bar{\rho}_X = \bar{\beta}_X + \bar{\alpha}_X$. To avoid trivial fixed points, we require $\bar{\beta}_H$ and $\bar{\beta}_L$ to be nonzero.

Note that while we require $\bar{\beta}_H$ and $\bar{\beta}_L$ to be nonzero, nothing prevents, say, $\beta_{H1} = \alpha_{H1} = 0$. Here, one daughter would always be indistinguishable from the parent, which recreates the case from the main paper and some models of sexual reproduction. Similarly, if we want a more typical asexual reproduction model, setting the two daughters to have identical rates will do that for you.

Also, while instances of extreme mutation can be interesting, we shall also require $\bar{\beta}_H$ and $\bar{\beta}_L$ to be less than 1 to avoid some pathological regimes. For example, if $\bar{\beta}_H > 1$, then $r_B \to \infty$ would unintuitively cause the high fitness population to go extinct, since h decreases per division.

On a division event, the total population of high fitness (H) cells can change by +2, +1, 0, -1, depending on the identity of the parent and the offspring. For example, if a low fitness (L) cell divides, and the first daughter mutates into a H cell and the second daughter is a nonmutated L cell, this would cause +1 new H cells. We notate this event as $(L \to H^* L)$, where the $*$'s identify the daughters as mutants, and the

order tracks the identity of the daughters. Using this notation, we can write down all the H-changing events as

$$\begin{aligned}
P_{+2} =& \ (L \to H^*H^*)\,, \\
P_{+1} =& \ (H \to HH) + (H \to H^*H) + (H \to HH^*) + (H \to H^*H^*) \\
&+ (L \to H^*L) + (H \to LH^*) + (L \to H^*L^*) + (H \to L^*H^*)\,, \\
P_{-1} =& \ (H \to L^*L^*) + (H \to \text{Death})\,.
\end{aligned}$$

Given that h is the fraction of the population which is H cells and ℓ is the fraction which is L cells, we can write the above events as probabilities to get

$$\begin{aligned}
\partial_t h &= 2P_{+2} + P_{+1} - P_{-1} \\
&= r_B h Z_B \left[(1-\rho_{H1})(1-\rho_{H2}) + \alpha_{H1}(1-\rho_{H2}) + (1-\rho_{H1})\alpha_{H2} + \alpha_{H1}\alpha_{H2} - \beta_{H1}\beta_{H2}\right] \\
&\quad + \ell Z_B \left[\beta_{L1}\beta_{L2} + \beta_{L1}(1-\rho_{L2}) + (1-\rho_{L1})\beta_{L2} + \alpha_{L1}\beta_{L2} + \beta_{L1}\alpha_{L2}\right] - (h/r_D)Z_D,
\end{aligned}$$

where Z_B and Z_D are defined as before. This simplifies into

$$\partial_t h = r_B h Z_B(1-\bar{\beta}_H) + \ell Z_B \bar{\beta}_L - (h/r_D)Z_D.$$

This is identical to the $\partial_t h$ equation in the single mutation case (section III in the main body), so all the results from that section would apply with the substitution $\beta_H \to \bar{\beta}_H$ and and the like. In particular,

$$h = \frac{1}{1 + r_B g(r_B r_D)}.$$

If we are careful about accounting cells in the same manner as about, we retrieve familiar ancestry ratios f_H and f_L

$$\begin{aligned}
f_H &= \frac{\bar{\alpha}_H}{\bar{\alpha}_H + \bar{\beta}_L g(r_B r_D)}, \\
f_L &= \frac{\bar{\beta}_H}{\bar{\beta}_H + \bar{\alpha}_L g(r_B r_D)}.
\end{aligned}$$

We also get familiar equations for the number of mutations within strains, with

$$\begin{cases}
\partial_t h_m = (1-\bar{\rho}_H)\left(h_m r_B Z_B\right) + \bar{\alpha}_H\left(h_{m-1} r_B Z_B\right) + \bar{\beta}_L\left(\ell_{m-1} Z_B\right) - \frac{h_m}{r_D} Z_D, \\
\partial_t \ell_m = (1-\bar{\rho}_L)\left(\ell_m Z_B\right) + \bar{\alpha}_L\left(\ell_{m-1} Z_B\right) + \bar{\beta}_H\left(h_{m-1} r_B Z_B\right) - \ell_m Z_D.
\end{cases}$$

Hence, all results from the main body also apply to the general case of asymmetrically mutating daughters.

Appendix B Infinite fitness edge cases

Throughout this paper we considered only finite fitnesses, but on various heat plots, we include the case of $r_B = \infty$ and $r_D = \infty$. Here we will discuss the approach for each. Moreover, for the sake of completeness, we will allow both daughters to be mutants, as in appendix A.

B.1 Limit of large positive selection.

In the case of infinite birth fitness, only high fitness cells divide. So $P_{BH} = P_B$ and $P_{BL} = 0$ so long as $h > 0$. In such as case,

$$\begin{aligned}
P_{+2} &= 0, \\
P_{+1} &= (H \to HH) + (H \to H^*H) + (H \to HH^*) + (H \to H^*H^*) \\
&= P_{BH}\left(1 - \bar{\beta}_H + \beta_{H1}\beta_{H2}\right), \\
P_{-1} &= (H \to L^*L^*) + (H \to \text{Death}) \\
&= P_{BH}\beta_{H1}\beta_{H2} + P_{DH}.
\end{aligned}$$

This means the average growth rate of high-fitness cells becomes

$$\partial_t h = \frac{1}{2}\left(1 - \bar{\beta}_H\right) - \frac{h}{r_D} Z_D.$$

The steady state can be solved from here to get

$$h = \frac{r_D(1 - \bar{\beta}_H)}{r_D(1 - \bar{\beta}_H) + \bar{\beta}_H}.$$

Similarly, the equations for the mutant populations also get simplified, with

$$\begin{aligned}
\partial_t h_m &= \frac{h_m}{2h}(1 - \bar{\rho}_H), \\
\partial_t \ell_m &= \frac{h_{m-1}}{2h}\bar{\beta}_H - Z_D \ell_m.
\end{aligned}$$

While simpler, the structure is similar to that in section V in the main body, so we can just take the results from there and take the limit of $r_B \to \infty$ to quickly get the results. Notably, we have

$$\begin{aligned}
\lim_{r_B \to \infty} Z_B &= 0 \\
\lim_{r_B \to \infty} r_B g &= \frac{\bar{\beta}_H}{r_D(1 - \bar{\beta}_H)}, \\
\lim_{r_B \to \infty} Z_B/g &= \frac{r_B}{2h}\frac{1 - \bar{\beta}_H}{\bar{\beta}_H},
\end{aligned}$$

$$\lim_{r_B \to \infty} r_B Z_B = \frac{1}{2h}.$$

As applied to the building blocks of equations for $\partial_t \mathcal{M}_k$ and $\partial_t \mathcal{S}_k$ (from section V in the main body), we have

$$\begin{aligned}
\tau &= r_D \frac{1-\beta_H}{2h}, \\
M_M &= 0, \\
M_S &= \ell \frac{\bar{\rho}_H}{\bar{\beta}_H}, \\
S_M &= 0, \\
S_S &= \frac{\bar{\alpha}_H}{2h}.
\end{aligned}$$

Although some terms go to zero, this does not disrupt the rest of the analysis, leading to the same conclusion about the skew approaching zero.

B.2 Limit of large negative selection.

Although the limit of large positive selection was routine, a little more care must be taken with large negative selection. For example, trying to naively take the $r_D \to \infty$ limit to the results of the main body's section V leads to the timescale τ diverging.

Because of the high level of negative selection, so long as any L cells exist, no H cells will die. In fact, an attempt to write an equation for h yields

$$\partial_t h = \ell Z_B \bar{\beta}_L + h(1-\bar{\beta}_H) Z_B.$$

Notably, this is strictly positive so long as $\bar{\beta}_H < 1$. So, absent finite-size errors, $h \to 1$ is assured.

Because there is no low-fitness population worth talking about, we have that the H population is the full population, so $\mathcal{H}_k = \mathcal{M}_k$. In particular, we have

$$\partial_t h_M = \frac{\bar{\alpha}_H}{2}(h_{m-1} - h_m),$$

where the $1/2$ term comes from the fact that at steady state, $P_B = P_D = 1/2$. We therefore have

$$\partial_t \mathcal{M}_k = \frac{\alpha_H}{2} \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{M}_w.$$

This is recursively solvable, recalling that $\mathcal{M}_0 = 1$. Letting $\gamma := \bar{\alpha}_H/2$, the first three orders are given by

$$\begin{aligned}
\mathcal{M}_1 &= \gamma t + \theta_1^M, \\
\mathcal{M}_2 &= \gamma^2 t^2 + (2\gamma\theta_1^M + \gamma)t + \theta_2^M,
\end{aligned}$$

$$\mathcal{M}_3 = \gamma^3 t^3 + 3\gamma^2(\theta_1^M + 1)t^2 + \gamma(3\theta_2^M + 3\theta_1^M + 1)t + \theta_3^M.$$

Therefore, the variance is given by

$$\begin{aligned} C_2 &= \mathcal{M}_2 - (\mathcal{M}_1)^2 \\ &= \gamma t + \theta_2^M - \left(\theta_1^M\right)^2, \end{aligned}$$

and the third central moment is given by

$$\begin{aligned} C_3 &= \mathcal{M}_3 - 3\mathcal{M}_2\mathcal{M}_1 + 2(\mathcal{M}_1)^3 \\ &= \gamma t + \theta_3^M - 3\theta_2^M\theta_1^M + 2(\theta_1^M)^3. \end{aligned}$$

That is to say, the mean, variance, and third central moment all grow with at the exact same rate of $\gamma = \bar{\alpha}_H/2$. Therefore, the skew goes as

$$\text{Skew} = \frac{C_3}{(C_2)^{3/2}} = \frac{\gamma t + \text{const.}}{(\gamma t + \text{const.})^{3/2}} \to 0.$$

So, just as in the main case, we have no long-term skew appearing.

Appendix C The Matrix Exponential Solution

In the main body, we got equations for the moments of L and H populations

$$\begin{aligned} \partial_t \mathcal{H}_k = &\left[(1-\beta_H)\, r_B Z_B - \frac{Z_D}{r_D}\right]\mathcal{H}_k + \beta_L Z_B \mathcal{L}_k \\ &+ r_B \alpha_H Z_B \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{H}_w \quad + \beta_L Z_B \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{L}_w \qquad \text{(C1)} \\ \partial_t \mathcal{L}_k = &\left[(1-\beta_L)\, Z_B - Z_D\right]\mathcal{L}_k \quad + r_B \beta_H Z_B \mathcal{H}_k \\ &+ r_B \beta_H Z_B \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{H}_w \quad + \alpha_L Z_B \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{L}_w. \qquad \text{(C2)} \end{aligned}$$

Let us vectorize the system in equations (C1) and (C2) by letting $\overrightarrow{m} = (\mathcal{H}_1, \mathcal{L}_1, \mathcal{H}_2, \mathcal{L}_2, \mathcal{H}_3, \mathcal{L}_3)^T$, where we truncate to $k = 3$, since our goal is to find the skew. Therefore, the dynamics are given by

$$\partial_t \overrightarrow{m} = J\overrightarrow{m} + \overrightarrow{v},$$

where the odd entries of $\overrightarrow{v}$ are $hr_B\alpha_H Z_B + \beta_L Z_B \ell$, and the even entries are $hr_B\beta_H Z_B + \alpha_L Z_B \ell$. Meanwhile, J is a block matrix of the form:

$$J = \begin{pmatrix} J_D & 0 & 0 \\ 2J_S & J_D & 0 \\ 3J_S & 3J_S & J_D \end{pmatrix},$$

with

$$J_D = \begin{pmatrix} r_B Z_B(1-\beta_H) - Z_D/r_D & \beta_L Z_B \\ r_B Z_B \beta_H & Z_B(1-\beta_L) - Z_D \end{pmatrix},$$

and

$$J_S = \begin{pmatrix} r_B \alpha_H Z_B & \beta_L Z_B \\ r_B Z_B \beta_H & \alpha_L Z_B \end{pmatrix}.$$

So long as the determinant of J is nonzero, this means that the moments are, in principle, exactly solvable, via

$$\overrightarrow{m}(t) = e^{Jt}(\overrightarrow{m}(0) + J^{-1}\overrightarrow{v}) - J^{-1}\overrightarrow{v}.$$

The determinant of J is decided by J_D. So calculating $|J_D|$ gives

$$|J_D| = r_B Z_B^2(1-\beta_H)(1-\beta_L) - r_B Z_B Z_D(1-\beta_H) - Z_B Z_D(1-\beta_L)/r_D + Z_D^2/r_D - r_B\beta_H\beta_L Z_B^2.$$

If we calculate at population steady-state ($n = N$), then we have

$$\frac{Z_B}{Z_D} = 1 + 2\left(r_D^{-1} - r_B\right) h Z_B,$$
$$\frac{Z_D}{Z_B} = 1 + 2\left(r_D^{-1} - r_B\right) h Z_D.$$

Therefore,

$$\begin{aligned}
|J_D| Z_B^{-1} Z_D^{-1} &= r_B \frac{Z_B}{Z_D}\left[(1-\beta_H)(1-\beta_L) - \beta_H\beta_L\right] - r_B(1-\beta_H) - (1-\beta_L)\frac{1}{r_D} + \frac{1}{r_D}\frac{Z_D}{Z_B} \\
&= r_B(1-\beta_H-\beta_L)\left(1 + 2(r_D^{-1} - r_B)hZ_B\right) + \frac{1}{r_D}\left(1 + 2(r_D^{-1} - r_B)hZ_D\right) - r_B(1-\beta_H) - (1-\beta_L)\frac{1}{r_D} \\
&= r_B(1-\beta_H) - r_B\beta_L + 2r_B(1-\beta_L-\beta_H)\left(r_D^{-1} - r_B\right) h Z_B + \frac{1}{r_D} \\
&\quad + 2\left(r_B - r_D^{-1}\right)\frac{h}{r_D} Z_D - r_B(1-\beta_H) - (1-\beta_L)\frac{1}{r_D} \\
&= \left(r_D^{-1} - r_B\right)\left(\beta_L + 2r_B(1-\beta_L-\beta_H)hZ_B - 2hZ_D/r_D\right) \\
&= \left(r_D^{-1} - r_B\right)\left(2r_B hZ_B(1-\beta_H) - 2hZ_D/r_D + 2\ell Z_B\beta_L - 2\ell Z_B\beta_L + \beta_L - 2r_B\beta_L hZ_B\right) \\
&= \left(r_D^{-1} - r_B\right)\left(2\partial_t h + \beta_L\left[1 - 2Z_B(r_B h + \ell)\right]\right) \\
&= \left(r_D^{-1} - r_B\right)\left(2\partial_t h + \beta_L\left[1 - 1\right]\right),
\end{aligned}$$

and so we get

$$|J_D| = 2\left(r_D^{-1} - r_B\right) Z_B Z_D \partial_t h.$$

As the system approaches a phenotypic equilibrium, we have $\partial_t h \to 0$. So if we want to find long-time dynamics (or worse, stationary state dynamics), then this vectorized formulation is ill suited for numerical and analytical examination.

Appendix D Recusion for Mutant Moments

In this section, we will produce the recursion equations that define the growth of the moments of the mutant distribution $\mathcal{M}_k$, and give the leading-order results for the central moments C_a.

Specifically, we have the following system of equations

$$\partial_t \mathcal{M}_k = -\tau \mathcal{M}_k + \mathcal{S}_k + M_M \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{M}_w + M_S \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{S}_w, \tag{D3}$$

$$\partial_t \mathcal{S}_k = S_M \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{M}_w + S_S \sum_{w=0}^{k-1} \binom{k}{w} \mathcal{S}_w, \tag{D4}$$

where

$$\mathcal{M}_k := \mathcal{H}_k + \mathcal{L}_k, \tag{D5}$$

$$\mathcal{S}_k := (r_B Z_B \beta_H/\ell)\mathcal{H}_k + (Z_B \beta_L/h)\mathcal{L}_k, \tag{D6}$$

for all $k \geq 1$.

By hypothesis, we assume the solutions take the form of polynomials, with

$$\mathcal{S}_k = \sum_{w=0}^{k} q_w^{Sk} t^w + \sum_{w=0}^{k-2} p_w^{Sk} t^w e^{-\tau t}, \tag{D7}$$

and

$$\mathcal{M}_k = \sum_{w=0}^{k} q_w^{Mk} t^w + \sum_{w=0}^{k-1} p_w^{Mk} t^w e^{-\tau t}, \tag{D8}$$

where q_w^{Mk}, p_w^{Mk}, q_w^{Sk}, and p_w^{Sk} are coefficients to be determined, and we take coefficients with index $w < 0$ to be zero.

Just as in the main text, the $k = 1$ step is easy to verify. We start with the $\mathcal{S}_1$ equation, and substitute appropriate values for $\mathcal{M}_0$ and $\mathcal{S}_0$ based on their definitions in (D5) and (D6), and get

$$\begin{aligned}
\partial_t \mathcal{S}_1 &= S_M \mathcal{M}_0 + S_S \mathcal{S}_0 \\
&= S_M(h + \ell) + S_S(r_B Z_B \beta_H h/\ell + Z_B \beta_L \ell/h) \\
&= S_M + S_S \tau.
\end{aligned}$$

For future notational convenience, we define

$$\mu := S_M/\tau + S_S. \tag{D9}$$

Therefore, we have $\mathcal{S}_1(t) = \tau\mu t + \theta_1^S$, where we use $\theta_k^S := \mathcal{S}_k(t=0)$ and $\theta_k^M := \mathcal{M}_k(t=0)$.

Similarly, we have

$$\begin{aligned}\partial_t \mathcal{M}_1 = & -\tau\mathcal{M}_1 + \mathcal{S}_1 + M_M\mathcal{M}_0 + M_S\mathcal{S}_0 \\ = & -\tau\mathcal{M}_1 + q_1^{S1}t + (M_M + \tau M_S) + q_0^{S1}.\end{aligned}$$

Integrating gives us

$$\mathcal{M}_1 = \mu t + \left(M_M + \tau M_S + \theta_1^S\right)/\tau + \left(\theta_1^M - (M_M + \tau M_S + \theta_1^S)/\tau\right) e^{-\tau t},$$

so to leading order, the mean mutation count grows at a rate $\mu := S_M/\tau + S_S$.

With a base established, we can create the inductive step to establish a recursion relation for the various coefficients. By subbing equations (D8) and (D7) into (D4), we get

$$\partial_t \mathcal{S}_{a+1} = \sum_{k=0}^{a} d_k^{S(a+1)} t^k + \sum_{k=0}^{a-1} c_k^{S(a+1)} t^k e^{-\tau t}$$

where

$$d_k^{S(a+1)} = \sum_{w=k}^{a} \binom{a+1}{w} \left(S_M q_k^{Mw} + S_S q_k^{Sw}\right) \tag{D10}$$

$$c_k^{S(a+1)} = (a+1)S_M p^{Ma} + S_M \sum_{w=k}^{a-2} \binom{a+1}{w+1} p_k^{M(w+1)} + S_S \sum_{w=k}^{a-2} \binom{a+1}{w+2} p_k^{S(w+2)} \tag{D11}$$

where we used lemma 3 in appendix E to collect terms appropriately, and we take the convention that $\sum_{w=x}^{y} z_w = 0$ when $x > y$.

If we plug these into the expressions in lemma 1 in appendix E, we get

$$\mathcal{S}_{a+1} = \sum_{k=0}^{a+1} q_k^{S(a+1)} t^k + \sum_{k=0}^{a-1} p_k^{S(a+1)} t^k e^{-\tau t}$$

with

$$q_k^{S(a+1)} = \frac{1}{k} d_{k-1}^{S(a+1)}, k \geq 1 \tag{D12}$$

$$q_0^{S(a+1)} = \theta_{a+1}^S + \sum_{w=0}^{a-1} \frac{w!}{\tau^{w+1}} c_k^{S(a+1)} \tag{D13}$$

$$p_k^{S(a+1)} = -\sum_{w=k}^{a-1} \frac{w!}{k!\tau^{w+1}} c_k^{S(a-1)}. \tag{D14}$$

By a similar procedure, we have that

$$\partial_t \mathcal{M}_{a+1} + \tau \mathcal{M}_a = \sum_{k=0}^{a+1} d_k^{M(a+1)} t^k + \sum_{k=0}^{a-1} c_k^{M(a+1)} t^k e^{-\tau t}$$

with

$$d_k^{M(a+1)} = q_k^{S(a+1)} + \sum_{w=k}^{a} \binom{a+1}{w} \left(M_M q_k^{Mw} + M_S q_k^{Sw} \right) \tag{D15}$$

$$c_k^{M(a+1)} = p_k^{S(a+1)} + M_M(a+1) p_k^{Ma} + \sum_{w=k}^{a-2} \left[\binom{a+1}{w+1} p_k^{M(w+1)} M_M + \binom{a+1}{w+2} p_k^{S(w+2)} M_S \right]. \tag{D16}$$

If we plug these into the expressions in lemma 2 in appendix E, we then get

$$\mathcal{M}_{a+1} = \sum_{k=0}^{a+1} q_k^{M(a+1)} t^k + \sum_{k=0}^{a} p_k^{M(a+1)} t^k e^{-\tau t}$$

with

$$q_k^{M(a+1)} = \sum_{w=k}^{a+1} \frac{w!}{k!} \frac{(-1)^{w-k}}{\tau^{w-k+1}} d_w^{M(a+1)} \tag{D17}$$

$$p_k^{M(a+1)} = \frac{1}{k} c_{k-1}^{M(a+1)}, k \geq 1 \tag{D18}$$

$$p_0^{M(a+1)} = \theta_{a+1}^M + \sum_{w=0}^{a+1} w! \frac{(-1)^{w+1}}{\tau^{w+1}} d_w^{M(a+1)}. \tag{D19}$$

If we insert (D16) and (D15) into equation (D17) and evaluate at $k = a+1$, we get

$$q_{a+1}^{M(a+1)} = \frac{1}{\tau} q_{a+1}^{S(a+1)}. \tag{D20}$$

Similarly, if we insert (D11) and (D10) into equation (D12) and evaluate at $k = a+1$, we find

$$q_{a+1}^{S(a+1)} = S_M q_a^{Ma} + S_S q_a^{Sa}. \tag{D21}$$

Putting equations (D20) and (D21) together gives the following closed forms via induction, using the fact that $\mu = S_S + S_M/\tau$:

$$q_a^{Ma} = \mu^a, \tag{D22}$$
$$q_a^{Sa} = \tau\mu^a. \tag{D23}$$

If we instead evaluate (D12) and (D17) at $k = a$, we get

$$q_a^{S(a+1)} = \frac{a+1}{2}\tau\mu^a + \frac{a+1}{a}\left(S_M q_{a-1}^{Ma} + S_S q_{a-1}^{Sa}\right),$$
$$q_a^{M(a+1)} = \frac{1}{\tau}q_a^{S(a+1)} + \frac{\mu^a}{\tau}(a+1)\left(M_M + M_S\tau - \mu\right).$$

Naturally, we can insert one into the other to find

$$q_a^{S(a+1)} = \frac{a+1}{a}\mu q_{a-1}^{Sa} + \mu^{a-1}(a+1)\left(\frac{\mu\tau}{2} + S_M\left[M_M/\tau + M_S - \mu/\tau\right]\right).$$

If we then solve this via the lemma 4 in appendix E, we get for $a \geq 1$

$$q_{a-1}^{Sa} = a\mu^{a-1}q_0^{S1} + a(a-1)\mu^{a-2}\left(\mu\tau/2 + S_M\left[M_M/\tau + M_S - \mu/\tau\right]\right).$$

Therefore, the subleading coefficient for the ath moment is

$$q_{a-1}^{Ma} = \frac{a}{\tau}\mu^{a-1}q_0^{S1} + a\mu^{a-2}\left[\frac{(a-1)\mu}{2} + \left(\frac{S_M}{\tau}(a-1) + \mu\right)\left(\frac{M_M}{\tau} + M_S - \frac{\mu}{\tau}\right)\right]. \tag{D24}$$

To recap, equation (D22) is the leading order coefficient of the growth of the ath moment of mutation counts, and (D24) is the subleading coefficient. This gives us enough information to calculate the behavior of the central moments, which characterize the shape of the mutation count histogram.

The ath central moment is given by

$$\begin{aligned} C_a &= \langle(m - \mathcal{M}_1)^a\rangle \\ &= \langle\sum_{k=0}^{a}\binom{a}{k}m^k{\mathcal{M}_1}^{a-k}(-1)^{a-k}\rangle \\ &= \sum_{k=0}^{a}\binom{a}{k}\mathcal{M}_k{\mathcal{M}_1}^{a-k}(-1)^{a-k}. \end{aligned}$$

We take $a \geq 2$, since the first central moment is definitionally zero.

If we substitute in the polynomial forms of the $\mathcal{M}_k$ terms and collect all the exponential terms under $\mathcal{E}$, we get

$$C_a = \sum_{k=0}^{a}\binom{a}{k}\left(\sum_{w=0}^{k}q_w^{Mk}t^w\right)\left(q_0^{M1} + \mu t\right)^{a-k}(-1)^{a-k} + \mathcal{E}$$

If we collect terms, this takes on the following form:

$$C_a = \sum_{b=0}^{a} t^b \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} \sum_{w=\max(0,b-k)}^{\min(b,a-k)} f_{a,k}(w, b-w) + \mathcal{E}, \tag{D25}$$
$$f_{a,k}(w,r) = \binom{a-k}{w} \left(q_0^{M1}\right)^{a-k-w} q_r^{Mk} \mu^w,$$

where the bounds on the sum are from the first argument of f needing to be less than $a-k$, the second has to be less than k, and their sum must be b.

Via the general form of the moments $\mathcal{M}_a$ (Equation (D8)), we know the leading term of the central moments grows at most as t^a. So to get the proper coefficient, we evaluate the $b = a$ term to get

$$\begin{aligned}
C_a[t^a] &= \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} \sum_{w=\max(0,a-k)}^{\min(a,a-k)} f_{a,k}(w, a-w) \\
&= \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} f_{a,k}(a-k, a-(a-k)) \\
&= \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} \binom{a-k}{a-k} \left(q_0^{M1}\right)^{a-k-(a-k)} q_k^{Mk} \mu^{(a-k)} \\
&= \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} q_k^{Mk} \mu^{(a-k)} \\
&= \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} \mu^k \mu^{(a-k)} \\
&= \mu^a \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} 1^k \\
&= \mu^a (1-1)^a.
\end{aligned}$$

Note the use of equation (D22). Therefore, we have that for all $a \geq 1$, then

$$C_a[t^a] = 0. \tag{D26}$$

In other words, every central moment grows at most as t^{a-1}.

To find the $a-1$ order coefficient, we pull the $b = a-1$ term from Equation (D25). Here,

$$C_a[t^{a-1}] = \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} \sum_{w=\max(0,(a-1)-k)}^{\min(a-1,a-k)} f_{a,k}(w, (a-1)-w)$$

$$= \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} \left(f_{a,k}(a-k-1,k) + f_{a,k}(a-k,k-1) \right)$$
$$= \sum_{k=0}^{a} (a-k) \binom{a}{k} (-1)^{a-k} q_0^{N1} \mu^{a-1} + \sum_{k=1}^{a} \binom{a}{k} (-1)^{a-k} q_{k-1}^{Nk} \mu^{a-k},$$

where we already inserted (D22). If we also insert (D24), we find

$$\begin{aligned} C_a[t^{a-1}] = q_0^{N1} \mu^{a-1} \sum_{k=0}^{a} \binom{a}{k} (-1)^{a-k} (a-k) + q_0^{S1} \frac{\mu^{a-1}}{\tau} \sum_{k=1}^{a} \binom{a}{k} (-1)^{a-k} k \\ + \mu^{a-2} \left[\frac{\mu}{2} + \frac{S_M}{\tau} \left(\frac{M_M}{\tau} + M_S - \frac{\mu}{\tau} \right) \right] \sum_{k=1}^{a} \binom{a}{k} (-1)^{a-k} k(k-1) \\ + \mu^{a-1} \left[\frac{M_M}{\tau} + M_S - \frac{\mu}{\tau} \right] \sum_{k=1}^{a} (-1)^{a-k} \binom{a}{k} k. \end{aligned}$$

While this is unpleasant to look at, all the sums simplify to various Kronecker deltas $\delta_{j,k}$. Keeping in mind $a \geq 2$ and eliminating the appropriate deltas, this simplifies to

$$C_a[t^{a-1}] = \delta_{a,2} \left[\mu + \frac{2S_M}{\tau} \left(\frac{M_M}{\tau} + M_S - \frac{\mu}{\tau} \right) \right]. \tag{D27}$$

This is nonzero when $a = 2$, and zero otherwise. This returns us to the results in the main text.

Appendix E Useful Lemmas

A number of elementary lemmas appear in our induction proofs, so for ease of reference they have been collected here.

Lemma 1. *Given the following equation for z,*

$$\dot{z} = -\tau z + \sum_{k=0}^{m} d_k t^k + \sum_{k=0}^{n} c_k t^k e^{-\tau t},$$

we have the solution

$$z(t) = \sum_{k=0}^{m} q_k t^k + \sum_{k=0}^{n+1} p_k t^k e^{-\tau t}$$

with

$$q_k = \sum_{w=k}^{m} d_w \frac{w!}{k!} \frac{(-1)^{w-k}}{\tau^{w-k+1}}$$
$$p_k = c_{k-1}/k, k \geq 1$$

$$p_0 = z(0) + \sum_{k=0}^{m} q_k (-1)^{k+1} \frac{k!}{\tau^{k+1}}.$$

Lemma 2. *Given the following equation for z,*

$$\dot{z} = \sum_{k=0}^{m} d_k t^k + \sum_{k=0}^{n} c_k t^k e^{-\tau t},$$

we have the solution

$$z(t) = \sum_{k=0}^{m+1} q_k t^k + \sum_{k=0}^{n} p_k t^k e^{-\tau t}$$

with

$$\begin{aligned} q_k &= d_{k-1}/k, k \geq 1 \\ q_0 &= z(0) + \sum_{k=0}^{n} d_k \frac{k!}{\tau^{k+1}} \\ p_k &= -\sum_{w=k}^{n} c_w \frac{w!}{k!} \frac{1}{\tau^{w-k+1}}. \end{aligned}$$

Lemma 3.

$$\begin{aligned} \sum_{k=0}^{m} \sum_{w=0}^{k} f(k,w) &= \sum_{k=0}^{m} \sum_{w=k}^{m} f(w,k) \\ \sum_{k=0}^{m} \sum_{w=0}^{k-1} f(k,w) &= \sum_{k=0}^{m-1} \sum_{w=k}^{m-1} f(w+1,k) \\ \sum_{k=0}^{m} \sum_{w=0}^{k-2} f(k,w) &= \sum_{k=0}^{m-2} \sum_{w=k}^{m-2} f(w,k) \end{aligned}$$

Lemma 4. *If $X_{a+1} = A_a X_a + B_a$, then*

$$X_a = X_1 \prod_{k=1}^{a-1} A_k + \sum_{k=1}^{a-1} B_k \prod_{w=k+1}^{a-1} A_w.$$