Sexual complexity as a screen for replication fidelity: the Everest effect

Patrick D. Shaw Stewart, Hungerford, Berkshire, UK

Abstract

Sexual reproduction is often accompanied by costly and very complex traits, from courtship displays and long-distance breeding migrations to gamete-recognition systems. Existing theories explain many such traits through Fisherian, handicap, good-genes, or condition-dependent processes, but they do not, however, fully address a dynamic problem in mutation-rate evolution: if environmental change favors mutator backgrounds during adaptation, what later restores high replication fidelity?

Here, I propose the Everest hypothesis, which treats short-term adaptation and replication fidelity as separate but interacting dimensions. Fidelity contributes to long-term fitness, but during environmental change, it can temporarily conflict with immediate adaptation because mutator alleles may generate beneficial variants and hitchhike with them. Roberts and Petrie's 2022 model showed that mate choice can favor elevated mutation rates and amplify the reproductive contribution of high-performing individuals, including males carrying beneficial mutations that arise on mutator backgrounds. Everest builds on this accelerator logic but adds the corresponding brake: as adaptation proceeds and the advantage of elevated mutation declines, complex reproductive systems can amplify selection for restored replication fidelity.

The proposed key mechanism is multigenic stacking. Many reproductive systems integrate numerous gene products, regulatory functions, and behavioral or molecular steps into a single reproductive test—such as a courtship display, migration-linked breeding sequence, or biochemical compatibility system—whose outcome is mating and fertilization or, alternatively, reproductive failure. Such serial biological tests can convert many small mutator-linked defects into large differences in mating or fertilization success, thereby strengthening selection against low-fidelity backgrounds and favoring antimutator or high-fidelity genotypes. Recombination can then allow offspring to retain adaptive alleles while recovering accurate replication machinery.

The hypothesis predicts a characteristic rise-and-fall cycle in mutation rate: mutators may increase during adaptation, but complex reproductive screening can later reduce mutation rates or mutational load without erasing adaptive gains. Everest is testable through experimental evolution, manipulation of reproductive screens, parent–offspring sequencing, and comparative analyses of taxa differing in reproductive complexity. It therefore complements existing theories of sexual selection by proposing that traits elaborated by those processes can be co-opted to speed the restoration of replication fidelity after the rise of adaptive mutators.

Keywords: sexual selection; mutation-rate evolution; good genes; drift-barrier hypothesis; mutation load; recombination; experimental evolution

---

Introduction: mutation rates, drift-barrier theory, and unstable environments

DNA replication is remarkably accurate but not perfect. Each generation introduces new mutations, including mutations in the replication machinery itself. Alleles that increase the error rate—mutator alleles—form a spectrum from severe to mild. Strong mutators are relatively easy to eliminate because they remain associated, at least initially, with the deleterious mutations they generate. Mild mutators pose a harder problem: each may impose selective effects too small for ordinary viability-fecundity selection to purge reliably, although their accumulation threatens the long-term

maintenance of complex genomes. (Viability-fecundity selection refers here to selection based on survival or reproductive output, apart from mate choice.)

Crucially, this spectrum is asymmetric. Mutator alleles can range from severe to mild, whereas antimutator alleles—those that improve replication fidelity—can, in practice, only be mild, because large single-step improvements in an already functional replication system are vanishingly rare except by reversion or compensation of a previous mutator mutation. Mechanisms that respond only to large fidelity differences will therefore eliminate strong mutators but fail to purge the mild mutators that erode fidelity or favor the mild antimutators that could restore it—yet long-term genome maintenance requires selection to act on both.

Replication fidelity is distinct from fitness. Fitness is a context-dependent measure of expected reproductive success; replication fidelity is a quantitative property of the replication machinery, measurable by sequencing as a per-base or per-genome mutation rate. For example, if a group of animals colonizes an island, one individual may instantly become fitter than its fellow colonists simply because it already carries alleles suited to the new environment, although its genetic composition and mutation rate are unchanged. Viability-fecundity selection acts directly on realized fitness; it detects mutation rate only indirectly, through harmful mutations that remain linked to mutator alleles. When populations are well adapted, most new mutations are deleterious and lower mutation rates tend to be favored, but genetic drift may allow suboptimal fidelity to persist when the benefit of further fidelity gains is very small.

A key implication of drift-barrier theory is that environmental change can shift the mutation-rate equilibrium toward higher mutation rates. Under drift-barrier [Sung et al., 2012; Lynch et al., 2023], selection reduces mutation rates only while improved fidelity confers an advantage large enough to overcome drift; equilibrium is reached when further gains are too small to be reliably favored. Environmental change can shift this mutation-rate equilibrium upward in two ways. First, novel conditions can change the distribution of fitness effects, increasing the proportion of beneficial mutations among new variasnts, as illustrated schematically in Figure 1, thereby making mutator alleles more likely to hitchhike during adaptation. Put another way, if selection against mutators normally depends on their linkage to the deleterious mutations they generate, then any shift in the distribution of mutation effects toward beneficial or weakly deleterious variants *must*, all else equal, weaken this purging mechanism and move the mutation-rate equilibrium upward. Second, environmental change often coincides with a reduction in effective population size, such as during colonization bottlenecks, thereby strengthening drift and reducing selection against mildly deleterious mutator alleles. As Duffy observed, a lineage “suddenly thrust into an environment that it is not well adapted to” may benefit from a slightly higher mutation rate [Duffy, 2018]. Environmental transitions are therefore expected to generate episodes in which mutator lineages rise substantially.

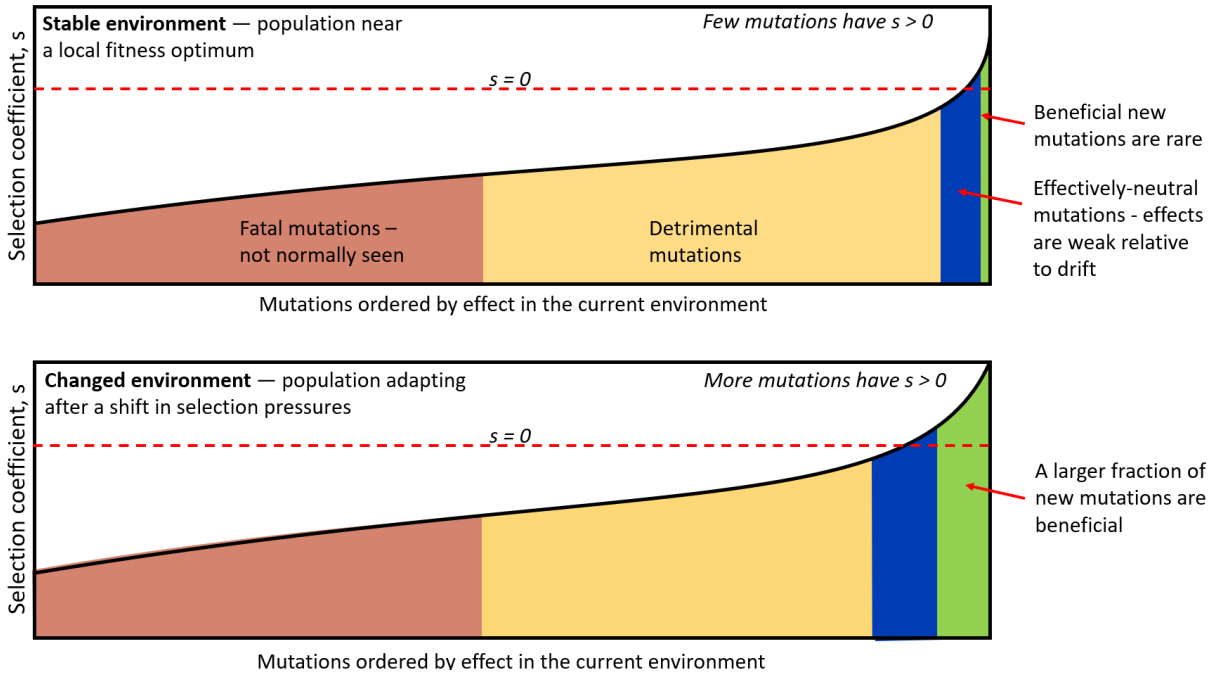


Figure 1. Mutation effects under different selection regimes. Mutations are grouped as fatal (brown, removed immediately), detrimental (beige, removed by purifying selection), effectively neutral (blue, fixed or lost by drift), and beneficial (green, fixed by positive selection). The dashed red line marks the neutrality threshold (s = 0). Top: In stable environments, populations are already well adapted, most readily available advantageous mutations have been fixed, and new beneficial mutations are rare; long-persisting lineages in relatively stable habitats, such as horseshoe crabs in shallow marine environments, illustrate this regime. Bottom: Environmental change shifts the distribution of fitness effects, so that some mutations that would previously have been neutral or deleterious become beneficial. Colonizing or radiating lineages, such as Darwin's finches in the Galápagos, illustrate this regime.

Figure 2 illustrates this scenario with an asexual stochastic individual-based simulation, detailed in Appendix S1. Environmental fluctuation repeatedly allows mutator lineages to rise during adaptation through weakened selection against mutators and hitchhiking with adaptive alleles; under plausible parameter values, these increases can be substantial.

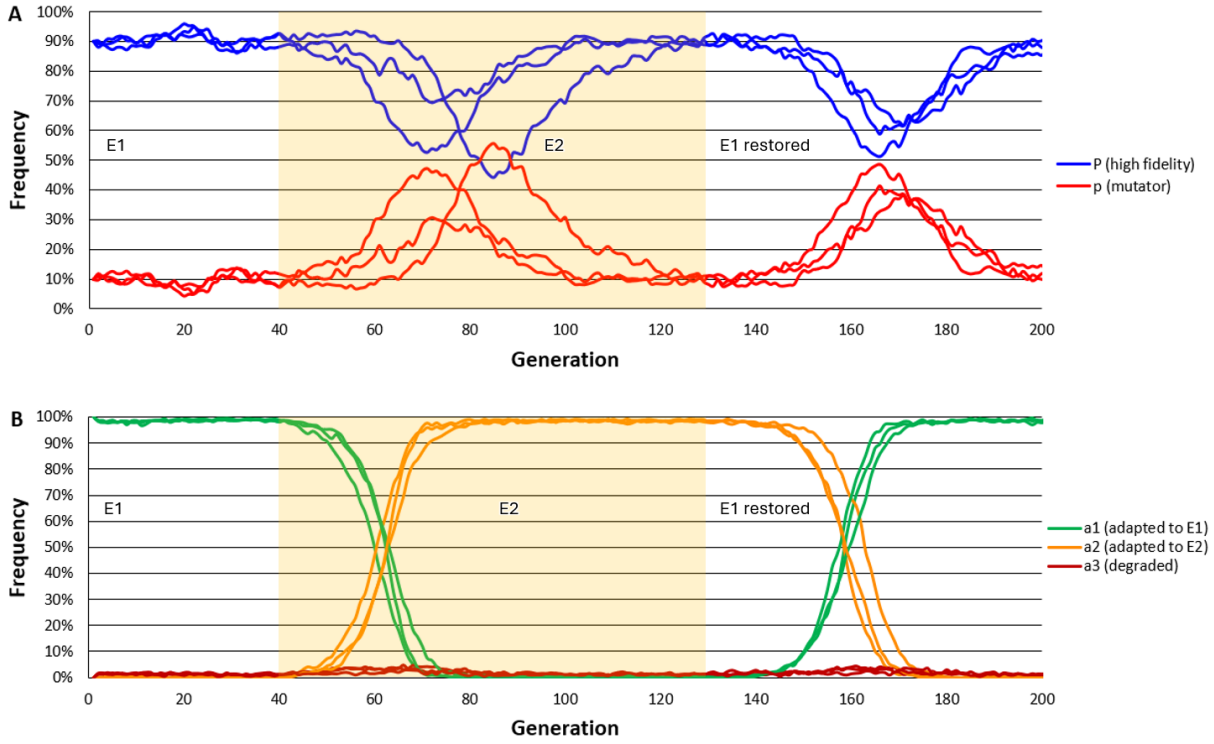


Figure 2. Individual-based simulation of mutator dynamics in an asexual population during environmental fluctuation. (A) Replication fidelity allele frequencies: mutator (*p*) frequency rises after each environmental transition. (B) Adaptation allele frequencies: adaptation succeeds in all runs regardless of polymerase background. Shading indicates Environment 2; unshaded regions indicate Environment 1. Transitions occur at generations 40 and 130. In 36 of 50 runs, peak mutator frequency exceeded 30%; full results are given in Figure S2 and Table S4.

This dynamic is consistent with empirical evidence from several systems, although mainly indirectly for multicellular taxa. Spontaneously arising mutator lineages repeatedly rose to high frequency in *E. coli* populations adapting to new environments [Sniegowski et al., 1997]. Environmental stress has been associated with genome-wide accumulation of de novo mutations and epimutations in *Arabidopsis* [Jiang et al., 2014]. Avian lineages with higher diversification rates show faster molecular evolution, including higher synonymous substitution rates, consistent with a link between diversification and mutation rate [Lanfear et al., 2010]. While some microorganisms use stress-induced mutagenesis via error-prone polymerases [Galhardo et al., 2007], the sexual animals and

plants Everest focuses on are more likely to experience shifts in mutation rates due to drift-barrier and hitchhiking dynamics [Lynch et al., 2023].

These considerations imply a fundamental tension. Over geological timescales, lineages require accurate replication to maintain complex genomes; over shorter timescales, however, elevated mutation rates can accelerate adaptation. Mutator alleles may therefore be transiently favored in unstable environments despite being harmful in the long run. Because mutations can arise in replisome genes, mutators may generate additional mutator alleles, increasing the risk of mutational meltdown unless some me xchanism limits their spread.

Sexual reproduction offers a possible resolution. Some lineages may be well adapted but low-fidelity, while others retain accurate replication machinery but are less well adapted. Recombination can generate offspring combining both advantages, but this process will be favored only if reproductive systems make the consequences of low fidelity visible to selection when immediate fitness differences are small. The Everest hypothesis proposes that the elaborate accompaniments of sexual reproduction are often co-opted to provide this function: reproductive complexity is therefore not merely ornamental but part of a broader system of genomic quality control, as developed in the following sections.

From mutator rise to fidelity restoration

The central question is therefore: how can sexual lineages make subtle differences in replication fidelity visible to selection, allowing higher-fidelity lineages to contribute disproportionately to future generations? The Everest framework (Figures 3 and 4) makes three distinctive conceptual contributions to understanding this problem. First, it treats adaptation and replication fidelity as separable evolutionary dimensions that can trade off and temporarily move in opposite directions. Second, it proposes an accelerator-and-brake cycle in mutation-rate evolution. Roberts and Petrie showed by modeling that strong mate choice, especially in lekking systems, can accelerate the spread of mutator backgrounds during adaptation by concentrating reproduction in high-performing individuals [Petrie & Roberts, 2007; Roberts & Petrie, 2022]. Everest asks how fidelity is subsequently restored. Third, it proposes multigenic stacking as the mechanism that can apply the corresponding brake: many gene products, regulatory steps, developmental processes, and behavioral or molecular functions can be integrated into a single reproductive test whose outcome is successful mating or fertilization—or, alternatively, reproductive failure. Such tests can convert many individually minor mutator-generated defects into a large difference in reproductive success. Recombination between adapted, lower-fidelity genotypes and high-fidelity partners can then produce offspring that combine adaptation with restored fidelity.

The Everest effect

The Everest effect is the amplification of small differences in replication fidelity into large differences in reproductive success. The name refers to climbing Mount Everest as a test of integrated performance: reaching the summit is a single readout of many systems working together under demanding conditions. A courtship display, a migration, or a fertilization sequence can work analogously, combining many biological components into one reproductive outcome. In this respect, an Everest test resembles a serial circuit: each added component both enlarges the target that mutation can strike and adds a link whose failure can compromise the whole outcome. This distinguishes Everest from broader condition-dependent signaling models, which usually treat elaborate traits as graded outputs of overall condition, with many physiological inputs contributing in parallel. Everest instead emphasizes a serial architecture in which damage that viability selection would barely register can be amplified into reproductive failure./784

The mechanistic argument depends on three features of complex reproductive systems. First, many lineages reproduce through systems whose elaboration substantially exceeds what fertilization or direct viability-fecundity benefits alone can explain. Second, such elaboration enlarges the mutational target and stacks many components into a single reproductive test. Third, disrupting any of those components often re/duces mating or fertilization success. If these features truly characterize the reproductive systems in question, they have an important consequence. Under the population-genetic conditions described in the Introduction—where environmental change can transiently increase mutator frequency through hitchhiking and weakened selection against mutators (Figure 2)—they imply an unavoidable statistical coupling between replication fidelity and

reproductive success. Lower-fidelity lineages generate more mutations, some of which disrupt components of the reproductive test and reduce mating or fertilization success. Enlarging the mutational target and stacking more mutation-sensitive components into a shared reproductive outcome should strengthen this coupling. The empirical question is therefore not whether some coupling exists, but whether it is strong and consistent enough to influence mutation-rate evolution. A core Everest prediction (Figure 3A) is that, all else equal, populations with greater reproductive complexity should maintain higher equilibrium replication fidelity than populations with simpler systems, because complex screens impose stronger penalties on mutator-generated damage.

An example illustrates the contrast with ordinary viability selection. A slightly compromised heart valve might reduce a non-migratory bird's lifetime reproductive output by 5–10%, with much of the effect masked by environmental variation; in a closely related migratory bird, the same defect could prevent the completion of a -demanding pre-breeding migration, reducing reproductive output to zero. Reproductive complexity can thus turn weakly expressed mutational damage into decisive reproductive failure—this is the Everest effect.

The same logic applies to antimutator alleles. By amplifying the reproductive consequences of newly accumulated damage, Everest tests can strengthen selection in favor of higher-fidelity genotypes. Both processes—purging mutators and favoring antimutators—are necessary because no screen is perfect: over macroevolutionary timescales, some mutators must slip through Everest testing and lineage sorting, and purging alone cannot restore the fidelity they erode; positive selection on antimutators is necessary.

This framing helps reconcile two apparently opposite effects of sexual selection on mutation rates. During environmental change, mate choice may act as an accelerator, favoring high-performing individuals even when their adaptive variants arise on mutator backgrounds. Once environments stabilize, however, the same reproductive architecture can act as a brake: elevated mutation rates continue to generate harmful mutations, and complex reproductive traits make the resulting damage disproportionately visible. Sexual selection can therefore act sequentially—first helping adaptation to proceed, then helping to restore fidelity when mutation load becomes the greater threat.

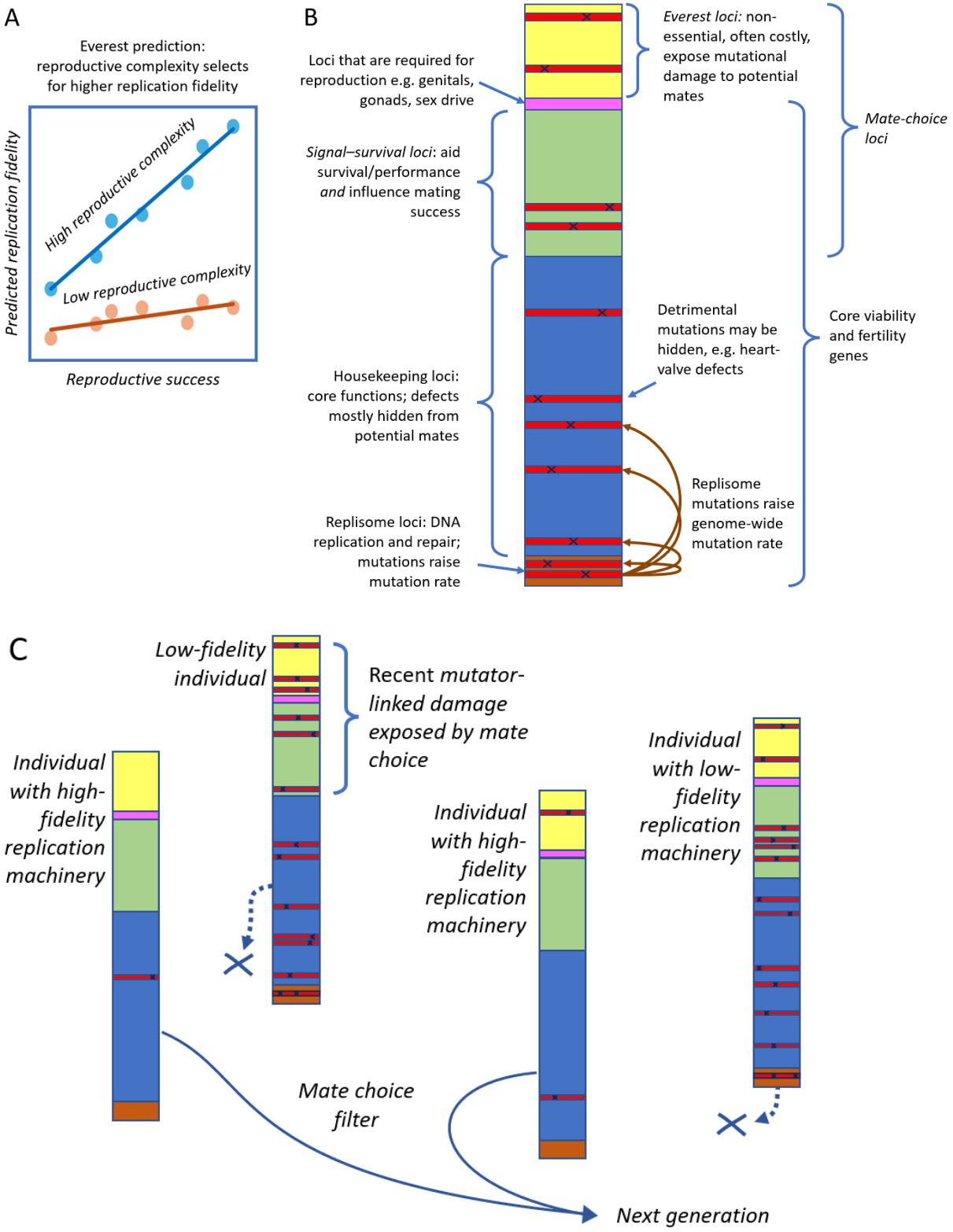


Figure 3. Everest hypothesis: predicted patterns, genomic architecture, and mechanism. Complex mate-choice systems can make mutator-linked damage visible to selection, coupling reproductive success to genome-wide genetic integrity. (A) Predicted equilibrium pattern. In stable environments, populations with high-complexity mate-choice systems (blue; many multigenic Everest traits) are predicted to maintain higher equilibrium replication fidelity than populations with low-complexity systems (brown). Complex reproductive screens impose stronger penalties on mutator-linked damage, so high reproductive success should be associated with higher replication fidelity, especially in high-complexity systems. Points and trend lines are schematic. (B) Loci by biological role. Yellow: Everest loci—non-essential, often costly, and conserved partly to expose mutational damage. Green: signal–survival loci—influencing both mating and survival. Blue: housekeeping loci—core viability and fertility functions, largely hidden from mates. Brown: replisome loci—replication and repair machinery. Replisome mutations increase the genome-wide

mutation rate, including at mate-choice loci, where their effects become apparent. Red X's mark mutations. (C) Individual-level mechanism. Each vertical bar represents an individual. Individuals carrying low-fidelity DNA replication machinery generate more new mutations over generations. The resulting damage remains associated with the low-fidelity background at first. Effects at mate-choice loci can be exposed by the mate-choice filter and reduce mating success. Individuals carrying high-fidelity replication machinery are therefore more likely to mate and transmit that machinery.

To clarify the mechanism, I distinguish several classes of loci. A "locus" is any genomic region—including coding, intronic, or regulatory sequences—where variation alters phenotype. *Mate-choice loci* affect mating or fertilization success. Within this class:

- *Signal–survival loci* contribute to both mating success and survival/performance, for example through strength, endurance, or cognition.
- *Everest loci* have effects that are non-essential for survival or fertilization—often costly or risky—but elaborated by sexual selection partly to make mutational damage highly visible.

These categories are positions on a continuum rather than sharp classes; what matters for fidelity screening is how many loci are integrated into a single readable test.

Elaborate displays provide a clear illustration of the proposed multigenic stacking. A peacock's train is not merely a large ornament: its effect depends on controlled feather growth, precisely positioned pigmentation and iridescence, bilateral symmetry, motor control, display behavior, and timing. Bird-of-paradise displays are similarly multigenic, integrating plumage structure, coloration, choreography, balance, stamina, sensory processing, and learning into a single stereotyped assessment.

Figure 3B orders locus types by function: Everest loci (yellow) and signal–survival loci (green) form a mate-choice sub-genome; housekeeping loci (blue) represent hidden core functions; and replisome loci (brown) encode replication machinery. Figure 3C shows the individual-level mechanism: error-prone replisomes generate an excess of new mutations across the genome, some of which affect mate-choice loci and reduce mating success (during the period before recombination breaks down the association between those mutations and the low-fidelity background).

Table 1 positions Everest relative to major frameworks for sexual traits, recombination, and mutation-rate dynamics. These include Fisherian and handicap models of sexual selection, classic accounts of recombination and mutation load—Hill–Robertson interference [Hill & Robertson, 1966], Muller's ratchet [Muller, 1964], and Kondrashov's "hatchet" [Kondrashov, 1988]—and Red Queen models of host–parasite coevolution [Hamilton et al., 1990]. The table summarizes which questions each framework addresses, not their overall explanatory power. The table is intended as an overview; fuller comparisons can be found in the section "Relations to previous work" and in Appendix S3.

Table 1. The position of Everest in relation to major frameworks for the evolution of sexual traits, recombination, and mutation-rate dynamics.

Framework	*Fidelity explicit*	*†Complexity-as-screen*	*‡Two-axis: fidelity can decrease while adaptation increases*	*Diagnostic tests*	*Selection contributions*
Everest	☑	☑	☑	☑(E,C,G)	M
Fisherian runaway	—	—	—	◐(C)	S
Handicap/indicator	—	◐	—	◐(E,C)	S
Good genes / condition / genic capture	—	◐	—	◐(C,G)	S+V
Red Queen (parasites)	—	—	—	☑(E,C)	V
Mutation-rate evolution	☑	—	◐	☑(E,G)	V
Kondrashov's "hatchet" (purging mutation load)	◐	—	◐	☑(E,G)	V
Hill–Robertson (linkage)	—	—	—	☑(E)	V
Muller's ratchet	—	—	—	☑(E)	V+L

†*Complexity-as-screen:* reproductive complexity evolves partly to increase sensitivity to mutational damage. Everest's mechanism is conjunctive—many components stacked into a single test, so a single disabling mutation can cause failure—thereby amplifying genome-wide damage in ways a graded, condition-dependent signal cannot. ◐ marks frameworks compatible with the idea but not predicting that mating traits become complex in order to reveal damage.

‡*Two-axis:* short-term adaptation and fidelity are treated as distinct and potentially conflicting properties rather than components of a single fitness measure.

Symbol key:

☑ = Central/explicit in that framework

◐ = Compatible or implicit/secondary (can be accommodated, but not a main focus)

— = Not addressed / not a prediction of the framework

Diagnostic test codes ("Diagnostic tests" column)

E = Experimental evolution/selection experiments

C = Comparative (across taxa/populations)

G = Genomic signature (sequence-based; mutation spectrum/rate, DNA replication/repair genes)

Selection contribution codes ("Selection contributions" column)

S = Sexual selection (mate choice/competition)

V = Viability-fecundity selection (non-mating)

L = Lineage-level effects (persistence/extinction)

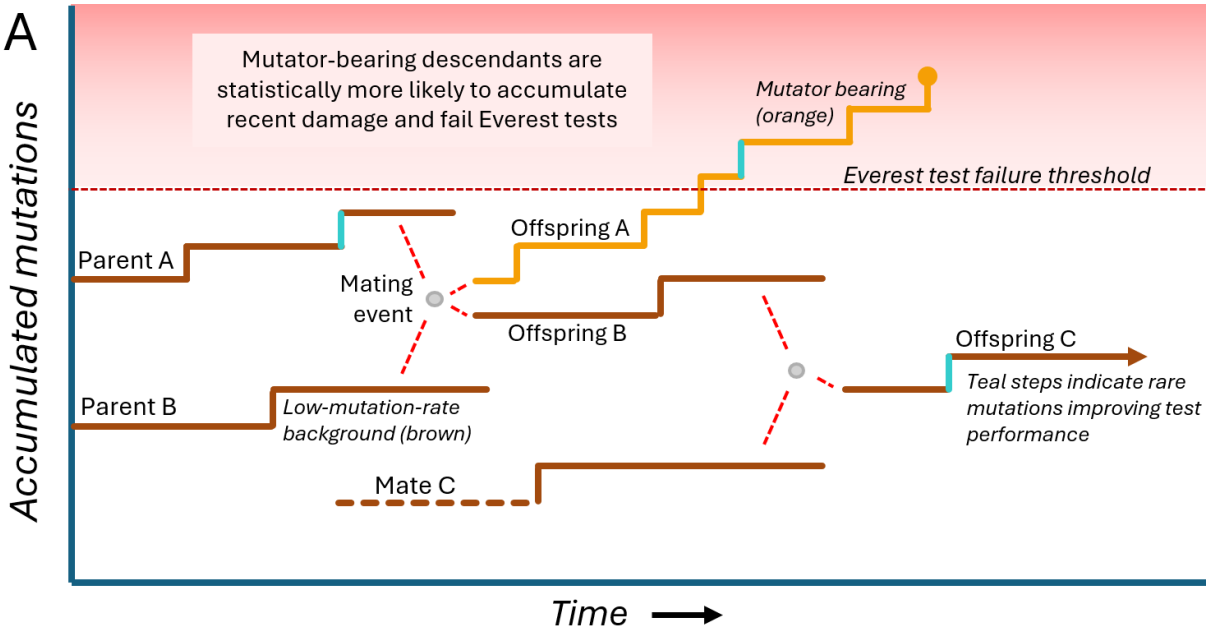
M = Mixed (more than one of S/V/L contributes)

Everest loci serve two functions. First, they expand the mutational target: adding non-essential components to reproductive traits increases the DNA regions where mutations can affect mate choice. Second, they can combine their own mutation-sensitive effects with existing signal–survival traits in a single reproductive test. A courtship display, migration-linked breeding schedule, or molecular compatibility system may test many underlying capacities that also contribute to fitness, such as sensory accuracy, motor coordination, metabolic regulation, learning, endurance, timing, and biochemical recognition.

While viability-fecundity selection should favor robustness and redundancy, sexual selection and lineage sorting may favor demanding tests where fragility cannot be avoided, precisely because they discriminate genetic integrity.

Critically, Everest tests must be shared widely enough within an interbreeding population to have statistical power. If preferences are fragmented—some individuals focusing on plumage, others on migration performance, others on biochemical compatibility—the Everest effect is weakened. However, Everest does not require a single universal test. It predicts convergence on one or a few dominant screens within each mating network, with the dominant screen varying across environments and regions. Where populations face different ecological demands, such as migratory versus resident strategies, different screens may predominate and contact zones may show mixtures of criteria.

Fidelity concerns mutation rate, not merely the total number of mutations an individual carries. A mutation rate is a slope—mutations accumulated per generation—so it cannot be read from a single phenotype in isolation. Selection on fidelity must therefore operate through repeated testing across generations. Figure 4 illustrates this logic across scales. At the parent–offspring level (panel A), parents who pass a demanding test establish a baseline; descendants carrying a mutator allele are statistically more likely to accumulate excessive recent damage and fail the same test. At the population level (panels B and C), strong selection can initially favor mutator lineages that adapt quickly but accumulate mutations unsustainably; recombination with remaining or immigrant high-fidelity genotypes can later generate well-adapted lineages with restored fidelity, tracing a rise-and-fall trajectory in adaptation–fidelity space. I refer to this process—recombination between adapted, low-fidelity genotypes and high-fidelity partners that generates offspring that inherit both adaptive alleles and accurate replication machinery—as *Everest recombination*. The orange trajectory in Figure 4 represents the accelerator phase described by Roberts and Petrie, in which sexual selection can favor mutator backgrounds during adaptation [Petrie & Roberts, 2007; Roberts & Petrie, 2022]. The blue trajectory represents the Everest extension: reproductive screening after recombination can restore high fidelity without erasing adaptive gains.



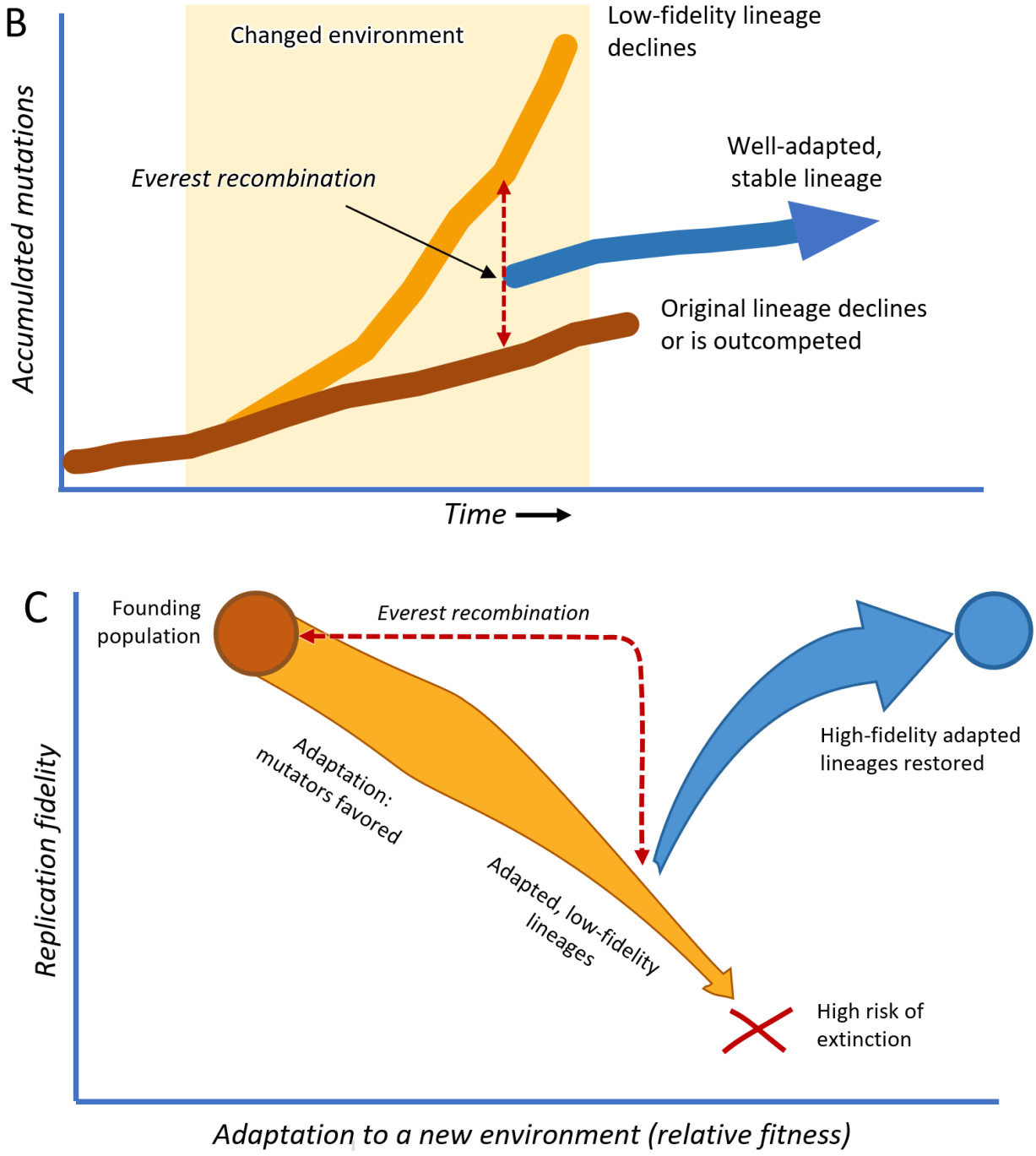


Figure 4. Everest predicts that mate choice can expose mutator-linked damage and favor recombination that restores high fidelity without erasing adaptation. (A) Parent–offspring schematic of mutation accumulation through time. Low-mutation-rate parents produce offspring that differ in subsequent mutation accumulation: one acquires a mutator allele (orange), generates mutations more rapidly, and is more likely to accumulate damaging changes that cause Everest-test failure and reduce mating success; another retains a low mutation rate (brown), remains below the failure threshold, and can produce low-mutation-rate descendants. The red dashed line represents the accumulated damaging effects at which the test fails. Teal steps denote rare mutations that improve test performance; dashed connectors indicate parentage. (B) Population-level mutation dynamics. During a changed environment (shading), a mutator lineage (orange) adapts quickly but accumulates mutations unsustainably and later declines; a low-mutation-rate lineage (brown) adapts more slowly and may decline or be outcompeted. Everest recombination—recombination between adapted low-fidelity genotypes and high-fidelity partners—followed by reproductive screening, yields a lineage (blue) that combines adaptation with restored fidelity. (C) The same scenario plotted in adaptation–fidelity space. Higher positions on the vertical axis indicate higher replication fidelity and therefore lower mutation rates. Mutator lineages move toward higher adaptation but lower fidelity (orange trajectory, red X); Everest recombination produces recovered lineages that retain adaptation while restoring high fidelity (blue). Curves are qualitative, not fitted to data.

The population dynamics underlying the recovery in Figure 4B and C warrant emphasis. During environmental change, low-fidelity mutator lineages may adapt quickly, rise to high frequency, and temporarily outcompete other lineages. However, their ongoing mutational load may erode fitness even in the new environment, eventually causing their numbers to decline. This reduces pressure on residual high-fidelity lineages, allowing their numbers to rebound and creating favorable conditions for Everest recombination. During colonization or metapopulation exchange, immigration can also replenish high-fidelity genotypes, providing further opportunities for Everest recombination without sacrificing local adaptation.

The hypothesis makes at least four testable predictions. First, under strong selection in a new environment, sexual populations should often evolve well-adapted but higher-mutation-rate lineages. Second, Everest recombination should be able to restore low mutation rates without sacrificing adaptation. Third, greater sexual and reproductive complexity should correlate with stronger Everest effects—lower mutation rates, reduced mutational load, or greater lineage persistence—relative to otherwise similar groups lacking such complexity. Fourth, where populations with different dominant screens interbreed, the Everest effect should be weakened; all else equal, such transition zones should exhibit higher mutation rates or mutational load than stable, well-screened populations on either side. Appendix S2 gives a quantitative argument showing how reproductive complexity can convert small per-locus mutation-risk differences into appreciable selection against mild mutators. I show there that when many mutation-sensitive loci contribute to a shared mating or fertilization screen, the implied selection coefficient against mild mutators can reach several percent.

The mate-selection strategies envisaged here can also reveal genetic defects beyond DNA replication. Complex behaviors and displays may also be sensitive to weakly acting mutations in housekeeping genes, including those involved in ribosomal function, cell cycle control, mitochondrial metabolism, transcription, translation, and so on.

Relations to previous work

Before turning to natural phenomena that Everest might help explain, it is useful to compare the hypothesis with existing frameworks for sexual traits, recombination, and mutation-rate evolution.

Everest is closest to good-genes and condition-dependent signaling models [Hamilton & Zuk, 1982; Rowe & Houle, 1996], which treat elaborate traits as indicators of genetic quality or genome-wide mutation load. It differs by specifying both a narrower target and a different top-level architecture. Whereas “condition” is a broad composite property, replication fidelity is a separable and measurable variable—the rate at which new mutations are generated. Condition-dependent models generally treat many physiological and developmental inputs as contributing in parallel to a graded signal, so strength in some components may partly compensate for weakness in others. Everest instead proposes that reproductive complexity can organize many mutation-sensitive components into a predominantly serial, conjunctive, and largely non-compensatory test. The underlying biology may contain both serial and parallel processes, but the top-level reproductive outcome depends on the successful completion of successive stages. In a peacock, for example, even construction of the train is highly sequential: feather buds must be positioned precisely to generate the train’s extraordinary symmetry; particular feathers must then grow to the appropriate lengths and forms; and multiple gene products must be expressed with precise spatial and temporal coordination to produce the two-dimensional eyespot patterns. The male must subsequently survive, reach and position himself at the lek, display the train appropriately, and elicit female acceptance. Failure at any critical stage can sharply reduce or eliminate reproductive success, and superior performance elsewhere may not compensate. Everest therefore emphasizes multigenic stacking into a single reproductive readout as opposed to the compensatory aggregation of overall condition.

Everest complements rather than replaces Fisherian and handicap frameworks [Fisher, 1930; Zahavi, 1975]. Traits elaborated by runaway or handicap dynamics can be co-opted as fidelity-linked reproductive screens—a peacock’s train may simultaneously reflect Fisherian selection, honest condition signaling, and sensitivity to genome-wide mutational damage. Everest does not require that reproductive complexity originate as a fidelity screen. Once such complexity has been co-opted

for fidelity screening, however, this additional function may contribute to its maintenance or further elaboration through individual-level selection. Over longer timescales, lineages with more effective screens may also be more likely to persist, producing lineage sorting even where fidelity screening played no role in the traits' original evolution.

A key distinction from pure handicap models is that a handicap requires only that a signal be costly enough to preclude falsification, whereas Everest predicts that complexity can acquire an additional adaptive role as a fidelity-sensitive screen: multigenic traits amplify mutational damage into visible reproductive failure in a way that simple costly traits cannot.

On mutation-rate evolution, Everest proposes an extension of drift-barrier logic: demanding reproductive screens translate small fidelity differences into large differences in mating or fertilization success. Petrie and Roberts describe the accelerator phase, in which mate choice can transiently favor mutator backgrounds during adaptation [Petrie & Roberts, 2007; Roberts & Petrie, 2022]. Everest adds the brake: recombination with high-fidelity partners, followed by reproductive screening, can restore fidelity after adaptation. Fuller comparison with these and other frameworks appears in Appendix S3 and Table 1.

Natural phenomena that the Everest hypothesis may help to explain

The Everest hypothesis is not proposed as an alternative to existing explanations for the origins of particular reproductive traits. Rather, it suggests that once demanding, multigenic reproductive traits exist, they may acquire a secondary role as fidelity-sensitive screens, regardless of how they first evolved. The following examples illustrate four broad settings in which such screening could operate: intense mate choice, demanding performance before reproduction, molecular or gametic compatibility, and developmental precision. They are intended to show the kinds of systems that might produce Everest effects and generate testable predictions, not to provide a comprehensive survey.

(1) **Lekking**. In lekking systems, males aggregate at display sites where females choose mates freely, usually without receiving resources or parental care. Lekking is therefore an extreme form of sexual selection and remains one of its persistent puzzles. The costs to males are obvious, but the costs to females can also be substantial: female Galápagos marine iguanas incur physiological costs while searching and assessing males on leks [Vitousek et al., 2007], female black grouse with higher body mass visit more males than lighter females [Rintamäki et al., 1995], and female pronghorn reduce mate-search effort during dry summers when body condition is compromised [Byers et al., 2006].

Everest interprets lekking as a high-intensity version of the accelerator/brake logic. During environmental change, intense female choice can accelerate adaptation by concentrating reproduction on high-performing males [Petrie & Roberts, 2007; Roberts & Petrie, 2022], including individuals carrying beneficial mutations that arose on mutator backgrounds. Once the environment stabilizes, however, the same system can act as a brake: elevated mutation rates are expected to affect the many traits required for lek success—survival, stamina, display, coordination, competitive position, and attractiveness—thereby allowing female choice to impose steep reproductive penalties on mutator lineages.

This framing offers a possible contribution to resolving the lek paradox, in which strong directional choice is expected to erode genetic variation in male quality and eventually make choosiness self-defeating [Kirkpatrick & Ryan, 1991; Rowe & Houle, 1996]. Roberts and Petrie help explain how variation can be regenerated during adaptation; Everest adds the corresponding purging phase, in which low-fidelity backgrounds are later exposed and selected against. Lekking may therefore be especially effective at driving the rise-and-fall dynamic predicted by Everest.

Everest also predicts that females may relax costly lek-based assessment under ecological stress, when signals of current adaptation become relatively more valuable and the costs of assessment rise. This is consistent with the reduced mate-search effort reported in pronghorn and black grouse.

(2) **Long-distance migration.** Arctic terns, Atlantic salmon, and monarch butterflies undertake demanding journeys that require coordinated navigation, physiology, endurance, and timing

(Figure 5). Everest does not claim that migration originated as a fidelity screen; migration has clear ecological functions. Once established, however, long-distance migration can act as a severe repeated performance test. Migration, therefore, differs from lekking in that it is expected mainly to act as a brake on elevated mutation rates rather than an accelerator. Lineages with elevated mutation rates are predicted to be more likely, over generations, to accumulate defects in one or more of these systems and fail before breeding. Persistent migration may therefore impose sustained downward pressure on mutation rates.

Monarch butterflies add a further complication because their migration is a multi-generation relay. Northerly-breeding insects must transmit alleles that control overwintering behavior in conditions that they themselves never experience. Such systems suggest testable predictions: all else equal, migratory lineages are expected to show lower mutation rates or lower mutational load than related non-migratory lineages. Gene flow from migratory into non-migratory populations might often reduce load, although this advantage could temporarily reverse during environmental change if non-migratory populations acquire locally adaptive variants more rapidly.

(3) **Molecular recognition systems** may extend the logic of Everest beyond animal courtship or female choice. In many organisms, reproductive success depends on complex biochemical interactions between gametes, mating types, or male and female tissues. These systems may create fidelity-sensitive screens through fertilization success rather than behavioral choice.

Flowering plants provide a useful example. Pollen grains function as mobile haploid organisms whose success depends on germination, pollen tube growth through maternal tissue, and precise guidance to the ovule. These steps involve successive cell-cell interactions and multilayered signaling pathways between pollen, pistil, and ovular tissues. Mutations that impair receptors, ligands, signaling components, growth dynamics, or guidance responses can reduce fertilization success before zygote formation. Pollen-pistil systems can therefore act as molecular Everest tests: they can impose reproductive penalties on lineages that generate more mutations, and therefore more damaging hits, in the molecular machinery required for successful fertilization.

Similar logic may apply to gamete-recognition systems in broadcast-spawning marine organisms, corals, fungi, protists, and animals in which fertilization success is shaped by sperm competition or cryptic female choice. In vertebrates, sperm success depends not only on motility but also on capacitation, passage through female reproductive tract filters, interaction with the zona pellucida or equivalent egg coats, and species-specific sperm–egg recognition and fusion. In each case, the relevant point is not that Everest explains the origin of molecular compatibility systems. Rather, once such systems exist, their complexity and specificity can couple replication fidelity to fertilization success.

Figure 5. Examples of demanding performance, courtship display, and reproductive recognition as multigenic fidelity-sensitive screens under Everest. Top row: long-distance migrants—Arctic tern, Atlantic salmon, and monarch butterfly—undertake journeys requiring coordinated navigation, physiology, endurance, and timing; successful migration may therefore provide a single reproductive readout of many underlying systems functioning together. Middle row: elaborate avian displays illustrate a similar principle of multigenic stacking. The male superb bird-of-paradise display and the peacock's lek-associated train each integrate multiple components—including plumage structure, coloration, symmetry, posture, coordination, stamina, and female assessment—into a single outcome: mating success. Bottom row: flowers illustrate developmental precision and pollinator-mediated screening. Symmetry is one component of a consistent, recognizable floral form that pollinators can learn to recognize and repeatedly visit [Møller & Eriksson, 1995; Chittka & Raine, 2006]. Coral spawning and fungal mating-type interactions illustrate reproductive filtering through molecular recognition, gamete compatibility, and fertilization. Everest interprets demanding performance, elaborate display, floral form, and reproductive recognition as potential filters through which small defects across many loci may be amplified into reduced reproductive success. These examples illustrate Everest's logic and generate testable predictions; they are not intended as a systematic survey.

Several other reproductive systems may impose related fidelity-sensitive filters but are not developed here as separate examples; as above, the point is not that Everest explains their origin, but that traits shaped by other processes may be co-opted as reproductive screens. Elaborate bird displays and complex songs, for example, depend on coordinated development, learning, neuromuscular control, and performance. Preferences for symmetry, or greater reproductive success of more symmetrical individuals, have been reported in humans evaluating potential partners, as well as in birds, insects, and flowers [Rhodes et al., 2001; Bańbura, 2005; Koshio et al., 2007]. Species-specific conformity of appearance in birds and other animals may provide a related opportunity for stabilizing selection around shared developmental templates, especially where both sexes display complex plumage or patterning; individual tits, jays, gulls, and similar birds can be

difficult for human observers to distinguish. Under Everest, visible deviations from the expected form could reduce reproductive success if they are penalized during mate choice. Floral complexity offers a plant analog, with pollinators rather than mates imposing the screen. Mutations that deform flowers, alter scent or color, or otherwise disrupt pollination routines may reduce pollinator visitation. Because most angiosperms lack an early-segregated germline, mutational damage visible in floral tissues is expected to be correlated with mutational damage in gametes. Floral complexity could therefore act as a pollinator-mediated fidelity test. Appendix S4 considers possible analogs in asexual organisms, where mate choice is absent, but complex performance traits or life-cycle bottlenecks may still filter mutational damage. Appendix S5 considers pregnancy loss as a possible post-fertilization fidelity screen in viviparous animals.

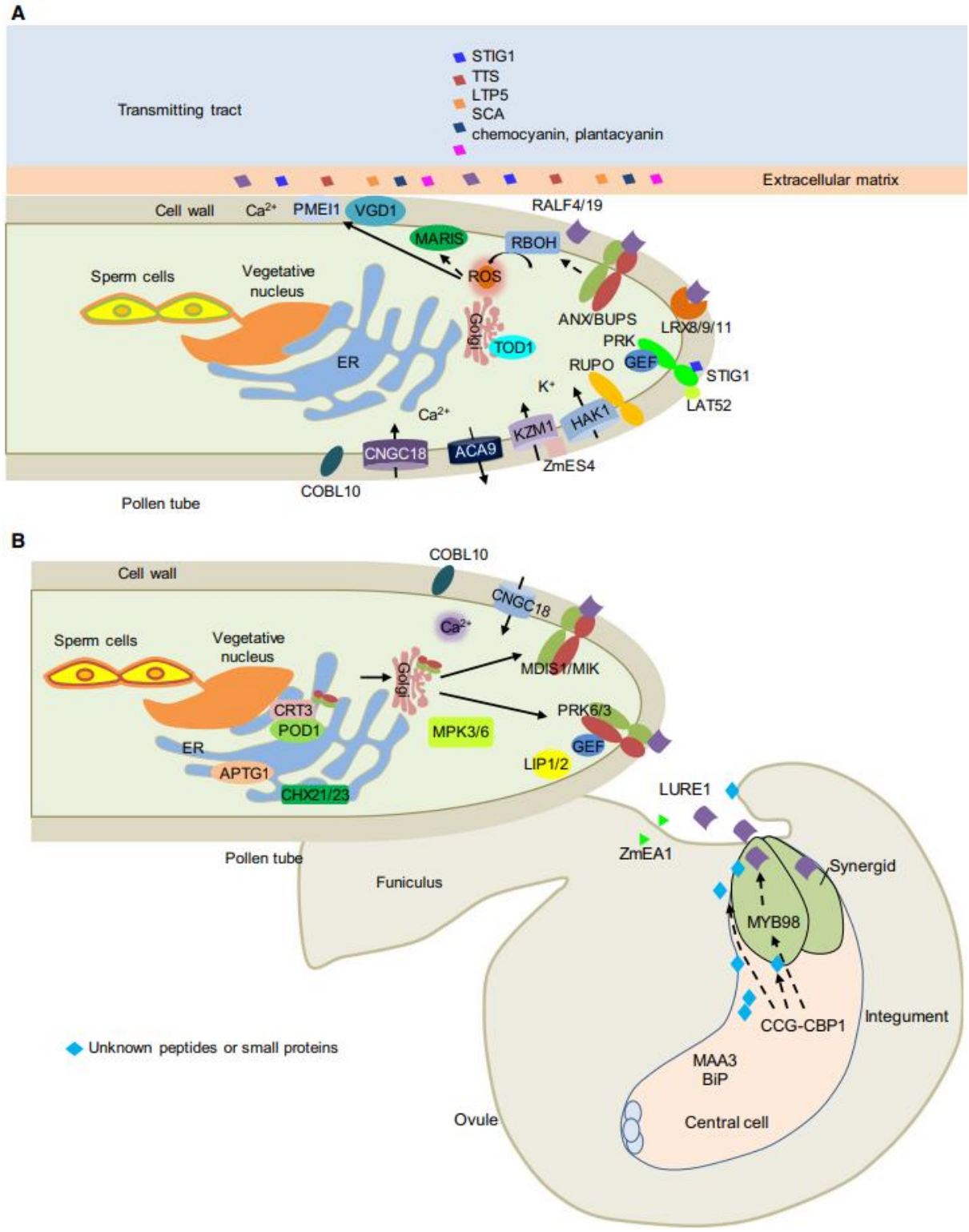


Figure 6. Pollen tube guidance as an example of molecular reproductive screening under Everest. Fertilization in flowering plants requires pollen tube growth through maternal tissue and precise navigation to the ovule, mediated by successive cell–cell interactions and multilayered biochemical

signaling [Li et al., 2018]. Pollen grains function as mobile haploid organisms whose success depends on many gene products. This complexity could impose a molecular reproductive screen, making fertilization success more sensitive to mutational damage and thereby illustrating how elaborate fertilization systems could act as fidelity-linked screens. Figure reproduced from Li HJ, Meng JG, Yang WC, "Multilayered signaling pathways for pollen tube growth and guidance," Plant Reproduction 31, 31–41 (2018), by permission of Springer Nature.

Levels of selection: individual selection and lineage sorting

The Everest hypothesis operates through individual-level selection, not group selection. Individuals with higher replication fidelity tend to perform better on Everest tests and gain a direct reproductive advantage; individuals carrying mutator alleles tend to perform worse and leave fewer descendants. This is the standard population-genetic argument that mutators are purged through linkage to the harmful mutations they generate [Kimura, 1967]. Everest does not change that logic but amplifies the selective signal until selection that would otherwise be too weak to overcome drift can purge mild mutators and favor mild antimutators. No individual sacrifices fitness for group benefit, and no group-level selection coefficient is required.

This perspective has a lineage-level consequence: lineages lacking effective fidelity screens are expected to accumulate greater mutational load and face elevated extinction risk, not because rival groups outcompete them but because they may collapse from within. The same logic may help explain how costly Everest tests, such as the peacock's train, can persist despite their burden: their viability costs may be offset by the evolutionary advantage of filtering out low-fidelity backgrounds. On this view, Everest tests may be more than a refinement of sexual selection: they may contribute to the long-term survival of complex organisms in a fluctuating world. This is lineage sorting, not group selection.

Limitations of existing experimental evidence

(1) **Experimental studies consistent with fidelity filtering.** A few experimental evolution studies provide indirect support for the Everest hypothesis, although it has not yet been tested directly. In flour beetles, sexual selection reduced extinction risk and purged deleterious mutations [Lumley et al., 2015]. In seed beetles, males from sexually selected lines transmitted smaller induced mutation loads following germline DNA damage, suggesting a link to germline maintenance [Baur & Berger, 2020]. In hermaphroditic freshwater snails, sexual selection similarly contributed to the purging of deleterious mutations across multiple generations [Noël et al., 2019]. Collectively, these studies show that sexual selection can reduce mutation load, but a direct test of the Everest hypothesis would require linking success in demanding reproductive traits to independently measured germline mutation rates—for example, through parent–offspring whole-genome sequencing.

(2) **Experimental studies that seem relevant to the Everest hypothesis without directly testing it.** Many experimental studies in sexual selection fall into this category. Carotenoid-supplementation experiments in zebra finches altered immune function and sexual attractiveness [Blount et al., 2003], but addressed current condition and ornament expression rather than heritable mutation rate or replication fidelity. Field experiments in common yellowthroats that link ornamentation to oxidative DNA damage likewise concern physiological maintenance rather than the transmission of germline mutations [Freeman-Gallant et al., 2011]. Long-term experimental evolution studies under altered mating systems are also only indirectly relevant. In *Drosophila melanogaster*, enforced monogamy altered genome-wide gene expression after many generations [Hollis et al., 2014], and experimental hyper-promiscuity produced evolutionary responses to the altered sexual environment [Perry et al., 2016]. In *Drosophila pseudoobscura*, long-term manipulation of sexual selection generated substantial genomic divergence between mating regimes [Wiberg et al., 2021]. These studies demonstrate that sexual selection can reshape traits, gene expression, and genomes over time, but they do not test the narrower Everest claim that demanding reproductive traits function as fidelity-linked screens by making mutator-linked damage visible to selection.

Experimental and observational tests of the Everest hypothesis

(1) Experimental evolution in laboratory and captive populations. A variety of sexual organisms could be used to test the Everest hypothesis in laboratory evolution experiments, including protists, plants, insects (such as *Drosophila*, flour beetles, or seed beetles), fish, wild birds, and captive and farm mammals and birds. The design in Figure 7A and Box 1 asks whether strong selection affects replication fidelity (L2 versus L1) and whether recombination with wild-type strains can restore it (L4 versus L3), with whole-genome sequencing likely needed for unambiguous interpretation; similar approaches could be extended to zoo populations. Figure 7B maps the predicted trajectories of each line onto adaptation–fidelity space, with the L3–L4 comparison testing whether recombination can restore fidelity while preserving adaptation. In all such designs, replicated lines and controls for effective population size, inbreeding, and demographic bottlenecks would be essential, because these factors can also affect mutation load.

(2) Sequencing existing laboratory-evolved lines. A complementary approach would be to revisit laboratory lineages already evolved under altered mating systems—such as enforced monogamy, hyper-promiscuity, reduced sexual selection, or altered mate choice. Whole-genome sequencing of archived or repeatable lines could test whether relaxation or redirection of reproductive filtering altered mutation rates, mutational loads, or allele frequencies in replication and repair genes. Experiments such as Perry et al.'s hyper-promiscuous *Drosophila* study are useful models because they show that genetic manipulation of the socio-sexual environment can rapidly reshape reproductive traits, but the authors did not directly measure replication fidelity [Perry et al., 2016].

(3) Experimental manipulation of cricket song. A workable investigation into the role of Everest loci could use crickets by disabling a major target of mate choice: the male calling song. In one line of a genetically mixed population captured before mating, males would be muted by applying non-toxic glue or filler to the file and scraper on their forewings, while control males would receive a sham treatment on non-stridulatory regions of the forewings, leaving their song intact. After several generations, whole-genome sequencing of parent–offspring trios (or pooled samples) from each line could be used to estimate per-generation mutation rates and changes in replication and repair genes. The Everest hypothesis predicts that the song-muted line should more readily retain mutator alleles or accumulate a greater mutational load than the control, because removing the screen relaxes reproductive filtering.

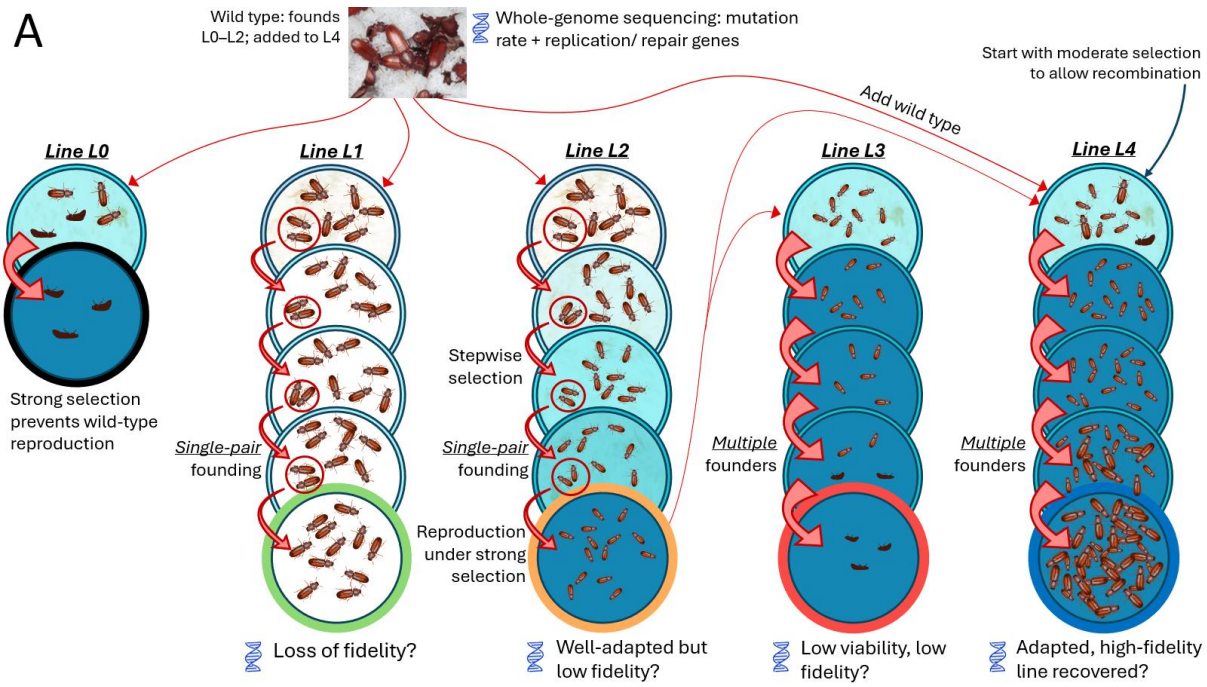
(4) Extension to birds and mammals. It would be valuable to extend experimentation to birds and mammals more broadly. The experimental approach of manipulating ornament elaboration has been established in widowbirds [Andersson, 1982] and peafowl [Petrie & Halliday, 1994], among others. Replicated populations could be maintained for multiple generations under conditions in which females either can or cannot use key ornamental traits in mate choice, for example by repeatedly shortening or obscuring those traits. Germline mutation rates could then be estimated from parent–offspring sequencing and compared between treatments. Everest predicts that populations in which ornament-based screening is disabled should evolve higher germline mutation rates than otherwise comparable control populations.

(5) Experiments with lekking animals. Lekking species might seem attractive for testing Everest because highly successful males are easy to identify in some systems, such as black grouse. However, the predicted relationship between lek success and replication fidelity is context-dependent. Highly successful males are expected to show lower mutation rates or lower mutational load only where the environment has been stable long enough for lekking to operate mainly as a fidelity filter. During environmental change, successful males may instead include individuals from mutator backgrounds carrying recent beneficial variants. This makes lekking a potentially informative but difficult test case. Moreover, if lekking already functions as a strong fidelity filter, surviving males may differ only subtly in fidelity, reducing statistical power. Any paternal signal will also be diluted because mothers contribute equally to offspring replication fidelity. Nevertheless, useful opportunities may exist. In black grouse and other intensively studied lekking species, long-term field datasets with archived genetic samples may enable tests for associations between male attractiveness, environmental context, and sequence variation in DNA replication and repair

genes. Similar approaches may be feasible in other lekking or lek-like systems, including tephritid fruit flies, lek-forming *Drosophila*, and fiddler crabs.

(6) Experimental manipulation of pollination in flowering plants. A parallel test could be conducted in a plant that is easy to grow and manipulate experimentally. A genetically mixed population would be split into two lines: in a restricted pollination line, each maternal plant would be hand-pollinated with pollen from randomly chosen single donors, while in a control line, plants would be left to open pollination, allowing pollen from many donors to compete and interact with the pistil. After several generations in a common garden, lines would be compared by whole-genome sequencing with attention to replication and repair genes. The Everest hypothesis predicts that removing pollen competition and reducing pistil-level screening should allow greater accumulation of deleterious mutations (and possibly reduce vigor) in the restricted line.

(7) Comparative and population-genetic tests of migration. The hypothesis predicts that long-distance migrants should have lower mutation rates than comparable non-migratory relatives—a prediction that can be approached in several complementary ways. First, existing ringing and mark–recapture survival data, combined with phylogenetic comparisons of migratory and resident lineages, could test whether migratory species tend to be longer-lived than closely related non-migratory counterparts, since longevity is one possible consequence of more effective mutator purging. Second, and more directly, whole-genome sequencing of parent–offspring trios from migratory and non-migratory populations of partially migratory species—such as European robins—could identify de novo mutations, estimate per-generation germline mutation rates, and test whether migration is associated with lower mutation rates. Third, genome-wide data from transitional areas where migratory and resident populations interbreed could be used to ask whether genetic introgression is asymmetric—that is, whether alleles move more often from one population into the other than vice versa. After controlling for population size, demography, dispersal opportunity, and geographic structure, Everest predicts that alleles from migratory populations should more often enter resident populations than the reverse. Such a pattern would not by itself prove higher replication fidelity in migrants, but it would provide a testable population-genetic signature consistent with migration acting as a demanding fidelity-linked screen.



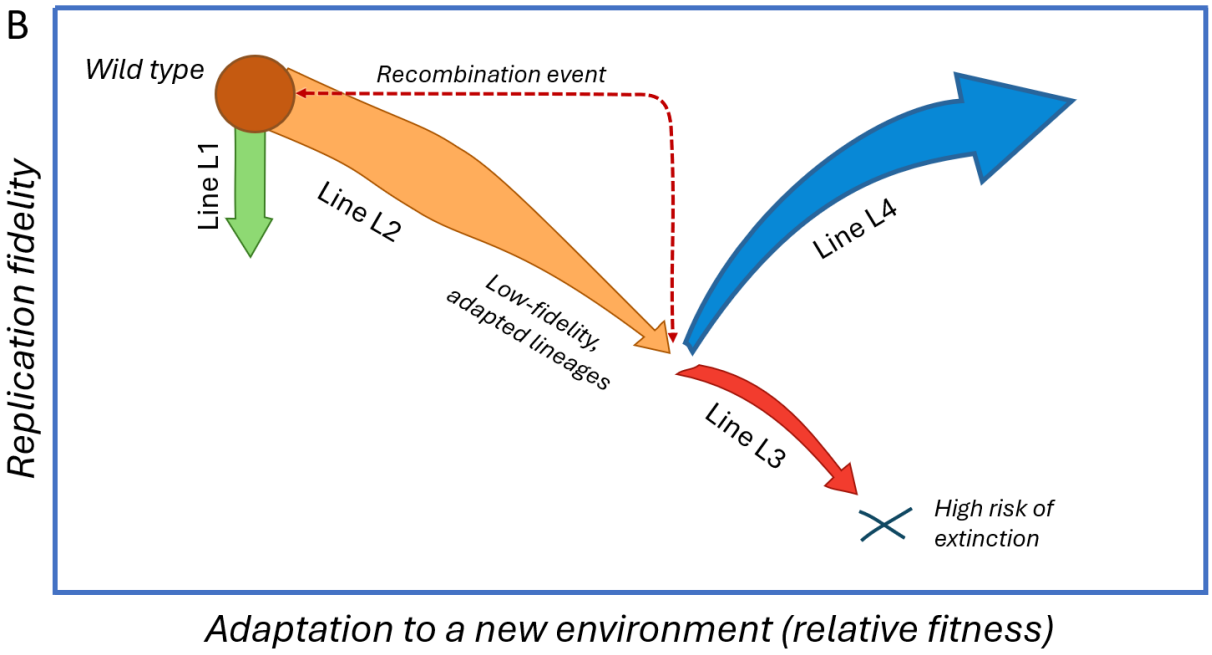


Figure 7. Experimental test of predicted mutation-rate dynamics and recombination-mediated recovery in a sexual model system such as flour beetles. (A) Experimental design. Populations are exposed to increasingly severe selective environments, shown by darker blue circles. L0 identifies conditions that just prevent wild-type reproduction. In L1 and L2, each generation is founded by a single pair to limit genetic mixing between independently evolving lineages: L1 is a bottleneck control without strong selection, whereas L2 experiences stepwise increasing selection and is predicted to permit the rise of mutator genotypes with reduced replication fidelity. From L2, two lines are derived. L3 continues strong selection without backcrossing, but with multiple individuals transferred each generation, testing whether the mutation rate remains elevated or increases, and whether this is accompanied by increased extinction risk. L4 introduces wild type to allow recombination, then continues strong selection by transferring multiple individuals each generation, testing whether recombination can generate descendants that combine retained adaptation with a lower directly measured mutation rate. Whole-genome sequencing of parent–offspring trios or pedigrees would estimate the mutation rate and detect changes in replication and repair genes. Because wild-type introgression can also complement defects elsewhere in the genome and produce heterosis or demographic rescue, improved viability in L4 would not by itself demonstrate restoration of replication fidelity. (B) Predicted trajectories in adaptation–fidelity space, using the same axes as Figure 4C. The wild-type population (brown) begins with high fidelity but low adaptation. L1 (green) represents bottlenecking without strong selection; L2 (orange), adaptation under stepwise selection with limited genetic mixing, yielding adapted but lower-fidelity lineages; L3 (red), continued strong selection without backcrossing, predicted to result in further fidelity loss and high extinction risk; and L4 (blue), recombination with wild type followed by strong selection, predicted to restore fidelity while maintaining adaptation. The key test is whether descendants in L4 retain adaptation while showing lower directly measured mutation rates than those in L2 and L3. This comparison tests whether recombination-mediated recovery is possible; it does not, by itself, establish that reproductive screening caused that recovery. Genomic sequencing and the other experimental approaches discussed in the main text provide complementary tests of these trajectories and their underlying mechanisms.

Box 1. Experimental readouts and predictions for the design shown in Figure 7

Measurements

1. Per-site germline mutation rate in the ancestral wild-type stock and in lines L1–L4, estimated from de novo mutations identified by whole-genome sequencing of parent–offspring trios or pedigrees
2. Candidate functional changes and allele frequencies in replication and repair genes, including DNA polymerases and their associated factors; where candidate mutator alleles are identified in L2, their frequencies and associations with directly measured mutation rates would be followed through L3 and L4
3. Genome-wide ancestry and allele-frequency changes, allowing changes at candidate replication and repair loci to be compared with the overall contribution of the introduced wild-type genome
4. Line viability and population dynamics under strong selection (extinction risk, population size, body size)

Null expectations: no detectable mutation-rate effect

If adaptation to the selective environment is independent of mutation-rate evolution:

- Wild type and L1, the bottleneck control, should show similar mutation rates and no consistent changes in replication or repair genes beyond drift
- L2 and L3 may become better adapted to the selective environment but should show no systematic increase in directly measured mutation rate or enrichment of functional mutator alleles
- Wild-type introgression may initially improve viability or lower mutation rate in L4 through direct inheritance, complementation, or heterosis; these effects alone would not demonstrate selection for restored fidelity

Patterns supporting the mutation-rate and recombination components of the Everest hypothesis

- L2 shows a higher directly measured mutation rate than the ancestral wild-type stock and L1, accompanied by candidate functional changes or allele-frequency shifts in replication or repair genes
- L3 shows persistence or further elevation of mutation rate and retention or spread of identified candidate mutator alleles; increased mutational load and reduced viability would provide supporting, but not independently diagnostic, evidence
- L4 produces descendants that retain adaptation while showing lower directly measured mutation rates and loss or replacement of candidate mutator alleles identified in L2

Patterns that would challenge the hypothesis

- L2 and L3 show no consistent increase in mutation rate or functional changes in replication or repair genes relative to controls
- L4 fails to produce descendants combining retained adaptation with lower mutation rates, despite the opportunity for recombination with high-fidelity genotypes

---

Conclusions

The Everest hypothesis offers three linked conceptual contributions to evolutionary theory, specifically to understanding the evolution of replication fidelity in sexual populations. First, it separates short-term adaptation from replication fidelity as distinct evolutionary dimensions. Second, it proposes an accelerator-and-brake dynamic in mutation-rate evolution. Third, it identifies multigenic stacking in complex reproductive systems as a mechanism that can strengthen the brake. Together, these ideas address how sexual populations may restore replication fidelity after adaptation without losing adaptive gains.

The first contribution is a two-dimensional framework for mutation-rate evolution. Short-term adaptation and replication fidelity are distinct but interacting properties: they can trade off and may temporarily move in opposite directions. Evolutionary trajectories should therefore be analyzed by tracking adaptation and fidelity separately, rather than collapsing both into a single measure of fitness.

The second contribution is an accelerator-and-brake dynamic. After environmental change, beneficial mutations may become more common or more valuable, while demographic bottlenecks can weaken

selection against mild mutators. Mutator alleles can therefore rise by hitchhiking with adaptive variants, and strong mate choice may intensify this accelerator phase by concentrating reproduction on a few high-performing individuals, including those on mutator backgrounds [Petrie & Roberts, 2007; Roberts & Petrie, 2022]. Everest adds the corresponding brake: as adaptation proceeds and the benefits of elevated mutation wane, the continuing costs of mutational damage become more important. Effective reproductive screens can then favor antimutator or high-fidelity backgrounds, driving mutation rates back down.

The third contribution identifies multigenic stacking as the mechanism that strengthens the brake. Complex reproductive systems can organize many mutation-sensitive developmental, physiological, behavioral, or molecular components into a predominantly conjunctive test with a single reproductive outcome—an Everest test. Because success depends on completing many successive steps, failure at a critical stage may not be offset by strong performance elsewhere. This relatively non-compensatory architecture, analogous to a serial circuit, can amplify small defects into large differences in reproductive success, strengthening selection against mild mutators and favoring antimutator or high-fidelity backgrounds when ordinary viability–fecundity selection is too weak to do so reliably.

By exposing fidelity-linked differences during mate choice, gamete competition, or fertilization, Everest tests can increase the reproductive contribution of high-fidelity genotypes and thus their opportunities to recombine with adapted genotypes. The resulting descendants may retain adaptive alleles while recovering more accurate replication machinery. This process, termed Everest recombination, can allow sexual populations to preserve recent adaptive gains without the burden of an elevated mutation rate.

Everest complements rather than replaces Fisherian, handicap, good-genes, and condition-dependent models [Fisher, 1930; Zahavi, 1975; Hamilton & Zuk, 1982]. Traits elaborated through these processes can later be co-opted as Everest tests that screen replication fidelity. Its distinctive claim is not that these traits require a new explanation of their origin, but that their complexity can acquire a secondary role in amplifying mutator-linked damage and promoting the restoration of replication fidelity.

At longer timescales, this individual-level filtering can produce lineage sorting. Although complex reproductive screens such as burdensome ornaments can be costly, lineages that expose mutator-linked damage more effectively may recover and maintain higher fidelity and persist longer than comparable lineages lacking effective screens. This is lineage sorting rather than group selection: the immediate mechanism remains differential reproductive success among individuals, and differential lineage persistence is a downstream consequence.

Everest's fidelity-screening logic can extend beyond organisms with overt behavioral mate choice to sexual lineages in which reproductive filtering occurs through gamete competition, fertilization, or molecular recognition. Examples include broadcast spawners, flowering plants, and fungi, in which sperm performance, pollen-tube growth, gamete compatibility, or mating-type recognition can determine reproductive success. In each case, the coordinated operation of many gene products can create effective (complex) Everest tests without overt behavioral choice.

In general, the Everest effect should be strongest when demanding, multigenic criteria are widely shared within an interbreeding population, and weaker when those criteria are simple, fragmented or inconsistent.

The hypothesis is testable and falsifiable (Figure 7). Experimental evolution under strong selection can test whether mutator lineages rise during adaptation and whether recombination or backcrossing restores fidelity without erasing adaptive gains. Manipulating reproductive traits in animals and plants can test whether removing or weakening critical screens allows mutator alleles or mutational load to accumulate, while parent–offspring sequencing can determine whether success in demanding screens is associated with lower germline mutation rates. Comparative studies can test whether, after controlling for demography and life history, lineages with more demanding reproductive screens show lower mutation rates, reduced mutational load, or greater long-term persistence.

The elaborate accompaniments of sexual reproduction—including demanding migrations, intricate courtship displays, and complex fertilization systems—may serve functions beyond their immediate contributions to reproductive success. By distinguishing replication fidelity from fitness, predicting an accelerator-and-brake dynamic in mutation-rate evolution, and identifying multigenic tests with a single reproductive readout as a mechanism of fidelity screening, the Everest hypothesis proposes that these traits may also contribute to genomic quality control. More broadly, Everest suggests that sexual reproduction may help complex organisms colonize and persist in unstable environments by permitting rapid adaptation during environmental change while promoting the restoration of replication fidelity afterward. The central empirical questions are how often such filtering occurs, how strongly it acts, and how much it contributes to the persistence and adaptability of sexual lineages.

References

Andersson M. Female choice selects for extreme tail length in a widowbird. *Nature*. 1982 Oct 28;299(5886):818-20. https://doi.org/10.1038/299818a0

Bańbura, J. Sexual selection in the Swallow Hirundo rustica - A review. *Acta Universitatis Lodziensis, Folia Biologica et Oecologica* 2 2005 Jan;2:57-69. https://doi.org/10.18778/1730-2366.02.04

Baur J, Berger D. Experimental evidence for effects of sexual selection on condition-dependent mutation rates. *Nature Ecology & Evolution*. 2020 May;4(5):737-44. https://doi.org/10.1038/s41559-020-1140-7

Blount JD, Metcalfe NB, Birkhead TR, Surai PF. Carotenoid modulation of immune function and sexual attractiveness in zebra finches. *Science*. 2003 Apr 4;300(5616):125-7. https://doi.org/10.1126/science.1082142

Byers JA, Byers AA, Dunn SJ. A dry summer diminishes mate search effort by pronghorn females: evidence for a significant cost of mate search. *Ethology.* 2006 Jan;112(1):74-80. https://doi.org/10.1111/j.1439-0310.2006.01127.x

Chittka L, Raine NE. Recognition of flowers by pollinators. *Current Opinion in Plant Biology*. 2006 Aug 1;9(4):428-35. https://doi.org/10.1016/j.pbi.2006.05.002

Duffy S. Why are RNA virus mutation rates so damn high? *PLOS Biology*. 2018 Aug 13;16(8):e3000003. https://doi.org/10.1371/journal.pbio.3000003

Fisher RA. *The Genetical Theory of Natural Selection.* Oxford: Clarendon Press; 1930.

Freeman-Gallant CR, Amidon J, Berdy B, Wein S, Taff CC, Haussmann MF. Oxidative damage to DNA related to survivorship and carotenoid-based sexual ornamentation in the common yellowthroat. *Biology Letters*. 2011 Jun 23;7(3):429-32. https://doi.org/10.1098/rsbl.2010.1186

Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a stress response and the regulation of evolvability. *Critical Reviews in Biochemistry and Molecular Biology*. 2007 Sep-Oct 1;42(5):399-435. https://doi.org/10.1080/10409230701648502

Hamilton WD, Axelrod R, Tanese R. Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences*. 1990 May;87(9):3566-73. https://doi.org/10.1073/pnas.87.9.3566

Hamilton WD, Zuk M. Heritable true fitness and bright birds: a role for parasites? *Science.* 1982 Oct 22;218(4570):384-7. https://doi.org/10.1126/science.7123238

Hill WG, Robertson A. The effect of linkage on limits to artificial selection. *Genetical Research*. 1966 Dec;8(3):269-94. https://doi.org/10.1017/S0016672300010156

Hollis B, Houle D, Yan Z, Kawecki TJ, Keller L. Evolution under monogamy feminizes gene expression in Drosophila melanogaster. *Nature Communications*. 2014 Mar 18;5(1):3482. https://doi.org/10.1038/ncomms4482

Jiang C, Mithani A, Belfield EJ, Mott R, Hurst LD, Harberd NP. Environmentally responsive genome-wide accumulation of de novo Arabidopsis thaliana mutations and epimutations. *Genome Research*. 2014 Nov 1;24(11):1821-9. https://doi.org/10.1101/gr.177659.114

Kimura M. On the evolutionary adjustment of spontaneous mutation rates. *Genetics Research*. 1967 Feb;9(1):23-34. https://doi.org/10.1017/S0016672300010284

Kirkpatrick M, Ryan MJ. The evolution of mating preferences and the paradox of the lek. *Nature.* 1991 Mar 7;350(6313):33-8. https://doi.org/10.1038/350033a0

Kondrashov AS. Deleterious mutations and the evolution of sexual reproduction. *Nature*. 1988 Dec 1;336(6198):435-40. https://doi.org/10.1038/336435a0

Koshio C, Muraji M, Tatsuta H, Kudo SI. Sexual selection in a moth: effect of symmetry on male mating success in the wild. *Behavioral Ecology.* 2007 May 1;18(3):571-8. https://doi.org/10.1093/beheco/arm017

Lanfear R, Ho SY, Love D, Bromham L. Mutation rate is linked to diversification in birds. *Proceedings of the National Academy of Sciences*. 2010 Nov 23;107(47):20423-8. https://doi.org/10.1073/pnas.1007888107

Li HJ, Meng JG, Yang WC. Multilayered signaling pathways for pollen tube growth and guidance. *Plant Reproduction*. 2018 Mar;31(1):31-41. https://doi.org/10.1007/s00497-018-0324-7

Lumley AJ, Michalczyk Ł, Kitson JJ, Spurgin LG, Morrison CA, Godwin JL, Dickinson ME, Martin OY, Emerson BC, Chapman T, Gage MJ. Sexual selection protects against extinction. *Nature*. 2015 Jun 25;522(7557):470-3. https://doi.org/10.1038/nature14419

Lynch M, Ali F, Lin T, Wang Y, Ni J, Long H. The divergence of mutation rates and spectra across the Tree of Life. *EMBO Reports.* 1995 May;73(1):15-22. https://doi.org/10.15252/embr.202357561

Møller AP, Eriksson M. Pollinator preference for symmetrical flowers and sexual selection in plants. *Oikos.* 1995 May 1:15-22. https://doi.org/10.2307/3545720

Muller HJ. The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 1964 May 1;1(1):2-9. https://doi.org/10.1016/0027-5107(64)90047-8

Noël E, Fruitet E, Lelaurin D, Bonel N, Segard A, Sarda V, Jarne P, David P. Sexual selection and inbreeding: Two efficient ways to limit the accumulation of deleterious mutations. *Evolution Letters*. 2019 Feb 1;3(1):80-92. https://doi.org/10.1002/evl3.93

Perry JC, Joag R, Hosken DJ, Wedell N, Radwan J, Wigby S. Experimental evolution under hyper-promiscuity in Drosophila melanogaster. *BMC Evolutionary Biology*. 2016 Jun 16;16(1):131. https://doi.org/10.1186/s12862-016-0699-8

Petrie M, Halliday T. Experimental and natural changes in the peacock's (Pavo cristatus) train can affect mating success. *Behavioral Ecology and Sociobiology*. 1994 Sep;35(3):213-7. https://doi.org/10.1007/BF00167962

Petrie M, Roberts G. Sexual selection and the evolution of evolvability. *Heredity*. 2007 Apr;98(4):198-205. https://doi.org/10.1038/sj.hdy.6800921

Rhodes G, Yoshikawa S, Clark A, Lee K, McKay R, Akamatsu S. Attractiveness of facial averageness and symmetry in non-Western cultures: In search of biologically based standards of beauty. *Perception*. 2001 May;30(5):611-25. https://doi.org/10.1068/p3123

Rintamäki PT, Alatalo RV, Höglund J, Lundberg A. Mate sampling behaviour of black grouse females (Tetrao tetrix). *Behavioral Ecology and Sociobiology*. 1995 Sep;37(3):209-15. https://doi.org/10.1007/BF00176719

Roberts G, Petrie M. Sexual selection for males with beneficial mutations. *Scientific Reports*. 2022 Jul 23;12(1):12613. https://doi.org/10.1038/s41598-022-16002-y

Rowe L, Houle D. The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 1996 Oct 22;263(1375):1415-21. https://doi.org/10.1098/rspb.1996.0207

Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of E. coli. *Nature*. 1997 Jun 12;387(6634):703-5. https://doi.org/10.1038/42701

Sung W, Ackerman MS, Miller SF, Doak TG, Lynch M. Drift-barrier hypothesis and mutation-rate evolution. *Proceedings of the National Academy of Sciences*. 2012 Nov 6;109(45):18488-92. https://doi.org/10.1073/pnas.1216223109

Vitousek MN, Mitchell MA, Woakes AJ, Niemack MD, Wikelski M. High costs of female choice in a lekking lizard. *PLOS ONE*. 2007 Jun 27;2(6):e567. https://doi.org/10.1371/journal.pone.0000567

Wiberg RAW, Veltsos P, Snook RR, Ritchie MG. Experimental evolution supports signatures of sexual selection in genomic divergence. *Evolution Letters*. 2021 5, Jun 1;5(3):214-229. https://doi.org/10.1002/evl3.220

Zahavi A. Mate selection—a selection for a handicap. *Journal of Theoretical Biology.* 1975 Sep 1;53(1):205-14. https://doi.org/10.1016/0022-5193(75)90111-3

The Everest hypothesis: Supplementary Material

This Supplementary Material contains: Appendix S1, an individual-based stochastic simulation showing how environmental change can elevate mutator allele frequency through hitchhiking; Appendix S2, a quantitative multi-locus analysis of reproductive screening; Appendix S3, an expanded discussion of Everest in relation to previous work; Appendix S4, a discussion of adaptation–fidelity tradeoffs beyond sexual organisms; and Appendix S5, a speculative discussion of pregnancy loss as a possible distributed fidelity-sensitive screen.

Appendix S1. An Individual-Based Simulation of Mutator Dynamics During Environmental Change

Overview

This section presents an individual-based stochastic simulation that shows how environmental change can elevate mutator allele frequency via mutator hitchhiking. The model tracks 1,024 individuals across 200 generations, each carrying alleles at two loci: a polymerase locus that determines replication fidelity and an adaptation locus that determines fitness in the current environment. The simulation incorporates genetic drift, probabilistic selection, and stochastic mutation events, allowing outcomes to vary across runs. The model is intentionally asexual: permanent linkage between the polymerase and adaptation loci produces a strong hitchhiking signal when adaptive mutations arise on mutator backgrounds. In sexual populations, recombination would weaken individual hitchhiking events by decoupling p from the beneficial mutations that carry it to high frequency—but it would also weaken the subsequent purging of mutators after environmental change, increasing rather than reducing the need for Everest screening. An interactive Excel spreadsheet implementing this model is available as Supplementary File S1, allowing readers to regenerate runs by pressing F9.

Model Description

The model tracks a haploid asexual population of 1,024 individuals across 200 generations. Each individual carries alleles at two loci (genomic positions):

1. *Polymerase locus (P/p):* determines replication fidelity. P = high-fidelity polymerase (low mutation rate); p = low-fidelity mutator polymerase (high mutation rate).
2. *Adaptation locus (A):* determines fitness in the current environment. Three alleles are possible: a1 (adapted to environment E1), a2 (adapted to environment E2), and a3 (degraded, irreversible dead-end allele with low fitness in all environments).

The critical asymmetry is that a1↔a2 transition rates and a1/a2→a3 degradation rates both depend on which polymerase allele is carried: P individuals mutate at rate μ_P per locus per generation, while p individuals mutate at rate μ_p, which is 50-fold higher. Transitions to a3 are irreversible. The polymerase locus itself flips between P and p at a low background rate μ_pol. This is not intended to reflect a specific biological mechanism; it is a modeling convenience that maintains a continuous background frequency of both alleles (~10% p at equilibrium), preventing either allele from being permanently lost by drift.

Dynamics

Each generation proceeds in three steps:

Step 1 – Selection: Each generation, pairs of individuals compete in probabilistic tournaments. The probability that individual A defeats individual B is fitnessA / (fitnessA + fitnessB), where fitness is determined by the adaptation locus and current environment (Table S1). The winner reproduces; this is repeated until 1,024 offspring are produced.

Step 2 – Mutation: Each offspring's adaptation locus mutates with probability determined by its polymerase allele. A mutation at the adaptation locus is deleterious (→a3) with probability given by the degradation probability, and adaptive (a1→a2 or a2→a1) otherwise. Transitions to a3 are irreversible.

Step 3 – Environment: The environment switches from E1 to E2 at generation 40 and returns to E1 at generation 130, creating two successive episodes of environmental change.

Table S1. Parameters

Parameter	Symbol	Value	Description
Population size	N	1,024	Fixed throughout
Generations	—	200	Per simulation run
High-fidelity mutation rate	μ_P	0.002	Mutation rate at the A-locus for P individuals
Low-fidelity mutation rate	μ_p	0.1	Mutation rate at the A-locus for p individuals
Polymerase flip rate	μ_pol	0.01	Background P↔p rate
Deleterious:beneficial ratio	—	20:1	Ratio of degrading to adaptive mutations
Degradation probability	—	0.952	Calculated from that ratio
Environment 2 start	—	Generation 40	Onset of environmental change
Environment 2 end	—	Generation 130	Return to Environment 1

Table S2. Fitness values by genotype and environment

Adaptation allele	Environment E1	Environment E2
a1 (adapted to E1)	1	0.5
a2 (adapted to E2)	0.5	1
a3 (degraded)	0.1	0.1

Table S3. Derived transition rates by polymerase background

Transition	P background	p background
a1 → a2 (adaptive)	0.0001	0.0048
a1 → a3 (deleterious)	0.0019	0.0952
a2 → a1 (adaptive)	0.0001	0.0048
a2 → a3 (deleterious)	0.0019	0.0952
a3 → anything	0	0

Table S3 shows derived values that are not directly editable. The parameters above should be adjusted to update these values in the model.

Initial Conditions

All 1,024 individuals begin as a1 (adapted to E1). Approximately 90% carry P and 10% carry p, reflecting the equilibrium mutator frequency under the chosen polymerase flip rate in the absence of directional selection.

Results

The simulation was run 50 times with independent random seeds. Two single runs are shown in Figure S1, illustrating the two qualitatively distinct outcomes the model produces. For each run, the peak mutator (p) allele frequency reached during the first environmental switch (generations 40–130) was recorded; the resulting distribution across all 50 runs is shown in Figure S2, and individual values are listed in Table S4.

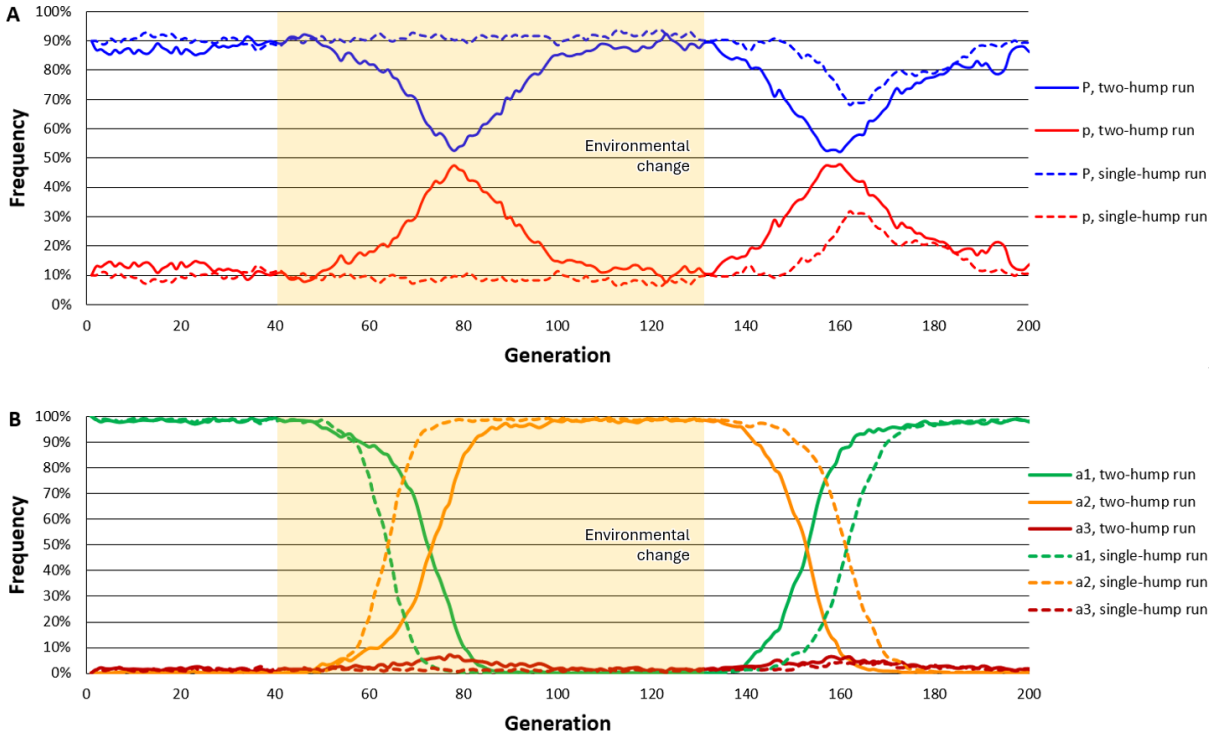


Figure S1. Example trajectories of the individual-based simulation, illustrating two qualitative outcomes summarized in Figure S2. Two runs are superimposed: a representative two-hump run (solid lines), in which the successful adaptive sweep arises on a mutator background at both environmental switches; and a representative single-hump run (dashed lines), in which the first adaptive sweep arises on a high-fidelity background, while the second arises on a mutator background. (A) Polymerase allele frequencies: P, high-fidelity polymerase, shown in blue; p, mutator polymerase, shown in red. When the successful adaptive allele arises on a p background, $p$hitchhikes upward with the spreading adaptive allele; when it arises on a P background, $p$remains near baseline. The single-hump outcome is therefore an expected stochastic result, not a failure of the model: the background that first supplies a successful adaptive variant can be either Por p, depending on mutation supply, starting frequency, degradation risk, and chance establishment. (B) Adaptation allele frequencies: a_1, adapted to E1, shown in green; a_2, adapted to E2, shown in orange; and a_3, degraded, shown in dark red. Adaptation succeeds in both runs; the runs differ in which polymerase background hosts the adaptive sweep, not in whether adaptation occurs. Recovery of high fidelity during stable phases reflects conventional selection against mutators through the adaptive locus; Everest predicts that this recovery should be amplified by multi-locus reproductive screens and steep mate-choice functions, as discussed in Appendix S2. Shading indicates the period of environmental change, generations 40–130. Parameter values are as in Tables S1–S3.

In this simulation, the recovery of high fidelity during stable phases is driven by conventional Kimura-style linkage logic: mutators are selected against because they remain linked to the deleterious mutations they generate at the adaptive locus. The Everest hypothesis predicts that this recovery should be stronger in sexual populations where many mutation-sensitive loci are stacked into shared reproductive screens, and where mate choice or fertilization success imposes a steeper reproductive penalty than ordinary viability–fecundity selection. Appendix S2 explicitly develops this quantitative

amplification, showing how increasing the number of mutation-sensitive loci and the steepness of the reproductive response can accelerate the recovery of high fidelity beyond that expected from conventional selection acting through ordinary adaptive or housekeeping loci.

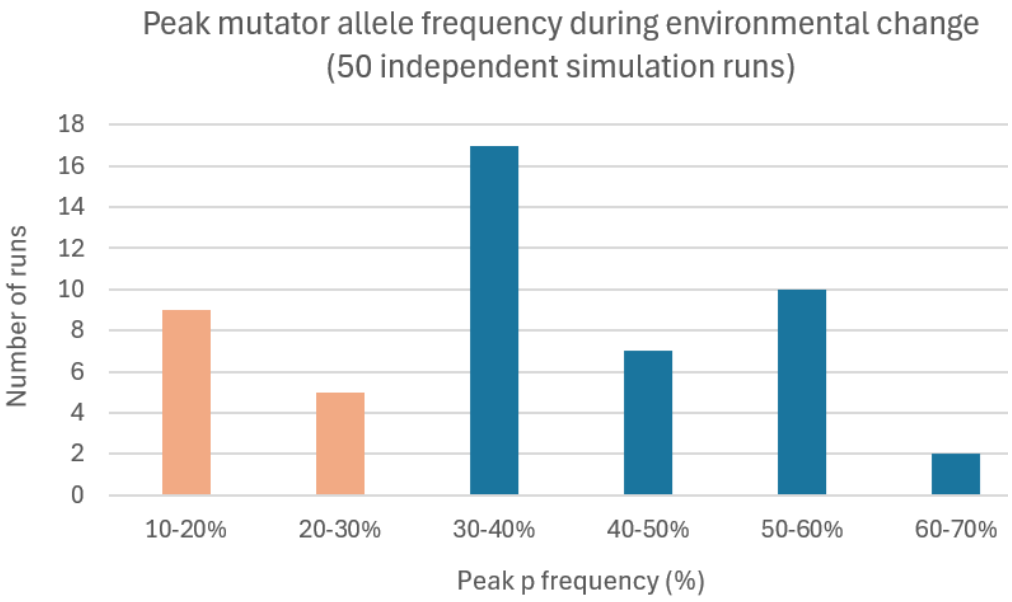


Figure S2. Distribution of peak mutator (p) allele frequency across 50 independent runs of the individual-based simulation. For each run, the maximum frequency reached by the mutator allele *p* during the first environmental switch (generations 40–130) was recorded; the 50 values are shown binned in 10% intervals. The distribution is bimodal, reflecting two qualitatively distinct outcomes. In 14 of 50 runs (orange bars, peak $p \leq 30\%$), the first adaptive $a_1 \rightarrow a_2$ mutation arose on a high-fidelity (P) background; the adaptive sweep carried no mutator with it, and peak *p* frequency remained at or near the pre-switch baseline of ~10%. In the remaining 36 of 50 runs (72%, blue bars, peak $p > 30\%$), the first adaptive mutation arose on a mutator (p) background; *p* hitchhiked with the spreading a_2 allele and rose substantially in frequency (mean peak across all 50 runs: 37.1%). The bimodality reflects the stochastic race between P- and p-background lineages to generate the first adaptive mutation: which background wins this race is decided early and largely determines the subsequent trajectory of mutator frequency.

Table S4. Peak mutator allele frequency recorded during 50 independent runs of the model during the first environmental switch (generations 40-130).

35.4%	43.1%	52.8%	33.0%	14.5%
33.0%	38.3%	34.0%	38.8%	39.3%
48.2%	44.3%	35.4%	27.9%	41.9%
35.3%	11.8%	38.1%	15.1%	21.7%
65.2%	64.3%	39.9%	51.0%	31.4%
44.1%	43.3%	52.7%	53.3%	24.8%
56.4%	50.0%	37.9%	31.7%	13.2%
37.4%	13.2%	56.4%	25.0%	52.4%
56.1%	57.9%	13.9%	14.6%	55.2%
14.9%	32.0%	27.0%	36.9%	19.5%

The distribution of peak p frequency was bimodal (Figure S2), reflecting the two outcomes shown in Figure S1. In 36 of 50 runs (72%), the adaptive sweep occurred predominantly on a mutator background (Figure S1, solid lines), driving peak p frequency above 30% during the first switch (mean peak across all runs: 37.1%). In the remaining 14 runs, adaptation proceeded on a high-fidelity (P) background during the first switch (Figure S1, dashed lines), and mutator frequency remained near the pre-switch baseline of ~10%. This bimodal pattern reflects the stochastic race between P-background and p-background individuals to generate the first adaptive a_2 mutation: when a p-

background individual wins, it hitchhikes with the spreading a_2 allele, substantially elevating mutator frequency; when a P-background individual wins, mutator frequency remains low despite the elevated mutation rate of p individuals.

The absence of mutator hitchhiking at the first switch in the single-hump run is an expected stochastic outcome, not a failure of the model. Although mutator-background individuals generate adaptive mutations at a higher per-individual rate, high-fidelity individuals are more numerous at baseline and are less likely to degrade to a_3. The successful adaptive mutation can therefore sometimes arise and establish on a P background, leaving p near its baseline frequency during that switch. The minority outcome should become progressively less common as the number of loci required for adaptation increases, because multi-locus adaptation gives mutator backgrounds more opportunities to generate the required variants. Since real adaptive episodes often involve many loci, the single-locus model likely underestimates mutator hitchhiking in nature. The 72% hitchhiking rate observed here is therefore best read as a conservative lower bound under these parameters.

Parameter Justification

Parameter values were chosen to be within the range observed in microbial experimental evolution systems, where mutator dynamics are best documented. The 50-fold difference between μ_P and μ_p is consistent with that observed in mutator strains of bacteria and RNA viruses [Sniegowski et al., 1997]. The background mutator frequency of ~10% is consistent with observations of mutator frequencies in natural and experimental bacterial populations [Sniegowski et al., 1997]. The deleterious:beneficial ratio of 20:1 is more favorable to beneficial mutations than most empirical estimates for stable environments, but is appropriate here as the model represents conditions of active environmental change, during which the proportion of beneficial mutations is expected to be substantially elevated [Lynch et al., 2023]. The population size of 1,024 is smaller than most natural populations; small Ne is conservative in the context of this model, as it represents conditions where selection against mutators is weakest, and Everest screening therefore matters most.

Model Limitations

This model deliberately simplifies several aspects of real biological systems: (1) only a single adaptation locus is modelled, whereas real adaptive episodes likely require changes at multiple loci, which would amplify the mutator hitchhiking effect; (2) the population is asexual, which maximises hitchhiking through permanent linkage — in sexual populations recombination partially decouples the polymerase and adaptation loci, weakening individual hitchhiking events but creating the problem that Everest screening is proposed to solve; (3) the model represents a single class of loci rather than a whole genome; (4) population size is fixed. Despite these simplifications, the model captures the essential dynamics: environmental change creates conditions under which mutator alleles rise substantially in frequency through hitchhiking with adaptive sweeps, demonstrating that elevated mutation rates during environmental change are a predictable population-genetic consequence rather than a theoretical possibility.

Methods and use of AI

This model was designed and implemented by the author. The model structure was developed and the model expanded to more individuals and generations with the help of Claude AI (Anthropic).

Appendix S2. Quantitative logic of multi-locus reproductive screening

The Everest hypothesis follows the conventional argument that mutator alleles can be purged because they remain linked to the deleterious mutations they generate. Its additional claim is that reproductive complexity can strengthen this purging by integrating many mutation-sensitive loci into a shared mating or fertilization screen. The amplification can be quantified with a simple probability argument.

Let n be the number of mutation-sensitive loci contributing to a reproductive screen, and let d be the per-locus probability of functionally relevant mutational damage over a defined interval. Assuming that damage events are independent, the probability that no component is damaged is $(1 - d)^n$. Therefore, the probability that at least one component is impaired is:

$$P \text{ (at least one impaired component)} = 1 - (1 - d)^n.$$

When d is small, this is approximately:

$$1 - (1 - d)^n \approx nd.$$

Thus, even small per-locus differences in mutation risk can become appreciable when many loci contribute to the same reproductive screen. For example, if high-fidelity genotypes have $d = 0.001$ and mild mutators have $d = 0.002$, the difference at one locus is only 0.001. Across 100 loci, however, the probability that at least one component is impaired is $1 - (0.999)^{100} = 0.095$ for high-fidelity genotypes and $1 - (0.998)^{100} = 0.181$ for mild mutators. The absolute difference in expected impairment is therefore nearly nine percentage points.

The evolutionary consequences of this nine-point differential also depend on the mating system. If expected reproductive success is written schematically as

$$W = V \times M,$$

where V is viability–fecundity performance and M is mating or fertilization success, then ordinary viability–fecundity selection acts mainly through V, whereas Everest tests act mainly through M. In weakly selective mating systems, impairment may cause only a modest reduction in M. In strongly selective systems, especially leks, the function can be much steeper: small declines in display, stamina, coordination, competitive position, or recognition can sharply reduce mating success, even to zero.

For example, if impairment of any one component reduces mating success by 50% in a steep mating system, the expected mating success of high-fidelity genotypes in the example above is $1 - 0.5\,(0.095) = 0.9525$, whereas that of mild mutators is $1 - 0.5\,(0.181) = 0.9095$. The implied selection coefficient against mild mutators is therefore $s = 1 - 0.9095/0.9525 \approx 0.045$, large enough, under simple constant-selection assumptions, to drive substantial mutator decline over tens of generations.

Thus, even a small per-locus difference in mutation risk can generate a selection coefficient of several percent when many loci are stacked into a steep reproductive screen. Everest therefore predicts that mutators should be removed more strongly when reproductive screens involve many mutation-sensitive components, and when small defects cause large drops in mating success.

This calculation is intentionally simple: real reproductive screens will include components with unequal effect sizes, partial redundancy, and correlated damage. These complications alter the details, but not the central expectation: reproductive screens that integrate many mutation-sensitive components should amplify small differences in replication fidelity into appreciable differences in mating or fertilization success, thereby strengthening selection against mild mutators and favoring the maintenance or recovery of high fidelity.

Appendix S3. Everest in relation to previous work

Everest builds on several existing lines of theory concerning sexual selection, recombination, mutation load, and mutation-rate evolution. This section places Everest alongside those frameworks in the same order as Table 1 in the main text, highlighting overlaps and key points of contrast.

Sexual selection frameworks

Fisherian runaway

Fisherian runaway explains the evolution of exaggerated ornaments through positive feedback between preference and display: attractive traits gain mates, and preferences for those traits spread because they produce attractive offspring [Fisher, 1930]. In its pure form, this process need not track any external measure of quality, so elaborated traits may be costly or arbitrary with respect to viability. Everest is compatible with runaway elaboration but suggests that such traits can later be co-opted as fidelity screens when their development depends on many mutation-sensitive loci.

Handicap and indicator models

Amotz Zahavi's handicap principle and related indicator models treat costly traits as honest signals of condition or underlying genetic state [Zahavi, 1975]. Everest is compatible with the general idea that mate choice can reveal otherwise hidden biological differences, but focuses on a specific property that is especially difficult to assess directly: replication fidelity. Everest also proposes a reason why some reproductive traits may become more complex than fertilization, signaling, or viability alone require: added complexity can increase sensitivity to mutational disruption, strengthening selection against error-prone lineages.

Good genes, condition dependence, and genic capture

Good-genes and genic-capture models treat elaborate traits as signals of genome-wide mutation load: accumulated deleterious mutations reduce condition, and condition-dependent ornaments are therefore expected to be less well formed when load is high [Hamilton & Zuk, 1982; Rowe & Houle, 1996]. Everest is close to these models in spirit, because both frameworks allow complex sexual traits to integrate genetic effects from many loci. The key difference is architectural. In genic-capture models, many genetic and physiological inputs usually feed in parallel into the broad composite state called condition, which then affects trait expression in a graded way. Everest instead emphasizes a more sequential logic, closer to a serial circuit: a reproductive display, migration-linked breeding schedule, or molecular compatibility system may depend on a chain of required components, so failure of any critical link can compromise the whole reproductive outcome.

Everest therefore shifts the emphasis in three ways. First, it treats replication fidelity—the germline mutation rate—as a distinct target of selection, rather than subsuming all genetic quality into condition or short-term fitness. Second, it predicts that adaptation and fidelity can move in opposite directions during environmental change and recovery. Third, it extends the same logic beyond animal mate choice to cases where filtering occurs through fertilization success, gamete competition, or biochemical compatibility.

Long-term consequences of costly elaboration

A related contrast concerns the long-term fate of highly costly traits. Under Fisherian runaway or handicap interpretations, extreme ornaments and demanding behaviors are generally expected to impose net viability costs, so lineages that evolve them may be disadvantaged relative to less elaborated lineages. Everest predicts an additional possibility: if costly traits also function as effective fidelity screens, they may improve long-term persistence by purging mild mutators more efficiently. Lineage sorting may therefore favor some elaborated reproductive systems despite their short-term costs, particularly when they help restore replication fidelity after episodes of environmental change.

Red Queen and fluctuating selection

Red Queen models emphasize ongoing coevolution, often with parasites, as a driver of sex and recombination [Hamilton et al., 1990]. Everest is compatible with the view that fluctuating selection is important but identifies a different quantity at risk: not simply adaptive variation against changing antagonists, but replication fidelity itself. In Everest, environmental change can favor mutator backgrounds during adaptation, while recombination and reproductive screening help restore fidelity after adaptation has been achieved.

Mutation rate evolution

A separate body of literature focuses on mutation-rate evolution and on the population-genetic mechanisms by which mutator alleles spread or are removed. In this literature, selection against mutators is usually indirect: mutator alleles are removed because they remain linked to the deleterious mutations they generate. Everest accepts this Kimura-style linkage logic but proposes a mechanism that can strengthen it. Complex reproductive screens can stack many mutation-sensitive

components into a shared mating or fertilization test, so that small differences in replication fidelity generate larger differences in reproductive success.

Roberts and Petrie extend this mutation-rate logic to sexual selection, showing that mate choice can transiently support higher mutation rates when beneficial mutations are available [Roberts & Petrie, 2022]. Everest accepts this accelerator logic but adds the corresponding brake: once adaptation has been achieved, demanding reproductive screens can expose the damage generated by elevated mutation rates, allowing recombination and mate choice to restore high fidelity without erasing adaptive gains.

Mutation load and classic benefits of recombination

Kondrashov's "hatchet"

Kondrashov's "hatchet" emphasizes the purging of deleterious mutation load under strong selection, especially when deleterious mutations interact synergistically [Kondrashov, 1988; Kondrashov, 1993]. Everest shares the concern with mutation load but focuses on a different target: not only the load already present, but the mutation rate that generates it. Complex reproductive screens can make small differences in mutation rate visible by translating them into larger differences in mating or fertilization success.

Hill–Robertson interference

Hill–Robertson interference provides a classic rationale for recombination: linkage among selected loci can reduce the efficiency of selection, and recombination can relieve this interference [Hill & Robertson, 1966]. Everest is compatible with this logic but emphasizes a different step in the process. Recombination can separate adaptive alleles from mutator backgrounds, but selection must still favor the offspring that inherit both adaptation and high fidelity. Everest proposes that complex reproductive screens help do this by making small differences in replication fidelity visible through mating or fertilization success.

Muller's ratchet and lineage-level decay

Muller's ratchet emphasizes the irreversible accumulation of deleterious mutations in finite asexual populations and the lineage-level consequences of mutational decay [Muller, 1964; Felsenstein, 1974]. Everest shares the broader concern with long-term genomic integrity but applies it to sexual lineages facing environmental change. In Everest, mutator backgrounds may rise during adaptation, but later threaten persistence unless high fidelity is restored. Mate choice and fertilization filters provide one route for this restoration: they can expose mutational disruption in complex traits and reduce the reproductive contribution of low-fidelity lineages.

This logic also builds on classic work on recombination and mutator alleles. Muller emphasized that recombination can bring together advantageous mutations arising in different individuals [Muller, 1932], while Kimura showed that mutator alleles can be indirectly selected against through linkage to the deleterious mutations they generate [Kimura, 1967]. Everest accepts both points but adds that reproductive screening can amplify this indirect selection by converting small differences in replication fidelity into larger differences in mating or fertilization success.

Genomic integrity and sex

Another line of work treats sex as a maintainer of genomic integrity. Hörandl and others argue that regular sex, recombination, and segregation can restore higher-quality genomes and prevent mutational decay, particularly in complex eukaryotes [Hörandl, 2009; Hörandl & Hadacek, 2013]. Everest aligns with this broad view but adds a specific filtering mechanism: mate choice, fertilization success, and related reproductive screens can expose mutational disruption in complex traits, thereby favoring lineages that maintain higher replication fidelity.

Drift-barrier theory

Drift-barrier theory addresses a different question: not how mutators are selected against, but why mutation rates do not evolve indefinitely toward higher fidelity. Selection can reduce mutation rates only while the incremental benefit of improved fidelity is large enough to overcome drift [Lynch et al., 2016]. Everest adopts this logic but proposes a mechanism that can shift the drift-barrier equilibrium toward higher fidelity. By stacking many mutation-sensitive components into shared reproductive screens and routing their effects through steep mating or fertilization success functions, reproductive complexity can amplify small fidelity differences into larger differences in reproductive success. In this sense, Everest does not replace drift-barrier theory; it proposes one way in which sexual lineages may strengthen selection for fidelity and push mutation rates closer to the lower bound set by molecular constraints.

The hypothesis integrates three ideas that are often treated separately: (i) replication fidelity as an independent dimension of variation, distinct from short-term fitness; (ii) the possibility that adaptation and fidelity can move in opposite directions during environmental change; and (iii) reproductive "excess" complexity as an integrity-sensitive screen that amplifies small fidelity differences by stacking many mutation-sensitive components into shared reproductive tests. For a concise summary of how conventional frameworks relate to Everest, and of the main diagnostic tests, see Table 1 in the main text.

Appendix S4. Adaptation–fidelity tradeoffs beyond sexual organisms

The adaptation–fidelity tradeoff emphasized by Everest should not be unique to sexually reproducing eukaryotes. Any replicating lineage must both adapt to changing conditions and maintain a functional genome, so mechanisms that transiently increase variation and later restore fidelity should be expected across the tree of life. Bacteria and viruses offer partial analogies: stress-induced mutagenesis, conjugation, recombination, and reassortment can all alter the balance between adaptation and fidelity. Conjugation systems such as the F factor in *E. coli* are especially suggestive, because they involve multiple plasmid-borne and host-dependent functions whose complexity may exceed the minimum required for DNA transfer. Crucially, these plasmids are replicated by host replication machinery, so an error-prone host background can generate disruptive mutations in the very genes required for plasmid transfer or maintenance. Like Everest traits, such complexity could increase the mutational target and make successful exchange more sensitive to replication errors, thereby biasing transfer toward higher-fidelity backgrounds. These are not Everest mechanisms in the strict sense, because they lack the regular, multigenic mate-choice screens emphasized here. They nevertheless support the broader premise that adaptation and fidelity can conflict, and that successful lineages require mechanisms for restoring genomic integrity after episodes of rapid change.

Appendix S5. Pregnancy loss as a possible distributed fidelity-sensitive screen

Mammalian reproduction provides a speculative example of how reproductive complexity could couple replication fidelity to reproductive success without any dedicated mechanism for detecting mutation rate. Successful mammalian pregnancy depends on a complex fetal–maternal interface involving implantation, trophoblast differentiation and invasion, decidual remodeling, immune tolerance, and early placentation; disruption of these processes is associated with implantation failure, miscarriage, and other pregnancy complications [Huang et al., 2023]. Everest would predict that, if elevated mutation rates increase the number of damaging variants in genes or regulatory elements involved in these processes, low-fidelity genotypes could experience increased reproductive loss due to developmental and placental failure. This argument is distinct from the well-established role of chromosomal abnormalities in early miscarriage, which are not themselves evidence of inherited mutator genotypes. The point is only that complex reproductive interfaces can

act as distributed screens: they need not measure mutation rates directly to impose reproductive penalties on genotypes that generate a greater burden of damaging variants.

References

Felsenstein J. The evolutionary advantage of recombination. Genetics. 1974 Oct 1;78(2):737-56. https://doi.org/10.1093/genetics/78.2.737

Fisher, RA. *The Genetical Theory of Natural Selection.* Clarendon Press, Oxford; 1930.

Hamilton WD, Zuk M. Heritable true fitness and bright birds: a role for parasites? Science. 1982 Oct 22;218(4570):384-7. https://doi.org/10.1126/science.7123238

Hamilton WD, Axelrod R, Tanese R. Sexual reproduction as an adaptation to resist parasites (a review). Proceedings of the National Academy of Sciences. 1990 May;87(9):3566-73. https://doi.org/10.1073/pnas.87.9.3566

Hill WG, Robertson A. The effect of linkage on limits to artificial selection. Genetics Research. 1966 Dec;8(3):269-94. https://doi.org/10.1017/S0016672300010156

Hörandl E. A combinational theory for maintenance of sex. Heredity. 2009 Dec;103(6):445-57. https://doi.org/10.1038/hdy.2009.85

Hörandl E, Hadacek F. The oxidative damage initiation hypothesis for meiosis. Plant reproduction. 2013 Dec;26(4):351-67. https://doi.org/10.1007/s00497-013-0234-7

Kimura M. On the evolutionary adjustment of spontaneous mutation rates. Genetics Research. 1967 Feb;9(1):23-34. https://doi.org/10.1017/S0016672300010284

Kondrashov AS. Deleterious mutations and the evolution of sexual reproduction. Nature. 1988 Dec 1;336(6198):435-40. https://doi.org/10.1038/336435a0

Kondrashov AS. Classification of hypotheses on the advantage of amphimixis. Journal of Heredity. 1993 Sep 1;84(5):372-87. https://doi.org/10.1093/oxfordjournals.jhered.a111358

Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, Foster PL. "Genetic drift, selection and the evolution of the mutation rate." *Nature Reviews Genetics.* 2016 Nov;17(11):704-14. https://doi.org/10.1038/nrg.2016.104

Lynch M, Ali F, Lin T, Wang Y, Ni J, Long H. The divergence of mutation rates and spectra across the Tree of Life. EMBO Reports. 2023 Oct 9;24(10):e57561. https://doi.org/10.15252/embr.202357561

Muller HJ. Some genetic aspects of sex. The American Naturalist. 1932 Mar 1;66(703):118-38. https://doi.org/10.1086/280418

Müller HJ. The relation of recombination to mutational advance. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1964 May 1;1(1):2-9. https://doi.org/10.1016/0027-5107(64)90047-8

Roberts G, Petrie M. "Sexual selection for males with beneficial mutations." *Scientific Reports*. 2022 Jul 23;12(1):12613. https://doi.org/10.1038/s41598-022-16002-y

Rowe L, Houle D. The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 1996 Oct 22;263(1375):1415-21. https://doi.org/10.1098/rspb.1996.0207

Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature.* 1997 Jun 12;387(6634):703–5. https://doi.org/10.1038/42701

Zahavi A. Mate selection—a selection for a handicap. *Journal of Theoretical Biology.* 1975 Sep 1;53(1):205-14. https://doi.org/10.1016/0022-5193(75)90111-3