Supplementary Information

1 Optimized driving rate constant

Our study establishes a relationship between the error correction efficiency of an enzyme (polymerase) and environmental stress. We quantify the ineffectiveness of the error correction process due to sudden environmental perturbations, and the eventual adaptation of the enzyme to the new environment. To regain high accuracy after a perturbation has occurred, enzymes evolve to adapt the newly required optimized driving rate constant, M_0, in the perturbed environment.

The optimized driving rate constant M_0 can be obtained by minimizing the error fraction, equation (3), w.r.t the driving rate constant, using

$$\frac{df}{dm} = 0$$

.

Assuming $K'_X = K'_Y$, $L'_X = L'_Y$ and $w = 0$, we get

$$M_0 = \frac{\{K_X \cdot K_Y \cdot L'_X \cdot (K'_X + L'_X)\}^{1/2}}{(K'_X + L'_X)} = \left\{ K_X \cdot K_Y \cdot \Big(\frac{L'_X}{K'_X + L'_X} \Big) \right\}^{1/2} = \left\{ K_X \cdot K_Y \cdot \Big(\frac{\frac{L'_X}{K'_X}}{1 + \frac{L'_X}{K'_X}} \Big) \right\}^{1/2}$$

$$\implies M_0 = \Big(\frac{q-1}{q} K_x K_y \Big)^{1/2}$$

Or,

$$M_0 = K_x \Big(\frac{q-1}{q\alpha} \Big)^{1/2}.$$

where,

$$q = 1 + \frac{L'_x}{K'_x} = 1 + \frac{L'_y}{K'_y}.$$

2 Calculation of minimum error fraction

$$f = \Big\{ \frac{m \cdot K'_Y + (m + K_Y) \cdot L'_Y}{m \cdot K'_X + (m + K_X) \cdot L'_X} \Big\} \cdot \frac{(K_X + m)(L_X + w)}{(K_Y + m)(L_Y + w)}$$

Applying the approximations $W \to 0$ and $K'_X = K'_Y$,

$$f = \Big\{ \frac{m + (m + K_Y) \cdot \frac{L'_Y}{K'_Y}}{m + (m + K_X) \cdot \frac{L'_X}{K'_X}} \Big\} \cdot \frac{(K_X + m) \cdot L_X}{(K_Y + m) \cdot L_Y}$$

$$\implies f = \Big\{ \frac{m + (m + K_Y) \cdot (q-1)}{m + (m + K_X) \cdot (q-1)} \Big\} \cdot \frac{(K_X + m) \cdot L_X}{(K_Y + m) \cdot L_Y}$$

$$\implies f = \Big\{ \frac{m + q \cdot m + q \cdot K_Y - m - K_Y}{m + q \cdot m + q \cdot K_X - m - K_X} \Big\} \cdot \frac{(K_X + m) \cdot L_X}{(K_Y + m) \cdot L_Y}$$

$$\implies f = \Big\{\frac{q\cdot m + (q-1)\cdot K_Y}{q\cdot m + (q-1)\cdot K_X}\Big\}\cdot\frac{(K_X+m)\cdot L_X}{(K_Y+m)\cdot L_Y}$$

$$\implies f = \Big\{\frac{m + \frac{(q-1)}{q}\cdot K_Y}{m + \frac{(q-1)}{q}\cdot K_X}\Big\}\cdot\frac{(K_X+m)\cdot L_X}{(K_Y+m)\cdot L_Y}.$$

Implementing the optimized driving rate equation, equation (17), in the above equation,

$$f_{min} = \Big(\frac{M_0 + \frac{M_0^2}{K_X}}{M_0 + \frac{M_0^2}{K_Y}}\Big)\cdot\frac{(K_X+M_0)\cdot L_X}{(K_Y+M_0)\cdot L_Y}$$

$$\implies f_{min} = \frac{K_Y}{K_X}\cdot\Big(\frac{M_0+K_X}{M_0+K_Y}\Big)^2\cdot\frac{L_X}{L_Y}$$

$$\implies f_{min} = \frac{K_Y}{K_X}\cdot\Big(\frac{K_X}{K_Y}\Big)^2\Big(\frac{1+\frac{M_0}{K_X}}{1+\frac{M_0}{K_Y}}\Big)^2\cdot\frac{L_X}{L_Y}$$

M_0 is the optimized driving rate constant that balances the discriminatory substrate flux between the two discriminating paths, step 1 and step 3, resulting in error minimization. This indicates the driving rate constant M_0 should be less than the $K_{X/Y}$ for effective discrimination through step 1. Otherwise, $[EY^*]$ will increase, promoting an increase in the rate of incorrect product formation, which increases the error fraction. Therefore, we assume $M_0 \lesssim K_X$ and $M_0 < K_Y$. This leads to

$$\Big(\frac{1+\frac{M_0}{K_X}}{1+\frac{M_0}{K_Y}}\Big)^2 \to 1.$$

Now, f_{min} will be,

$$f_{min} \approx \frac{K_X}{K_Y}\cdot\frac{L_X}{L_Y} = \alpha^2. \tag{1}$$

3 Quantification of the increased error rate due to perturbation

Values of the optimized driving rate constant and the error rate in the prevailing environment can be calculated using 'driving_function.m'and 'error_calculator.m':

- Driving rate constant = driving_function(f, K_x)
- Error rate = error_calculator(m, K_x)

Increased error rate due to the perturbation can be quantified with following algorithm (sub-script old and new represent the prevailing environment and perturbed environment).

- Optimized driving rate constant ($M_{0,old}$) = driving_function(f_{min}, $K_{x,old}$)
- Increased error rate (f') = error_calculator($M_{0,old}$, $K_{x,new}$)
- Increased amount = $\frac{f'}{f_{min}}$

The MATLAB functions 'driving_function.m' and 'error_calculator.m', which are used to compute the increment in the error rate, are available at:
`https://github.com/Sashik05/kinetic-proofreading-mutation-robustness.git`.

Algorithm 1: Simulation of trait evolution of the population with fixed coding region length under kinetic proofreading–based adaptation dynamics

Input: Initial population size $N_0 = 1000$; carrying capacity $N_{\text{cap}} = 2N_0$;
Proofreading constants $k_{\text{old}} = 42.0098$, $k_{\text{new}} = 1$;
Error threshold $f_{\text{th}} = 10^{-6}$; initial trait variance $\sigma = 10^{-3}$;
Genome length gen_len $= 10^6$ (bp); generations $T_{\max} = 2 \times 10^3$; replicates $R = 16$;
Other parameters: finite-difference $\epsilon = 10^{-8}$, logistic steepness $a = 5$.

Output: Per-generation time series (averaged over replicates) of mean trait $\bar{m}_t$, trait variance $\text{Var}(m_t)$, selection gradient β_t, observed adaptation rate $\Delta\bar{m}_t$, predicted rate $\text{Var}(m_t)\beta_t$, and final population distributions.

Precompute and initialization:

Compute allowable trait values and initial optimum: $m_{\text{values}} \leftarrow$ `driving_function`$(f_{\text{th}}, k_{\text{old}})$

$m_0 \leftarrow$ `err_fraction`(k_{old}); $m_0^{\text{new}} \leftarrow$ `err_fraction`(k_{new}); $m_{\text{opt}} \leftarrow m_0^{\text{new}}$

Preallocate storage arrays for each replicate: $\bar{m}_{r,t}$, $\text{Var}_r(m_t)$, $N_{r,t}$, $\beta_{r,t}$, and `FinalPop`

for $r \leftarrow 1$ **to** R ***in parallel*** **do**

  Initialize population traits: $M_{\text{pop}} \leftarrow m_0 + \sigma \cdot \mathcal{N}(0,1)^{N_0}$

  Initialize per-replicate local storage (vectors of length $T_{\max}$)

  for $t \leftarrow 1$ **to** $T_{\max}$ **do**

    if *M_{pop} is empty* **then**

      record NaNs and **continue**

    $N \leftarrow |M_{\text{pop}}|$

    $\bar{m} \leftarrow \text{mean}(M_{\text{pop}})$

    `// Environmental switch for error calculation`

    if $t \leq 100$ **then**

      $f_{\text{bar}} \leftarrow$ `error_calculator`$(\bar{m}, k_{\text{old}})$

    else

      $f_{\text{bar}} \leftarrow$ `error_calculator`$(\bar{m}, k_{\text{new}})$

    $\Delta m \leftarrow \text{gen\_len} \times f_{\text{bar}}$

    `// Mutation step size`

    store $\bar{m}$ in $\bar{m}_{r,t}$; store $\text{var}(M_{\text{pop}})$ in $\text{Var}_r(m_t)$; store N in $N_{r,t}$

    `// --- Compute theory-based selection gradient` β `---`

    $f_+ \leftarrow$ `error_calculator`$(\bar{m} + \epsilon, k_{\text{new}})$

    $f_- \leftarrow$ `error_calculator`$(\bar{m} - \epsilon, k_{\text{new}})$

    $\dfrac{df}{dm} \leftarrow (f_+ - f_-)/(2\epsilon)$

    $p_{\text{kill,bar}} \leftarrow f_{\text{bar}}/f_{\text{th}}$; $S_{\text{bar}} \leftarrow 1 - p_{\text{kill,bar}}$

    $p_{\text{off,bar}} \leftarrow \dfrac{\exp(aS_{\text{bar}})}{1 + \exp(aS_{\text{bar}})}$

    $W_{\text{bar}} \leftarrow S_{\text{bar}} \cdot p_{\text{off,bar}}$

    $\dfrac{d\ln W}{df} \leftarrow -\dfrac{1}{f_{\text{th}}(1 - f_{\text{bar}}/f_{\text{th}})} - \dfrac{a}{f_{\text{th}}}(1 - p_{\text{off,bar}})$

    $\beta \leftarrow \dfrac{d\ln W}{df} \cdot \dfrac{df}{dm}$

    store β in $\beta_{r,t}$

Algorithm 2: Birth–death and mutation dynamics per generation (Cont.)

```
// --- Birth{death and mutation for all individuals ---
Initialize empty offspring_list and death_mask
foreach i = 1, ..., N do
    m_i ← M_pop(i)
    if t ≤ 100 then
        f_i ← error_calculator(m_i, k_old)
    else
        f_i ← error_calculator(m_i, k_new)
    p_kill ← f_i / f_th
    if p_kill ≥ rand() or f_i < 0 then
        mark individual i for death; continue
    m_1 ← m_i − Δm;   m_2 ← m_i;   m_3 ← m_i + Δm
    if any of m_1, m_2, m_3 ≤ 0 then
        skip reproduction for this individual; mark for death; continue
    s_i ← 1 − p_kill;   p_off ← exp(a s_i) / (1 + exp(a s_i))
    if (p_off − 0.25) > rand() then
        n_off ← 2
    else
        n_off ← 1

    r_pick ← rand()
    if n_off == 1 then
        if r_pick < 1/3 then
            append m_1 to offspring_list
        else if r_pick < 2/3 then
            append m_2 to offspring_list
        else
            append m_3 to offspring_list
    else
        if r_pick < 1/3 then
            append [m_1; m_3] to offspring_list
        else if r_pick < 2/3 then
            append [m_1; m_2] to offspring_list
        else
            append [m_2; m_3] to offspring_list
    mark parent i for death (parents die after reproduction)
Remove individuals in death_mask from M_pop
If offspring_list not empty, append offspring to M_pop
if |M_pop| > N_cap then
    randomly sample N_cap individuals from M_pop
Store replicate results: m̄_{r,·}, Var_r(m·), N_{r,·}, β_{r,·}, and final M_pop
```

Algorithm 3: Postprocessing and plotting

Postprocessing:

Compute replicate-averaged mean trait:

$$\langle \bar{m}_t \rangle = \frac{1}{R} \sum_{r=1}^{R} \bar{m}_{r,t}$$

Compute variance: $\langle \mathrm{Var}(m_t) \rangle = \frac{1}{R} \sum_r \mathrm{Var}_r(m_t)$

Compute mean population size: $\langle N_t \rangle = \frac{1}{R} \sum_r N_{r,t}$

Estimate 99.9% confidence intervals (using $z = 3.291$)

Compute observed adaptation rate $\Delta \bar{m}_t = \bar{m}_t - \bar{m}_{t-1}$

Compute predicted rate (Fisher–Lande): $\mathrm{Rate}_{\mathrm{pred},t} = \langle \mathrm{Var}(m_t) \rangle \cdot \langle \beta_t \rangle$

Output:

Plots of: (i) population size vs generation; (ii) mean trait trajectory with CI and m_{opt}; (iii) raw adaptation rate $|\Delta \bar{m}_t|$; (iv) observed vs predicted rate; (v) histogram of final trait distribution.

Return: Averaged trajectories, CI, predicted rates, and final populations.

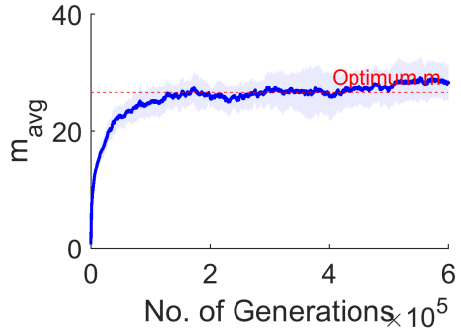
4 Enzyme evolution considering identical phenotypic change for entire population and individual phenotypic change

Evolutionary trajectory of m_{avg} of the population, by assuming that the change in driving rate constant (phenotypic change) in a generation to be uniform across the whole population is illustrated in Fig. S1a.

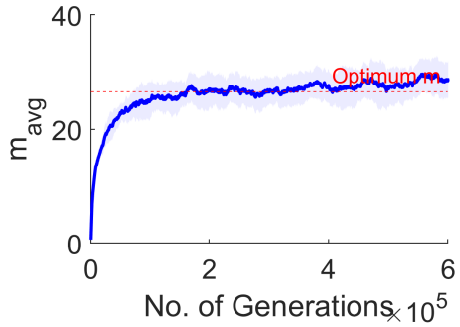
$$\Delta m = \frac{l \cdot \langle f \rangle}{\gamma}$$

Similarly, Fig. S1b demonstrates the m_{avg} evolution across generations, when the phenotypic change (driving rate constant) is different for each individual depending on their individual replication error rate.

$$\Delta m_i = \frac{l \cdot f_i}{\gamma}$$



(a) Phenotypic change is uniform across the population



(b) Phenotypic change varies for each individual

Supplementary Figure S1: *This figure illustrates the adaptive dynamics of the driving rate constant (evolving phenotype) in the perturbed environment. (a) Phenotypic change due to genomic variation is assumed to be uniform across the population. (b) Phenotypic change is different for each individual in a population across generations. However, the adaptive dynamics in both cases are similar.*

5 Comparison of standard adaptation rate with the observed adaptation rate

To quantitatively evaluate the consistency of our model, we define the following adaptation metrics. Here, we define the effective fitness per generation (W) of the evolving trait m. Effective fitness integrates the population's survival probability with its fecundity, thereby capturing both persistence and reproductive output. Operationally, W represents the expected number of individuals, including newly mutated variants, those contribute to the next generation. This measure thus provides a consistent quantitative basis for tracking the course of adaptation. This formulation follows established models of evolutionary prediction (Fisher's fundamental theorem and Lande's quantitative genetic framework [1, 2]).

$$W = SP_{\text{off}} \tag{2}$$

Here, S and P_{off} denote the mean selection coefficient and the mean offspring probability of the population in a given generation, respectively.

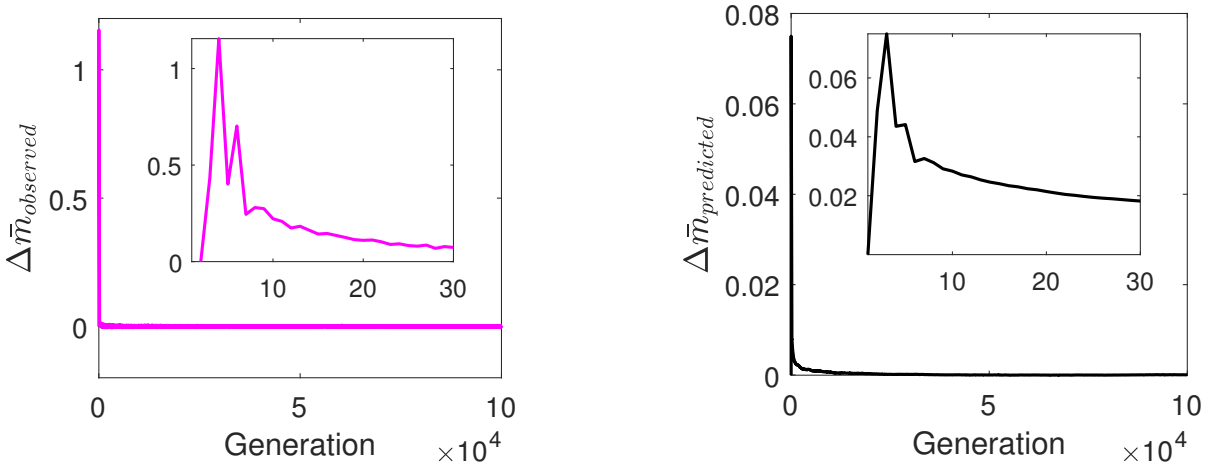
To further quantify evolutionary dynamics and to track the propagation of effective fitness across generations, we define two measures of adaptation rates: the standard and the observed. The standard adaptation rate follows the Fisher/Lande framework, where the expected change in the mean trait ($\Delta\bar{m}_{\text{standard}}$) is approximated by the product of trait variance, (Var(m)), and the selection gradient, β, which measures how fitness changes with the trait [1, 2, 3].

$$\Delta\bar{m}_{predicted} = \text{Var}(m)\beta, \qquad \text{where } \beta = \frac{\partial \log(\mathbf{W})}{\partial m} = \frac{\partial \log(\mathbf{W})}{\partial f} \cdot \frac{\partial f}{\partial m} \tag{3}$$

Conversely, the observed adaptation rate ($\Delta\bar{m}_{\text{observed}}$) reflects the change in the population mean trait across successive generations:

$$\Delta\bar{m}_{observed} = \Delta\bar{m} \tag{4}$$

Here, Var(m) denotes the trait variance, and $\Delta\bar{m}$ represents the generational change in the population mean trait value.



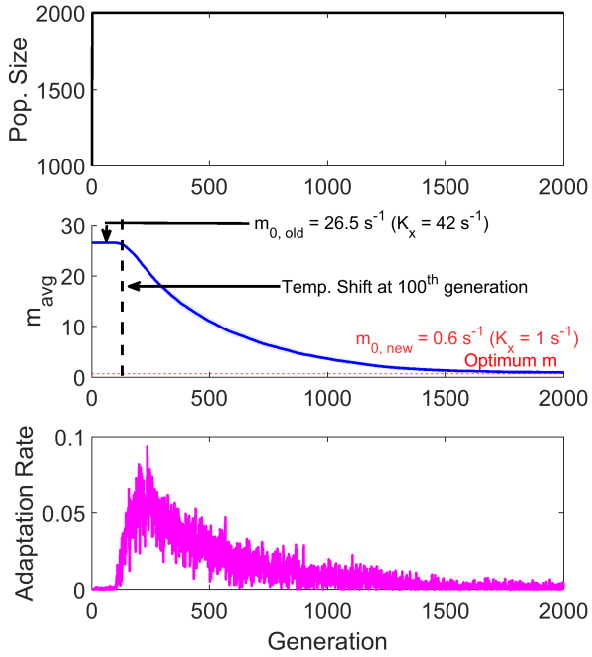
(a) (b)

Supplementary Figure S2: *Evolution of the adaptation rate across generations. (a) The observed adaptation rate ($\Delta\bar{m}$observed, magenta curve) and (b) the predicted adaptation rate ($\Delta\bar{m}$predicted, black curve) exhibit closely aligned trajectories, validating the consistency of the model. In both cases, the adaptation rate rises sharply in the early generations, reflecting the sudden increase in replication error following the environmental (temperature) shift. As the population mean trait approaches the new optimum, the adaptation rate progressively declines, driven by the reduced mutation rate at steady state.*

As shown in Fig. S2, both adaptation rate measures follow the same evolutionary trend: an initial sharp rise reflecting the sudden error increase imposed by the temperature shift, followed by a gradual decline as mutations become less frequent and the population mean converges toward the new optimum ($m_{0,\text{new}}$). This consistent pattern in adaptation rates validates the assumptions of our model.

6 Adaptation Dynamics Following Temperature Decline

Following the algorithm in '*Evolutionary Framework*', we introduced a perturbation of a temperature shift of $\Delta T = 42\ K$ from 335 K ($m_{0,old} = 26.5\ s^{-1}$, $K_x = 42\ s^{-1}$) to 293 K ($m_{0,new} = 0.6\ s^{-1}$, $K_x = 1\ s^{-1}$) on a population with initial size 10^3 and coding region size as 10^6 bp.



Supplementary Figure S3: *The figure illustrates the adaptation dynamics of a population with a coding-region size of* 10^6 *bp following a temperature shift of* $\Delta T = 42$ K, *from 335 K* ($m_{0,\text{old}} = 26.5\ \text{s}^{-1}$, $K_x = 42\ \text{s}^{-1}$) *to 293 K* ($m_{0,\text{new}} = 0.6\ \text{s}^{-1}$, $K_x = 1\ \text{s}^{-1}$). *The black curve shows population growth, the blue curve tracks adaptation toward the new optimal driving-rate constant, and the magenta curve depicts the adaptation rate across generations, representing the change in the phenotypic driving-rate constant per generation.*

7 Choice of Selection Co-efficient

To associate replication error rates with survival probability, the individual elimination probability is defined as

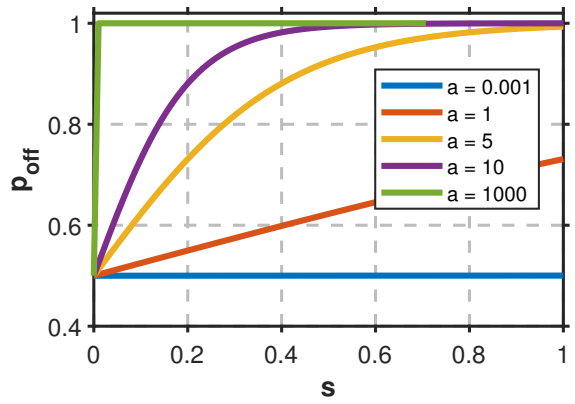
$$P_{\text{kill}} = \frac{f_i}{f_{\text{th}}}, \tag{5}$$

where f_i is the individual error rate (Eq. 3) and $f_{th} = 10^6$ is the threshold beyond which survival becomes unlikely.

The probability of generating viable offspring is expressed as

$$P_{\text{off}} = \frac{e^{as}}{1 + e^{as}}, \qquad s = 1 - P_{\text{kill}}. \tag{6}$$

Where 's' is the selection coefficient and defined as the survival probability, while 'a' dictates the steepness of the offspring probability curve, discriminating between low and high-fitness individuals.



Supplementary Figure S4: *The figure illustrates how the offspring probability (P_{off}) varies with the selection coefficient (s) for different values of the steepness parameter a (as defined in Eq. 6). When a is very small or very large (blue and green curves), P_{off} loses its ability to discriminate between high- and low-fitness individuals. Intermediate values of a, particularly around $a = 5$, provide the strongest discrimination and therefore yield a more biologically meaningful mapping between fitness and reproductive probability.*

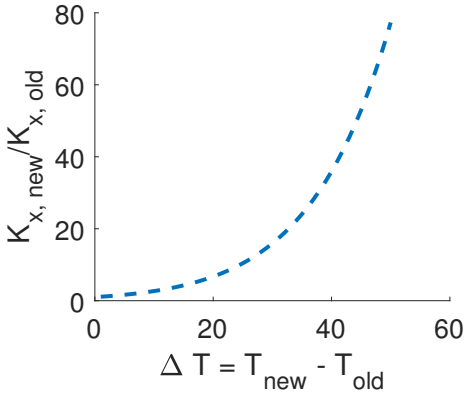
Fig. S4 illustrates how the steepness parameter a shapes the relationship between offspring probability (P_{off}) and the selection coefficient (s). When a is very small (e.g., $a = 0.001$), $P_{\text{off}} = 0.5$ for all values of s, indicating a complete loss of discriminatory power between low- and high-fitness individuals. Conversely, when a is very large, P_{off} saturates at 1 across the full range of s, again eliminating meaningful selection. In contrast, intermediate values of the steepness parameter produce a realistic fitness-to-offspring mapping. As shown in Fig. S4, $a = 5$ yields desirable behavior: as $s \to 0$, $P_{\text{off}} \to 0.5$, and as $s \to 1$, $P_{\text{off}} \to 1$. Similar qualitative behavior holds for values of a near 5 (e.g., $a = 4, 6, 7$), all of which provide adequate discrimination between individuals of differing fitness.

8 Dependence of dissociation rate constant (K_x) with temperature shift

The relationship between dissociation rate constant (K_x) with temperature follows the transition state equation:

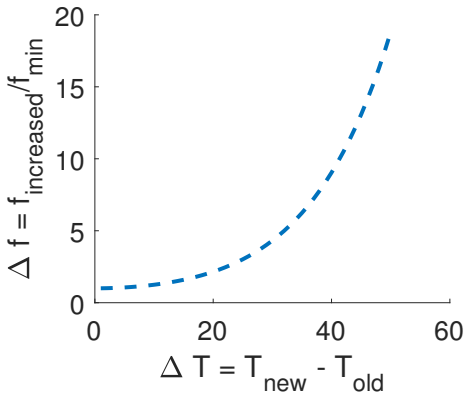
$$K_x = \frac{K_B T}{h} exp(-\frac{\Delta G^{\ddagger}}{RT})$$

Where $\Delta G^{\ddagger}$ is the free energy of activation, K_B is Boltzmann's constant ($K_B = 1.38 \times 10^{-23}\ JK^{-1}$), T is the absolute temperature, R is the universal gas constant ($R = 8.314\ Jmol^{-1}K^{-1}$) and h is Planck's constant ($h = 6.626 \times 10^{-34}\ Js$). Implementing these values and $K_x = 2s^{-1}$ at 293 K from [4, 5], we estimate $\Delta G^{\ddagger} = 70.023\ kJmol^{-1}$.



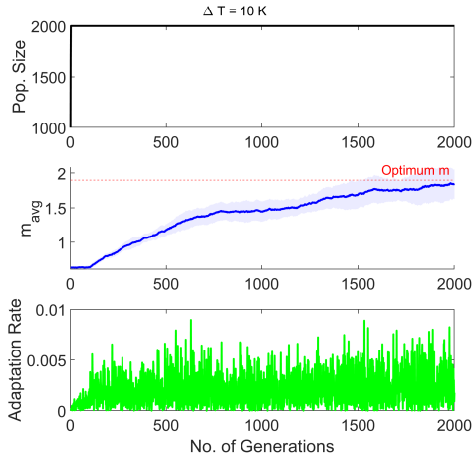
Supplementary Figure S5: *The figure illustrates how the dissociation rate constant varies with temperature. The x-axis denotes the change in temperature (ΔT), and the y-axis shows the corresponding fold-change in the dissociation rate constant.*

9 Dependence of increased error rate with temperature shift



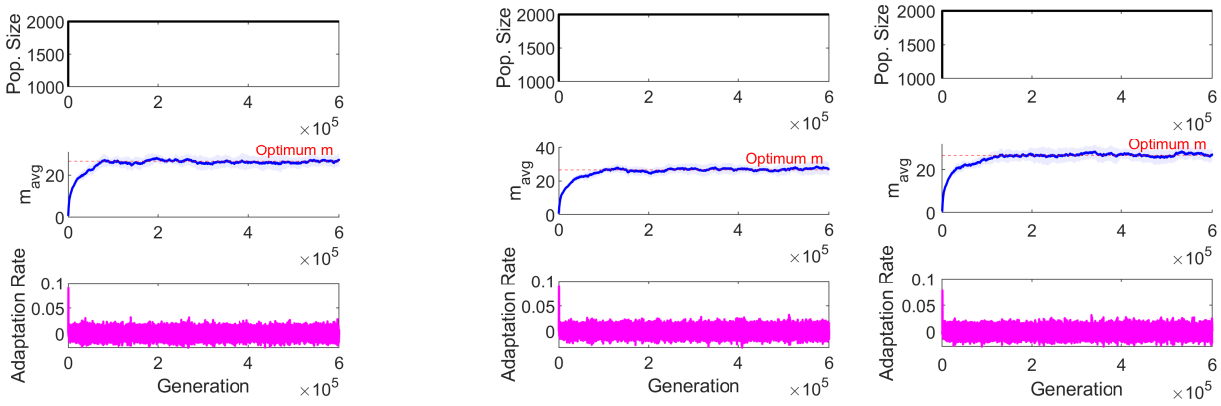
Supplementary Figure S6: *The figure illustrates the increase in replication error rate as a function of temperature shift. The x-axis denotes the magnitude of the temperature change (ΔT), and the y-axis shows the corresponding increase in error rate.*

10 Evolution under a temperature shift of 10K



Supplementary Figure S7: *The figure illustrates the evolutionary trajectory of enzyme population under a temperature shift of 10 K.*

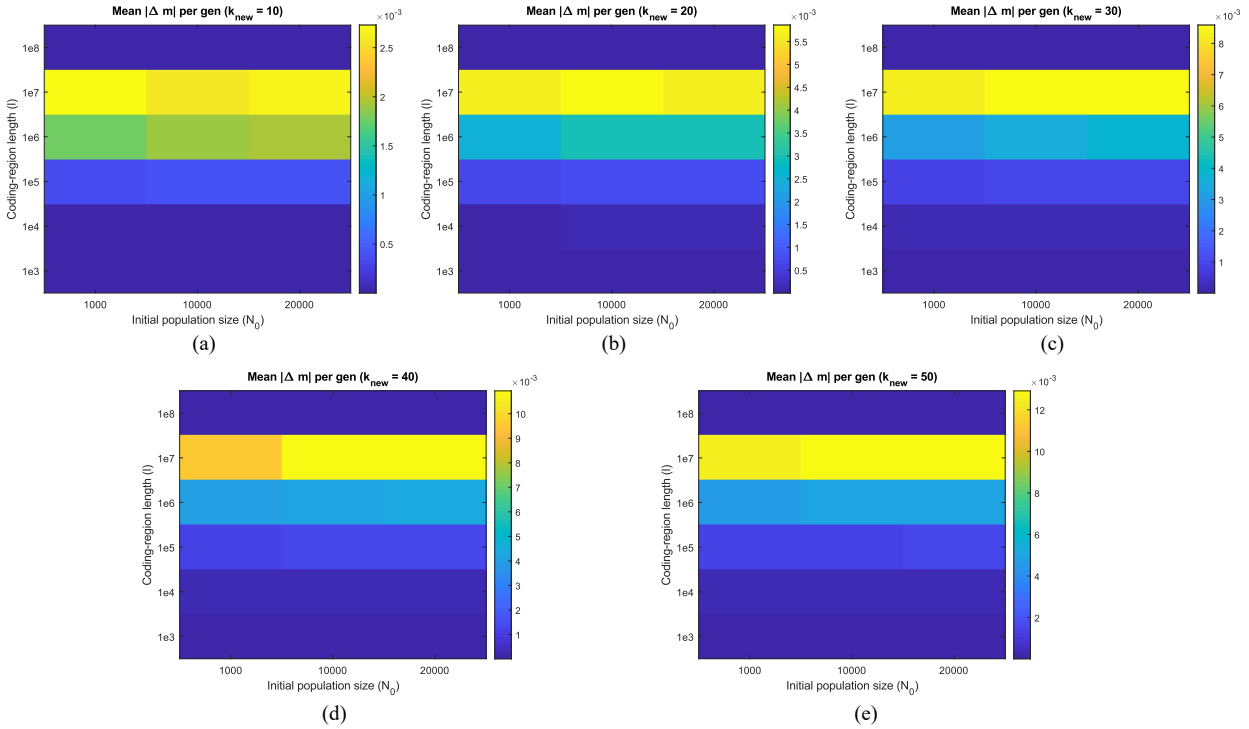
11 Evolution for different P_{off} boundaries



(a) P_{off} boundary at 0.65 (b) P_{off} boundary at 0.85 (c) P_{off} boundary at 0.95

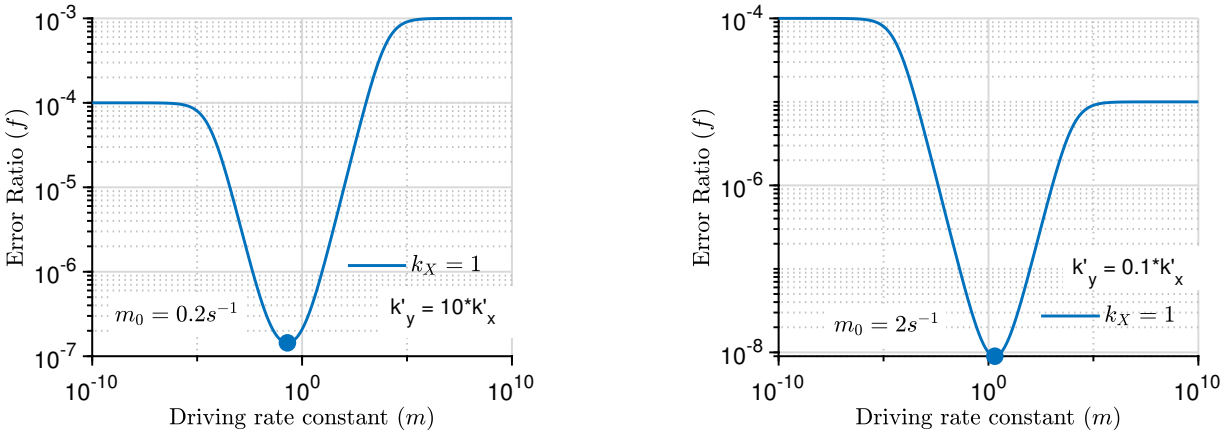
Supplementary Figure S8: *The figure illustrates the adaptive dynamics of a population with an initial size of* 10^3 *subjected to an environmental perturbation corresponding to a 42 K temperature shift, evaluated under three different stochastic offspring-probability thresholds* (P_{off})*. Although these boundaries influence the* rate *of adaptation, with higher thresholds generating more offspring and thereby enhancing selective filtering, they do not alter the qualitative outcome of the evolutionary trajectory. In (a) individuals produce two offspring when* $P_{off} \geq 0.65$*, otherwise one; in (b), the threshold is* $P_{off} \geq 0.85$*; and in (c),* $P_{off} \geq 0.95$ *triggers two-offspring production. These choices introduce stochasticity without changing the overall adaptive response.*

12 Adaptability of different populations at different temperatures



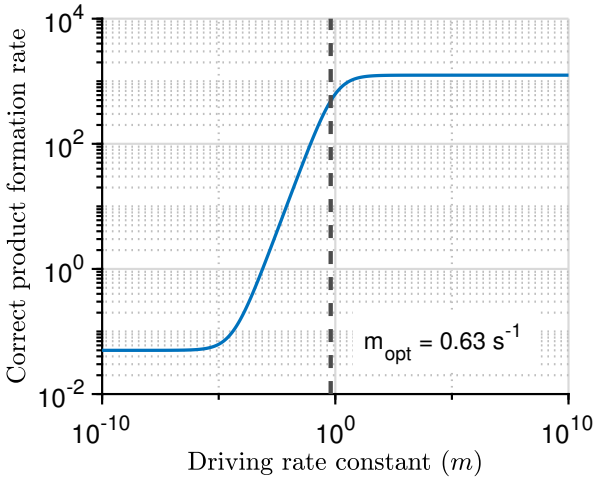
Supplementary Figure S9: *Heatmaps showing the adaptability of populations across combinations of initial population size and coding-region length under different magnitudes of environmental perturbation. Each panel displays the average per-generation change in the phenotypic driving rate constant (adaptability) for three initial population sizes* (10^3, 10^4, 2×10^4) *evaluated across coding-region lengths* 10^3, 10^4, 10^5, 10^6, 10^7, *and* 10^8 *bp. Panels correspond to increasing perturbation strength, parameterized by the perturbed dissociation rate constant* $K_{x,\mathrm{new}}$*:* ***(a)*** $K_{x,\mathrm{new}} = 10$ *at* $\Delta T = 24.5\,\mathrm{K}$*;* ***(b)*** $K_{x,\mathrm{new}} = 20$ *at* $\Delta T = 32.7\,\mathrm{K}$*;* ***(c)*** $K_{x,\mathrm{new}} = 30$ *at* $\Delta T = 37.7\,\mathrm{K}$*;* ***(d)*** $K_{x,\mathrm{new}} = 40$ *at* $\Delta T = 41.37\,\mathrm{K}$*;* ***(e)*** $K_{x,\mathrm{new}} = 50$ *at* $\Delta T = 44.27\,\mathrm{K}$*. The color bar indicates the mean adaptation rate, quantified as the average change in the driving rate constant per generation. Across all perturbation strengths, populations with coding-region lengths* 10^6 *and* 10^7 *bp consistently exhibit the highest adaptability.*

13 Variation of error rate with driving rate constant with discrimination in the association rates



Supplementary Figure S10: *The figure illustrates the variation of the error rate f with the driving rate constant m for a modest difference between the association rates (K'_x and K'_y) governing the attachment of a new dNTP to the polymerase active site. (a) f vs. m for $K'_y = 10\,K'_x$ with $K'_x = 250$ s^{-1}; the optimal driving rate constant m_o shifts from 0.6 s^{-1} to 0.2 s^{-1}. (b) f vs. m for $K'_y = 0.1\,K'_x$ with $K'_x = 250$ s^{-1}; the optimal driving rate constant m_o shifts from 0.6 s^{-1} to 2 s^{-1}. When the association rate of the incorrect substrate (Y) increases, $[EY] > [EX]$, a smaller optimal driving rate constant is required to balance the discriminatory flux between the two dissociation pathways (steps 1 and 3). Conversely, when the association rate of the incorrect substrate is lower than that of the correct one, $[EY] < [EX]$, the optimal driving rate constant increases to balance the discriminatory flux between the competing pathways.*

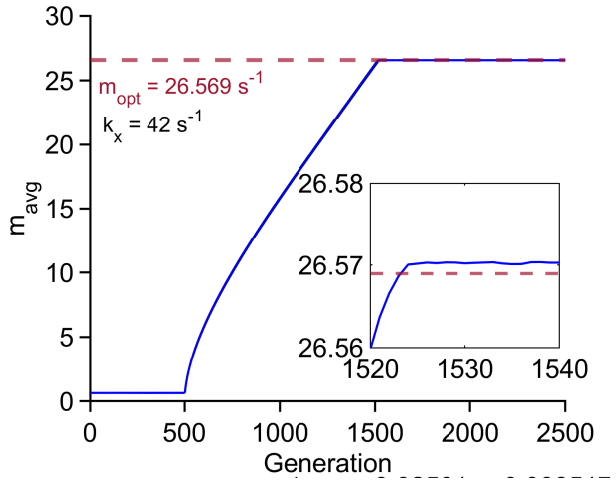
14 Correct product formation rate vs driving rate constant



Supplementary Figure S11: *The figure illustrates the variation of the correct product formation rate with the driving rate constant. For very small values of the driving rate constant, the speed remains constant. With further increase in the driving rate constant, the speed increases and attains saturation at a driving rate of* 30 s^{-1}. *The optimal driving rate constant lies near the saturation value of the speed. Here,* $K'_x = 250$ s^{-1}, $K_x = 1$ s^{-1}, $L_x = 0.2$ s^{-1}, *and* $L'_x = 10^{-2}$ s^{-1}.

15 Evolution under biased culling

The above evolutionary dynamics is modeled by implementing uniform culling without bias to maintain a finite population size. Under this neutral regulation, the final population exhibited a broader distribution of driving rate constants than the initial population (Fig. 3b), and the mean value m_{avg} oscillated around $m_{0,\text{new}}$ after reaching the adapted state in the perturbed environment (Fig. 3c). However, when biased culling was introduced, with viability as the primary and speed (correct product formation rate) as the secondary criterion, m_{avg} converged to a new value higher than $m_{0,\text{new}}$ (Fig. S12). Together with Fig. S11, this indicates that the evolved population favors higher speed while remaining within a viable error-rate range, i.e., near the minimum-error regime.



Supplementary Figure S12: *The figure illustrates the evolution of* m_{avg} *of the population following a perturbation from* $K_x = 1\ \text{s}^{-1}$ *to* $K_x = 42\ \text{s}^{-1}$*. A biased culling scheme is used to maintain a finite population size, where viability is given primary priority and the speed (correct product formation rate) secondary priority. We observe that* m_{avg} *reaches the adaptation state faster than under random culling and maintains an offset toward higher* m *values, indicating that the population favors speed while preserving viability.*

Kinetic Parameters

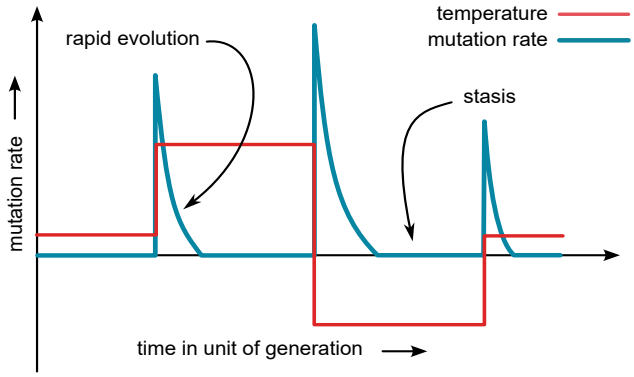
Parameter	Value	Unit
$K'_X = K'_Y$	250	s^{-1}
$L'_X = L'_Y$	10^{-2}	s^{-1}
L_X	0.2	s^{-1}
K_X	2	s^{-1}
$f_0 = \frac{K_X}{K_Y}$	10^{-4}	–
Energetic discrimination factor α	10^{-4}	–

Supplementary Table S1: Representative kinetic parameters used in the simulations.

16 Emergent Signatures of Punctuated Evolution

In the molecular evolution of enzymes involved in genomic processes such as DNA replication and protein synthesis, environmental stress leads to an increase in errors within these processes. Increased errors in one genomic process can propagate and amplify errors in others. For example, increased errors in DNA replication intensify the errors in protein synthesis, which is inherently error-prone due to its multistage and complex nature [6]. Consequently, the accumulation of such errors across genomic processes increases the mutational frequency at the organism level, causing the rapid evolution of organisms. Notably, such high genetic variations also often accompany speciation [7].

As enzymes adapt to reduce these errors, the overall mutational frequency stabilizes, allowing the organism to adjust to the perturbed environment and enter a phase of evolutionary stasis. This stasis is maintained until another abrupt environmental perturbation, such as a sudden temperature fluctuation, occurs. During such episodes, the elevated error rates in genomic processes drive rapid bursts of evolution. Thus, the interplay between rapid evolutionary change under elevated mutational frequencies and subsequent stabilization at minimal error rates through readjustment of optimal driving rate constant (see Fig. 3 and 4) reflects the signatures of punctuated equilibrium, as illustrated in Fig. S13.



Supplementary Figure S13: *Signature of punctuated equilibrium: Prolonged periods of evolutionary stasis interrupted by rapid bursts of evolutionary change due to sudden environmental perturbations. Temperature is considered here as a proxy for environmental variation, with the red line denoting the temperature changes in the environment over evolutionary timescales. Each abrupt change in temperature increases enzymatic error rates, elevating mutational frequency and driving rapid adaptation towards minimal-error enzymatic configurations. These transitions, shown by the blue curve, can result in the emergence of new traits or even speciation.*

Fig. (S13) illustrates that environmental stress, such as temperature fluctuations (red line), disrupts the periods of stasis, leading to rapid evolutionary changes in organisms. The temperature shift increases the error rate in genomic processes, thereby elevating the mutation rate (blue line). Because of this sudden increase in mutations, the organism evolves rapidly to adapt to the new environment. When adaptation lowers error rates, mutation rate levels off, and the system goes back to stasis until the next perturbation in the environment. This dynamic is similar to the adaptation-rate framework we talked about earlier, shown in Fig. 4. Rapid adaptation is shown by the steep rise in the adaptation rate curve, while stasis is shown by the slow decline as the population's mean trait gets closer to the new optimum.

References

[1] Ronald Aylmer Fisher. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press, 1999.

[2] Russell Lande. "Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry". In: *Evolution* (1979), pp. 402–416.

[3] Russell Lande and Stevan J Arnold. "The measurement of selection on correlated characters". In: *Evolution* (1983), pp. 1210–1226.

[4] Kinshuk Banerjee, Anatoly B Kolomeisky, and Oleg A Igoshin. "Elucidating interplay of speed and accuracy in biological error correction". In: *Proceedings of the National Academy of Sciences* 114.20 (2017), pp. 5183–5188.

[5] Isaac Wong, Smita S Patel, and Kenneth A Johnson. "An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics". In: *Biochemistry* 30.2 (1991), pp. 526–537.

[6] D Allan Drummond and Claus O Wilke. "The evolutionary consequences of erroneous protein synthesis". In: *Nature Reviews Genetics* 10.10 (2009), pp. 715–724.

[7] Stephen Jay Gould and Niles Eldredge. "Punctuated equilibria: the tempo and mode of evolution reconsidered". In: *Paleobiology* 3.2 (1977), pp. 115–151.