Supplementary Material - Evolutionary dynamics of an epigenetic switch in a fluctuating environment

Mariana Gómez-Schiavon[1,2,3] and Nicolas E. Buchler[2,3,4]

[1] *Program in Computational Biology & Bioinformatics, Duke University, Durham, NC 27708*

[2] *Center for Genomic & Computational Biology, Duke University, Durham, NC 27710*

[3] *Department of Biology, Duke University, Durham, NC 27708*

[4] *Department of Physics, Duke University, Durham, NC 27708*

(Dated: August 29, 2016)

I. SUPPORTING FIGURES

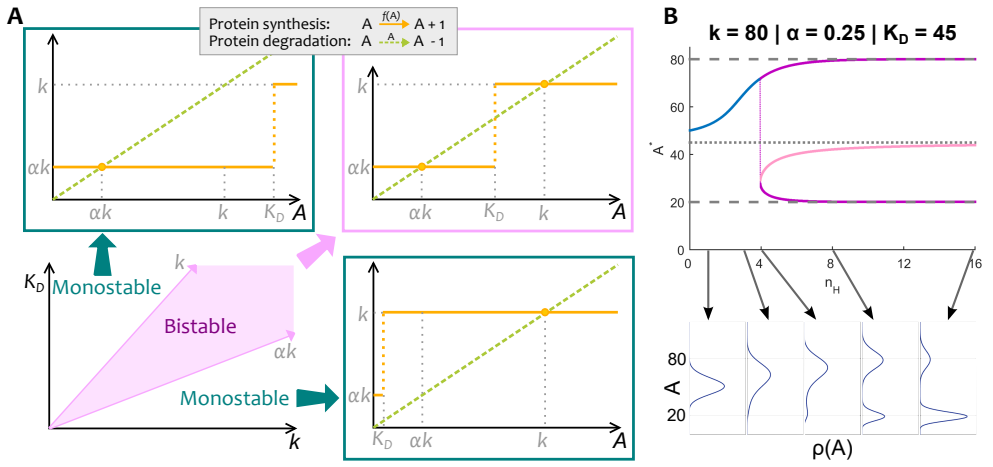

FIG. S1. **Biophysical parameters, deterministic steady state solutions and stochastic stationary distributions of protein levels.** (A) The effect of the maximum synthesis rate (k) and the affinity constant (K_D) over the deterministic steady state solutions of the protein expression (i.e. $\frac{dA^*}{d\tau} = f(A^*) - A^* = 0 \Leftrightarrow f(A^*) = A^*$ where $f(A) = k(\alpha + (1-\alpha)\frac{A^{n_H}}{A^{n_H}+K_D^{n_H}})$) in the limit of high Hill coefficients ($n_H \to \infty$). If $K_D < \alpha k$ the system is monostable HIGH with the protein expression steady state (A^*) equal to k; on the other hand, if $K_D > k$ then the system is monostable LOW with $A^* = \alpha k$. When $\alpha k \leq K_D \leq k$ is intermediate, these two steady states coexist and the system is bistable. (B) Bifurcation diagram of the protein steady states as the Hill coefficient (n_H) varies while keeping the rest of the biophysical parameters fixed. As n_H value increases, the system goes from monostable (blue dots) to bistable (violet and pink dots). As $n_H \to \infty$, the stable steady states monotonically approach their limiting values, αk and k (dashed gray lines), and the unstable steady state asymptotically approaches K_D (dotted gray line). We show a few examples of the stationary distribution of the protein expression for stochastic simulations with intrinsic biochemical noise (bottom). As n_H approaches the bifurcation point (where the system passes from being monostable to bistable) the stationary distribution becomes wider (i.e. the phenotype is more variable). In the bistable region, even if the two modes of the stationary distribution do not change much, their relative weights can be significantly affected by the value of the unstable steady state, as stochastic transitions from one stable mode to the other become more or less probable.

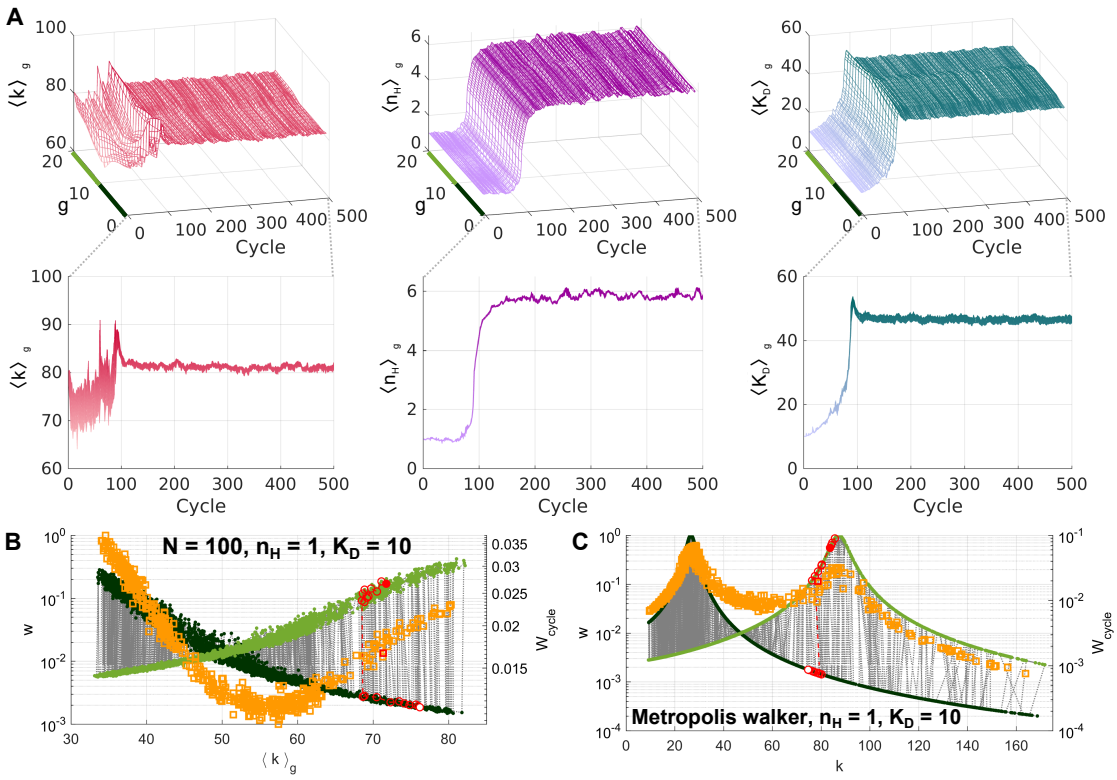

FIG. S2. **Counteracting selection pressure destabilizes the "no-response" genotype in a fluctuating environment.** (A) The average population genotype per generation $\langle\theta\rangle_g$ per cycle (with bottom projection to better display the range of values per cycle) for the simulation shown in Figure 3. The population rapidly evolves away from a "no-response" strategy (< 10 cycles) before steadily increasing the affinity constant (K_D), protein synthesis rate (k), and Hill coefficient (n_H) until the final adapted bistable genotype is reached. (B) To better understand the forces that destabilize the "no-response" strategy, we ran the same simulation and only mutated k ($n_H = 1$, $K_D = 10$ are fixed) with a smaller population size $N = 100$. We plot the population average per generation $\langle k\rangle_g$ and population fitness w (dark green for LOW, light green for HIGH); geometric mean fitness per cycle (W_{cycle}) is orange. A sample evolutionary trajectory over one cycle is shown in red, starting from the white-filled circle. Although selection tends to increase the average fitness each generation, our simulation shows that $\langle k\rangle_g$ is surprisingly unconstrained. (C) A single Metropolis walker (with deterministic gene expression dynamics, i.e. no noise) in the same fluctuating fitness landscape shows identical behavior. Here, a mutation (θ') occurs every generation ($u = 1$) and it is accepted (i.e. $\theta \leftarrow \theta'$) with probability $min(1, w^{(E)}(\theta')/w^{(E)}(\theta))$. The Metropolis walker shows that destabilization of "no-response" and unconstrained $\langle k\rangle_g$ arises because of fluctuating fitness landscapes and not from genetic drift or gene expression noise.

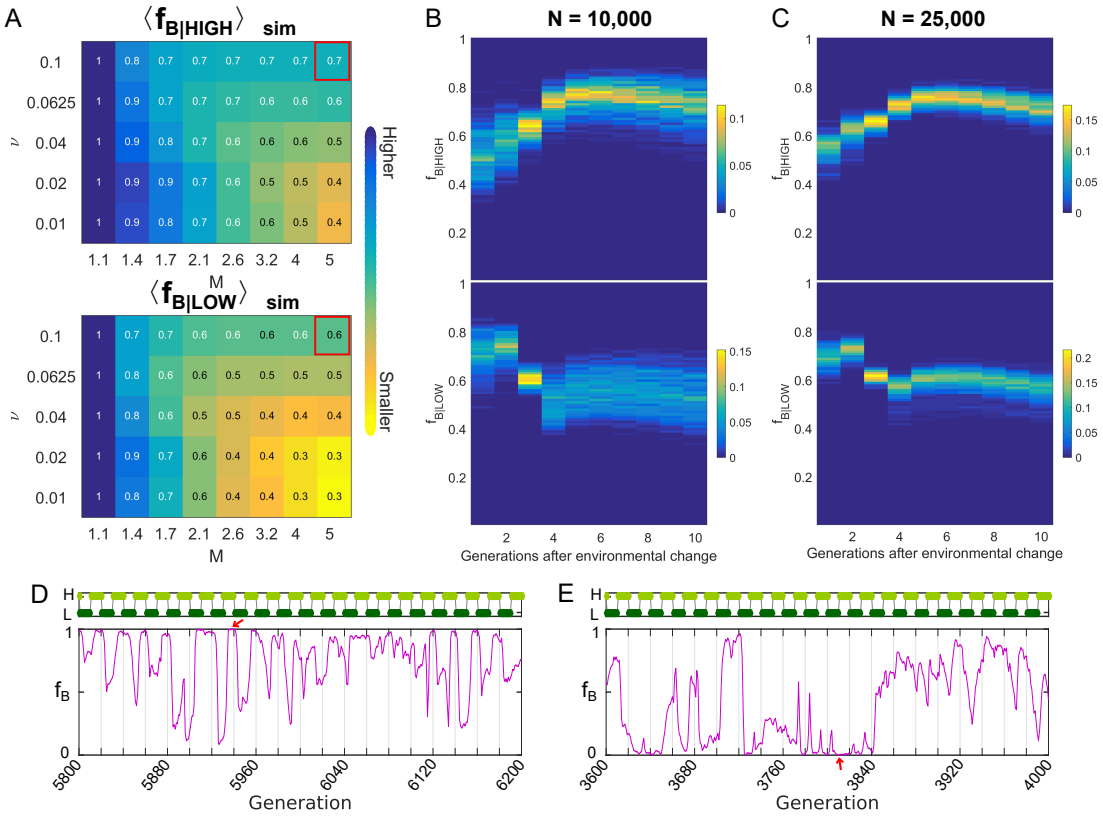

FIG. S3. **Coexistence of bistable and monostable sub-populations is an evolutionarily stable state.** (A) The colormaps show the average fraction of bistable individuals in the population $\langle f_{B|E}\rangle_{\text{sim}}$ in each environment (HIGH and LOW) of 10 independent evolutionary simulations for mutation step-size (M) and environmental fluctuation frequency (ν). Each simulation ran for 10,000 generations with evolutionary parameters $N = 10000$, $s_t = 6$, $u = 0.03$, and $k = 80$, $n_H = 6$, and $K_D = 45$ as the initial genotype (θ_1). (B) Density plot of $f_{B|E}$ for $\nu = 0.1$, and $M = 5$ (red box in A) as a function of generations after environmental change. Each column corresponds to the distribution of the $f_{B|E}$ over the entire simulation. (C) Increasing the population size to $N = 25000$ sharpened the observed trends of stable co-existence. (D,E) Two examples showing how the (D) monostable and (E) bistable subpopulations can become extinct (i.e. $f_B = 1$ and $f_B = 0$, respectively; red arrows) yet are re-established in the evolutionary simulation. Both examples were for evolutionary parameters $N = 630$, $s_t = 6$, $u = 0.03$, $\nu = 0.1$, and (D) $M = 1.7$, (E) $M = 5$. We deliberately decreased N because extinction events are more common in smaller populations.

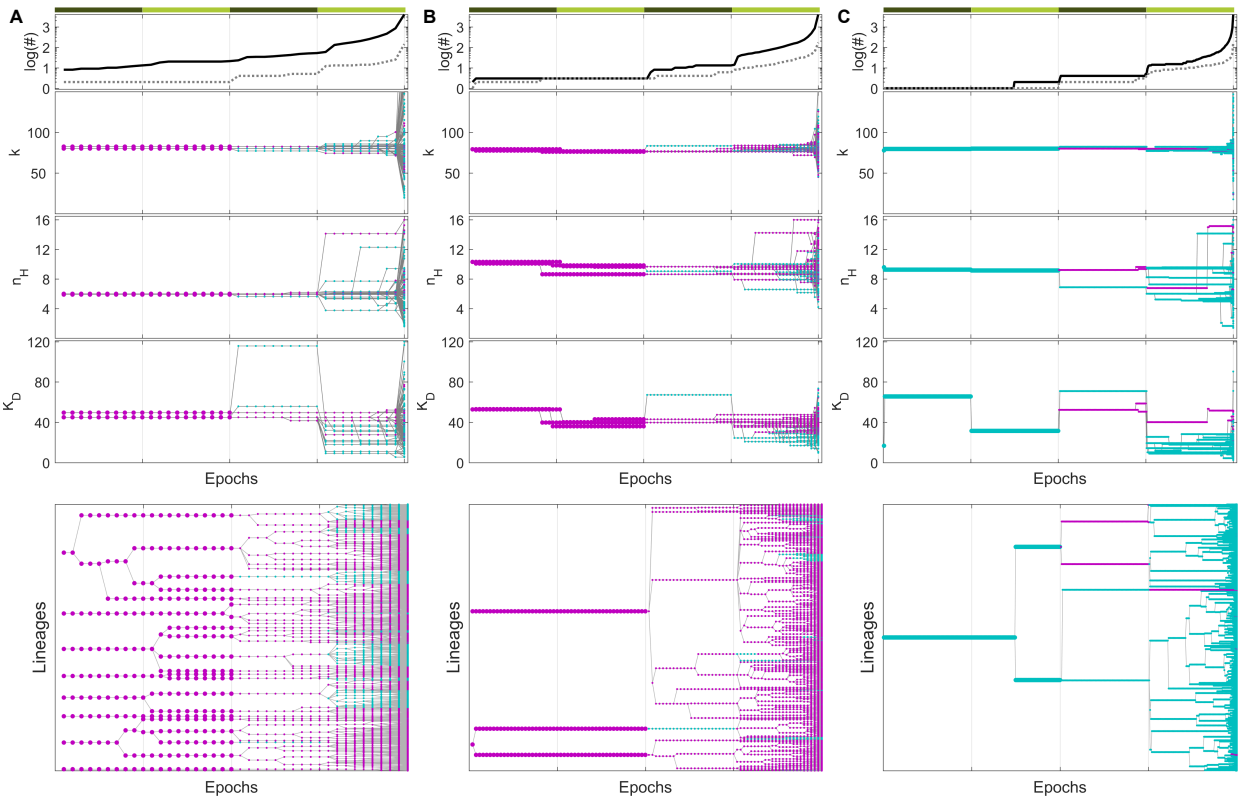

FIG. S4. **Lineage analysis of cells evolving in a fluctuating environment.** At the end of each cycle (LOW epoch + HIGH epoch), we analyzed the genealogy of cells over the past two cycles. All cells were classified based on the evolutionary strategy used by their 2-cycle ancestor over a full cycle (bigger dots). The top bars show the environmental state per epoch (dark green for LOW, light green for HIGH). We plot the number (#) of distinct lineages (solid line) and genotypes (dotted line) as a function of past generations on the top row. The middle rows plot the corresponding genotypes θ and the bottom shows the individual ancestral lineages. Ancestral genotypes can be bistable (violet) or monostable (blue). (A) Example of lineage analysis of cells that use epigenetic switching (ES) strategy for $\nu = 0.1$ and $M = 5$, i.e. their 2-cycle ancestors were fully bistable and persisted a full cycle without mutations. Note that there are distinct lineages with identical genotypes. (B) Example of lineage analysis of cells that use bistable adaptation (BA) strategy for $\nu = 0.04$ and $M = 2.1$, i.e. their 2-cycle ancestors were fully bistable but accumulated mutations over the next cycle. (C) Example of lineage analysis of cells that use genetic adaptation (GA) strategy for $\nu = 0.01$ and $M = 5$, i.e. their 2-cycle ancestors had monostable genotypes and accumulated mutations over the next cycle. In all cases, we used $N = 4000$, $s_t = 15$, and $u = 0.03$.

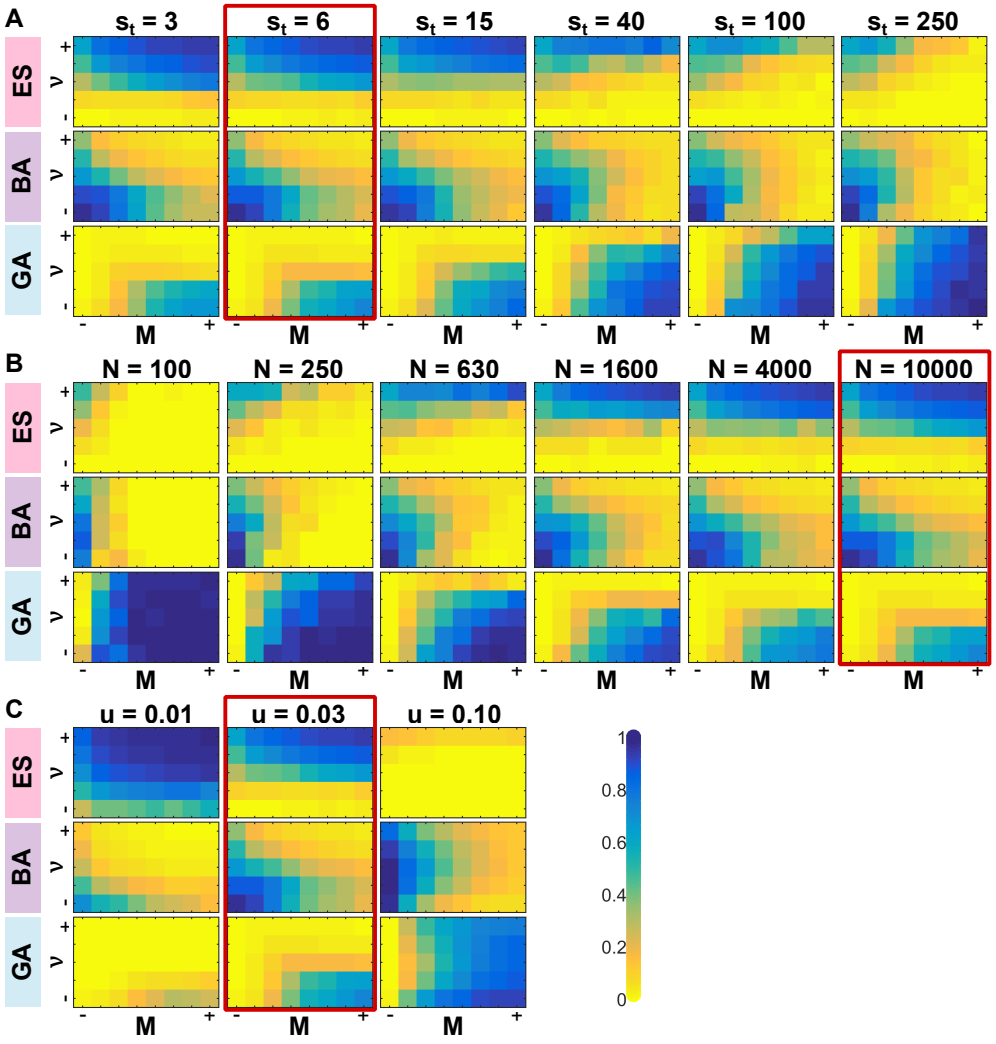

FIG. S5. **Increasing selection pressure or mutation rate favors genetic adaptation, where as increasing population size favors epigenetic switching.** Each color map shows the average fraction of parental lineages using each adaptation strategy (epigenetic switching, ES; bistable adaptation, BA; genetic adaptation, GA) for the same range of mutation step-size (M) and environmental fluctuation frequency (ν) as Figures 5-6. Evolutionary parameters used in main text ($s_t = 6$, $N = 10000$, $u = 0.03$) are highlighted in red boxes. (A) The effect of only changing the selection pressure (s_t) over three evolutionary replicas. (B) The effect of only changing the population size (N) over three evolutionary replicas. (C) The effect of only changing the mutation rate (u) over ten evolutionary replicas. All simulations ran 10000 generations with $k = 80$, $n_H = 6$, $K_D = 45$ as the initial genotype θ_1.

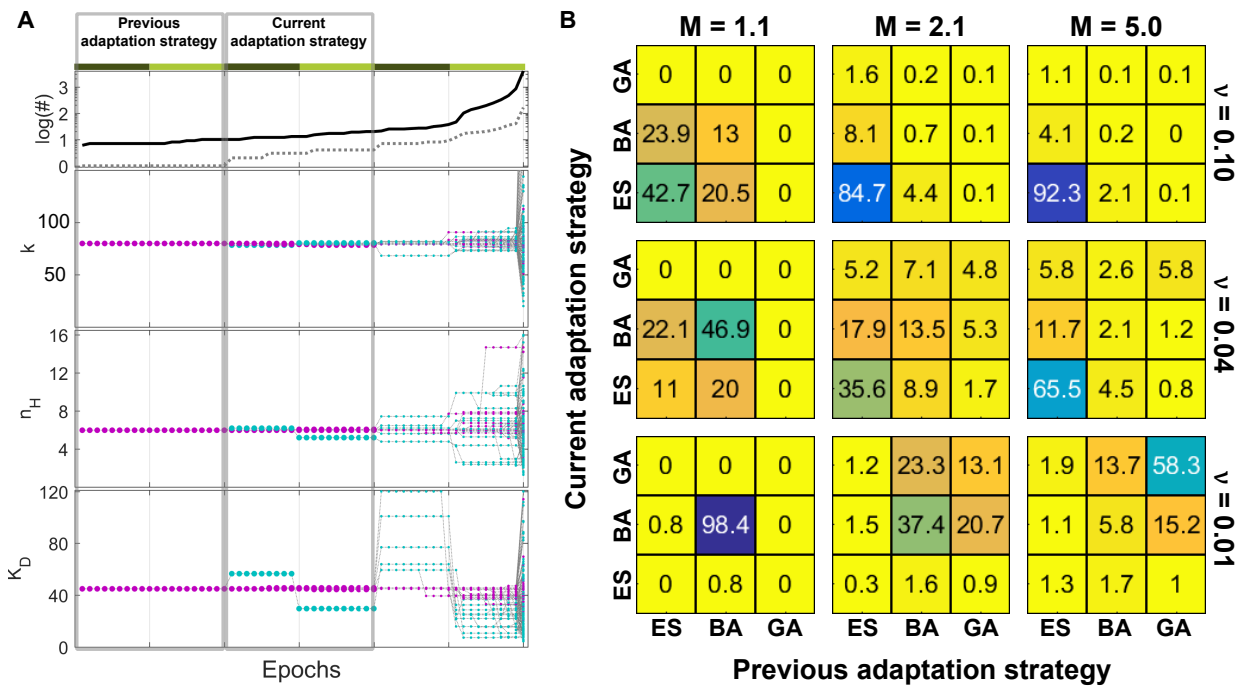

FIG. S6. **Transitions between adaptation strategies as a function of evolutionary parameters.** (A) An example showing how the current and previous adaptation strategies were defined for a surviving population. In this particular simulation, the population displayed multiple strategies in the ancestral lineage between the 2-cycle and 1-cycle ancestors (current adaptation strategy): 63.825% epigenetic switching (ES), 35.95% bistable adaptation (BA), and 0.225% genetic adaptation (GA). However, the ancestral lineage between the 3-cycle and 2-cycle ancestors (previous adaptation strategy) were 100% ES. (B) The color maps show the percentage of ancestral lineages that displayed one adaptation strategy (current adaptation strategy) and their adaptation strategy in the preceding ancestral lineage (previous adaptation strategy). These statistics were calculated for 10 evolutionary replicas for mutation step-size (M) and environmental fluctuation frequency (ν). Each simulation was run 10,000 generations with evolutionary parameters $N = 10000$, $s_t = 6$, $u = 0.03$, and $k = 80$, $n_H = 6$, and $K_D = 45$ as the initial genotype θ_1.

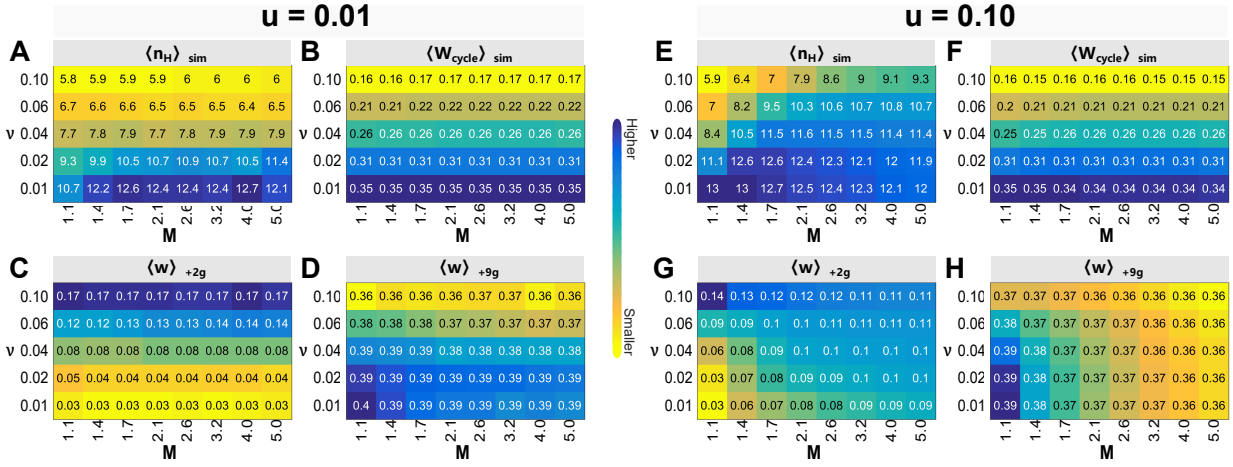

FIG. S7. **Increasing the mutation rate makes population fitness sensitive to mutation step-size (M) and affects qualitative trends in adaptation time and phenotypic robustness.** (A,E) Average population Hill coefficient $\langle n_H \rangle_{\text{sim}}$ of 3 evolutionary replicas for mutation step-size (M) and environmental fluctuation frequency (ν). (B,F) Average geometric mean fitness per cycle $\langle W_{cycle} \rangle_{\text{sim}}$ for the same simulations. (C,G) Average population fitness at the second generation $\langle w \rangle_{+2g}$ and (D,H) ninth generation $\langle w \rangle_{+9g}$ after an environmental transition. All simulations ran for 10,000 generations with identical evolutionary parameters and initial genotypes as in Figure 5-6, except for mutation rates (u) listed above each plot.

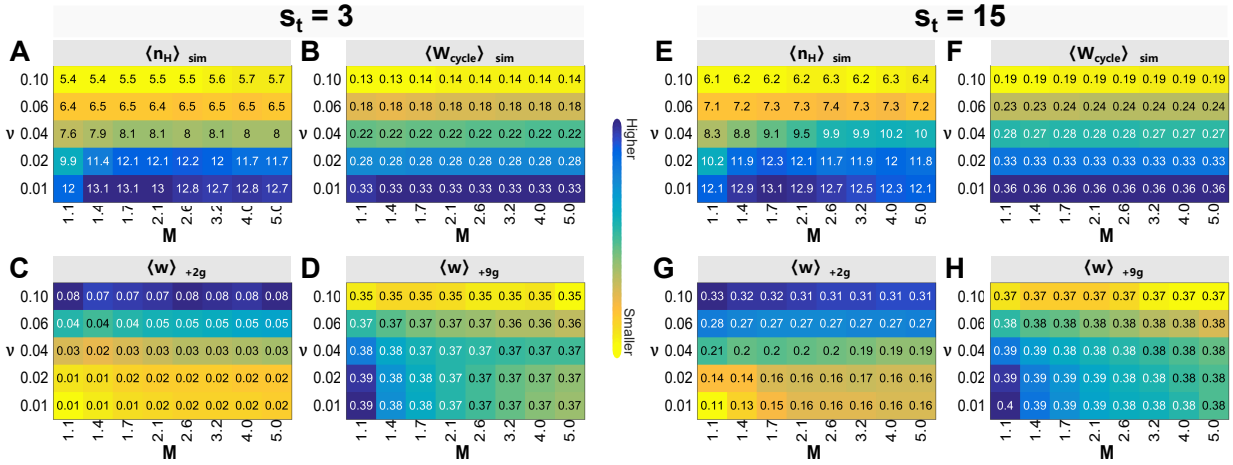

FIG. S8. **Increasing the selection pressure increases the population fitness without affecting the qualitative trends in adaptation time and phenotypic robustness.** (A,E) Average population Hill coefficient $\langle n_H \rangle_{\text{sim}}$ of 3 evolutionary replicas for mutation step-size (M) and environmental fluctuation frequency (ν). (B,F) Average geometric mean fitness per cycle $\langle W_{cycle} \rangle_{\text{sim}}$ for the same simulations. (C,G) Average population fitness at the second generation $\langle w \rangle_{+2g}$ and (D,H) ninth generation $\langle w \rangle_{+9g}$ after an environmental transition. All simulations ran for 10,000 generations with identical evolutionary parameters and initial genotypes as in Figure 5-6, except for selection pressure (s_t) listed above each plot.

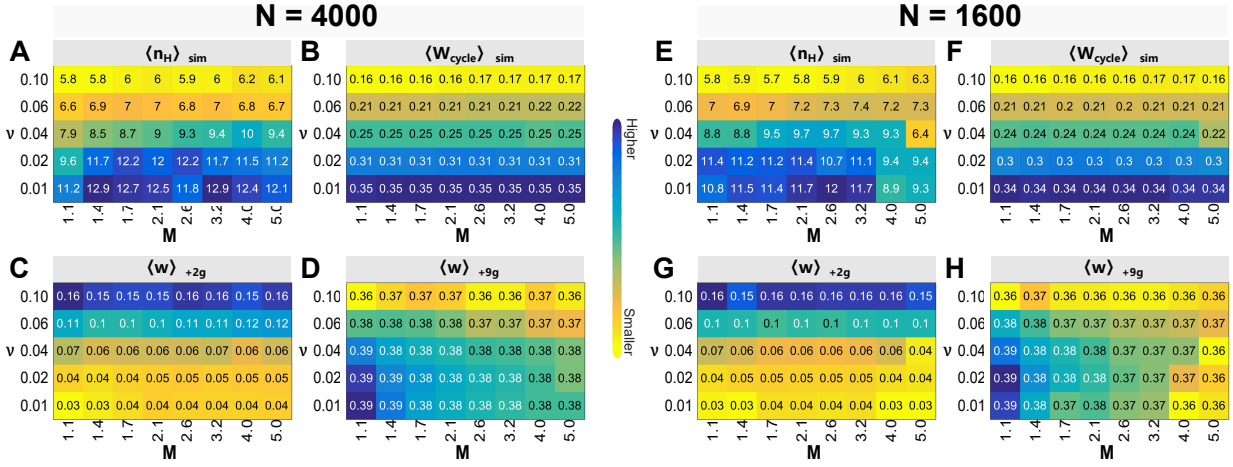

FIG. S9. **Decreasing the population size increases the variation between independent replicas, without affecting the qualitative trends in fitness, adaptation time and phenotypic robustness.** (A,E) Average population Hill coefficient $\langle n_H \rangle_{\text{sim}}$ of 3 evolutionary replicas for mutation step-size (M) and environmental fluctuation frequency (ν). (B,F) Average geometric mean fitness per cycle $\langle W_{cycle} \rangle_{\text{sim}}$ for the same simulations. (C,G) Average population fitness at the second generation $\langle w \rangle_{+2g}$ and (D,H) ninth generation $\langle w \rangle_{+9g}$ after an environmental transition. All simulations ran for 10,000 generations with identical evolutionary parameters and initial genotypes as in Figure 5-6, except for population size (N) listed above each plot.

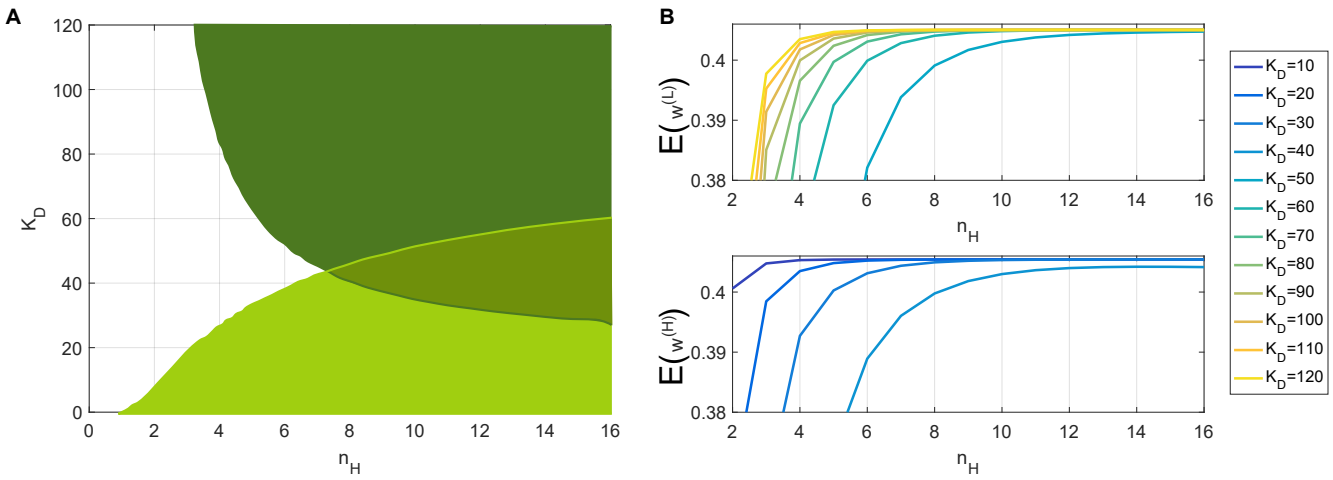

FIG. S10. **Increasing n_H increases the average population fitness in both environments.** (A) Contour plots as a function of biophysical parameters with fixed $k = 80$ where the steady state (A^*) for LOW are $A^{(L)} \pm 1\%$ (dark green) and for $HIGH$ are $A^{(H)} \pm 1\%$ (light green). (B) We analytically calculated the average fitness of an infinite, clonal population with steady-state protein distribution $\rho(A)$ given n_H, K_D and fixed $k = 80$. The normalized distribution $\rho(A)$ was estimated numerically for each set of biophysical parameters. We calculated the expected fitness in each environment by integration $E(w^{(E)}) = \Sigma_a w^{(E)}(a) \cdot \rho(a)$.

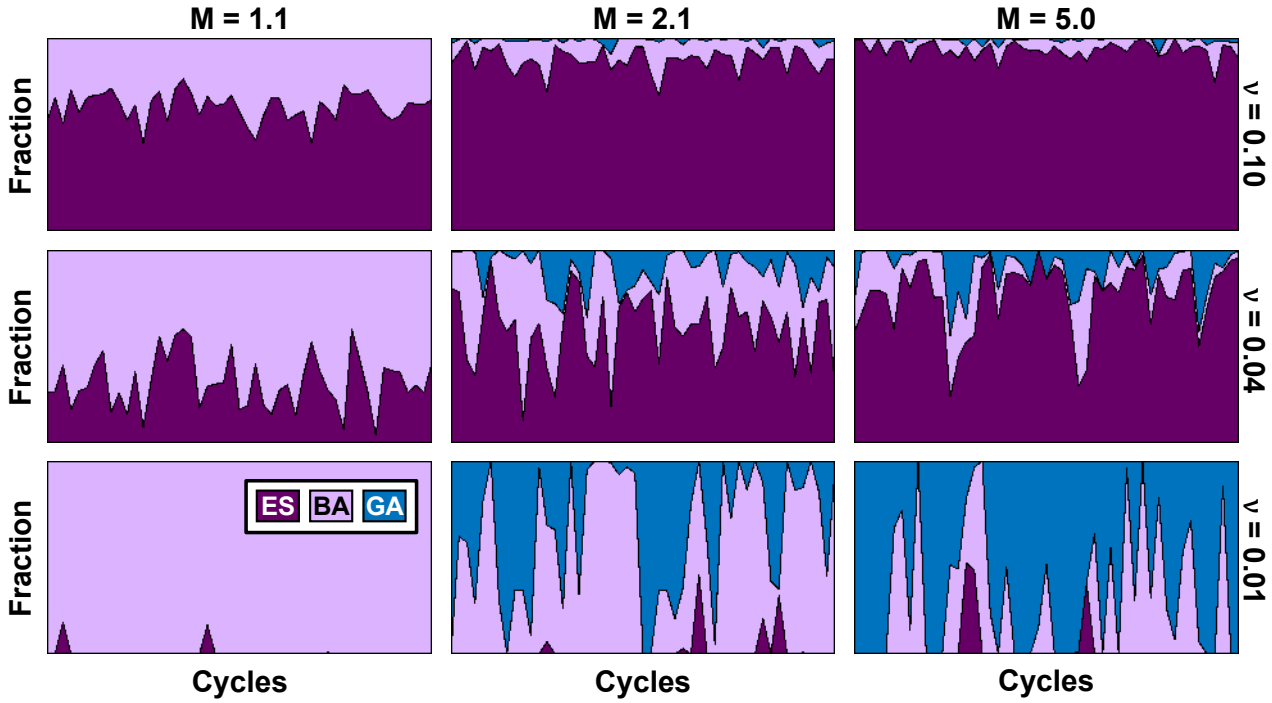

FIG. S11. **Different adaptation strategies co-exist and exhibit large fluctuations over evolutionary time.** Each plot shows the fraction of parental lineages with different adaptation strategies (epigenetic switching, ES; bistable adaptation, BA; genetic adaptation, GA) per environmental cycle for mutation step-size (M) and environmental fluctuation frequency (ν). In each case, only the last 50 cycles are shown. Each simulation ran 10,000 generations with evolutionary parameters $N = 10000$, $s_t = 6$, $u = 0.03$, and $k = 80$, $n_H = 6$, and $K_D = 45$ as the initial genotype θ_1.

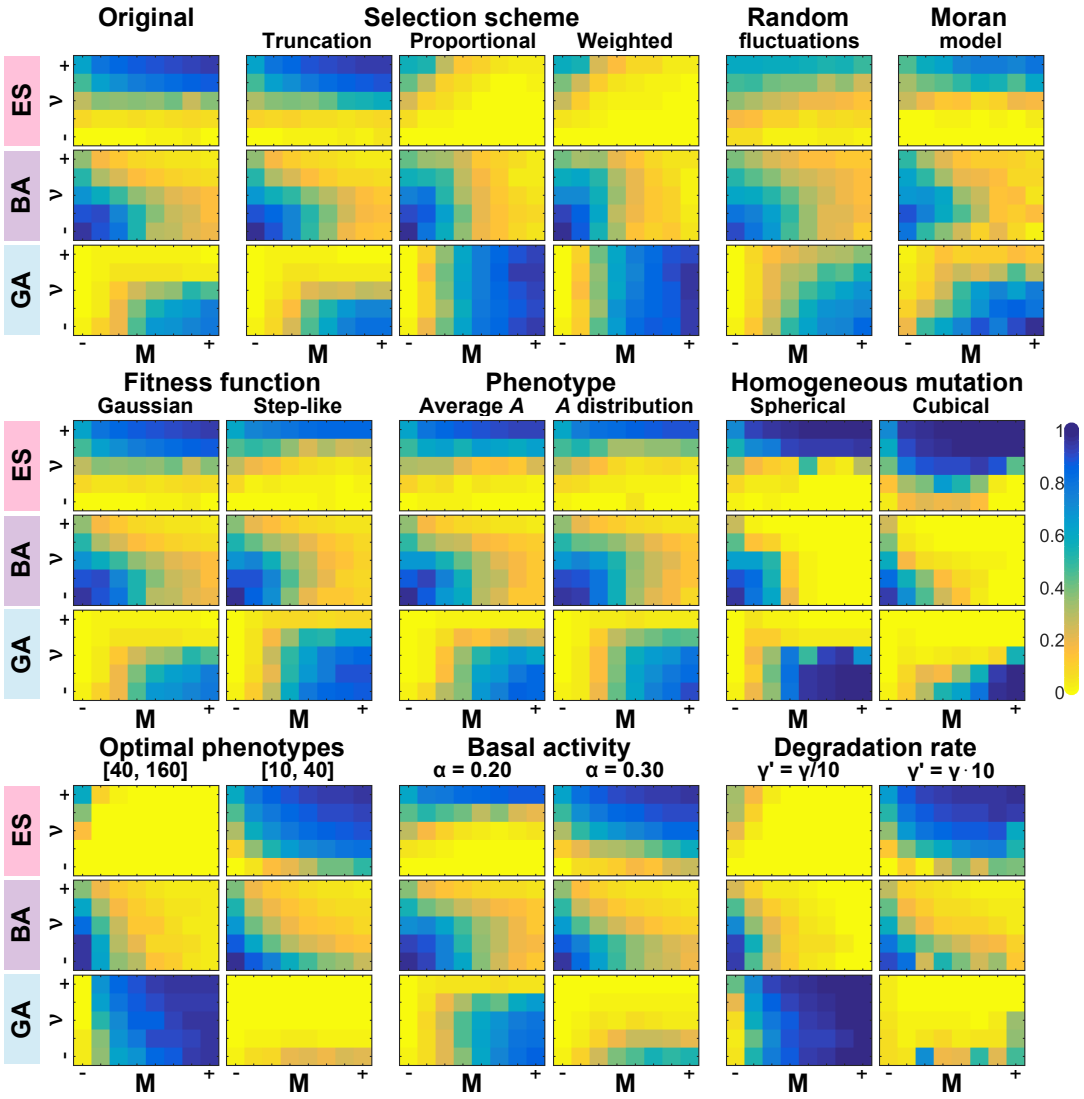

FIG. S12. **The same qualitative trends on the selection of adaptation strategies per evolutionary condition is maintained in a wide variety of alternative model assumptions.** Each color map shows the population average fraction of parental lineages using each adaptation strategy (epigenetic switching, ES; bistable adaptation, BA; genetic adaptation, GA) for the same range of mutation step-size (M) and environmental fluctuation frequency (ν) as Figures 5-6. Differences in assumptions or parameters are listed above each plot. All values are the average of 3 evolutionary replicas of simulations run 10,000 generations with $N = 4,000$, $s_t = 6$, $u = 0.03$ and $k = 80$, $n_H = 6$, and $K_D = 45$ as the initial genotype θ_1. The exceptions are the weighted and proportional selection schemes where the selection pressure (s_t) cannot be tuned, and the Moran model case which ran for 1,000 generations. When the basal activity (α) was changed, we adjusted the low optimal phenotype such that the ratio of $A^{(L)} = \alpha \cdot A^{(H)}$, where $A^{(H)} = 80$. See *Alternative assumptions details* for an explicit description of assumptions or parameters.

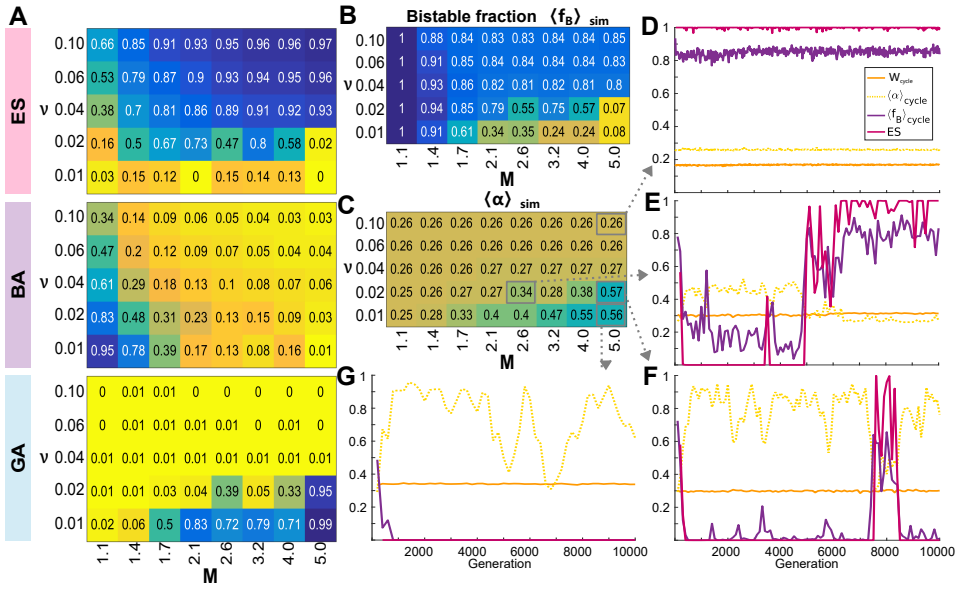

FIG. S13. **Allowing basal activity (α) to evolve does not qualitatively change our results.** (A) The color bar shows the average fraction of parental lineages using each adaptation strategy (epigenetic switching, ES; bistable adaptation, BA; genetic adaptation, GA) of 3 evolutionary replicas under the corresponding mutation step-size (M) and environmental fluctuation frequency (ν). Each simulation ran for 10,000 generations with evolutionary parameters $N = 10000$, $s_t = 6$, $u = 0.03$ and $k = 80$, $n_H = 6$, $K_D = 45$ and $\alpha = 0.25$ as the initial genotype (θ_1). The corresponding (B) average bistable fraction ($\langle f_B \rangle_{\text{sim}}$) and (C) average α ($\langle \alpha \rangle_{\text{sim}}$). For some examples, the dynamics over time for the geometric mean fitness per cycle (W_{cycle}), the average basal activity ($\langle \alpha \rangle_{cycle}$), and the average bistable fraction ($\langle f_B \rangle_{cycle}$) per cycle, as well as the fraction of parental lineages using ES as the adaptation strategy per cycle, are shown: (D) $\nu = 0.1$ and $M = 5$; (E) $\nu = 0.02$ and $M = 2.6$; (F) $\nu = 0.0.02$ and $M = 5$; and (G) $\nu = 0.01$ and $M = 5$.

II. ALTERNATIVE ASSUMPTIONS DETAILS

- *Truncation selection scheme*: Only a certain fraction of the best individuals can be selected, each with the same probability. Blickle & Thiele (1995) calculated the truncation fraction that result in the same selection strength as a given tournament size. We estimate that $s_t = 6$ corresponds to a 0.24 truncation fraction. We used this fraction in our Truncation selection simulation.

- *Proportional selection scheme*: The probability of an individual to be selected is proportional to its fitness value. Selection pressure cannot be tuned.

- *Weighted selection scheme*: A random individual is picked from the population and is cloned into the new population if a uniformly distributed random number (from the interval [0,1]) is below its fitness. Selection pressure cannot be tuned.

- *Random environmental fluctuations*: The environment fluctuates randomly between the two possible states with mean frequency ν.

- *Moran model*: Reproduction and death events are treated as stochastic events allowing overlapping generations. At each time step, an individual is chosen for reproduction using the defined tournament selection scheme, and an individual is randomly chosen from the population for death to keep the population size N fixed. N time steps occur in the previously defined lifespan time, such that the reproduction rate (and then mutation) is equivalent to the original model. These simulations are considerably more expensive computationally, so we used a shorter simulation time (only 1,000 generations).

- *Gaussian fitness function*:

$$\omega_g^{(E)}(A) = e^{-\frac{(A-A^{(E)})^2}{2\sigma^2_{(E)}}} \tag{S1}$$

  where $\sigma^2_{(E)}$ is equal to the width in the Lorentzian fitness function (v^2).

- *Step-like fitness function*:

$$\omega_s^{(E)}(A) = \begin{cases} 1 & \text{if } (A - A^{(E)})^2 \leq 2\sigma^2_{(E)} \\ 0, & \text{otherwise} \end{cases} \tag{S2}$$

  where $\sigma^2_{(E)}$ is equal to the width in the Lorentzian fitness function (v^2).

- *Average A as phenotype*: Use the life-time average protein number to assign its fitness score to each individual in the population.

- *A distribution as phenotype*: Use the life-time protein number distribution to calculate the average fitness score for each individual in the population.

- *Homogeneous spherical mutation*: Same mutation scheme as the one described in the paper, but substituting r in Eqs (5-7) with $\sqrt[3]{r}$.

- *Homogeneous cubic mutation*: Drawing three uniformly distributed random values between -1 and 1 ($r_i \sim U(-1, 1)$) for each biophysical parameter p_i, and update their value as follows:

$$p_i' \leftarrow p_i \cdot M^{r_i}. \tag{S3}$$