Supplemental Information

Phenotypic Diversity Using Bimodal and Unimodal Expression of Stress Response Proteins

Javier Garcia-Bernardo and Mary J. Dunlop

Biophysical Journal

Supporting Material

Phenotypic Diversity Using Bimodal and Unimodal Expression of Stress Response Proteins

Javier Garcia-Bernardo[1] and Mary J. Dunlop[2,*]

[1]Department of Computer Science and [2]School of Engineering, University of Vermont, Burlington, Vermont

*Correspondence: mjdunlop@uvm.edu

Supplementary Information

Supplementary Methods

Weak stressor

We distinguished between two strengths of the stressor: strong, where the cells with low levels of protein expression are killed in the presence of high concentrations of the stressor and weak, where the stressor stops cell division for cells with low protein expression. The growth rate for a weak stressor is calculated by

$$\lambda_{P,S} = \begin{cases} \frac{c_{th} - c(P,S)}{c_{th}} & \text{if } c(P,S) < c_{th} \\ 0 & \text{if } c(P,S) \geq c_{th}. \end{cases} \tag{1}$$

Fitness function

We assume fast switching rates, which produces identical distributions of phenotypes in low and high stress conditions. At equilibrium, the growth of a population depends on the fraction of cells at each protein level. The ratio of final to initial number of cells in one generation is

$$N_1/N_0 = \sum_P x_P 2^{\lambda_{P,S}}, \tag{2}$$

where N_0 and N_1 are the initial number of cells and the number of cells after one generation (1). x_P is the fraction of the population in the protein state P and $\lambda_{P,S}$ is the growth rate of cells in the protein state P and stress level S. The values of P correspond to the protein levels $\{0 - 100, 100 - 200, ..., 9900 - 10000\}$ molecules. Because the population stays in equilibrium, the total growth is

$$N_t/N_0 = \left(\sum_P x_P 2^{\lambda_{P,S}} \right)^t, \tag{3}$$

where t is the time spent in equilibrium conditions measured in generations. Note that the population grows if the value of the sum is greater than 1.

For the case with two environments, the ratio of final to initial cells is determined by

$$N_t/N_0 = \left(\sum_P x_P 2^{\lambda_{P,S_l}} \right)^{t_l} \cdot \left(\sum_P x_P 2^{\lambda_{P,S_h}} \right)^{t_h}, \tag{4}$$

where t_l and t_h are the times spent in low and high stress conditions. S_l and S_h are the low and high stress conditions. Instead of calculating the product of ratios, we record the logarithm of the products of ratios

$$\log\left(N_t/N_0\right) = t_l \log\left(\sum_P x_P 2^{\lambda_{P,S_l}} \right) + t_h \log\left(\sum_P x_P 2^{\lambda_{P,S_h}} \right). \tag{5}$$

As the fitness function in our differential evolution algorithm, we use

$$R = \frac{\log\left(N_t/N_0\right)}{t_l + t_h}, \tag{6}$$

which corresponds to the geometric growth rate. In the special case where no cells survive ($\lambda_{P,S_h} = -1$ for all values of P), we set R to -1. When some cells survive, R evaluates to a number between 0 and 1, where 0 indicates that the population is not growing at all and 1 indicates that it is growing optimally.

Sensing case

When cells are able to sense the environment, they can adapt their distribution in high stress conditions. The switching rates between protein states are on the order of the doubling time. Thus, we assumed one generation is required to sense and adapt to the environment.

The log of the ratio of final to initial number of cells in this case is calculated as

$$\log\left(N_t/N_0\right) = t_l \log\left(\sum_P x_P 2^{\lambda_{P,S_l}} \right) + \log\left(\sum_P x_P 2^{\lambda_{P,S_h}} \right)$$
$$+ (t_h - 1) \log\left(\sum_P y_P 2^{\lambda_{P,S_h}} \right), \tag{7}$$

where x_P is the fraction of cells in protein level P adapted to the low stress environment and y_P is the fraction of cells adapted to the high stress environment. We obtained the distribution of y_P

2

values by evolving the optimal unimodal distribution in fixed high stress conditions. The first term in this equation describes the population in low stress, the middle term describes the population during the one generation after the transition from low to high stress, and the final term describes the population during the time in high stress.

Differential evolution algorithm

We used the Python differential evolution code (2), available online at http://www1.icsi.berkeley. edu/~storn/code.html. We modified it to use the Python module *numpy*, which allows for fast operations on vector objects. Differential evolution finds optimal or near-optimal solutions by iteratively improving a set of candidate solutions based on their fitness. For example, in the 2γ case, differential evolution evolves the six parameters a_1, b_1, w_1, a_2, b_2, and w_2.

Initialization of the population

The population was set to 40 trial vectors, following recommendations for good performance of the differential evolution algorithm from (3).

For the 1γ distribution, the initial population was set by creating vectors $\{a, b\}$ in the range $a = [0.5, 100]$, $b = [10, 400]$. For the full algorithm, we used bounds $[0.5, 100]$ for a and $[10, 4000]$ for b to allow for parameters in the range of experimentally derived values from (4). The difference between the initialization ranges and bounds were empirically found to provide better convergence of the algorithm.

For the 2γ distribution, the initial population was created in the range $a_1 = [0.5, 2]$, $b_1 = [10, 400]$, $w_1 = [0, 1]$, $a_2 = [2, 100]$, $b_2 = [10, 400]$, and $w_2 = [0, 1]$; with bounds $a_1 = [0.5, 100]$, $b_1 = [10, 4000]$, $w_1 = [0, 10]$, $a_2 = [0.5, 100]$, $b_2 = [10, 4000]$, and $w_2 = [0, 10]$. The difference between the initialization ranges allows the two gamma distributions to start with proteins distributed between low and high protein expression levels to allow the evolution of bimodality.

For the case with no restriction on the protein distribution, we evolved the fraction of the population at each protein level, $p(P)$, directly with an initialization and bound range $[0, 1]$. The

range has little effect on performance, since the final $p(P)$ values are normalized to $\sum_P p(P) = 1$. Given the large number of parameters, we initialized the population with 400 trial vectors, which we found to improve convergence.

Evolution and parameters

For each trial vector x_i in the population, a mutant vector v_i was created using the DE/rand/1/bin method (2), which combines three random trial vectors in the population, x_{r_1}, x_{r_2}, and x_{r_3}, according to

$$v_i = x_{r_1} + F \cdot (x_{r_2} - x_{r_3}), \tag{8}$$

where the multiplier F was chosen randomly from the range $[0.5, 2]$, since these values have been shown to improve convergence, especially when the fitness function is noisy (3).

If the population had not converged after 50 generations (see termination description below), the algorithm was switched from DE/rand/1/bin to DE/best/1/bin (2), and the mutant vector was created as

$$v_i = x_{best} + F \cdot (x_{r_1} - x_{r_2}), \tag{9}$$

where x_{best} is the trial vector with highest fitness in the population. This change in the method allows for a more refined search in the latter parts of the evolution.

The mutant vector v_i is combined with the ith trial vector, x_i, in the population using 90% of the values in v_i and 10% of the values in x_i, i.e. the crossover rate is 0.9, as suggested by Price and Storn in the case of parameter dependence (3). The resulting vector, u_i, substitutes x_i for the set of trial vectors if it allows for a higher fitness value than x_i.

Termination

The algorithm was allowed to run for 500 generations, or until convergence. Convergence was defined as being achieved if the difference between the mean fitness of the population and the fitness of all candidate solutions was within 10^{-10} fitness units.

Supplementary Text

Fitness function

After a transition time, and while the environment is constant, the distribution of protein states in a population is at equilibrium (5, 6). The equilibrium depends on the growth rates $\lambda_{P,S}$ and the switching rates between protein states (6). Here, we evolve the distribution of protein levels in the population that yields the highest growth rate. Therefore, we indirectly evolve the switching rates that produce that equilibrium. Furthermore, we assumed that the switching rates were on the order of the doubling time of a cell, which produces fast transition times, consistent with memory observed in the level of proteins *in vivo* (7, 8). One concern is that the distribution of protein levels may change during the time in high stress, i.e. fewer cells may leave the resistant state than were leaving before the stress appeared. However, this effect is only pronounced when the switching rates are low compared with the differences in growth rates between cells (6). For our simulations, the maximum difference in growth rates is one order of magnitude below the doubling time of the cell, and therefore we assumed that the fraction of cells that leave the resistant state is constant through the different environmental conditions.

Differences in growth rates and relationship to time for one population to overtake another

Provided the population is growing, the growth rate is normalized to be between $0 - 1$, where the value 1 corresponds to one cell division in the fastest possible cell division time.

The growth of cells is given by the equation

$$N_t = N_0 2^{R \cdot t},$$

where N_t is the final number of cells, N_0 is the initial number of cells, t is time that the population has been growing, and R is the growth rate.

The benefit of sensing or bimodality is measured in points as the difference in growth rate multiplied by 100, with units (generations)$^{-1}$. For example, the ratio of cells between two populations 1 point apart is given by

$$\frac{2^{(x+0.01)t}}{2^{(x)t}} = 2^{0.01 \cdot t}.$$

Therefore, the time until the more fit condition represents 90% of the population (in other words, the ratio of more to less fit cells is 90:10, or 9) corresponds to

$$
\begin{aligned}
9 &= 2^{0.01 \cdot t} \\
\log_2(9) &= 0.01 \cdot t \\
t &= 317 \text{ generations.}
\end{aligned}
$$

For *E. coli*, a generation corresponds to about 38 minutes in rich media (9), and therefore 317 generations is 8.4 days.

Supplementary Figures

Supporting Citations

1. Monod, J., 1949. The Growth of Bacterial Cultures. *Annual Review of Microbiology* 3:371–394.

2. Storn, R., and K. Price, 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* 11:341–359.

3. Price, K., R. M. Storn, and J. A. Lampinen, 2006. Differential evolution: a practical approach to global optimization. Springer Science & Business Media.

4. Taniguchi, Y., P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, 2010. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)* 329:533–8.

5. Thattai, M., and A. van Oudenaarden, 2004. Stochastic gene expression in fluctuating environments. *Genetics* 167:523–30.

6. Visco, P., R. J. Allen, S. N. Majumdar, and M. R. Evans, 2010. Switching and growth for microbial populations in catastrophic responsive environments. *Biophysical Journal* 98:1099–108.

7. Rosenfeld, N., J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz, 2005. Gene regulation at the single-cell level. *Science* 307:1962–5.

8. Charlebois, D. A., and M. Kaern, 2012. What all the noise is about: the physical basis of cellular individuality. *Canadian Journal of Physics* 923:919–923.

9. Reshes, G., S. Vanounou, I. Fishov, and M. Feingold, 2008. Timing the start of division in E. coli: a single-cell study. *Physical Biology* 5:046001.
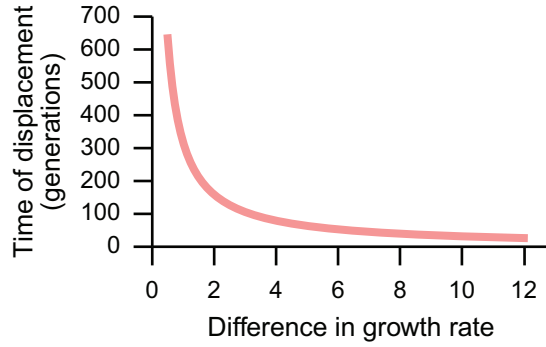
Figure S1: Relationship between difference in growth rate and time for population displacement. The time required for a fast growing population to displace a slow growing one is plotted as a function of the difference in growth rate between the two populations. Displacement is defined as achieved when the more fit condition represents 90% of the population. See Supplementary Text for additional discussion.
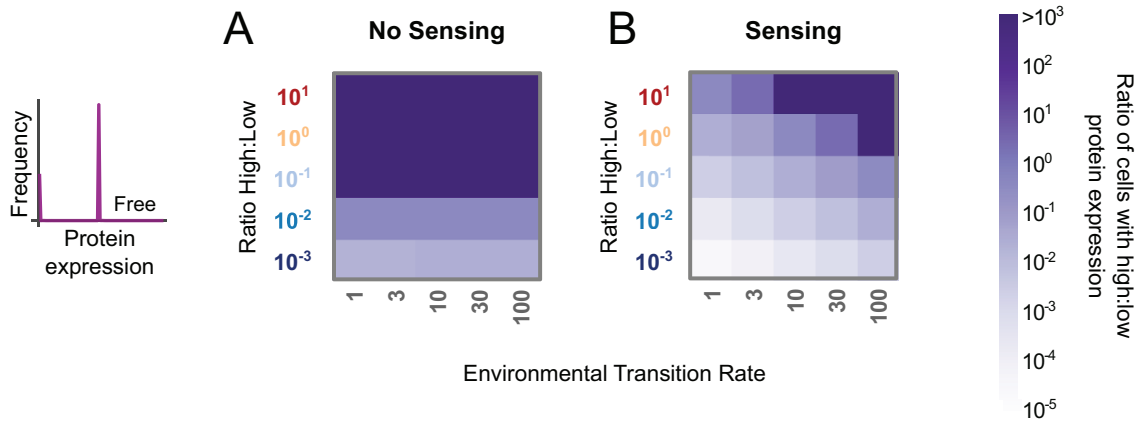

Figure S2: Bimodality is evolved in many conditions, even when there are no restrictions on the shape of the protein distribution. The distribution of proteins is allowed to evolve freely, with no restrictions on its shape. An example of the solutions is shown in the cartoon on the left. (A, B) The ratio of cells with high protein expression is plotted as a function of the environmental conditions for the (A) no sensing and (B) sensing case. Unimodal (dark purple) and bimodal (light purple) distributions are evolved.
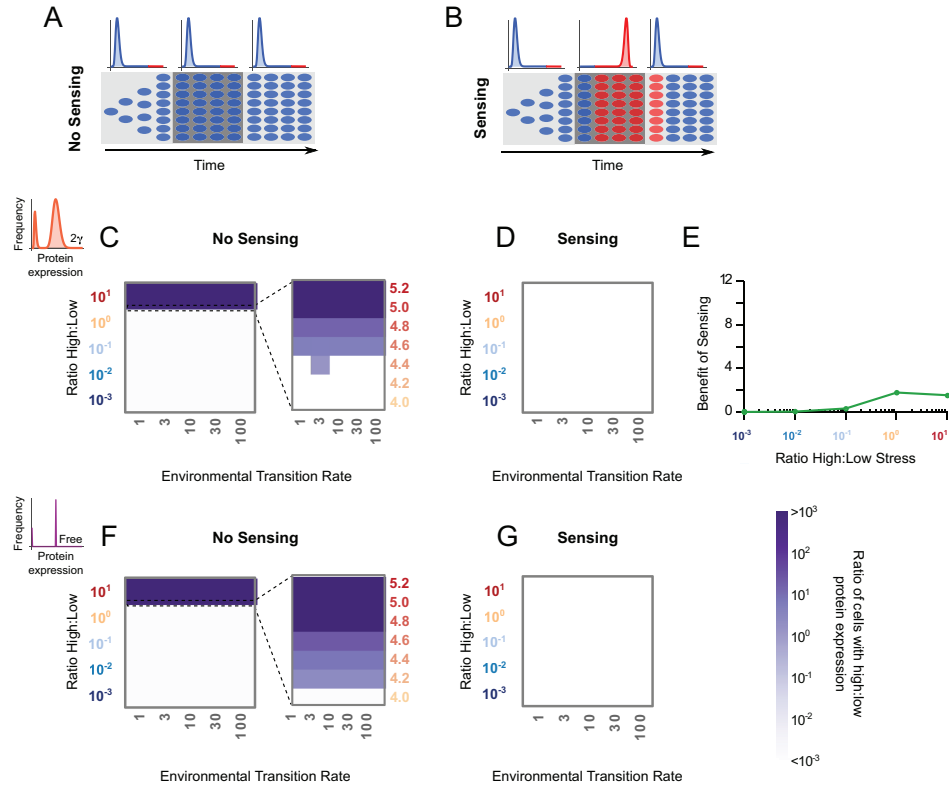
Figure S3: Bimodality is not generally evolved under weak stressors. (A) Environments vary between low and high stress. Without sensing, the population can be composed of cells with low levels of protein expression (histograms). Under low stress, the population grows well. Under high stress, it stays latent. (B) With sensing, all cells sense and adapt to the current environment after one generation. (C–D) Simulations use the 2γ restriction. Ratio of cells with high to low protein expression for (C) no sensing and (D) sensing populations. Dark purple colors indicate unimodal distributions with high protein expression. Light purple is bimodal. White is a unimodal distribution with low protein expression. Inset shows the very small region where bimodality is evolved. (E) The benefit of sensing is plotted as a function of the ratio of high to low stress. The benefit is measured as the difference in growth rate between the sensing and no sensing populations (Supplementary Text). These simulations use an environmental transition rate of 10. Corresponding plots to (C–D) are shown in (F-G) for the case with no restrictions when evolving the protein distribution.
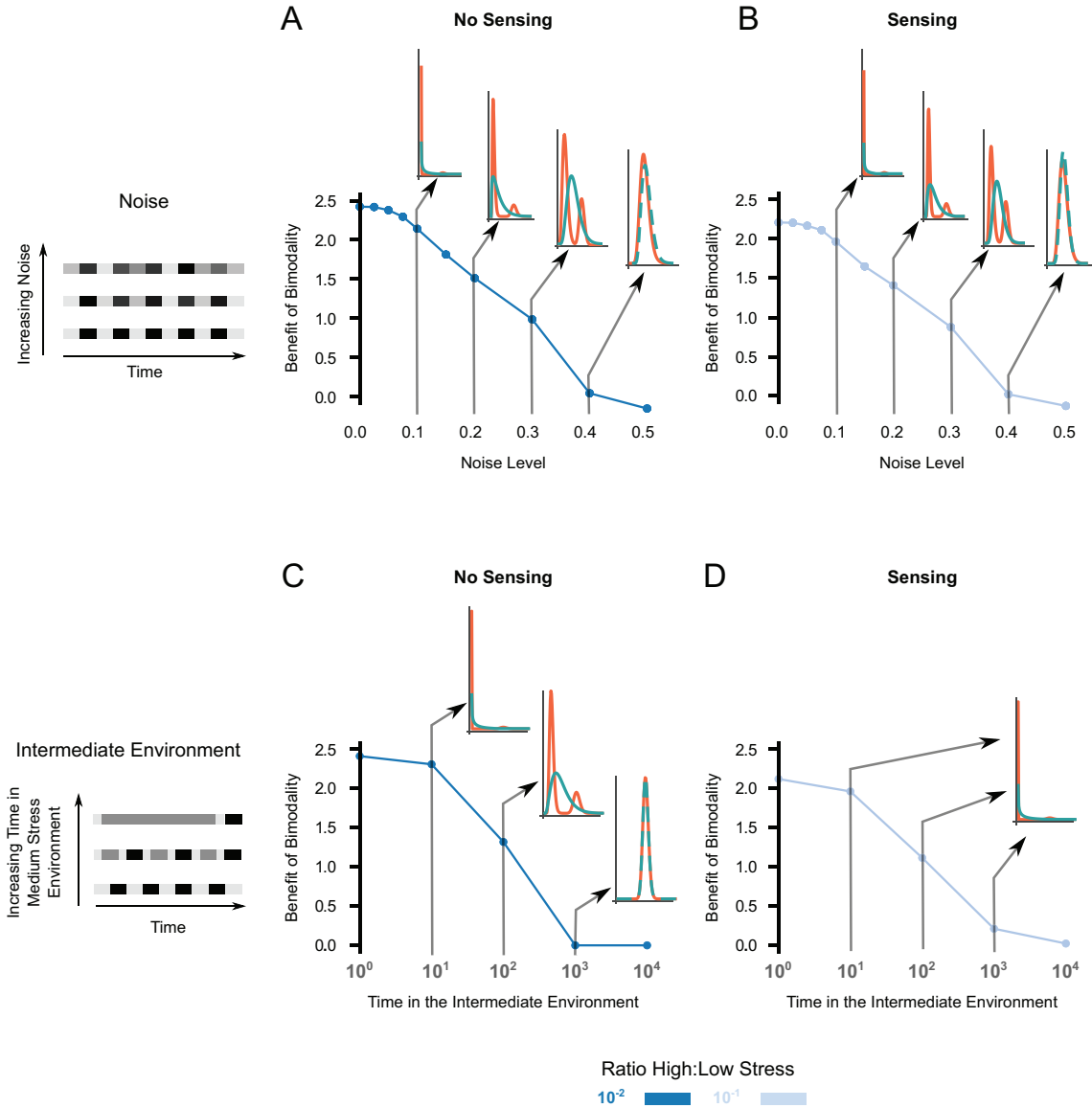
Figure S4: The benefit of bimodality decreases as noise or time in the intermediate environment is increased. Examples of the distributions evolved with the 1γ and 2γ restrictions are shown (histograms) for increasing levels of noise in the (A) no sensing and (B) sensing cases. (C, D) The distributions evolved with the 1γ and 2γ cases are shown for increasing time in the intermediate, medium stress environment for the (C) no sensing and (D) sensing cases. For the sensing case, the evolved distributions are identical for all intermediate environment times, but the benefit decreases as the time spent in the intermediate environment becomes a larger fraction of the whole simulation time.
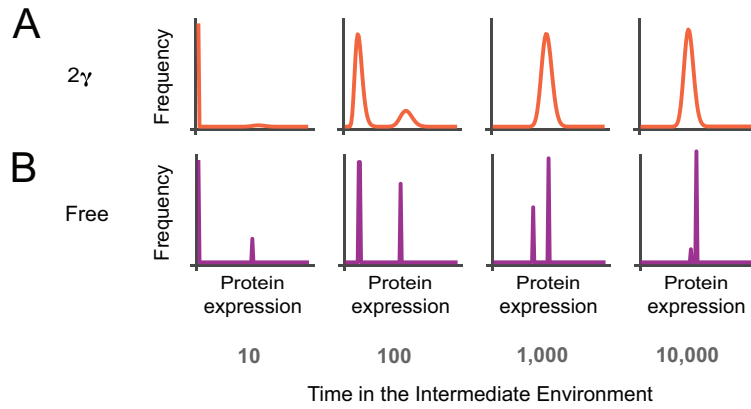
10

Figure S5: Bimodality, not trimodality, is evolved in the environment with three stress levels (low, medium, and high). The distribution of protein levels is plotted for solutions obtained using the (A) 2γ requirement and (B) the case with no restrictions on protein distribution shape. For all simulations, the time in high and low stress is 10 generations, while the number of generations in medium stress is listed below each figure panel. Results for other stress ratios are similar. Note that trimodal distributions are never evolved for any stress conditions.
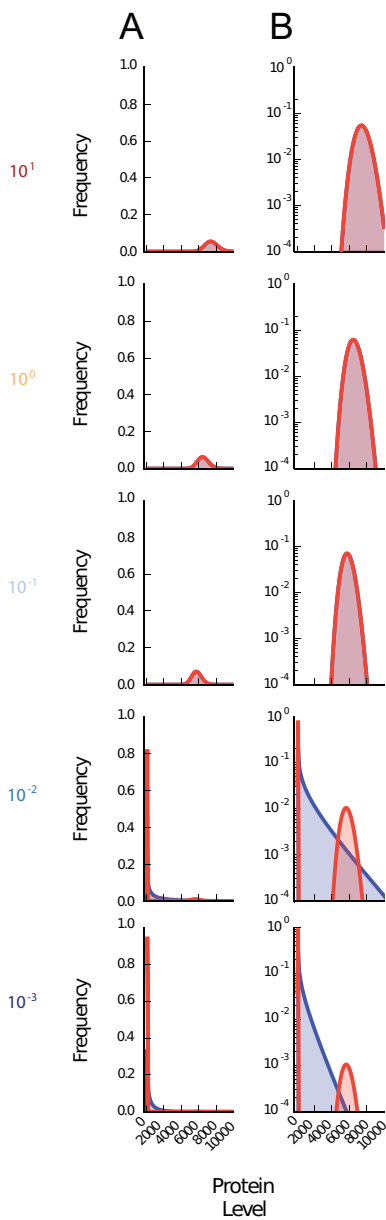
Figure S6: Comparison of evolved unimodal and bimodal distributions. Distribution of protein levels in (A) linear and (B) logarithmic scales for different ratios of high to low stress for an environmental transition rate of 10. Blue shows the 1γ case; red shows the 2γ case.
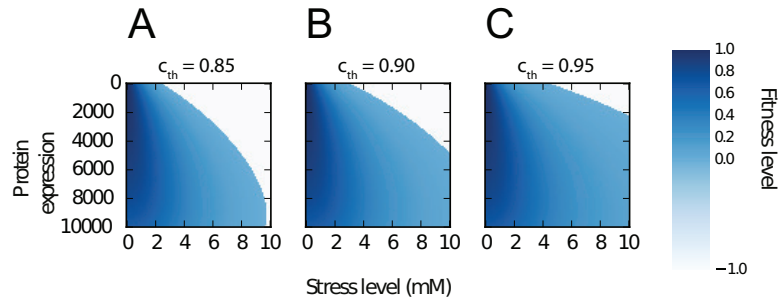
Figure S7: Fitness landscape for different threshold values. The value of c(P,S) in terms of the phenotypes P and stress levels S is plotted for different values of the threshold c_{th}: (A) 0.85, (B) 0.9, and (C) 0.95. Changes in the threshold only affect the front where the cells die. Note that the lightest area corresponds to conditions where cells are dead.