nature portfolio

https://doi.org/10.1038/s41559-022-01797-w

## Supplementary information

# Low protein expression enhances phenotypic evolvability by intensifying selection on folding stability

In the format provided by the authors and unedited

**Supplementary material**

# Low protein expression enhances phenotypic evolvability by intensifying selection on folding stability

Shraddha Karve[1,5], Pouria Dasmeh[1,2], Jia Zheng[1,3], Andreas Wagner[1,3,4,5]

**Affiliations**

[1]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.

[2]Swiss Institute of Bioinformatics, Quartier Sorge-Batiment Genopode, Lausanne, Switzerland.

[3]Westlake Institute for Advanced Study, Hangzhou, China.

[4]The Santa Fe Institute, Santa Fe, New Mexico, USA.

[5]Stellenbosch Institute for Advanced Study (STIAS), Wallenberg Research Centre at Stellenbosch University, Stellenbosch 7600, South Africa

[5]Ashoka University, NH 44, Rajiv Gandhi Education City, Sonipat, 131029, India.

*Correspondence to: Andreas Wagner (email: andreas.wagner@ieu.uzh.ch).

**Table S1**: Average growth rates for the four replicate H and L populations in each generation. For these measurements, we grew all replicate H and L populations from each generation in 96-well plates on $LB^{Kan}$ for 24h. We recorded the optical density at 600nm ($OD_{600}$) every ~20minutes, and used this data to determine the growth rate of each population using GrowthRates[1]. *p*-values are Bonferroni-corrected and based on unpaired t-tests comparing H and L populations every generation. H populations have significantly higher growth rates in generations 2 and 3. The data show that the misfolding toxicity of highly expressed GFP is not likely to play a major role in our experiments, because otherwise H populations would show lower growth rates than L populations.

| Generation | H | L | *P* |
|---|---|---|---|
| 1 | 0.011273 | 0.010253 | 0.44 |
| 2 | 0.010425 | 0.008885 | 0.0036 |
| 3 | 0.01105 | 0.00917 | 0.024 |
| 4 | 0.010358 | 0.008868 | 0.1398 |
| 5 | 0.010468 | 0.008955 | 0.24 |
| 6 | 0.011328 | 0.009428 | 0.066 |

**Table S2:** Cyan fluorescence of H and L populations in each generation of directional evolution

| Population | Generation | Fold change in cyan fluorescence | Absolute cyan fluorescence |
|---|---|---|---|
| H1 | 1 | 1.17 | 246 |
| H2 | 1 | 1.14 | 241 |
| H3 | 1 | 1.83 | 343 |
| H4 | 1 | 1.08 | 233 |
| L1 | 1 | 1.60 | 181 |
| L2 | 1 | 1.24 | 157 |
| L3 | 1 | 1.22 | 156 |
| L4 | 1 | 1.63 | 183 |
| H1 | 2 | 1.03 | 226 |
| H2 | 2 | 4.59 | 749 |
| H3 | 2 | 1.24 | 257 |
| H4 | 2 | 3.26 | 554 |
| L1 | 2 | 8.69 | 658 |
| L2 | 2 | 5.46 | 441 |
| L3 | 2 | 2.62 | 250 |
| L4 | 2 | 11.25 | 830 |
| H1 | 3 | 1.09 | 234 |
| H2 | 3 | 4.62 | 754 |
| H3 | 3 | 0.48 | 145 |
| H4 | 3 | 3.68 | 616 |
| L1 | 3 | 7.06 | 548 |
| L2 | 3 | 6.30 | 497 |
| L3 | 3 | 9.53 | 714 |
| L4 | 3 | 12.68 | 926 |

| Population | Generation | Fold change in cyan fluorescence | Absolute cyan fluorescence |
|---|---|---|---|
| H1 | 4 | 0.03 | 78.6 |
| H2 | 4 | 0.90 | 207 |
| H3 | 4 | 2.41 | 429 |
| H4 | 4 | 1.29 | 264 |
| L1 | 4 | 5.45 | 440 |
| L2 | 4 | 8.75 | 662 |
| L3 | 4 | 0.52 | 109 |
| L4 | 4 | 7.31 | 565 |
| H1 | 5 | 3.26 | 554 |
| H2 | 5 | 5.23 | 843 |
| H3 | 5 | 5.54 | 889 |
| H4 | 5 | 3.79 | 631 |
| L1 | 5 | 11.03 | 815 |
| L2 | 5 | 23.28 | 1638 |
| L3 | 5 | 14.74 | 1064 |
| L4 | 5 | 14.35 | 1038 |
| H1 | 6 | 2.26 | 406 |
| H2 | 6 | 4.14 | 683 |
| H3 | 6 | 1.22 | 254 |
| H4 | 6 | 2.58 | 453 |
| L1 | 6 | 13.05 | 951 |
| L2 | 6 | 20.88 | 1477 |
| L3 | 6 | 15.67 | 1127 |
| L4 | 6 | 17.43 | 1245 |

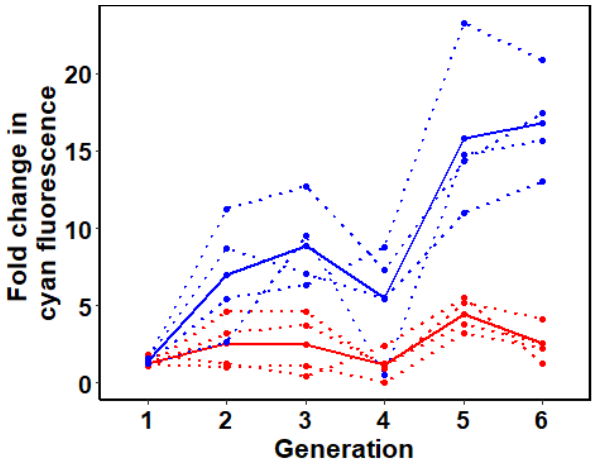


**Figure S1: Low expression level results in faster adaptation towards the new phenotype.** Fold change in cyan fluorescence, compared to the ancestor during six generations of directed evolution in four high expression ('H', red) and four low expression populations ('L', blue). Bold lines represent an average trajectory over the four replicates. L populations evolved faster than H populations ($p$=0.0005, n=4, general linear model, Methods).

**Table S3:** Average fold change in green and cyan fluorescence in high frequency variants engineered into the ancestral H and L genetic background. Data is based on two replicate fluorescence measurements for each engineered variant and for the ancestor (Methods).

| | Cyan fluorescence | | Green fluorescence | |
|---|---|---|---|---|
| | H | L | H | L |
| A65S | 3.43407 | 4.137543 | 0.092806 | 0.094614 |
| D133N | 0.89378 | 5.97646 | 0.93282 | 4.232363 |
| S147P | 2.539378 | 2.838488 | 0.911046 | 0.886537 |
| K156R | 1.374349 | 5.421649 | 1.33432 | 3.948028 |
| K166N | 1.378042 | 6.525172 | 1.247993 | 4.28698 |
| S175G | 5.028223 | 8.641581 | 3.990286 | 5.57894 |
| T203S | 1.238521 | 4.012543 | 0.774058 | 0.099869 |
| A206T | 1.114903 | 6.984794 | 1.09536 | 4.968531 |
| G232A | 1.480591 | 6.724141 | 1.450302 | 4.755498 |

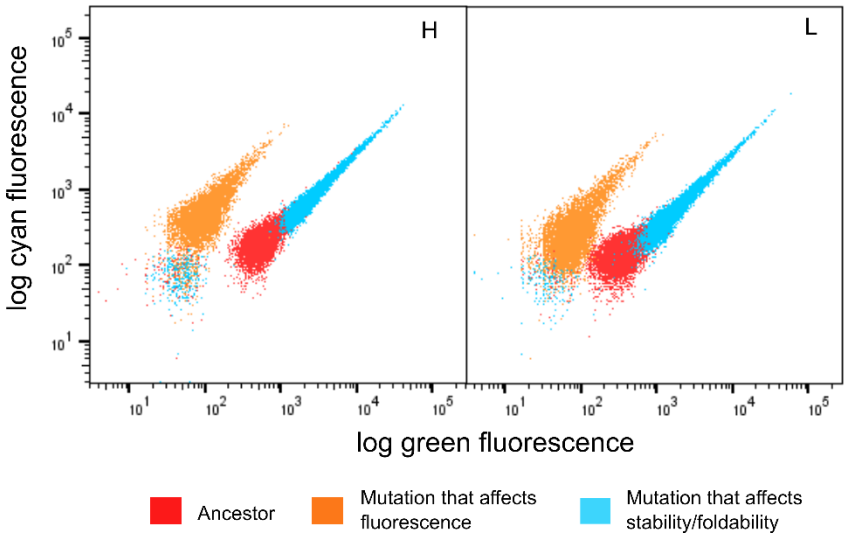


**Figure S2: Characterization of individual high frequency variants using two-fluorescence plots.** The horizontal and vertical axes show logarithmically (base 10) transformed green fluorescence ($\lambda_{ex}$=488nm, $\lambda_{em}$=530±15nm) and cyan fluorescence ($\lambda_{ex}$=405nm, $\lambda_{em}$=510±25nm) respectively. The two panels show the logarithm of green and cyan fluorescence in the H (left) and L (right) populations. The red cloud denotes the ancestral fluorescence, the orange cloud shows the fluorescence of the neo-functionalizing (color-shifting) mutation A65S[2] and the light blue cloud denotes the fluorescence of the mutation S175G which improves stability/foldability of GFP[3]. Note the shift to higher cyan and lower green fluorescence for the orange cloud (neofunctionalizing mutation), and enhancement of both green and cyan fluorescence for the light blue cloud.

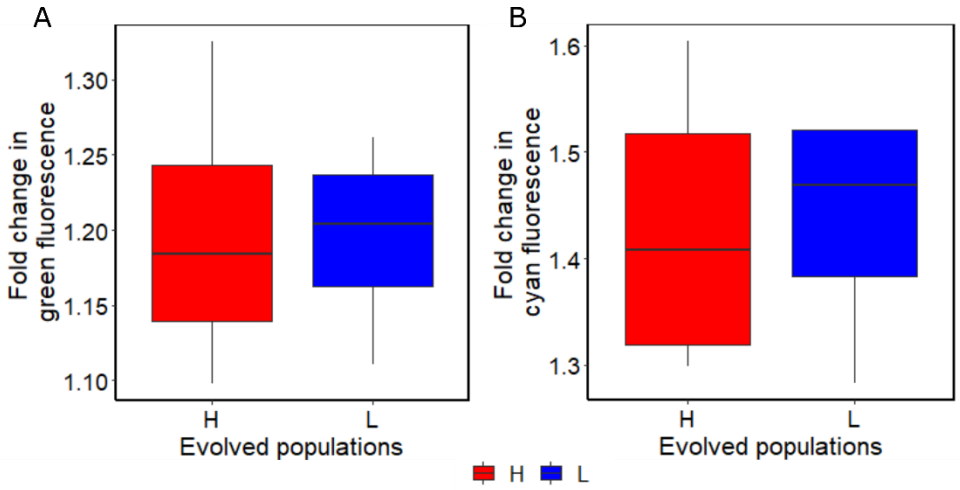


**Figure S3: Fold change in green and cyan fluorescence after three rounds of accumulation of cryptic variation.** Average fold change in green (panel A) and cyan (panel B) fluorescence during three generations of directed evolution under weak stabilizing selection did not differ significantly between H and L populations (unpaired t-tests, n= 3, $p$ = 0.77, Cohen's d= 0.32±2.28, t= -0.39, 95%CI= -0.83,0.69, df=2 and n= 3, $p$ = 0.99, Cohen's d= 0.04±2.26, t= 0.05, 95%CI= -0.54, 0.56, df=2 for green fluorescence; n= 3, $p$ = 0.43, Cohen's d=0.77±2.35, t=-0.95, 95%CI= -1.34,0.85, df=2 and $p$ = 0.55, Cohen's d=0.56±2.31, t=-0.69, 95%CI= -1.01,0.73, df=2.11 for cyan fluorescence of H and L populations respectively). Green fluorescence was statistically indistinguishable similar between L (blue) and H (red) populations after the last round of stabilizing selection (unpaired t test, n=4, $p$=0.96, Cohen's d= 0.03±1.72, t= 0.04, 95%CI= -0.14,0.15, df=2) and changed on average ~1.2-fold relative to the ancestor in both L and H populations (panel A). Cyan fluorescence was statistically indistinguishable between L (blue) and H (red) populations after the last round of stabilizing selection (unpaired t test, n=4, $p$=0.91, Cohen's d= 0.04±1.64, t= -0.1, 95%CI= -0.23,0.21, df=2). Fluorescence changed on average ~1.4-fold in both L and H populations after the last round of stabilizing selection (panel B). In both the panels the box represents the interquartile range while the solid line represents the median. The whiskers extend to the lowest and the highest values in the dataset.

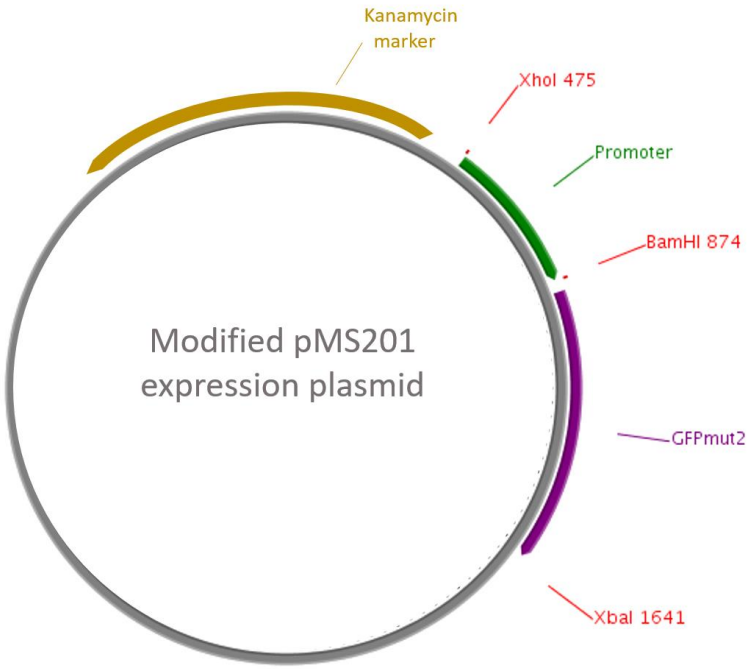


**Figure S4: Map of the modified plasmid pMS201 used in this study with the relevant regions.** The plasmid carries a kanamycin marker. The promoter region, both for high (H) or low ( L) expression, is flanked by XhoI and BamHI restriction sites. GFP is flanked by the BamHI site, which lies downstream of the promoter at the 5′ end, and by an XbaI site at the 3′ end.

**Table S4**: Mean and standard deviation of expression in absolute fluorescence units, measured in triplicate, for the ancestral H (acpP) and L (argW) promoters.

| H (acpP) | | L (argW) | |
|---|---|---|---|
| **Mean** | **Standard deviation** | **Mean** | **Standard deviation** |
| 1005 | 125.3 | 358 | 201 |
| 1041 | 122.1 | 349 | 208.2 |
| 1065 | 125.3 | 358 | 255.6 |

**Table S5:** Primer sequences

| Name | Sequence | Purpose |
|---|---|---|
| GFP_5F_synth | CGGGATCCTNTNGATTTAAGAAGGAGATATACATATGAGTAAAGGAGAAGAA | For the construction of expression plasmid |
| GFP_3R_synth | GCTCTAGAGCTTGCATGCCTG | |
| Mut_5F | GGATCCTTTAAGAAGGAGATATACAT | For mutagenesis |
| Mut_3R | CTTTCGTTTTATTTGATGCCTCTAGA | |
| Quiet_5F | gcagtcgaacatgtagctgactcaggtcacGGGCGTTAAGCAGGAAGAAGTTAC | For attaching universal sequence |
| Noisy_5F | gcagtcgaacatgtagctgactcaggtcacGATATAACGAGCCCCTCCTAAG | |
| Reverse | tggatcacttgtgcaagcatcacatcgtagCAGTCTTTCGACTGAGCCTTTC | |
| 1F | CAGCAGATCATGTCGAgcagtcgaacatgtagctgactcaggtcac | Primers with unique barcodes |
| 2F | TACGATCGTAGCTGCTgcagtcgaacatgtagctgactcaggtcac | |
| 3F | GTACACGCTGTGACTAgcagtcgaacatgtagctgactcaggtcac | |
| 4F | ACAGTGCGCTGTCTATgcagtcgaacatgtagctgactcaggtcac | |
| 5F | TATCAGCACGACATGCgcagtcgaacatgtagctgactcaggtcac | |
| 6F | ACTGCGAGATACACACgcagtcgaacatgtagctgactcaggtcac | |
| 7F | CATACATCGCGCAGTAgcagtcgaacatgtagctgactcaggtcac | |
| 8F | CACATATCAGAGTGCGgcagtcgaacatgtagctgactcaggtcac | |
| 9F | GATAGCTGCTAGCTGAgcagtcgaacatgtagctgactcaggtcac | |
| 10F | TCACTGTGTGTGTCTGgcagtcgaacatgtagctgactcaggtcac | |
| 11F | GATCTGTCGTGAGCGTgcagtcgaacatgtagctgactcaggtcac | |
| 12F | TATCTGAGCGCGAGCAgcagtcgaacatgtagctgactcaggtcac | |
| 13F | CTCACGTACGTCACACgcagtcgaacatgtagctgactcaggtcac | |
| 14F | CGTGTCGCGCATATCTgcagtcgaacatgtagctgactcaggtcac | |
| 15F | TGTGCACGACAGCAGTgcagtcgaacatgtagctgactcaggtcac | |
| 16F | ACGTCAGCACTGCTCTgcagtcgaacatgtagctgactcaggtcac | |

| | | |
|---|---|---|
| 17F | GACTGCACATGCACGAgcagtcgaacatgtagctgactcaggtcac | |
| 18F | CTACGAGACAGATCGCgcagtcgaacatgtagctgactcaggtcac | |
| 1R | CGAGTAGACAGTGACTtggatcacttgtgcaagcatcacatcgtag | |
| 2R | AGTGCACGAGCATATGtggatcacttgtgcaagcatcacatcgtag | |
| 3R | CGTGTGATGTCTACAGtggatcacttgtgcaagcatcacatcgtag | |
| 4R | ACACGCGATCTAGTGTtggatcacttgtgcaagcatcacatcgtag | |
| 5R | AGCGCAGTGTATAGTGtggatcacttgtgcaagcatcacatcgtag | |
| 6R | GATGAGAGAGCTCTCTtggatcacttgtgcaagcatcacatcgtag | |
| 7R | TATGCAGTCTGTCGTCtggatcacttgtgcaagcatcacatcgtag | |
| 8R | TCACGCTCTGTCTACTtggatcacttgtgcaagcatcacatcgtag | |
| 9R | GCGATCTAGCTATGTCtggatcacttgtgcaagcatcacatcgtag | |
| 10R | GCTATACGCTCGATACtggatcacttgtgcaagcatcacatcgtag | |
| 11R | GTGACTGCGTGTCTAGtggatcacttgtgcaagcatcacatcgtag | |
| 12R | CGAGTGCTAGACGATGtggatcacttgtgcaagcatcacatcgtag | |
| 13R | ACGCACTATGACGTCGtggatcacttgtgcaagcatcacatcgtag | |
| 14R | CTGTCAGAGTAGCTCGtggatcacttgtgcaagcatcacatcgtag | |
| 15R | TCGCGATAGTCTCGCAtggatcacttgtgcaagcatcacatcgtag | |
| 16R | GATGCTCGAGTCGATCtggatcacttgtgcaagcatcacatcgtag | |
| 17R | GTGAGCGCAGTGAGTAtggatcacttgtgcaagcatcacatcgtag | |
| 18R | ACAGATGTCTGTGCGCtggatcacttgtgcaagcatcacatcgtag | |
| A206T_5F | CAATCTACCCTTTCGAAAGATCCCAACGAAAAGAGAGAC | For engineering single mutants |
| A206T_3R | GAAAGGGTAGATTGTGTGGACAGGTAATGGTTGTCTG | |
| A65S_5F | CACTACTTTCTCGTATGGTCTTCAATGCTTTGCGAG | |
| A65S_3R | AGACCATACGAGAAAGTAGTGACAAGTGTTGGCCAT | |
| D133N_5F | TTTAAAGAAAATGGAAACATTCTTGGACACAAATTGGAAT | |

| | | |
|---|---|---|
| D133N_3R | ATGTTTCCATTTTCTTTAAAATCAATACCTTTTAACTCG | |
| G232A_5F | GATTACACATGCCATGGATGAACTATACAAATAAATGTC | |
| G232A_3R | CATCCATGGCATGTGTAATCCCAGCAGCTGTTA | |
| K156R_5F | CATGGCAGACAGACAAAAGAATGGAATCAAAGTTAACTTC | |
| K156R_3R | ATTCTTTTGTCTGTCTGCCATGATGTATACATTGTGTG | |
| K166N_5F | TAACTTCAATATTAGACACAACATTGAAGATGGAAGCGT | |
| K166N_3R | GTGTCTAATATTGAAGTTAACTTTGATTCCATTCTTTTG | |
| K3R_5F | TACATATGAGTAGAGGAGAAGAACTTTTCACTGGAGTT | |
| K3R_3R | GTTCTTCTCCTCTACTCATATGTATATCTCCTTCTTAAA | |
| S147P_5F | CAACTATAACCCACACAATGTATACATCATGGCAGACAAAC | |
| S147P_3R | CATTGTGTGGGTTATAGTTGTATTCCAATTTGTGTCCA | |
| S175G_5F | TGAAGATGGAGGCGTTCAACTAGCAGACCATTATCAACA | |
| S175G_3R | AGTTGAACGCCTCCATCTTCAATGTTGTGTCTAATTTTGAA | |
| T203S_5F | TACCTGTCCTCACAATCTGCCCTTTCGAAAGATCCC | |
| T203S_3R | GCAGATTGTGAGGACAGGTAATGGTTGTCTGGTAAAAGG | |

82

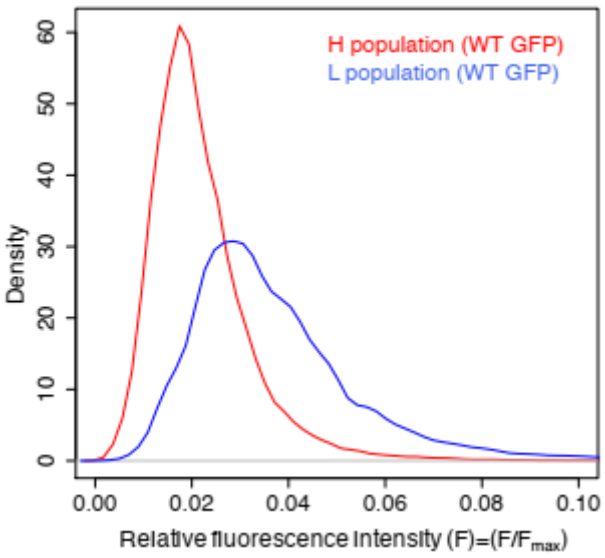


**Figure S5.** Relative fluorescence intensity of WT GFP in H population (in red) and L population (in blue). For both distributions, the relative fluorescence intensity is calculated by dividing all fluorescence values by the maximum fluorescence intensity in the population. The two distributions are statistically distinct ($p \sim 10^{-16}$, two-sided Kolmogorov-Smirnov test), and the H population has a low relative fluorescence intensity ($p < 10^{-16}$, two-sided Wilcoxon's rank-sum test, effect size calculated as $Z/\sqrt{N} = 0.75$).

**Supplementary information for the theoretical analysis and simulations of GFP evolution**

**Part 1: Fitness function**

To model the evolution of fluorescent proteins, we used the following fitness function, which equates the fitness of a GFP variant in our experiments to its fluorescence intensity. This intensity in turn results from various biophysical properties of GFP, as well as from instrument properties, as follows:

$$F_{protein} \equiv Fluorescence\ intensity = \mathrm{k}\emptyset\ [\,1 - 10^{-\varepsilon bC}] \qquad \text{(Equation S1)}$$

In Equation S1, *k* is an instrument proportionality constant, Ø is the quantum yield of the fluorescent protein, $\varepsilon$ is the extinction coefficient, *b* is the path length of the detector, and *C* is the concentration of protein in solution. Assuming $\varepsilon = 55900\ M^{-1}cm^{-1}$ for GFP 40, a 1 cm path length of the cuvette used to detect fluorescence, ~50000 sorted cells in a 500 µl volume of PBS buffer as in our experimental protocol, and a copy number of ~20 GFP-expressing pMSs201 plasmids per cell (*E.coli* promoter collection, GE healthcare Dharmacon), the product $\varepsilon bC$ calculates as ~ $55900\ M^{-1}cm^{-1} \times 1\ cm \times \frac{5\times10^{4}\times20\ copies}{5\times10^{-4}\ L} \times \frac{1mol}{6.02\times10^{23}} \sim 1.8 \times 10^{-10}$. This small value allowed us to expand $1 - 10^{-\varepsilon bC}$ as $1 - 1 + \varepsilon bC + O(\varepsilon bC^2) \sim \varepsilon bC$. This expansion helps to simplify Equation S1, which yields the following simpler instrument-independent fitness function:

$$F_{protein}(instrument - independent\ ) = \frac{F_{protein}}{\mathrm{k}\ I_0 b} = \varepsilon\emptyset C = \xi C \qquad \text{(Equation S2)}$$

In this expression, the product of the extinction coefficient (ε) and quantum yield (Ø) define the fluorescence output ξ of GFP. This quantity describes the brightness of a single GFP molecule when it absorbs light with a specific wavelength. The fluorescence intensity is then the product of the fluorescence output of a single GFP molecule and the concentration *C* of GFP in solution. To apply our fitness function to single cells, we assumed that $C = A \times f(\Delta G)$ where *A* is the basal expression of GFP in the cell, and f(Δ*G*) is a function that relates the fluorescence intensity of a single GFP molecule to its free energy of folding stability (Δ*G*). This gives rise to Equation S3 which is the cellular fitness function we use for fluorescence intensity:

$$F_{protein} = \xi_{protein}\, A\, f(\Delta G) \qquad \text{(Equation S3)}$$

As previously proposed and motivated by experimental data [4], we used a sigmoidal function to model the dependence of f(Δ*G*) on $\Delta G$ :

$$f(\Delta G_{mut}) \sim \frac{e^{-(\Delta G_{mut} - \Delta G_{WT})}}{1+e^{-(\Delta G_{mut} - \Delta G_{WT})}} \qquad \text{(Equation S4)}$$

In equation S4, $\Delta G_{WT}$ and $\Delta G_{mut}$ are the free energy of folding of wild-type (ancestral) GFP and of a given mutational variant. We used $\Delta G_{WT} = -12$ kcal/mol which is in the range of experimentally reported free energies of folding for GFP (~-10 to -16 kcal/mol)[5-7]. For mathematical brevity we refer to $f(\Delta G_{mut})$ as the *stability factor* of GFP molecules, to distinguish it from protein stability, $\Delta G$.

**Part 2: The theoretical relationship between selection strength, protein stability, and expression level in GFP populations**

To better understand how selection strength and expression level can affect the stability of selected GFP molecules in each generation, we first use a simple theoretical model. In our experiments, we selected cells for survival whose fluorescence intensity lay in the top 0.05% of the population. Mathematically, we represent this relative fluorescence intensity as $F_{rel,}$ which equals

$$F_{rel} = \frac{F_{max} - F_{sel}}{F_{max}} = 1 - \frac{F_{sel}}{F_{max}}\,, \qquad \text{(Equation S5)}$$

where $F_{max}$ denotes the maximum fluorescence intensity among all cells in a population and $F_{sel}$ denotes the fluorescence intensity of the cells that survive selection. The criterion that a cell has to fulfill in order to be selected then becomes

$$F_{sel} > F_{max}\,(1 - F_{rel}) \qquad \text{(Equation S6)}$$

Note that our strong selection implies a small value of $F_{rel}$ ($F_{rel}$ =0.0005) because only those cells survive whose fluorescence ($F_{sel}$) is very close to $F_{max}$. Replacing fluorescence intensity in Equation S6 with its molecular determinants from Equation S3 yields

$$\left[\xi_{protein}\, A\, f(\Delta G)\right]_{sel} > \left[\xi_{protein}\, A\, f(\Delta G)\right]_{max} (1 - F_{rel}) \qquad \text{(Equation S7)}$$

Here, the subscripts *max* and *sel* denote cells with maximum fluorescence, and cells that are selected for survival. If GFP is equally abundant in all cells (i.e., $A_{max} = A_{sel}$), and the fluorescence output of all selected molecules is the same, $((\xi_{protein})_{max} = (\xi_{protein})_{sel})$, Equation S7 simplifies to

$$\frac{f(\Delta G)_{sel}}{f(\Delta G)_{max}} > 1 - F_{rel} \qquad \text{(Equation S8)}$$

We refer to the ratio $\frac{f(\Delta G)_{sel}}{f(\Delta G)_{max}}$ as the normalized GFP stability factor, that is the stability factor of surviving protein molecules divided by the maximal stability factor among all proteins in the population.

Equation S8 states that in the absence of any variation in GFP abundance in a population, the strength of selection (1-$F_{rel}$) determines a lower bound for this normalized stability factor, i.e., only GFP proteins above a given value of $\frac{f(\Delta G)_{sel}}{f(\Delta G)_{max}}$ will survive selection. The stronger selection becomes (the smaller $F_{rel}$ becomes), the larger the value of $\frac{f(\Delta G)_{sel}}{f(\Delta G)_{max}}$ that is necessary for a cell to survive selection.

We next examine the more realistic case that GFP abundance varies among cells in a population, leaving the main assumption that $\xi_{protein}$ is the same for all selected variants in place. In this case, we can rewrite Equation S7 as

$$\frac{f(\Delta G)_{sel}}{f(\Delta G)_{max}} \times \frac{A_{sel}}{A_{max}} > (1 - F_{rel}) \qquad \text{(Equation S9)}$$

Equation S9 governs the relationship between protein stability and abundance that GFP molecules surviving selection must have. That is, the product of the normalized stability factor and the (analogously defined) normalized abundance of surviving GFP molecules must be higher than the strength of selection (1-$F_{rel}$). In

sum, this theoretical model illustrates the idea of stability compensation, that is, GFP variants with low stability can survive selection, as long as their low stability is compensated by increased protein abundance.

**Part 3: Details of evolutionary simulations**

**3.1. Overview of the simulation model**

We simulated the evolution of GFP populations towards cyan fluorescence by developing a mechanistic model that relates the fluorescence intensity of a GFP expressing cell to the biophysical properties of GFP, namely its fluorescence output, stability, and cellular abundance (Equation $S_3$). Figure S6 illustrates the structure of this model. Any one cell expresses GFP molecules whose fluorescence output, folding stability, and cellular abundance can vary among cells in the same population (Figure S6A). Only protein stability and fluorescence output can change by mutation and are thus evolvable. In contrast, the abundance of GFP can vary among individuals in a population due to gene expression noise, but we do not allow it to be affected by a mutation. This is motivated by our experimental design, which does not subject the GFP promoter to mutagenesis. Our simulations assume a population with a constant size of 10000 cells, all of which initially encode wild-type GFP with an initial fixed stability and protein fluorescence output (Figures S6B-C, also see sections 3.2 and 3.3).

We then start the first round of mutation and selection. For the mutation step, we assume that individual GFP-coding genes experience one point mutation per round of mutation, which corresponds roughly to the amino acid mutation rate of ~0.8 in our experiments. Each such mutation changes the stability and fluorescence output of a GFP variant according to distributions estimated from a library of GFP mutants and from FoldX stability calculations, as detailed in sections 3.2 (Equation S10) and 3.3 (Equation S11, Figure S6B-C). Because we model the evolution of a population of 10000 cells, we repeat this mutational step 10'000 times for ancestral GFP. We convert the thermodynamic stabilities of each resulting variant (in kcal/mol), to a stability factor that ranges from 0 to 1, corresponding to minimum and maximum stability in our model (Equation $S_4$).

We allow cells that carry a GFP variant with a given fluorescence output and stability to vary in their expression level due to gene expression noise. To model such noise, we randomly assign GFP abundance

values to each of the 10000 cells that are sampled at random from a log-normal distribution with mean $\mu$ and standard deviation $\sigma$, such that $\sigma = 0.3\mu$. This scaling is motivated by an experimentally observed relationship between the mean and the standard deviation of gene expression driven by 1523 *E. coli* promoters[8] (Figure S6F). We then calculate the fluorescence intensity of each cell by multiplying the values of fluorescence output, the stability factor and GFP abundance. The end result is a population of 10000 cells that differ in their GFP fluorescence intensities because of variation in these three factors (Figure S6D).

We next describe how we model selection in a way that allows us to keep population size constant. Of the 10000 cells in a modeled population, we select those cells for survival whose fluorescence intensity lies in a top fixed percentile of the population (Figure S6G). For example, to select the top 1% of fluorescing cells we allow the top 100 fluorescing cells from this population to survive. To create a population of survivors that is equal in size to that of the pre-selection population, we repeat the selection process for replicates of our original population of 10000 cells until we collect 10000 survivors. For example, if we target the top 1% of cells to survive selection, we repeat selection for 100 replicate populations of 10000 cells to collect 10000 (=100*100) survivors. To generate a replicate population, we use bootstrapping, i.e., we sample 10000 value with replacement from the pre-selection distribution of each GFP property (protein stability, abundance and fluorescence output). This process mimics the selection of cells by fluorescent activating cell sorting (FACS).

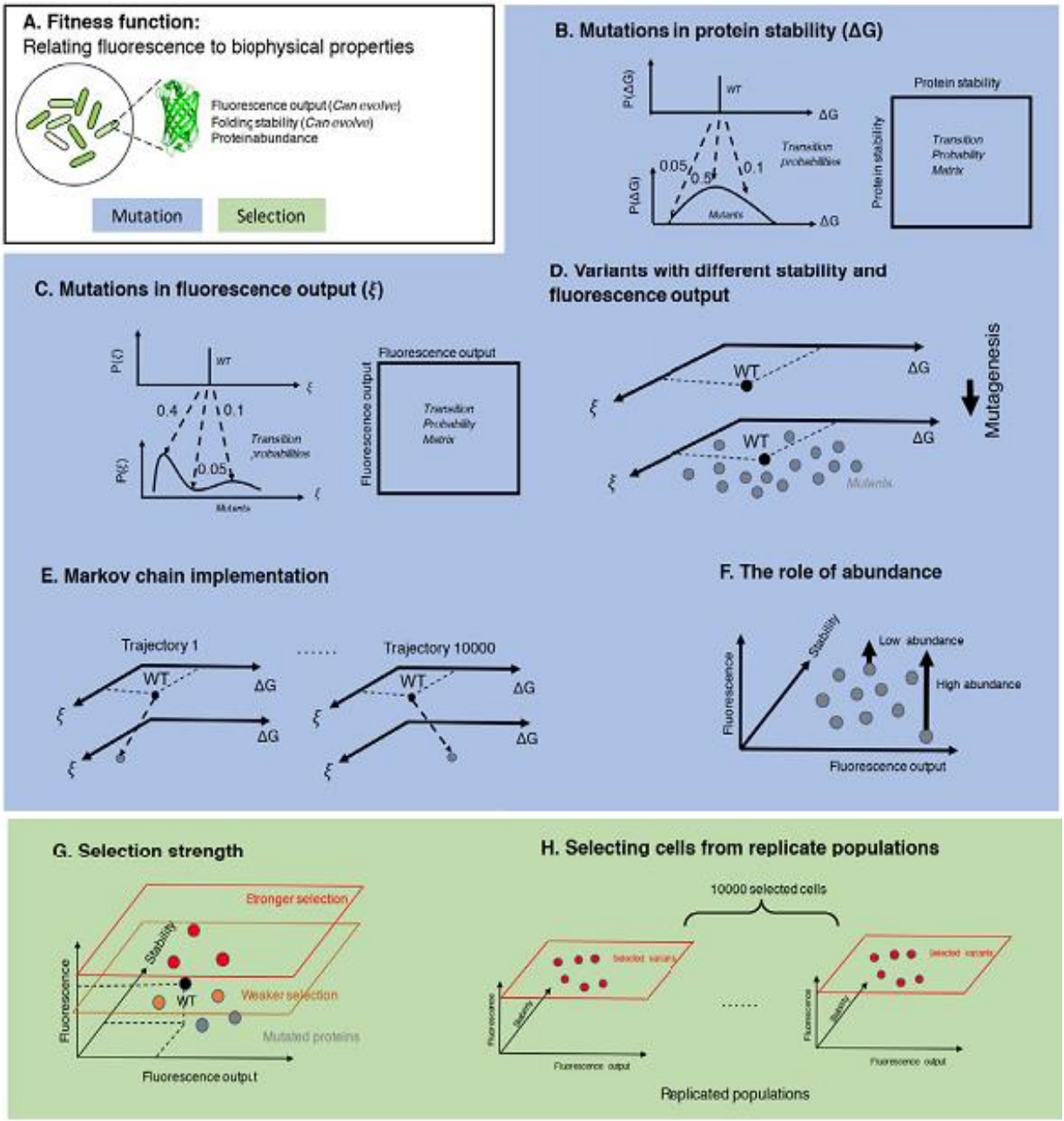


**Figure S6. The principles behind our simulation model: A.** We simulate the evolution of the fluorescence intensity of GFP populations. Each cell expresses GFP molecules whose fluorescence output and folding stability can be altered by mutations. GFP abundance may vary among cells due to gene expression noise, but we do not allow it to evolve. Evolution proceeds through cycles of mutation and selection. The details of mutation and selection cycles are shown in blue and green boxes, respectively. We partition the range of sensible phenotypic values for both folding stability and fluorescence output into $n$=100 equidistant bins, ranging from the lowest value of the property of interest to its highest value. We refer to these bins as states

in our model. **B.** To model the effect of mutations on protein stability, we sample a distribution of the stability of single point mutations calculated by the free energy estimator, FoldX. We assume that the stability of a pre-mutated protein (shown here as WT for 'wild-type'), can change to different values with different transition probabilities (Equation S13). Calculating these transition probabilities for all possible stability values yields a transition probability matrix. **C.** To model the effect of single mutations on fluorescence output, we sample an empirically known distribution of the fluorescence intensity of a GFP library with single point mutations of our ancestral GFP genotype (Equation S12). Similar to protein stability, the fluorescence output of a pre-mutated protein can change to different values, which yields a matrix based of mutational transition probabilities. **D.** Collectively, mutations in folding stability and fluorescence output generate variants with different values of these properties in the population. **E.** To simulate the mutation process efficiently, we use a formalism akin to Markov chains to simulate 10000 evolutionary trajectories starting from ancestral GFP. We advance the 10000 trajectories by one mutational step and record the resulting states, which correspond to 10000 values for protein stability and fluorescence output. **F.** Protein abundance can change the fluorescence intensity of our simulated cells. Because of gene expression noise, some cells will have a higher fluorescence intensity than others, even if they express GFP variants with identical stability and fluorescence output. **G.** We implement selection by letting a fixed percentile of cells with high fluorescence survive. Stronger selection leads to the survival of molecules with higher stability and fluorescence output. **H.** To keep the population size constant in our simulations, we create replicates of pre-selected population and selected 10,000 cells from these replicate populations. The surviving cells express different values for protein stability and fluorescence output that we use as starting points for the second round of evolution.

### 3.2. Calculating stability distribution of single GFP mutants and the stability factor

To model how mutations affect protein stability, we used FoldX[11] to calculate the folding free energy change ($\Delta\Delta G$) of all single point mutations, based on the crystal structure of GFP (PDB ID: 1EMA). These computationally predicted values had a bi-Gaussian distribution, and we used the following equation to model this distribution:

$$P(\Delta\Delta G) = P_1 \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\left(\frac{\Delta\Delta G-\mu_1}{2{\sigma_1}^2}\right)} + (1-P_1)\frac{1}{\sigma_2\sqrt{2\pi}} e^{-\left(\frac{\Delta\Delta G-\mu_2}{2{\sigma_2}^2}\right)} \quad \text{(Equation S10)}$$

We estimated the parameters of this distribution by fitting mixture models to the distributions of $\Delta\Delta G$ for single point mutations (estimated by FoldX) using the mixtools package in R[12]. This procedures yielded $\mu_1 = 0.54, \sigma_1 = 0.98,\ \mu_2 = 2.05,\ \sigma_2 = 1.91$, P1=0.49, and P2=1-P1=0.51.

### 3.3. Calculating the fluorescence intensity distribution of single GFP mutants

To generate protein variants with different fluorescence intensities and to simulate fluorescence evolution realistically, we need to know the distribution of the effect of mutations on fluorescence intensity and protein stability in GFP. We evaluated the distribution of fluorescence intensities from the logarithmically transformed intensities of a library of single point GFP mutants[10]. These intensities have a bimodal distribution whose first mode is centered on ~1.52 ($\log_{10}$ intensity units) which corresponds to non-fluorescent (or 'dark') GFP variants. The second mode is centered on ~ 3.37 and mainly represents the nearly-neutral variants where fluorescence intensity is close to that of WT GFP. We fitted the following bi-lognormal distribution to these experimentally measured intensities:

$$P(\log(\xi)) = P_1 \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\left(\frac{\log(\xi)-\mu_1}{2{\sigma_1}^2}\right)} + (1-P_1)\frac{1}{\sigma_2\sqrt{2\pi}} e^{-\left(\frac{\log(\xi)-\mu_2}{2{\sigma_2}^2}\right)} \qquad \text{(Equation S11)}$$

We estimated the parameters of Equations S11 by fitting mixture models to the distributions of fluorescence intensities and folding stabilities, using the mixtools package in R[12]. This procedure yielded $\mu_1 = 1.52,\ \sigma_1 = 0.18, \mu_2 = 3.37,\ \sigma_2 = 0.37$, $P_1$=0.23 and $P_2$=1-$P_1$=0.77.

### 3.4. Markov chain implementation of the mutation step

To simulate how mutation changes fluorescence output and folding stability, we used a formalism akin to discrete time Markov chains, which describes how single mutations change these properties. To this end, we first partitioned the range of sensible phenotypic values for both fluorescence output and folding stability into *n*=100 equidistant bins, ranging from the lowest value of the property of interest to its highest value. These bins correspond to the possible states of fluorescence output and stability of individual proteins expressed in a cell.

More specifically, we allowed protein stability to range between $\Delta G$ =-20 kcal/mol and $\Delta G$=0 kcal/mol, and we subdivided this range into bins of width 0.2 kcal/mol. This provides a wide dynamic range for changes in protein stability. For example, the average stability of *E.coli* proteins is ~ -7.1 kcal/mol, with 99% of proteins having stabilities higher than -4 kcal/mol[13]. Similarly, we allowed fluorescence output to range between log($\xi$)=0 to log($\xi$)=10 in bins of width 0.1 log fluorescence intensity. This range also covers the naturally observed range of log(fluorescence intensity), which lies between ~1 to ~5 log units[4]. During a simulation, the fluorescence output and stability of a GFP molecule expressed in a cell is in one of these *n* states, which we refer to as $n_{stab}$ and $n_{FL}$. For instance, if the stability of a GFP variant is in state $n_{stab}$=1, it lies in the bin covering the interval between -20 kcal/mol and -19.8 kcal/mol). Likewise, if fluorescence output is in state $n_{FL}$ =1, the logarithm of its value lies in the 0 to 0.1 bin.

At the beginning of a simulation, we initialize all members of a population (10000 cells) to be in the state of the ancestral GFP molecule that our experiments started with. The measured stability of ancestral GFP is ~ -12 kcal/mol and corresponds to $n_{stab}$=60. The experimentally observed log fluorescence intensity of ancestral GFP is ~ 3, which corresponds to $n_{FL}$=30.

To subject this ancestral GFP to mutation, we need to calculate the post-mutation states $n_{stab}$ *and* $n_{FL}$ of GFP for all 10000 cells. For this purpose, we need to know the transition probabilities between all possible protein stability ($n_{stab}$) and fluorescence output states ($n_{FL}$). We next describe how to calculate these transition probabilities.

To model mutational changes in protein stability, we calculated the transition probability between different stability values as a result of a single mutation with the following equation, which is identical to equation S10 but centered on the stability of pre-mutated protein ($\Delta G_i$):

$$P_{ij}^{Stab} = P_1 \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\left(\frac{(\Delta G_j - \Delta G_i) - \mu_1}{2{\sigma_1}^2}\right)} + (1 - P_1) \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\left(\frac{(\Delta G_j - \Delta G_i) - \mu_2}{2{\sigma_2}^2}\right)} \qquad \text{(Equation S13)}$$

From this equation, a single mutation will most likely decrease the stability of the pre-mutated protein ( $\Delta G_i$) by ~1.27 kcal/mol (the average of the two modes centered on $\mu_1$=0.54 and $\mu_2$=2.05 kcal/mol, with P1=0.49 and P2=0.1).

For fluorescence output, we calculated the transition probability of the $i^{th}$ state to the $j^{th}$ state as a result of a single point mutation with the following equation:

$$P_{ij}^{FL} = P_1 \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\left(\frac{(\log(\xi_j)-\log(\xi_i))-\mu_1}{2{\sigma_1}^2}\right)} + (1-P_1)\frac{1}{\sigma_2\sqrt{2\pi}} e^{-\left(\frac{(\log(\xi_j)-\log(\xi_i))-\mu_2}{2{\sigma_2}^2}\right)} \quad \text{(Equation S12)}$$

This equation is identical to Equation S11 except that its probability densities is centered on $\log(\xi_i))$, i.e., the GFP's fluorescence output of the pre-mutated state. The equation reflects the assumption that the mutation-induced change in GFP's fluorescence output ($\log(\xi_i)$) follows a bimodal distribution. The first mode ($\mu_1$) of fluorescence output is centered on fluorescence intensities that are indistinguishable from a negative control population (log10 intensity =$1.52 \pm 0.18$). These are non-fluorescing ('dark') molecules. The second mode ($\mu_2$), is centered on the state of interest ($\log(\xi_i)$), and represents nearly-neutral variants (i.e., variants whose fluorescence are slightly higher and lower than that of pre-mutated GFP). Biologically speaking, single point mutations in any GFP molecule can create two sub-populations, one sub-population that is non-fluorescent and the other sub-population whose fluorescence is very similar to the pre-mutated GFP molecule. The empirically estimated values of $P_1$=0.23 and $P_2$=0.77 show that any GFP molecule can become non-fluorescent with a probability of ~0.23 or stay fluorescent with a probability of ~0.77.

We used Equations S12 and S13 to generate two 100x100 transition probability matrices, $M_{ij}^{stab}$ and $M_{ij}^{FL}$ for protein stability and fluorescence output, respectively. We note that in these matrices, the sum of transition probabilities in each row equals 1, i.e.,

$$\sum_{j=1}^{n} P_{ij}^{FL} = \sum_{j=1}^{n} P_{ij}^{Stab} = 1 \quad \text{(Equation S14)}$$

This means that a mutation will cause the stability or fluorescence of the pre-mutated GFP to either remain unchanged, or to become changed to one of the other 99 (=100-1) states. If a Markov chain is in the $i^{th}$ state, the next state is sampled from a multinomial distribution whose probabilities are expressed by the $i^{th}$ row of the corresponding transition probability matrix.

### 3.5. Modeling protein abundance

To model the variation of protein abundance in each population, we assumed that the standard deviation of protein abundance ($\sigma_A$) scales linearly with the average protein abundance ($\mu_A$) as $\sigma_A \sim 0.3\ \mu_A$. This scaling is

derived from experimental data on the relationship between mean and standard deviation of fluorescence intensity for 1523 *E. coli* promoters studied by Silander et al.[9] (Figure S7). Note that the mean protein abundance distinguishes the H and L populations, so that GFPs in H populations are on average three-fold more abundant than their counterparts in L populations.

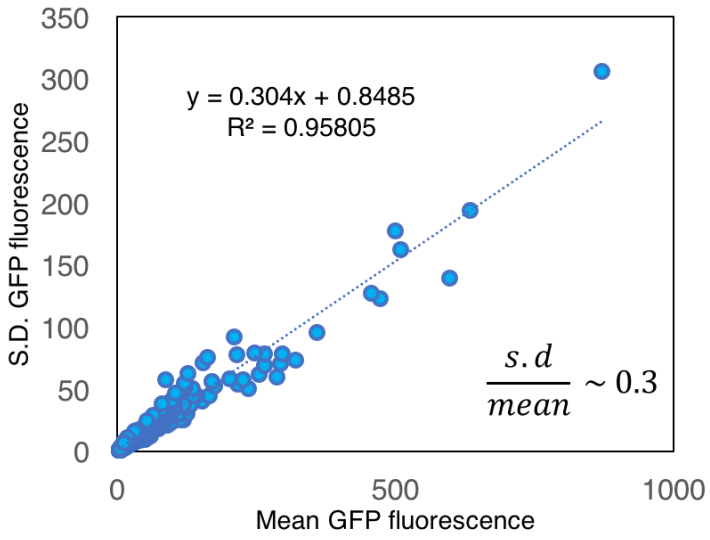


**Figure S7:** The relationship between the mean and standard deviation of fluorescence intensity of GFP expressed from 1523 natural *E. coli* promoters[9]. We used this relationship to model protein abundance variation caused by gene expression noise in our evolving GFP populations.

### 3.6. Calculating the fluorescence intensity of simulated cells

We calculated the fluorescence intensity of a cell containing a given GFP variant using equation S3, which depends on the values of GFP protein abundance (A), fluorescence output (ξ), and folding stability ($\Delta G$). We used the randomly assigned abundances for our 10000 cells in the population as described in section 3.2, and obtained the values of fluorescence output and stability from the post-mutation values as described in section 3.5.

### 3.7. Implementing selection

We allowed only those cells to survive into the next generation whose fluorescence intensity lay in a fixed percentile of the population's fluorescence distribution. In the simulations we used 4 different percentiles, i.e., 0% (no selection), 10%, 50%, and 90% to find out how selection and expression affect the accumulation of mutations with destabilizing effects.

### 3.8. Simulating cycles of mutation and selection

To initialize our simulations, we assumed the following initial state of our populations: $n_{Stab}$=60, and $n_{FL}$=30. Just in our experiments, we then applied six cycles of mutation and selection, except that the simulations allowed us to simulate different strengths of selection (as described in section 3.6).

After the selection step of any given round, we calculated the average stability and fluorescence output of the surviving GFP variants.

To keep population size constant in our simulations, we created replicates of pre-selected population and selected 10,000 cells from these replicate populations. The number of such replicate populations depends on the stringency of selection. For example, if only the top 1% of cells from pre-selected population are to survive selection, we needed to generate 100 replicate populations to obtain the same number of surviving cells as our population size prior to selection. We generated the GFP stability and fluorescence output values for the cells in a replicate population by sampling with replacement 10000 values from the pre-selection distribution of protein stability and fluorescence output. In addition, we assigned protein abundance values to cells of each replicate populations from the log-normal distribution described in section 3.5.

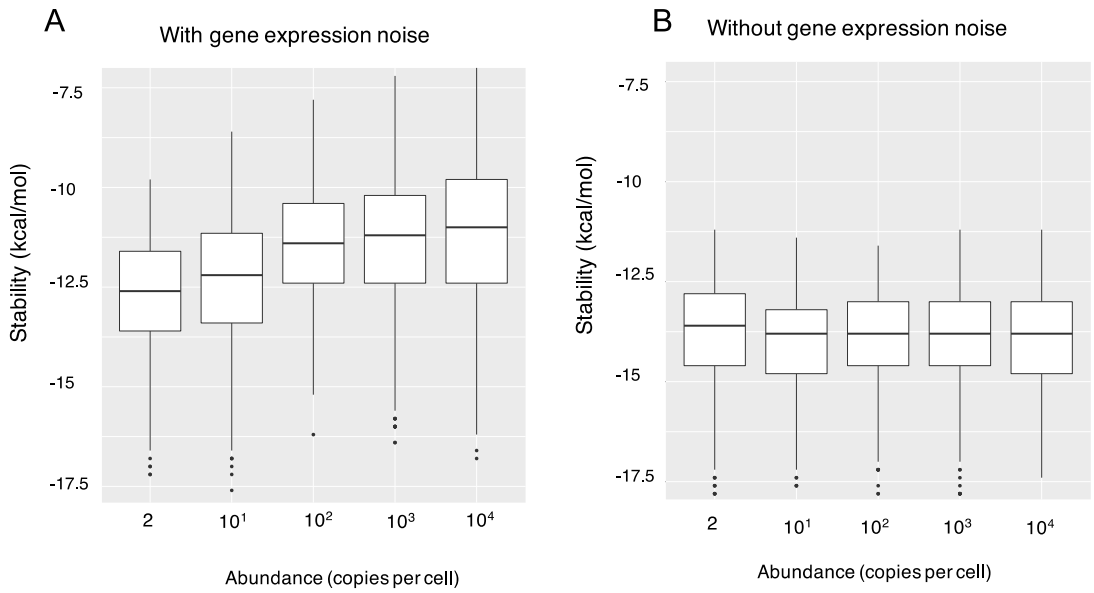


**Figure S8.** The stability of selected GFP molecules for different mean protein abundance levels in the simulations with (**A**) and without (**B**) gene expression noise (protein abundance distribution). In panel A, protein abundance distribution for each population is modeled by a log-normal distribution where the mean expression level is indicated on the x-axis, and the standard deviation ($\sigma_A$) scales linearly with the average protein abundance ($\mu_A$) as $\sigma_A \sim 0.3\ \mu_A$. In panel B, all cells have the same mean expression level indicated on the x-axis, with zero standard deviation. For both panels, we evolved simulated GFP populations with a selection strength of 0.9, i.e., only 10% of GFP molecules survive selection. We selected n=10,000 independent cells from 10 replicate pre-selection populations each having 100,000 cells. The boxes in the box plots range from the first quantile to the third quantile. The median is represented by a line across the box. Whiskers show 1.5 times of the interquartile range. The circles located above the top whisker or below the bottom whisker are outliers whose values are either higher or lower than 1.5 times the interquartile range (third quartile – first quartile), respectively.

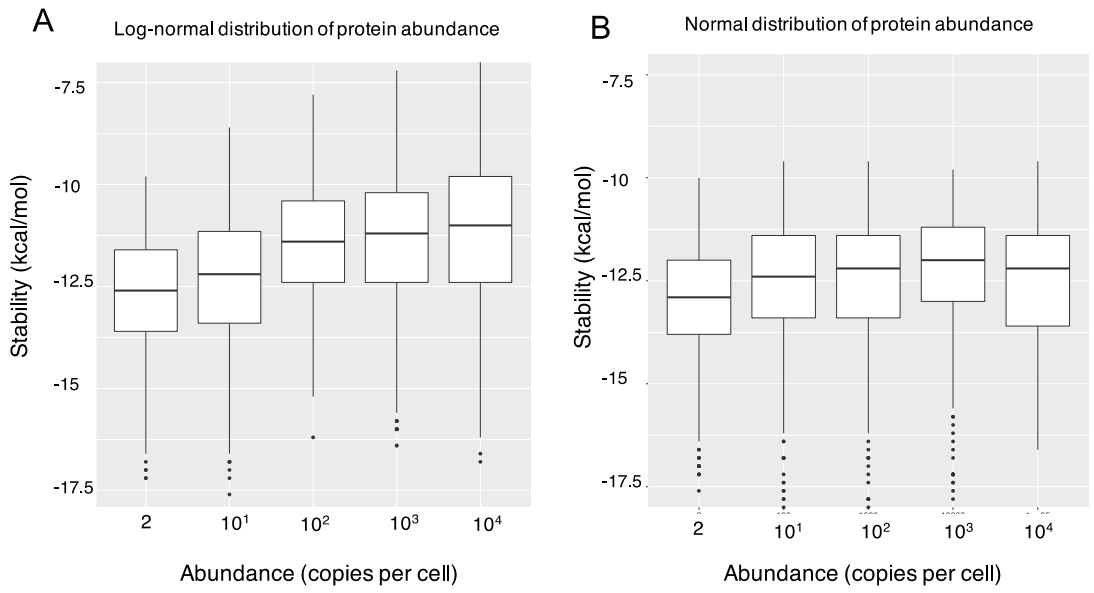


**Figure S9.** The stability of selected GFP molecules for different mean protein abundance levels in the simulations. **A.** Gene expression noise (protein abundance distribution) is modeled by a log-normal distribution. **B.** Gene expression noise (protein abundance distribution) is modeled by a normal distribution. For both panels the standard deviation of protein abundance ($\sigma_A$) scales linearly with the average protein abundance ($\mu_A$) as $\sigma_A \sim 0.3\ \mu_A$. For both panels, we evolved simulated GFP populations with a selection strength of 0.9, i.e., only 10% of GFP molecules survive selection. We selected n=10,000 independent cells from 10 replicate pre-selection populations each having 100,000 cells. The boxes in the box plots range from the first quartile to the third quartile. The median is represented by a line across the box. Whiskers show 1.5 times of the interquartile range. The circles located above the top whisker or below the bottom whisker are outliers whose values are either higher or lower than 1.5 times the interquartile range (third quartile – first quartile), respectively.

**Part 4: enforcing an association between variance and skewness of a gamma distribution of protein abundances**

A gamma distributions is defined by its 'shape' parameter ($\alpha$) and its 'rate' parameter ($\beta$). Its mean, variance and skewness are equal to $\frac{\alpha}{\beta}$, $\frac{\alpha}{\beta^2}$, and $\frac{2}{\sqrt{\alpha}}$. To increase mean, variance and skewness at the same time, we decided to decrease both $\alpha$ and $\beta$ parameters, such that the $\beta$ parameter decreases to a greater extent than the $\alpha$ parameter. Specifically, we used

$$\alpha = \frac{1}{i} \qquad \text{(Equation S15)}$$

$$\beta = \left(\frac{1}{i}\right)^2 \qquad \text{(Equation S16)}$$

where $i \in \{10, 100, 500, 1000, 5000\}$. Mean protein abundances for these values in our simulations equal 111, 852, 1424, and 11471, respectively.

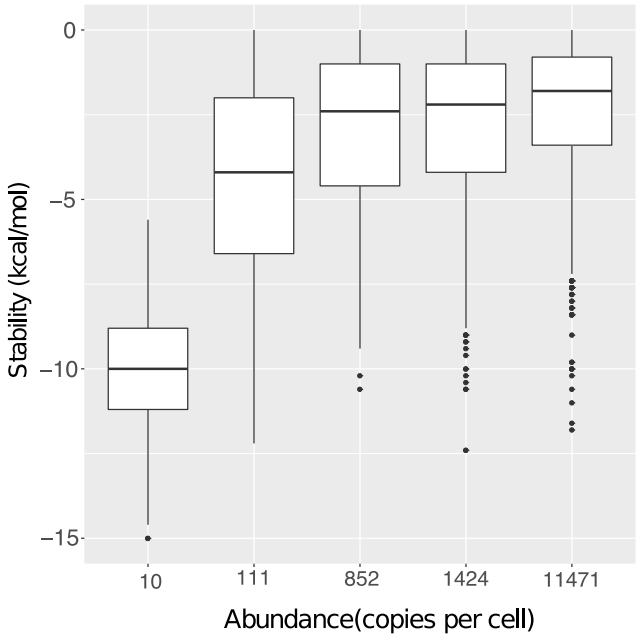


**Figure S10.** Stability of selected GFP molecules for different mean protein abundance levels in the simulations. We assume that protein abundances are gamma distributed with means indicated on the x-axis. We used Equations S15 and S16 to ensure that both variance and skewness of these distributions increase as their mean increases. For all population we evolved simulated GFP populations with a selection strength of 0.9, i.e., only 10% of GFP molecules survive selection. We selected n=10,000 independent cells from 10 replicate pre-selection populations each having 100,000 cells. The boxes in the box plots range from the first

quartile to the third quartile. The median is represented by a line across the box. Whiskers show 1.5 times of the interquartile range. The circles located above the top whisker or below the bottom whisker are outliers whose values are either higher or lower than 1.5 times the interquartile range (third quartile – first quartile), respectively.

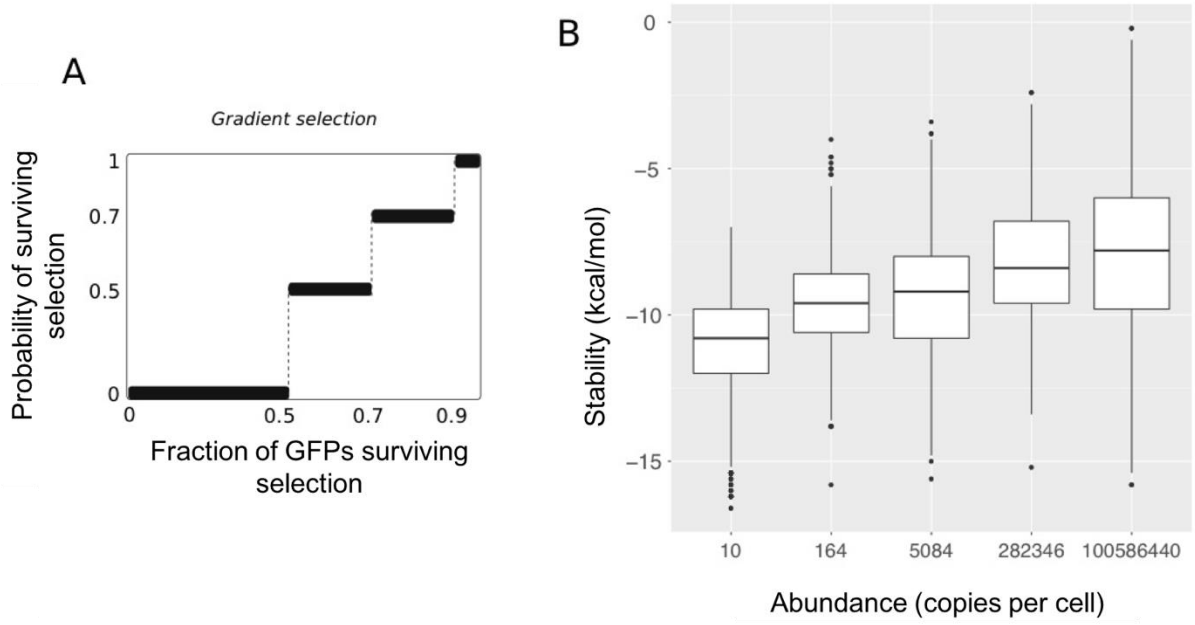


**Figure S11. A.** A probabilistic selection scheme where GFP molecules with different fluorescence intensities are allowed to survive with different probabilities. In this scheme, we allowed 100% of GFP molecules whose fluorescence intensity was in the top 10% of a population to survive, 70% of GFP molecules whose fluorescence intensity was in the top 30%, and 50% of molecules whose fluorescence intensity was in the top 50%. Selection strength on the x-axis shows the fraction of cells surviving selection with the probability of surviving selection on the y-axis. **B.** Stability of surviving GFP molecules for different mean protein abundances in simulations using the probabilistic selection scheme described in panel A. We selected n=10,000 independent cells from 10 replicate pre-selection populations each having 100,000 cells. The boxes

in the box plots range from the first quantile to the third quantile. The median is represented by a line across the box. Whiskers show 1.5 times of the interquartile range. The circles located above the top whisker or below the bottom whisker are outliers whose values are either higher or lower than 1.5 times the interquartile range (third quartile – first quartile), respectively.

**Part 5: Probabilistic selection based on a cell's fluorescence intensity**

To investigate whether selecting cells in proportion to fluorescence intensity can also lead to the accumulation of destabilizing mutations in H populations, we performed new simulations. Briefly, we generated two ancestral populations with 50,000 cells each to represent H and L populations. The distribution of protein abundance was log-normal in these populations, the only difference being that the H population had a higher mean and standard deviation compared to the L population. We used a mean abundance of 100 and 300 GFP molecules, and a standard deviation of 30 and 90 GFP molecules for the L and the H populations, respectively. We also used a Gaussian distribution of protein stabilities (ΔG) with ΔGmean=-12 kcal/mol (corresponding to that of WT GFP), and ΔGsd=2 kcal/mol. We chose these values from previously measured estimates of protein folding stabilities (see: PNAS 101.46 (2004): 16192-16197; Journal of molecular biology 369.5 (2007): 1318-1332). We then sampled 10,000 cells from these populations with replacement, such that the sampling probability increased linearly from the first to the 100th percentile of fluorescence intensity,

$$\mathrm{P(sampling\ from\ percentile\ i)} = \frac{\mathrm{i}}{\sum_{\mathrm{i}=1}^{100} \mathrm{i}} = \frac{\mathrm{i}}{5050}, \qquad \text{(Equation S17)}$$

and repeated this sampling procedure 1000 times. Each time, we recorded the average stability of the GFP molecules that survived selection in the simulated H and L populations. We then compared their average stabilities. This analysis shows that GFPs in the L population are more stable than in the H population (see figure S12, p ~ 10-16; Wilcoxon rank sum test).

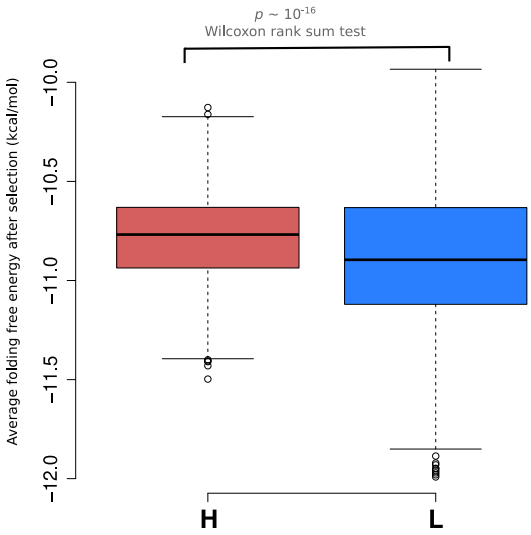


**Figure S12: Average folding stability of selected GFPs in H and L populations in 1000 simulations.** In each simulation, we sampled n=10,000 GFPs independently from an ancestral populations with 50,000 cells. The probability that a cell survives selection in these simulations was proportional to its fluorescence intensity (Equation S17). The boxes in the boxplots ranges from the first quantile to the third quantile. The median is represented by a line across the box. Whiskers show 1.5 times of the interquartile range. The circles located above the top whisker or below the bottom whisker are outliers whose values are either higher or lower than 1.5 times the interquartile range (third quartile – first quartile), respectively. We calculate the value of $p=10^{-16}$ from a two-sided Wilcoxon rank sum test with effect size calculated as $Z/\sqrt{N}=1.07$.

**Part 6: Checking the convolution of log-normal distributions**

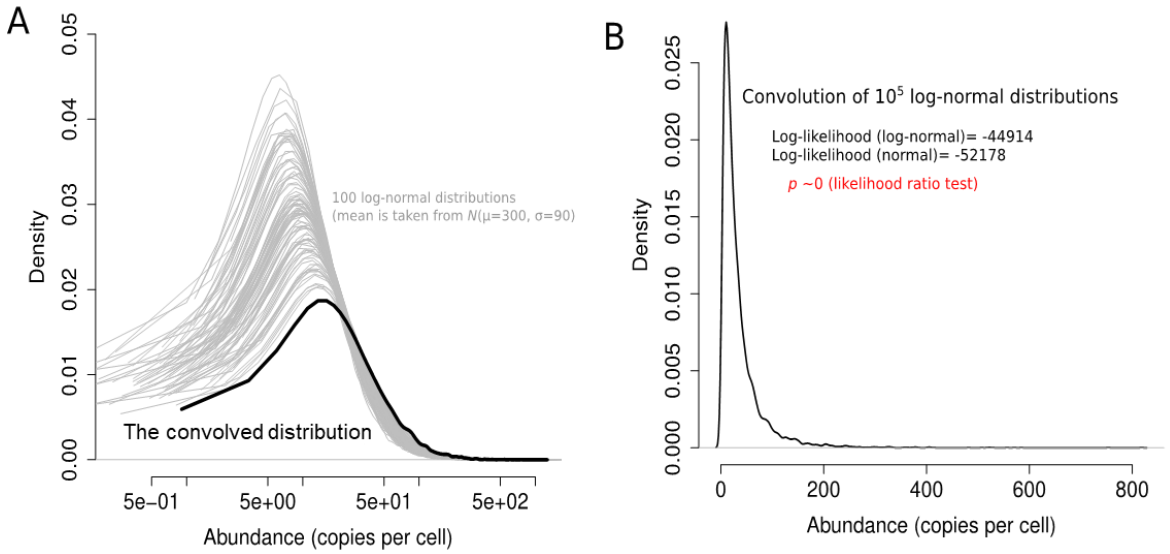


**Figure S13: The convolution of log-normal distributions is a log-normal distribution itself.** A) An example where gene expression noise is modelled by log-normal distributions. We generated 100 log-normal distributions (shown in grey) and show the convolution of all these 100 distributions in black. B) The convolution of $10^5$ log-normal distributions is log-normally distributed. We used a likelihood ratio test to compare the likelihoods of a normal and a log-normal distribution fitting the data. In both panels A and B, the mean and standard deviation of each log-normal distribution were taken from a normal distribution with μ=300 and σ=90 GFP molecules, respectively.

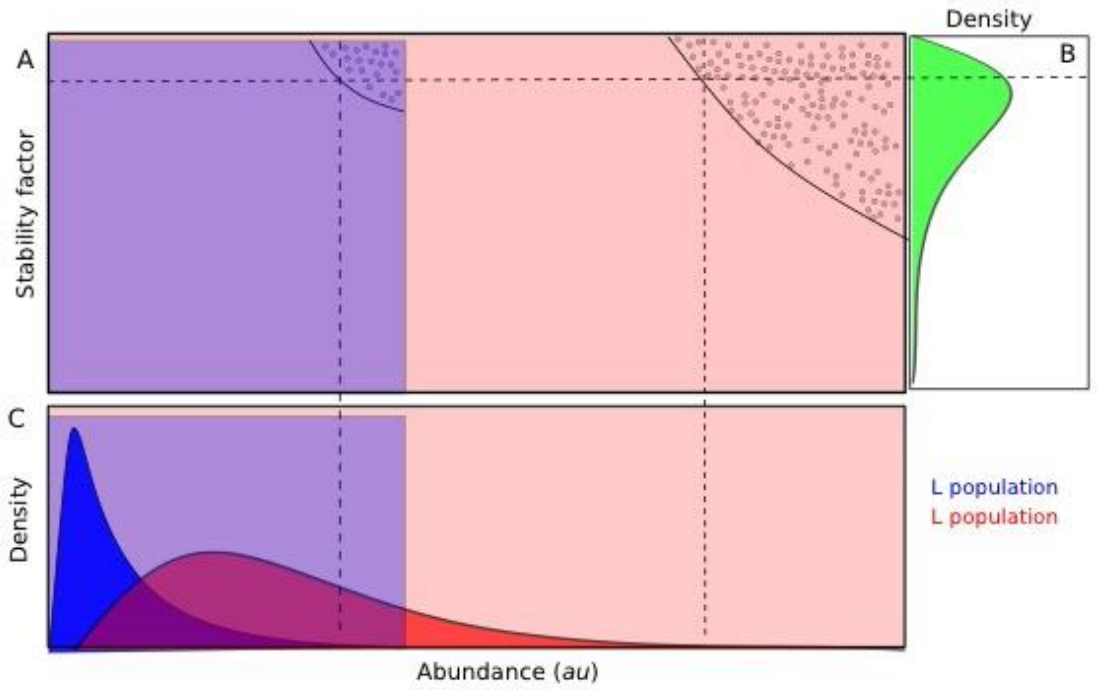


**Figure S14.** Increased skewness of relative GFP abundance in H populations can explain the retention of more destabilizing mutations in these populations compared to L populations. the figure shows the stability factor versus protein abundance (panel A), the density distribution of the stability factor (panel B), and the density distributions of protein abundances for L and H populations (panel C). We represent the stability of the L and H populations by the same distribution (shown in green), because the stability of generated mutants does not change with expression level. The distribution of protein abundances for L and H populations are shown in blue and red, respectively, in panel C. For both panels A and B, the blue and red rectangles show the ranges of protein abundance from zero to maximum protein abundance in L and H populations, respectively. Note the overlap between the two rectangles, which we indicated by moving the red rectangle slightly further up. The horizontal dashed line shows an example of a stability percentile of GFPs in cells surviving selection. The vertical dashed lines show the corresponding abundance percentiles of GFPs in surviving cells for the L and H populations. The reason that surviving cells are scattered in two isoclines rather than the confined rectangles between the dashed lines in each population is the synergy between stability and abundance, as described in the main text. For both populations, GFPs whose stability is lower than the specific percentile (the horizontal dashed lines) survive selection only if their abundances in host cells is higher than the selection threshold (percentile; vertical dashed lines). Conversely, the stability of surviving GFPs can be higher than the specific percentile (the horizontal dashed line) only if their abundances in host cells are lower than the selection threshold ( percentile; vertical dashed lines). In H populations, because of the log-normal distribution of protein abundance, the same percentile corresponds to an absolute expression level that is further away from the maximum protein abundance, compared to the L population. Stability-scarcity synergy then contributes to a lower stability of selected GFPs in H population than in L population.

**References**

1 Hall, B. G., Acar, H., Nandipati, A. & Barlow, M. Growth rates made easy. *Molecular biology and evolution* **31,** 232-238 (2014).

2 Cormack, B. P., Valdivia, R. H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173,** 33-38 (1996).

3 Siemering, K. R., Golbik, R., Sever, R. & Haseloff, J. Mutations that suppress the thermosensitivity of green fluorescent protein. *Current Biology* **6,** 1653-1663 (1996).

4 Sarkisyan, K. S. *et al.* Local fitness landscape of the green fluorescent protein. *Nature* **533,** 397-401 (2016).

5 Dietz, H. & Rief, M. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proceedings of the National Academy of Sciences* **101,** 16192-16197 (2004).

6 Huang, J.-r., Craggs, T. D., Christodoulou, J. & Jackson, S. E. J. J. o. m. b. Stable intermediate states and high energy barriers in the unfolding of GFP. **370,** 356-371 (2007).

7 Reddy, G., Liu, Z. & Thirumalai, D. J. P. o. t. N. A. o. S. Denaturant-dependent folding of GFP. **109,** 17832-17838 (2012).

8 Silander, O. K. *et al.* A genome-wide analysis of promoter-mediated phenotypic noise in Escherichia coli. *PLoS genetics* **8,** e1002443 (2012).

9 Silander, O. K. *et al.* A genome-wide analysis of promoter-mediated phenotypic noise in Escherichia coli. *PLoS Genet* **8,** e1002443 (2012).

10 Ferrer-Admetlla, A., Leuenberger, C., Jensen, J. D. & Wegmann, D. J. G. An approximate Markov model for the Wright–Fisher diffusion and its application to time series data. **203,** 831-846 (2016).

11 Schymkowitz, J. *et al.* The FoldX web server: an online force field. *Nucleic acids research* **33,** W382-W388 (2005).

12 Ghosh, K. & Dill, K. J. B. j. Cellular proteomes have broad distributions of protein stability. **99,** 3996-4002 (2010).

13 Ghosh, K. & Dill, K. Cellular proteomes have broad distributions of protein stability. *Biophysical journal* **99,** 3996-4002 (2010).