A Calculations over a Moving Window for NAI

Breaking  \( \eta(\vec{p}) \) into Components  
We define a window where we are evaluating the NAI of a gene as going from \( x \) to \( y \), inclusive. We break the \( \eta \) function into three parts, one part is before the window \((1, x – 1)\), the second part is within the window \((x, y)\), the third part is after the window, \((y + 1, n)\).

\[
\eta(\vec{p}) = \sum_{i=1}^{n} \beta_{i} \cdot \sigma_{i-1} \cdot (1 - p_{i}) + \beta_{n+1} \cdot \sigma_{n}
\]

\[
= \sum_{i=1}^{x-1} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \left( \prod_{k=i+1}^{x-1} \frac{1}{p_{k}} \times \prod_{k=x}^{y} \frac{1}{p_{k}} \times \prod_{k=y+1}^{n} \frac{1}{p_{k}} \right) + \sum_{i=x}^{y} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \left( \prod_{k=x}^{y-1} \frac{1}{p_{k}} \times \prod_{k=y+1}^{n} \frac{1}{p_{k}} \right) + \sum_{i=y+1}^{n} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \prod_{k=i+1}^{n} \frac{1}{p_{k}} + \beta_{n} \tag{1}
\]

\[
= \sum_{i=1}^{x-1} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \left( \sigma_{1, i} \times \prod_{k=x}^{y} \frac{1}{p_{k}} \times \sigma_{n,y} \right) + \sum_{i=x}^{y} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \left( \prod_{k=x}^{y-1} \frac{1}{p_{k}} \times \sigma_{n,y} \right) + \sum_{i=y+1}^{n} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \sigma_{n,i} + \beta_{n} \tag{2}
\]

Calculating \( \bar{\eta} \) and \( \text{Var}(\eta) \)

Mean \( \eta \):  
Given the amino acid sequence of a gene and assuming that the choice of codon at each position is independent, the expected cost-benefit ratio of a sequence which is allowed to vary over a window from \( i = x \) to \( y \) is given by,

\[
E(\eta_{x,y}) = \sum_{i=1}^{x-1} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \left( \sigma_{1, i} \times \prod_{k=x}^{y} \mathbb{E} \left[ \frac{1}{p_{k}} \right] \times \sigma_{n,y} \right) + \sum_{i=x}^{y} \beta_{i} \mathbb{E} \left[ \frac{1 - p_{i}}{p_{i}} \right] \left( \prod_{k=x}^{y-1} \mathbb{E} \left[ \frac{1}{p_{k}} \right] \times \sigma_{n,y} \right) + \sum_{i=y+1}^{n} \beta_{i} \left( \frac{1 - p_{i}}{p_{i}} \right) \sigma_{n,i} + \beta_{n} \tag{4}
\]
where, for notational convenience, we define

\[
\sigma_{i,j} = \begin{cases} 
\prod_{k=i+1}^{j} p_k & i < j \\
\prod_{k=j+1}^{i} \frac{1}{p_k} & i > j 
\end{cases}
\]  \tag{6}

Unlike calculating \( \eta \) for an entire gene, here the expectations are conditional only on the possible set of \( p_i \) values in \( \vec{p}(\vec{c}) \) from \( i = x \) to \( y \). As before, our expectations for \( p_i \) are taken over \( p \) values for a given set of synonymous codons.

**Var** \( (\eta_{x,y}) \): Beginning with Equation (1) gives,

\[
\text{Var} (\eta_{x,y}) = \sum_{i=1}^{x-1} \text{Var} \left( \beta_i \left( \frac{1-p_i}{p_i} \right) \left( \prod_{k=i+1}^{x-1} \frac{1}{p_k} \times \prod_{k=y+1}^{n} \frac{1}{p_k} \right) \right) + \sum_{i=x}^{y} \text{Var} \left( \beta_i \left( \frac{1-p_i}{p_i} \right) \left( \prod_{k=i+1}^{y} \frac{1}{p_k} \times \prod_{k=y+1}^{n} \frac{1}{p_k} \right) \right) + \sum_{i=y+1}^{n} \text{Var} \left( \beta_i \left( \frac{1-p_i}{p_i} \right) \prod_{k=i+1}^{n} \frac{1}{p_k} \right) \tag{7}
\]

\[
= \sum_{i=1}^{x-1} \left( \beta_i \left( \frac{1-p_i}{p_i} \right) \left( \prod_{k=i+1}^{x-1} \frac{1}{p_k} \times \prod_{k=y+1}^{n} \frac{1}{p_k} \right) \right)^2 \text{Var} \left( \prod_{k=i+1}^{y} \frac{1}{p_k} \right) + \sum_{i=x}^{y} \left( \beta_i \left( \frac{1-p_i}{p_i} \right) \prod_{k=i+1}^{y} \frac{1}{p_k} \right)^2 \text{Var} \left( \prod_{k=i+1}^{n} \frac{1}{p_k} \right) + \sum_{i=y+1}^{n} \left( \beta_i \left( \frac{1-p_i}{p_i} \right) \prod_{k=i+1}^{n} \frac{1}{p_k} \right)^2 \tag{8}
\]

For notational simplicity in writing the variance and covariance terms, we begin by defining

\[
X = \prod_{k=x}^{n} \left( \frac{1}{p_k} \right) \tag{9}
\]

\[
Y_i = \left( \frac{1-p_i}{p_i} \right) \prod_{k=i+1}^{y} \left( \frac{1}{p_k} \right) \tag{10}
\]

Again, given our assumption of independence of \( p \) values at different positions, it follows that

\[
E(X) = \prod_{k=x}^{n} \mathbb{E} \left[ \frac{1}{p_k} \right] \tag{11}
\]

\[
E(Y_i) = \mathbb{E} \left[ \left( \frac{1-p_i}{p_i} \right) \prod_{k=i+1}^{y} \mathbb{E} \left[ \frac{1}{p_k} \right] \right] \tag{12}
\]
and

\[ \text{Var}(\eta_{x,y}) = \text{Var} \left( \frac{\sigma_n}{\sigma_n} X \sum_{i=1}^{x-1} \beta_i \frac{1-p_i}{p_i} \sigma_{x-1,i} \sigma_{o,y} + \sigma_{o,y} \sum_{i=x}^{y} \beta_i Y_i \right) \]  
\[ = (\sigma_{o,y})^2 \left( \text{Var} \left( X \sum_{i=1}^{x-1} \beta_i \frac{1-p_i}{p_i} \sigma_{x-1,i} \sigma_{o,y} + \sum_{i=x}^{y} \beta_i Y_i \right) \right) \]  
\[ = (\sigma_{o,y})^2 \left( \sum_{i=1}^{x-1} \beta_i \frac{1-p_i}{p_i} \sigma_{x-1,i} \sigma_{o,y} \right)^2 \text{Var}(X) + y \sum_{i=x}^{y} \beta_i^2 \text{Var}(Y_i) + 2 \sum_{i=x}^{y} \beta_i \sum_{j=i+1}^{y} \beta_j \text{Cov}(Y_i, Y_j) \]  
\[ + 2 \left( \sum_{i=1}^{x-1} \beta_i \frac{1-p_i}{p_i} \sigma_{x-1,i} \sigma_{o,y} \right) \sum_{j=x}^{y} \beta_j \text{Cov}(X, Y_j) \]  

In the above argument, the variance terms follow a similar form to the full calculation.

\[ \text{Var}(X) = \prod_{j=x}^{y} \mathbb{E} \left[ \left( \frac{1}{p_j} \right)^2 \right] - \prod_{j=x}^{y} \left[ \mathbb{E} \left( \frac{1}{p_j} \right) \right]^2 \]  
\[ \text{Var}(Y_i) = \mathbb{E} \left[ \left( \frac{1-p_i}{p_i} \right) \prod_{j=i+1}^{y} \mathbb{E} \left( \frac{1}{p_j} \right)^2 \right] - \left( \mathbb{E} \left[ \frac{1-p_i}{p_i} \right] \prod_{j=i+1}^{y} \mathbb{E} \left[ \frac{1}{p_j} \right] \right)^2 \]  

Turning to the calculation of the two covariance terms in equation (15) we first note that, in general,

\[ \text{Cov}(Y_i, Y_j) = \mathbb{E} [ (Y_i - \mathbb{E}[Y_i]) (Y_j - \mathbb{E}[Y_j])] \]  
\[ = \mathbb{E} [Y_i Y_j] - \mathbb{E} [Y_i] \mathbb{E} [Y_j] \]  
\[ \text{Cov}(X, Y_j) = \mathbb{E} [X Y_j] - \mathbb{E} [X] \mathbb{E} [Y_j] \]  

When \( i < j \leq y \), it follows that

\[ \mathbb{E} [Y_i Y_j] = \mathbb{E} \left[ \frac{1-p_i}{p_i} \right] \prod_{k=i+1}^{j-1} \mathbb{E} \left[ \frac{1-p_k}{p_k} \right] \prod_{k=j+1}^{y} \mathbb{E} \left[ \frac{1}{p_k} \right]^2 \]  
\[ \mathbb{E} [X Y_j] = \mathbb{E} \left[ \frac{1-p_i}{p_j} \right] \prod_{k=x}^{j-1} \mathbb{E} \left[ \frac{1}{p_k} \right] \prod_{k=j+1}^{y} \mathbb{E} \left[ \frac{1}{p_k} \right]^2 \]
B  Model Selection for Distribution of $\eta$ across Synonymous Genotype Space

The fact that the size of the synonymous genotype spaces $S$ for the average gene is on the order of $3^{400} = 7.4 \times 10^{190}$ makes it impossible to survey the $\eta$ values across such a space completely. Instead we assume that the distribution of $\eta$ across $S$ can be approximated by a continuous distribution such as the Gamma, Weibull, Normal, and Log-Normal distribution. We evaluated the fit of $\eta$ values to each of these distributions based on how well they fit a random sampling of 1000 alleles from $S$ for each of 2000 randomly selected genes in the $S. cerevisiae$ genome using the Akaike Information Criteria (AIC).

On a per gene basis, the gamma distribution gave the lowest AIC values in 53.75% of the genes. The Normal, Log-Normal and Weibull distributions had the lowest AIC values in 36.1%, 10.15% and 0% of the genes, respectively. In terms of the combined dataset of 2000 genes, the gamma distribution had the AIC lowest score and the AIC differences $\Delta AIC_i$ values for the other distributions were 3886, 8078, and 375,226 for the Normal, Log-Normal, and Weibull distributions, respectively.

C  Skewness Reducing Transformation of $\eta$ Distribution

The specific parameters used in the transformation are based on the shape and scale parameters of the gamma distribution describing the distribution of $\eta$ values across the synonymous genotype space, i.e. $\alpha$ and $\beta$, respectively. Based on this transformation (PACE AND SALVAN, 1997), for a given allele the transformed $\eta_{obs}$ value and central moments of $\eta$ for its synonyms are,

$$\eta'_{obs} = 3 \left( \eta_{obs} - \eta_{min} \right)^{\frac{2}{3}} - 1$$

$$\bar{\eta}' = 3 \left( \beta^{-\frac{2}{3}} \frac{\Gamma(\alpha + \frac{1}{3})}{\Gamma(\alpha)} - 1 \right)$$

$$\text{Var} \left( \eta' \right) = \frac{9}{\Gamma(\alpha)^2} \beta^{-\frac{2}{3}} \left( \Gamma \left( \alpha + \frac{2}{3} \right) - \frac{\Gamma \left( \alpha + \frac{1}{3} \right)^2}{\Gamma(\alpha)} \right)$$

where the $'$ is used to distinguish the transformed from the untransformed terms.

D  Allele Substitution Model in CES

Our simulations follow the ideas developed in GILCHRIST (2007) where each locus has its own average protein production rate $\phi$. In this model, the marginal fitness effect of allele $i$ at that locus is $w(i) = w_i \propto \exp(-q\eta_i)$, where $q$ is a scaling term that relates energy expenditure and fitness. Note that, unlike $\phi$, $q$
does not vary between genes. This fitness function is consistent with the idea that any change in $\sim P/sec$ that an organism must expend to meet its target protein production rate $\phi$ caused by a change in $\eta$ will lead to a very small, but fixed proportional change in fitness. This ensures that the strength of selection for reducing energetic costs is consistent across all genes. We assume that new alleles of a gene are generated through a step wise mutation process where $\mu$ represents the per nucleotide mutation rate. The probability a new potentially invading allele $j$ will replace the resident allele $i$ is based on the relative fitness of the two alleles and the organism’s effective population size $N_e$. The exact substitution probabilities are calculated using the formulation presented in Sella and Hirsh (2005), i.e.

$$\pi(i \to j) = \frac{1 - \left(\frac{w_i}{w_j}\right)^2}{1 - \left(\frac{w_i}{w_j}\right)^{2N_e}} = \frac{1 - \exp\left[-2\phi q (\eta_i - \eta_j)\right]}{1 - \exp\left[-2N_e \phi q (\eta_i - \eta_j)\right]}.$$  

(26)

E Robustness of NAI to Parameter Uncertainty

In order to estimate the sensitivity of NAI scores to changes in parameter estimates, we calculated the sensitivity coefficient, $\Psi$ (Hamby, 1994) for each parameter. For instance, sensitivity coefficient, $\Psi$ for $b$ is defined as

$$\Psi = \frac{d\text{NAI}}{db} \frac{b}{\text{NAI}}. \quad (27)$$

In general, we find that the calculation of NAI scores is remarkably robust to uncertainty in the values underlying its calculation such as the background nonsense error rate $b$ ($\Psi=0.003$), the cost of ribosome initiation $a_1$ ($\Psi=0.001$), and the cost of peptide elongation $a_2$ ($\Psi=0.009$) (Supporting Figure S3). To understand NAI’s robustness, we return to our calculations of an allele’s cost-benefit ratio $\eta$. We begin by noting that while the probability of a nonsense error occurring somewhere along a transcript may be substantial, the actual probability per codon or unit time is quite small, on the order of $1 \times 10^{-4}$/codon or $1 \times 10^{-3}$/sec. If one performs a first order Taylor series expansion for $\eta$ as defined in Equations (1)-(5) we get.

$$\eta(\bar{p}) = (a_1 + a_2 n) + b \sum_{i=1}^{n} \frac{a_1 + a_2 (i - 1)}{c_i} + O(b^2).$$  

(28)
Based on this result we can calculate estimates of the first two moments of \( \eta \) as,

\[
\bar{\eta} \approx (a_1 + a_2 n) + b \sum_{i=1}^{n} (a_1 + a_2 (i - 1)) \mathbb{E} \left( \frac{1}{c_i} \right) \tag{29}
\]

\[
\text{Var} (\eta) \approx b^2 \sum_{i=1}^{n} (a_1 + a_2 (i - 1))^2 \text{Var}(Y_i) \tag{30}
\]

\[
= b^2 \sum_{i=1}^{n} (a_1 + a_2 (i - 1))^2 \left( \mathbb{E} [Y_i^2] - \mathbb{E} [Y_i]^2 \right). \tag{31}
\]

This approximation shows it is possible to factor out the background nonsense error rate \( b \) from all three terms used to calculate NAI: \( \eta_{\text{obs}} \), \( \bar{\eta} \) and \( \sqrt{\text{Var}(\eta)} \). Even after our Box-Cox transformations, the \( b \) we have factored out will cancel, thus explaining why NAI is relatively insensitive to changes in \( b \) so long as \( b \ll c_i \) for all codons. A similar result can be obtained with the elongation cost parameter \( a_2 \). Conceptually, increasing either term is similar to simply rescaling the \( \eta \) values for the synonymous set of alleles. Since NAI measures the adaptation of an allele relative to its coding synonyms, rescaling the \( \eta \) values across this space will have no effect on an allele’s relative position. We can explain NAI’s insensitivity to changes in \( a_1 \) by noting that the average gene has \( \sim 400 \) amino acids and so long as \( a_1 \) is not orders of magnitude greater than \( a_2 \), then \( a_2 (i - 1) \) will be greater than \( a_1 \) for most codon positions within an allele. Thus, changing \( a_1 \) also has little impact on the NAI value of an allele as well.

NAI values were also found to be robust to small changes in the estimates of elongation rates of codon. These sensitivity coefficients ranged in value from \( \Psi = -9 \times 10^{-4} \) to \( \Psi = 0.235 \) with their average value being 0.003 (Supporting Table S3). In general, slowly translating codons were more sensitive to changes in their elongation rates than codons with high elongation rates.