This software package contains codon compression tools as published in “Codon Compression Algorithms for Saturation Mutagenesis”

File contents:

Three Perl scripts:

* Prog.pl (Algorithm A)
* Script\_thread\_redun.pl (Algorithm B)
* Expand.pl

The Common module, Rules file and Math folder are necessary for the scripts to run.

Additionally, there are some organism-specific usage tables. Other tables for different organisms can be generated by the user, in the same format.

To run:

In Mac run from terminal, as Perl is preinstalled.

In PC use a Perl environment, like Padre (can be freely downloaded from: <http://padre.perlide.org/> ). After installation, open and run the Perl scripts from the Padre window.

Script details:

General comments:

In all scripts the user is asked to input the relevant usage table.

In the compression scripts the user is asked to input the amino acids to be removed. X indicates the three stop codons. If more than one, input the amino acids separated with commas, without spaces.

DYNAMCC\_R: This script is designed to explore the codon redundant space. It generates a list of compressed codons starting with the most highly used codons for every amino acid in the selected pool (the “Best List”). Then, it adds the next highly used codon and compresses the list again, and so on until it uses all available codons.

DYNAMCC\_0: This script is designed to compress codons of a given set of amino acids without redundancy. The user can control whether the compression is being made by codon Rank or Usage. Rank indicates the usage ranking threshold for codon inclusion. Usage allows the user to define the exact usage ratio threshold. The usage input maximum value is derived from the amino acid composition and is defined in the dialogue. The script also allows to add redundancy. This is EXTREMELY computationally intensive, and should be set to 0 for most applications. Finally, the user also can define the number of threads to run in parallel. During the run, the script reports on its progress. At the end of the run, the results are presented as a table of the amino acids and their chosen codons, and a list of the compressed codons.

- If the number of compressed codons is too high, use lower usage thresholds to increase the codon pool the script can choose from.

Expand.pl: This script allows the user to expand compressed codons to their non-compressed form.

To add usage tables, create a .txt file in the same format as the provided usage tables and place them within the same folder.