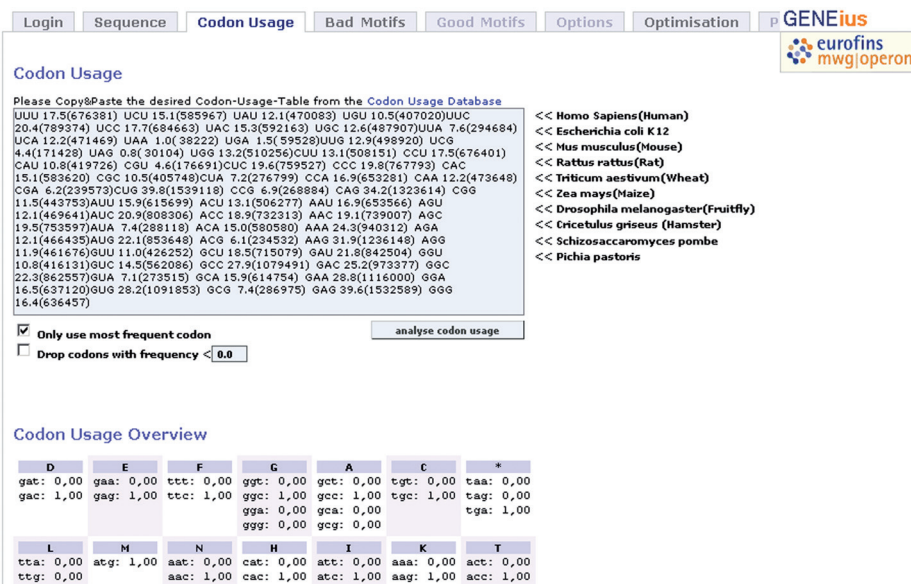


# New Software from Eurofins MWG Operon optimises Sequences for Gene Synthesis

Dr. Uwe Koehler & Dr. Tobias Wagner, Eurofins MWG Operon, Ebersberg



The screenshot shows the GENEius software interface. At the top, there are navigation tabs: Login, Sequence, Codon Usage (selected), Bad Motifs, Good Motifs, Options, and Optimisation. The main content area is titled 'Codon Usage' and contains a 'Please Copy&Paste the desired Codon-Usage-Table from the Codon Usage Database' section. This section lists various organisms and their codon usage tables, such as Homo Sapiens (Human), Escherichia coli K12, Mus musculus (Mouse), Rattus rattus (Rat), Triticum aestivum (Wheat), Zea mays (Maize), Drosophila melanogaster (Fruitfly), Cricetus griseus (Hamster), Schizosaccharomyces pombe, and Pichia pastoris. Below this list, there are checkboxes for 'Only use most frequent codon' (checked) and 'Drop codons with frequency < 0.0'. A button labeled 'analyse codon usage' is also present. At the bottom, there is a 'Codon Usage Overview' table showing the frequency of codons for each amino acid across different organisms.

D	E	F	G	A	C	*
gat: 0,00	gaa: 0,00	ttt: 0,00	ggt: 0,00	gct: 0,00	tgt: 0,00	taa: 0,00
gac: 1,00	gag: 1,00	ttc: 1,00	ggc: 1,00	gcc: 1,00	tgc: 1,00	tag: 0,00
			gga: 0,00	gca: 0,00		tga: 1,00
			ggg: 0,00	gcg: 0,00		
L	M	N	H	I	K	T
tta: 0,00	atg: 1,00	aat: 0,00	cat: 0,00	att: 0,00	aaa: 0,00	act: 0,00
ttg: 0,00		aac: 1,00	cac: 1,00	atc: 1,00	aag: 1,00	acc: 1,00

organism. The result after several optimisation steps, is a codon frequency that is nearly identical to the natural conditions. It is also possible to exclude certain codons with low frequency or to select a single codon per Amino acid. Another advantage of GENEius is the prevention of direct or indirect repeats and hairpin structures in the optimised sequence, which could cause problems during transcription or translation. Moreover, GENEius distributes the GC content in the optimised sequence evenly, thereby increasing the efficiency of transcription and translation.

Once a synthetic DNA fragment is synthesised, it must often be subcloned into a suitable vector using unique restriction sites. GENEius can avoid chosen restriction sites in the optimised sequence, ensuring that 5' and 3' cloning sites will be unique. Other user specific motifs e.g. splice donor and acceptor sites or transcription factor binding sites can be added to the list of „bad motifs“ and will be avoided in the optimised sequence. Unlike other codon adaptation programs, GENEius does not compromise codon adaptation quality when optimising for the other features.

Another very important feature is the possibility of introducing restriction sites or other „good motifs“ at certain positions within the final DNA sequence. GENEius screens for amino acid combinations in the protein sequence that allow introduction of e.g. an XhoI site (CTC GAG). Protein sequences LSS or PSR could be coded by NNC TCG AGN. If LSS or PSR is present in the protein sequence, introduction of an XhoI site at that position is easily possible. By using this feature, long protein coding DNA sequences can be designed in such a way that unique restriction sites are present in the newly synthesised DNA fragment e.g. for further subcloning of single protein domains. Once the sequence has been optimised, GENEius can also be used for designing the gene synthesis primers.

A light version of GENEius is available for free on our website. The full version of GENEius is exclusively used by scientists at Eurofins MWG Operon for optimising sequences for gene synthesis.

The demand for synthetic DNA fragments in the biological and pharmaceutical R&D is getting larger every day. The synthesis of DNA fragments is more frequently handed over to special service providers, both by the industry and academic institutions. This not only reduces time and cost, but also results in a better outcome from the experiments and processes. Advanced techniques and optimized processes enable a fast and reliable double stranded DNA synthesis. All synthetic sequences, from short DNA fragments of 100 bases to long genes with a couple thousand bases and even small genomes, can be synthesised in high quality by experts in a timely and cost efficient manner.

If gene sequences, for example from human origin, were to be used in a heterologous expression system, like E.coli, an optimization at DNA or amino acid level would be needed.

The Gene Synthesis Service of Eurofins MWG Operon has this optimisation as standard. In cooperation with BioLink GmbH, Planegg – Martinsried, Germany, Eurofins MWG Operon has developed a new software for this purpose. Several years of experience in gene synthesis was used to develop new sequence optimisation algorithms for the software. The resulting software is called GENEius, an intelligent optimisation software with enhanced functions that enable a perfect sequence adjustment.

A big advantage of GENEius is the automatic adaptation of codon usage, for instance to adapt DNA sequences for Protein Expression in a target organism. Researchers that use heterologous expression systems to express cDNA's often encounter low expression levels. A reason for that could be the change in codon preference from organism to organism, although the genetic code is the same in all organisms with minor exceptions. As 61 triplets code for only 20 amino acids, most of the amino acids are coded by 2, 3, 4 or even 6 triplets. This leads to different frequencies of codons in different organisms. For instance, the codons AGG and AGA used to code Arginine are used in a frequency of 21% in humans, whereas the same codons are used in a frequency of 2 and 4% in E.coli, respectively. Thus, if a gene from human origin is inserted into an E.coli expression system, the tRNA's of the codons AGG and AGA will be used up very quickly, yielding a very low level of expression.

## Adaptation of Codon Usage

The GENEius software from Eurofins MWG Operon uses DNA or amino acid sequences as a starting point and adjusts the coding sequence exactly to the codon usage of the destination