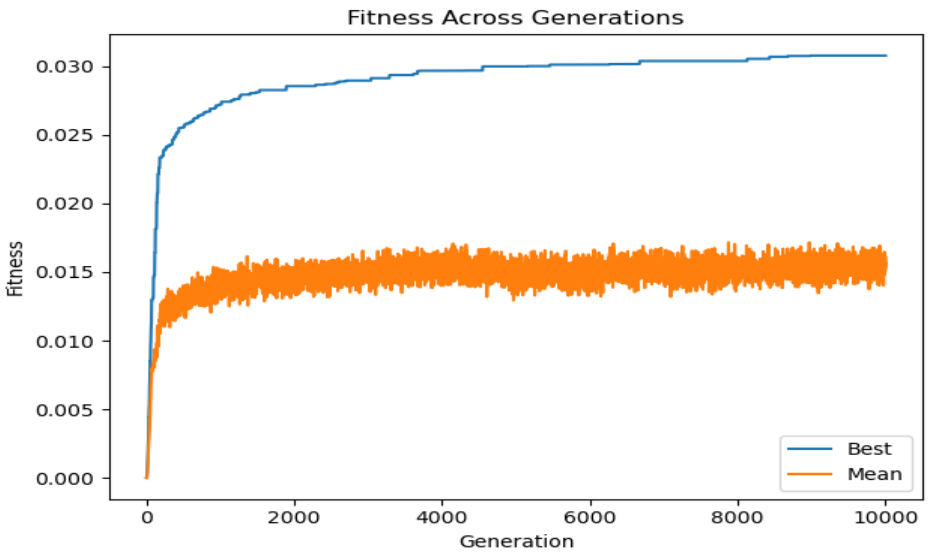
Codon Optimization Report

1. Genetic Algorithm Optimization Summary

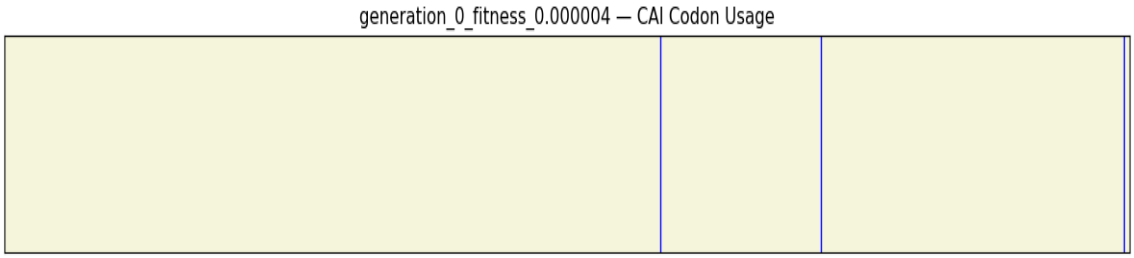
Total generations: 9999
Best composite fitness achieved: 0.030773

1.1 Fitness Progression

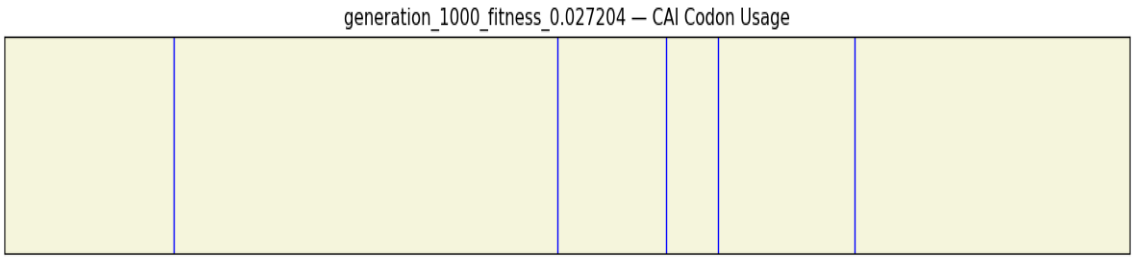


1.2 CAI Positional Plots Across Generations

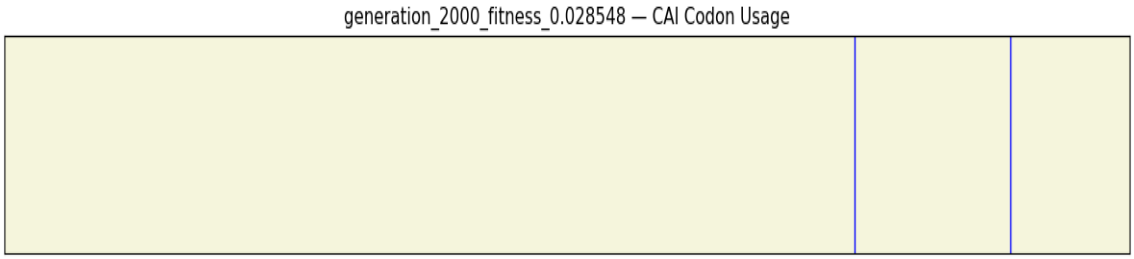
Generation 0



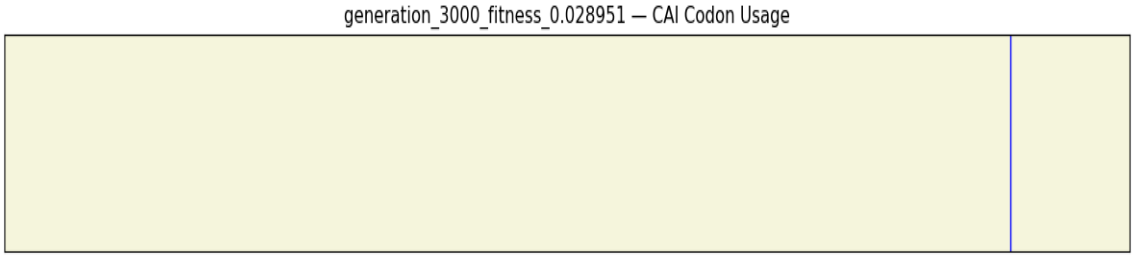
Generation 1000



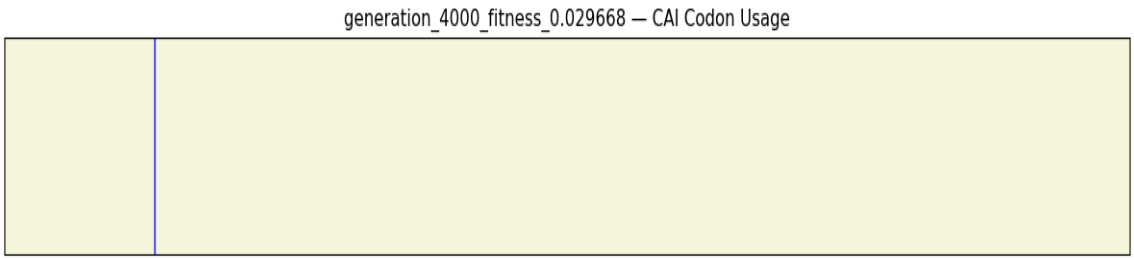
Generation 2000



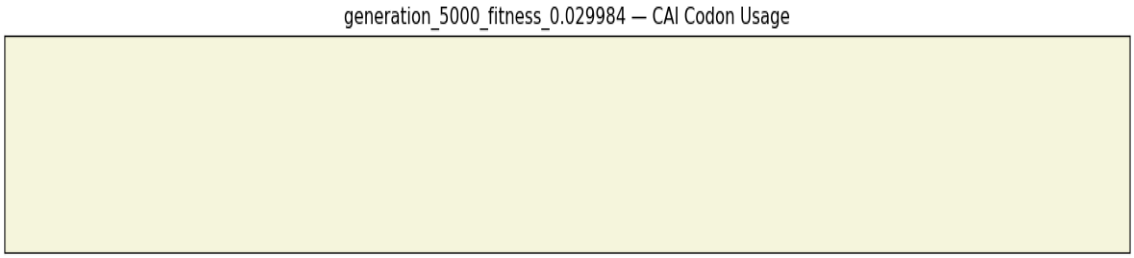
Generation 3000



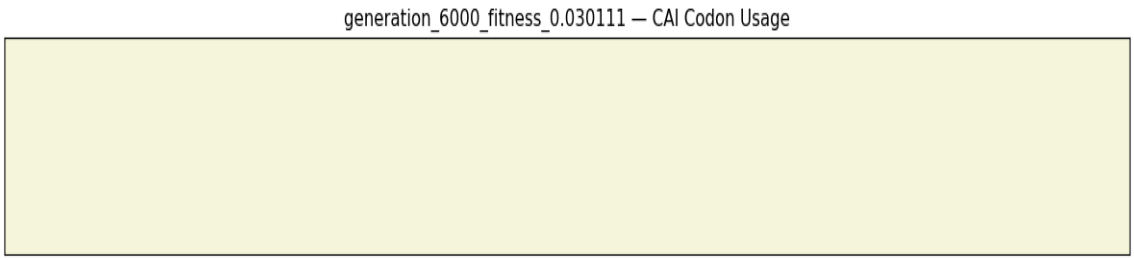
Generation 4000



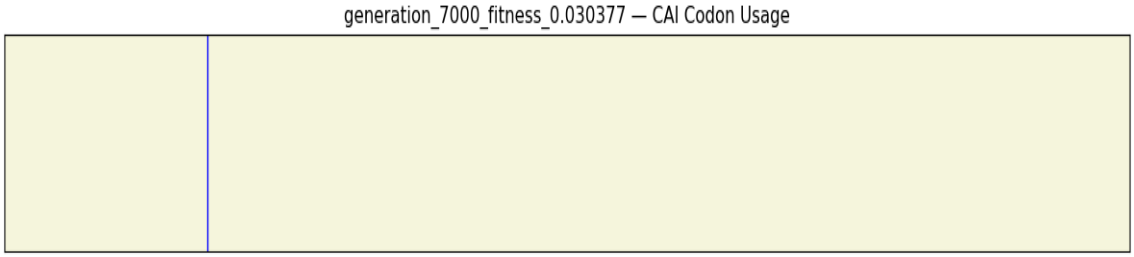
Generation 5000



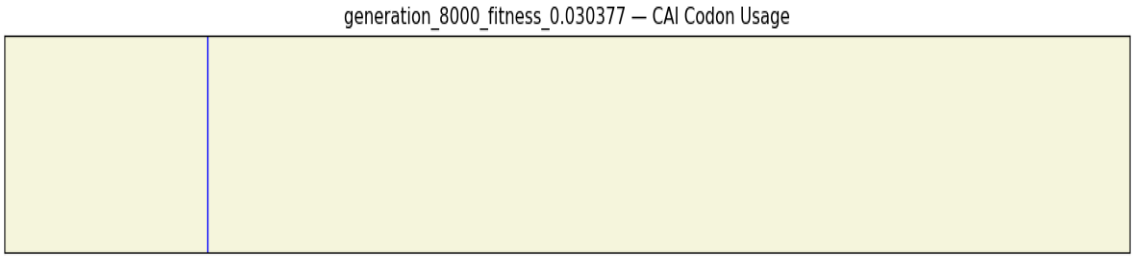
Generation 6000



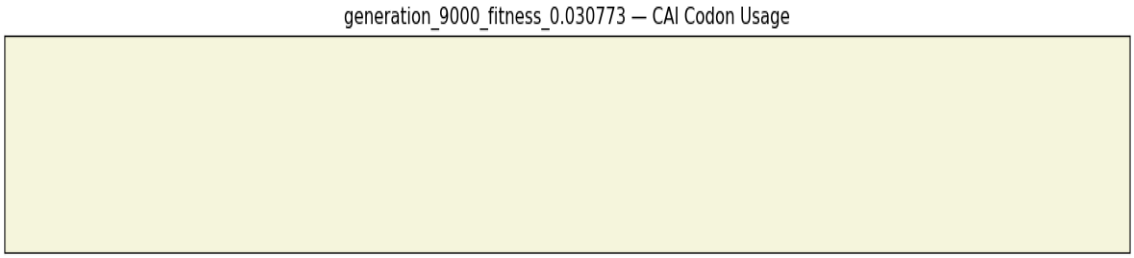
Generation 7000



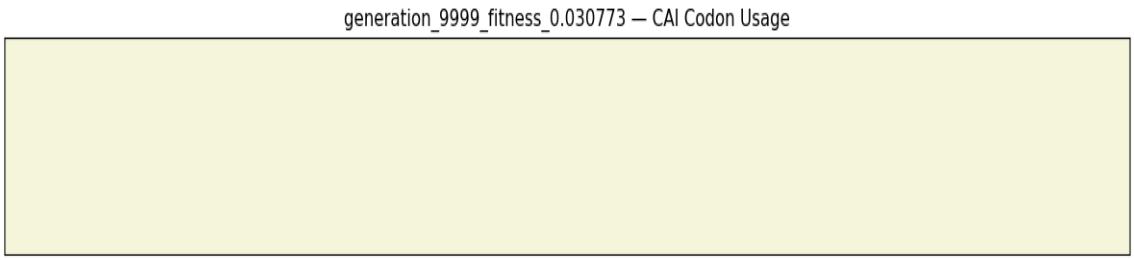
Generation 8000



Generation 9000



Generation 9999



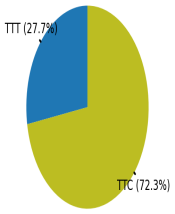
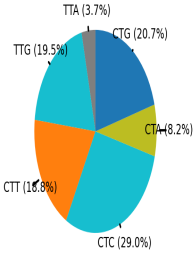
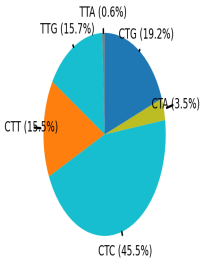
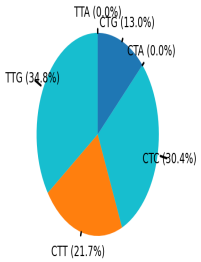
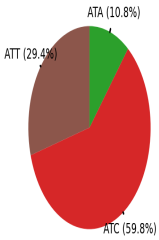
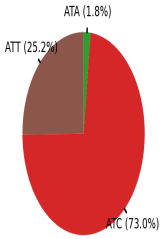
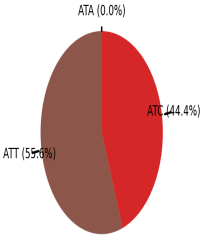
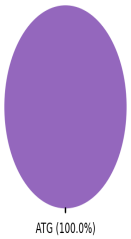
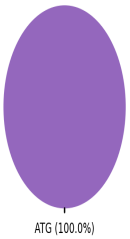
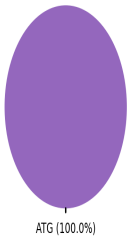
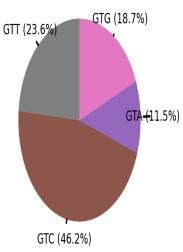
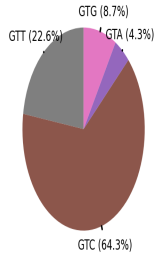
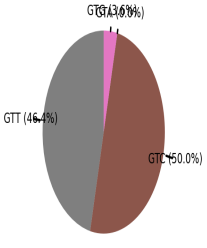
2. Final Optimized Sequence Metrics

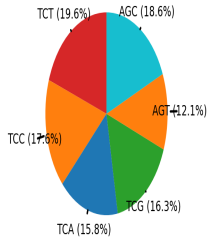
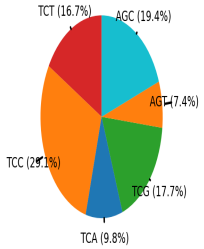
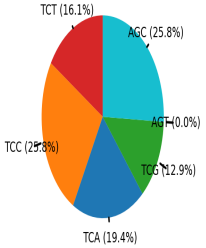
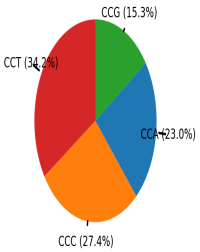
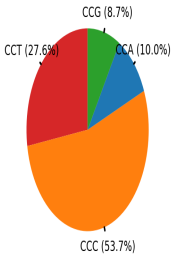
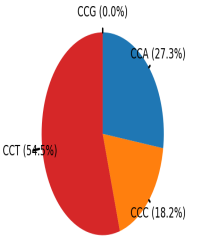
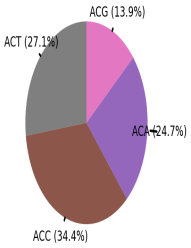
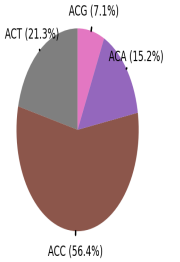
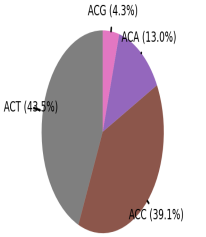
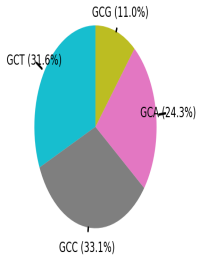
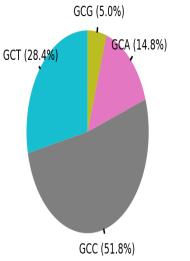
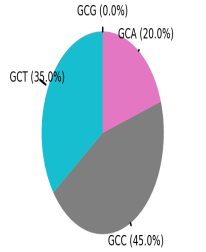
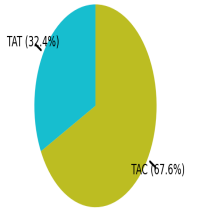
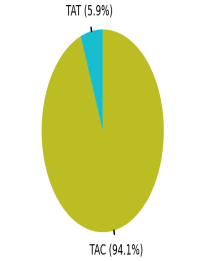
Sequence length: 1215 nt
GC content: 0.5259
GC3 content: 0.5654

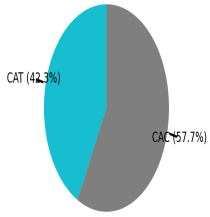
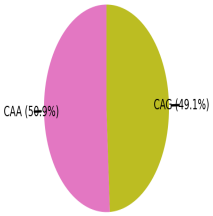
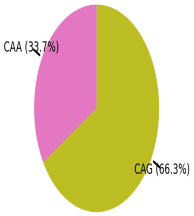
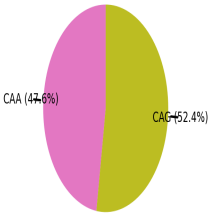
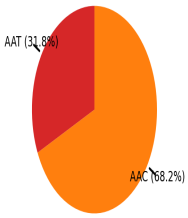
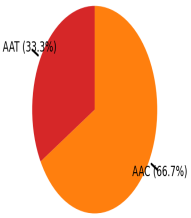
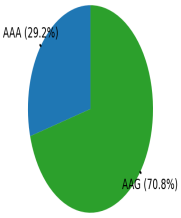
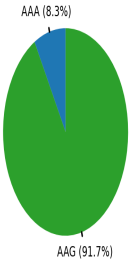
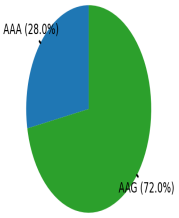
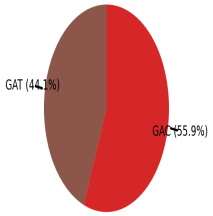
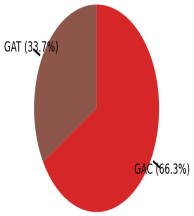
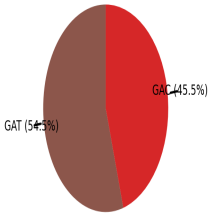
2.1 Final Optimized Sequence

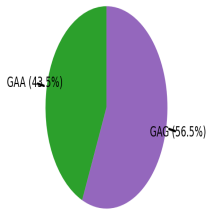
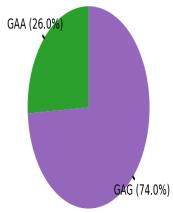
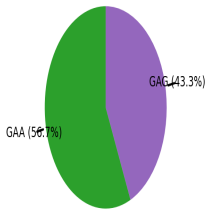
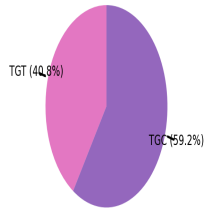
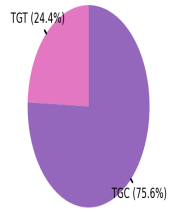
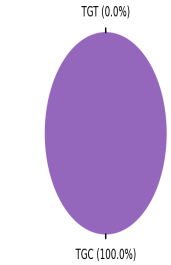
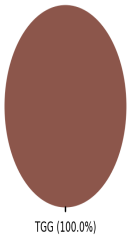
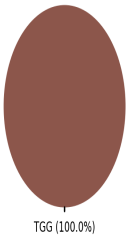
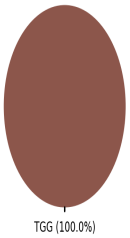
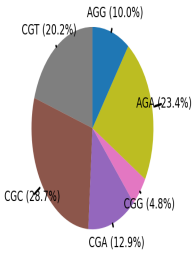
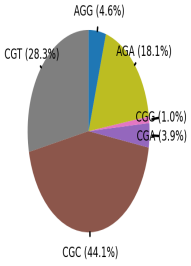
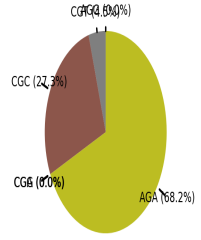
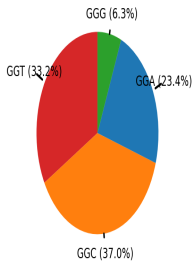
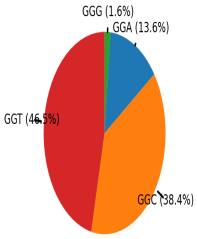
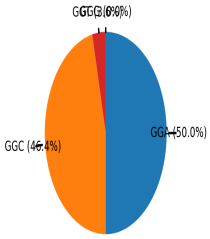
```
>alfa_nb_linker_mScarlet_optimized_final
ATGGTCTCCAAGGGCGAAGCTGTCATCAAGGAATTCATGAGATTCAAAGTCCACATGGAG
GGATCAATGAACGGCCACGAATTCGAGATTGAGGGAGAAGGAGAGGGAAGACCCTACGAA
GGCACCCAGACAGCAAAGCTGAAAGTCACAAAGGGCGGCCCTTTGCCATTTTCATGGGAT
ATTCTTAGCCCACAATTCATGTACGGCTCTAGAGCCTTTATTAAGCACCCTGCCGATATT
CCAGATTACTACAAGCAATCTTTTCCTGAAGGCTTCAAATGGGAGCGCGTTATGAATTTC
GAAGACGGAGGAGCAGTCACTGTGACTCAAGACACTAGCTTGGAAGATGGAACCCTGATC
TACAAAGTTAAATTGAGAGGAACTAACTTCCCACCCGATGGCCCTGTTATGCAGAAGAAG
ACTATGGGCTGGGAAGCCTCAACCGAGCGCCTCTACCCTGAAGATGGCGTCCTCAAGGGA
GATATTAAGATGGCTTTGAGACTTAAGGATGGAGGAAGATACCTCGCCGATTTCAAGACT
ACTTACAAGGCCAAGAAGCCAGTTCAGATGCCTGGAGCTTACAACGTCGACAGAAAATTG
GACATCACCTCTCACAATGAGGATTATACTGTCGTTGAGCAATACGAGCGCAGCGAAGGA
CGCCACAGCACCGGCGGCATGGACGAGCTCTACAAGCTCGAGGGCTCCGGACAGGGACCT
GGCTCCGGACAGGGATCCGGCTCCCCCGGATCTGGACAAGGCCCAGGCTCGGGCCAGGGA
TCAGGCCCCGGACAAGGCAGCGGACCTGGACAGGGCTCAATGGCCGAAGTCCAACTCCAG
GAGAGCGGCGGAGGACTTGTCCAACCTGGAGGCTCCTTGAGACTTTCCTGCACTGCCAGC
GGCGTTACAATCTCGGCTTTGAACGCCATGGCCATGGGCTGGTACAGACAGGCTCCTGGA
GAACGCAGAGTTATGGTCGCAGCTGTCTCCGAAAGAGGCAATGCAATGTACAGAGAATCG
GTTCAGGGAAGATTCACTGTTACCAGAGACTTCACGAACAAGATGGTCAGCCTGCAAATG
GACAACTTGAAGCCTGAAGACACCGCTGTTTACTACTGCCACGTCCTTGAGGACAGAGTT
GATTCATTCCATGATTACTGGGGTCAGGGTACCCAAGTTACCGTTTCTTCGAAACTCGAA
CGCCCTCACCGTGAC
```

2.2 Codon Usage Comparison

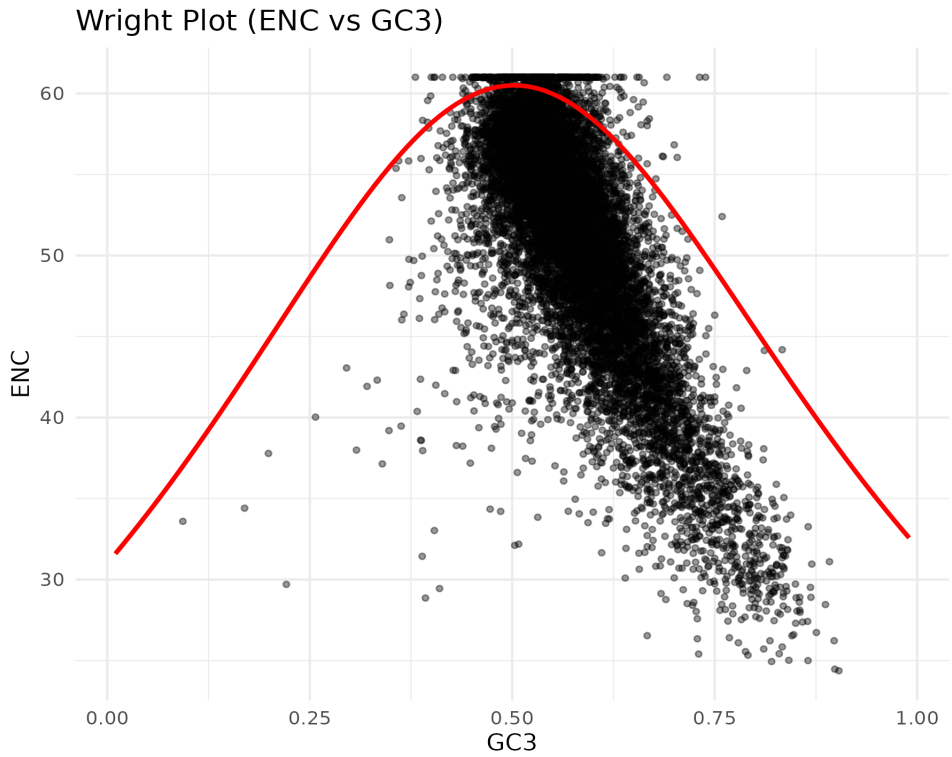
Residue	Genome	Ribosomal	Final
Phenylalanine Phe (F) TTT TTC			
Leucine Leu (L) TTA TTG CTT CTC CTA CTG			
Isoleucine Ile (I) ATT ATC ATA			
Methionine Met (M) ATG			
Valine Val (V) GTT GTC GTA GTG			

Residue	Genome	Ribosomal	Final
Serine Ser (S) TCT TCC TCA TCG AGT AGC			
Proline Pro (P) CCT CCC CCA CCG			
Threonine Thr (T) ACT ACC ACA ACG			
Alanine Ala (A) GCT GCC GCA GCG			
Tyrosine Tyr (Y) TAT TAC			

Residue	Genome	Ribosomal	Final
Histidine His (H) CAT CAC			
Glutamine Gln (Q) CAA CAG			
Asparagine Asn (N) AAT AAC			
Lysine Lys (K) AAA AAG			
Aspartate Asp (D) GAT GAC			

Residue	Genome	Ribosomal	Final
Glutamate Glu (E) GAA GAG			
Cysteine Cys (C) TGT TGC			
Tryptophan Trp (W) TGG			
Arginine Arg (R) CGT CGC CGA CGG AGA AGG			
Glycine Gly (G) GGT GGC GGA GGG			

3. Genome Codon Bias (Wright Plot)



4. Interpretation Notes

- Optimization balances translational efficiency and genome-wide codon conformity.
- GC content is constrained using a Gaussian penalty centered on genome average.
- Cryptic splice-site penalties were applied during genetic algorithm optimization.
- Final sequence should be validated experimentally before synthesis.