

GenRCA Rare Codon Analysis Report

Gene Name: *Insulin*

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

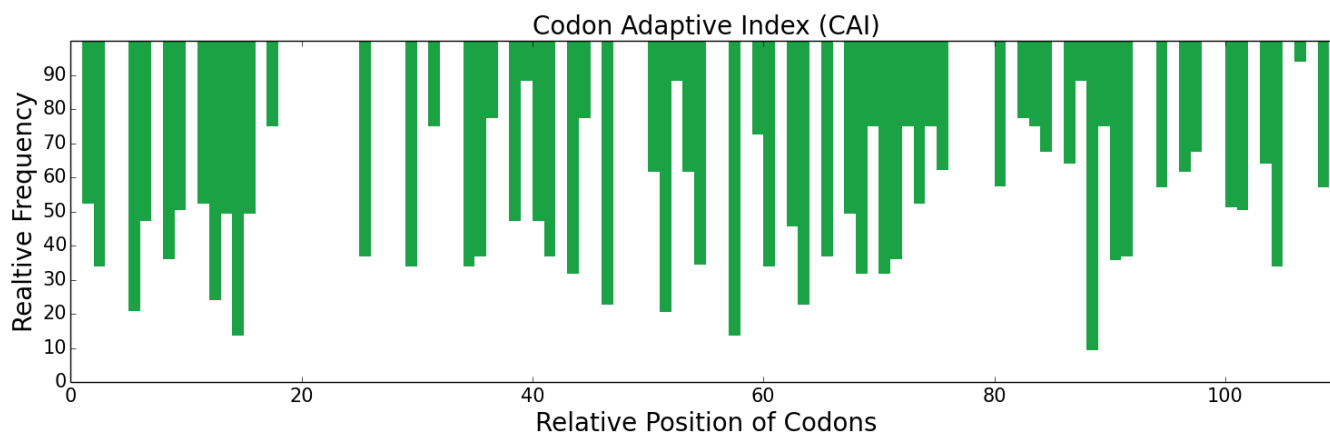
DNA Sequence Length: 330

DNA Sequence:

```
ATGGCACTATGGATGCGACTGTTGCCCCCTTTTGGCACTCTTAGCGTTATGGGGACCAGAT
CCAGCTGCTGCTTTTGTAAACCAACATCTATGTGGATCTCATCTAGTAGAGGCTCTGTAT
CTGGTATGTGGCGAGAGAGGGGTTTTTTTACACACCGAAAAACACGTAGAGAAGCGGAAGAC
CTACAAGTGGGGCAAGTAGAATTAGGCGGAGGCCCGGAGCAGGATCATTGCAACCATTG
GCCTTGGAGGGATCCTTGCAGAAACGGGGAATAGTAGAACAATGCTGTACATCCATTTGT
AGTCTTTACCAGCTAGAAAATTACTGCAAC
```

Protein Sequence:

```
MALWMRLPLLLALLALWGPDPAAAFVNQHLCSHLVEALYLVCGERGFFYTPKTRREAED
LQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSI CSLYQLENYCN
```



Negative CIS Elements	Negative repeat Elements
1	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.49	[11]
ENC	Effective Number of Codons	20~61	51.62	[2, 3]

RCBS	Relative Codon Bias Strength	≥ 0	0.46	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.78	[5]
CDC	Codon Deviation Coefficient	0~1	0.17	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.57	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.36	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.25	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.71	[11]
P	Codon Preference	≥ 1	1.25	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.42	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.63	[1]
CFD	Codon Frequency Distribution	0~1	0.07	[2]
FOP	Frequency of Optimal Codons	0~1	0.4	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.27	[5]
COUSIN18			1.64	
CBI	Codon Bias Index	-1~1	0.12	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.65	[7]
RCA	Relative Codon Adaptation	≥ 0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.46	[9]
B	Codon Usage Bias	0~2	0.51	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.4	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.48	[2]
P2	P2 Index	0~1	0.45	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.38	[1]
GC	GC Content		0.48	
GC1	GC Content at the First Position of Synonymous Codons		0.61	

GC2	GC Content at the Second Position of Synonymous Codons		0.45	
ENcp	Effective Number of Codon Pairs	20~61	61	[2]
CPS	Codon Pair Score	-1~1	0.08	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.74	[5]

Gene Name: Somatotropin

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

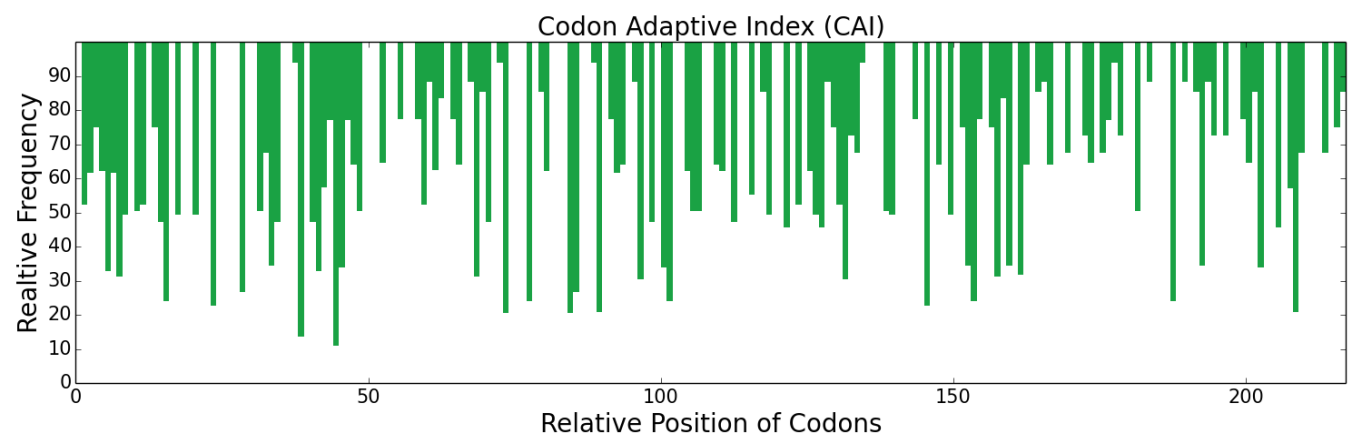
DNA Sequence Length: 651

DNA Sequence:

ATGGCAACAGGATCAAGGACAAGCTTATTGCTTGCATTTGGACTGCTCTGTTTACCATGG
TTACAAGAAGGGTCTGCTTTTCCAACGATTCCACTTTCCCGTCTGTTTGATAATGCGATG
CTGAGGGCCCAACGCCTACACCAGCTTGCTTTTGATACCTACCAAGAGTTTGAAGAGGCA
TATATCCCTAAGGAGCAGAAGTATAGCTTCCTGCAAAATCCGCAAACCTCTCTCTGTTTC
TCAGAATCTATTCCGACGCCATCTAATCGAGAAGAGACACAGCAAAAATCGAACCTGGAA
CTACTCAGAATTTCACTTCTTTTGATTGAGTCATGGCTGGAACCAGTCCAATTCTTAAGA
TCTGTGTTTGCAAACCTATTAGTGTATGGAGCATCGGACTCCAATGTTTACGATCTTTTA
AAGGATTTGGAGGAAGGGATTGAGACTTTAATGGGACGTCTCGAGGATGGAAGCCCTCGT
ACTGGCCAGATTTTCAAACAGACTTACTCCAAGTTTGACACCAACTCCCACAATGACGAT
GCTCTTTTGAAAACTACGGTCTCTTGTATTGTTTCCGTAAAGACATGGACAAGGTTGAG
ACCTTCCTAAGAATTGTGCAATGCCGATCCGTTGAAGGTTCTGTGGATTTC

Protein Sequence:

MATGSRTSLLLAFLGLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEA
YIPKEQKYSFLQNPQTSLCFSESIPTSPNREETQQKSNLELLLRISLLLIQSWLEPVQFLR
SVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDD
ALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference

RSCU	Relative Synonymous Codon Usage	0~6	1.2	[11]
ENC	Effective Number of Codons	20~61	55.89	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.27	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.51	[5]
CDC	Codon Deviation Coefficient	0~1	0.12	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.55	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.16	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.11	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.86	[11]
P	Codon Preference	≥ 1	1.13	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.24	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.68	[11]
CFD	Codon Frequency Distribution	0~1	0.07	[2]
FOP	Frequency of Optimal Codons	0~1	0.39	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.31	[5]
COUSIN18			0.09	
CBI	Codon Bias Index	-1~1	0.09	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.44	[7]
RCA	Relative Codon Adaptation	≥ 0	1.12	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.28	[9]
B	Codon Usage Bias	0~2	0.32	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.48	[11]
gtAI	Genetic tRNA Adaptation Index	0~1	0.54	[2]
P2	P2 Index	0~1	0.51	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.44	[11]
GC	GC Content		0.44	

GC1	GC Content at the First Position of Synonymous Codons		0.5	
GC2	GC Content at the Second Position of Synonymous Codons		0.38	
ENcp	Effective Number of Codon Pairs	20~61	30.42	[21]
CPS	Codon Pair Score	-1~1	0.04	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[51]

Gene Name: Erythropoietin

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

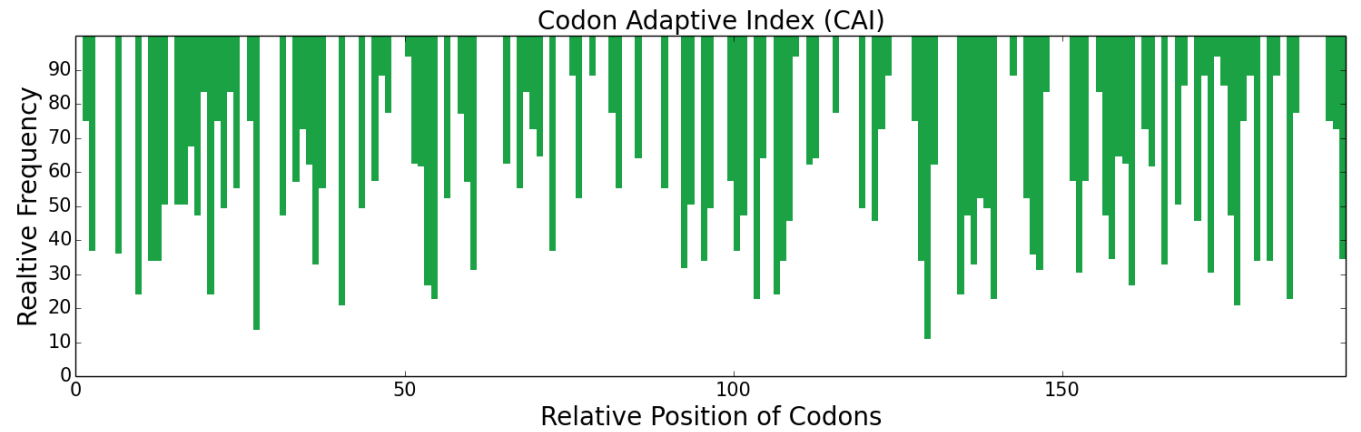
DNA Sequence Length: 579

DNA Sequence:

ATGGGAGTACATGAATGTCCCGCTTGGCTCTGGCTACTACTTTCTCTTCTTTCCCTGCCT
CTCGGATTACCTGTCTTGGGAGCGCCACCAAGACTGATTTGCGACTCAAGGGTCTTGGA
CGATACTTGTTAGAAGCCAAAGAGGCTGAAAAATATCACAAACGGGGTGTGCAGAACACTGC
AGCTTGAACGAAAACATCACTGTCCCTGACACCAAGGTAAACTTTTATGCATGGAAGA
ATGGAGGTGCGTCAACAGGCTGTTGAAGTCTGGCAAGGCCTTGCTCTATTATCTGAAGCC
GTACTGAGAGGGCAGGCTCTCCTAGTGAATTCTTCACAGCCATGGGAGCCATTGCAATTA
CATGTGGACAAAGCTGTTTCTGGACTACGCTCATTGACTACTCTCCTGAGGGCATTAGGG
GCTCAAAAAGAAGCAATAAGCCCTCCAGATGCTGCCTCGGCCGCTCCTCTGCGTACCATC
ACGGCTGACACATTTAGGAAGCTTTTCAGAGTGTATTCGAATTCCTGCGAGGAAACTA
AAGCTATATACTGGGGAGGCTTGTAAGAACTGGAGACCGT

Protein Sequence:

MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTGCAEHC
SLNENITVPDTKVNIFYAWKRMEVGGQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL
HVDKAVSGLRSLTTLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL
KLYTGEACRTGDR



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.22	[11]

ENC	Effective Number of Codons	20~61	59.88	[2 , 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.25	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.67	[5]
CDC	Codon Deviation Coefficient	0~1	0.14	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.57	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.12	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.11	[9 , 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.85	[11]
P	Codon Preference	≥ 1	1.14	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.29	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.64	[1]
CFD	Codon Frequency Distribution	0~1	0.07	[2]
FOP	Frequency of Optimal Codons	0~1	0.37	[3 , 4]
COUSIN59	Codon Usage Similarity Index	∞	0.18	[5]
COUSIN18			-0.26	
CBI	Codon Bias Index	-1~1	0.09	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.47	[7]
RCA	Relative Codon Adaptation	≥ 0	1.09	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.24	[9]
B	Codon Usage Bias	0~2	0.34	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.45	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.51	[2]
P2	P2 Index	0~1	0.64	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.42	[1]
GC	GC Content		0.49	
	GC Content at the First Position			

GC1	of Synonymous Codons		0.58	
GC2	GC Content at the Second Position of Synonymous Codons		0.46	
ENcp	Effective Number of Codon Pairs	20~61	61	[2]
CPS	Codon Pair Score	-1~1	0.04	[3] , [4]
Codon Volatility	Codon Volatility	0.5~1	0.73	[5]

Gene Name: Interferon alpha

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

DNA Sequence Length: 567

DNA Sequence:

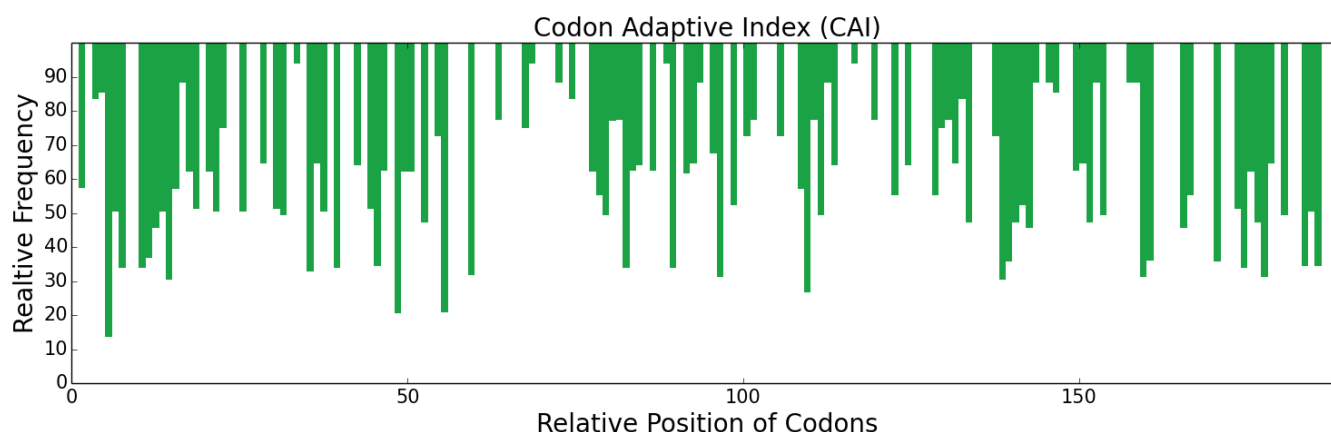
```

ATGGCCTCTCCTTTTCGCGCTTCTAATGGTTCTAGTAGTGCTTTCGTGCAAATCAAGTTGT
TCACTTGGATGTGATCTTCCAGAAACCCATAGTTTATAGATAATAGAAGGACCCTTATGCTA
TTGGCTCAGATGAGTCGTATCTCTCCGTCATCATGTCTGATGGACCGACATGATTTTGGC
TTTCCACAAGAGGAATTTGATGGAAATCAATTTCAAAAAGCTCCTGCTATTTTCAGTCTTA
CACGAGCTAATCCAGCAAATCTTTAATCTATTTACAACCAAAGATTCCAGCGCTGCATGG
GACGAGGATTTGTTGGACAAGTTTTGCACGGAGTTATATCAGCAATTGAATGATTTGGAG
GCTTGTGTCATGCAGGAAGAAAGAGTCGGAGAGACCCCTCTGATGAACGCTGACTCGATA
CTGGCAGTGAAAAAGTATTTTCAAGAATCACCCCTGTATTTAACTGAAAAGAAATATAGC
CCCTGTGCTTGGGAAGTGGTCAGAGCTGAAATAATGAGAAGTCTATCACTGAGCACCAAC
TTACAAGAACGTCTTCGTAGAAAGGAG
  
```

Protein Sequence:

```

MASPFALLMVLVVLVLSCKSSCSLGCDDLPEHSLDNRRTLMLLAQMSRISPSSCLMDRHDFG
FPQEEFDGNQFQKAPAIISVLHELIQQIFNLFTTKDSSAAWDEDLLDKFCTELYQQINDLE
ACVMQEERVGETPLMNADSIKAVKKYFRRITLYLTEKKYSPCAWEVVRAEIMRSLSLSTN
LQERLRRKE
  
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.33	[11]

ENC	Effective Number of Codons	20~61	48.55	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.35	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.79	[5]
CDC	Codon Deviation Coefficient	0~1	0.17	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.52	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.21	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.15	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.79	[11]
P	Codon Preference	≥ 1	1.18	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.4	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.69	[11]
CFD	Codon Frequency Distribution	0~1	0.02	[2]
FOP	Frequency of Optimal Codons	0~1	0.38	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.07	[5]
COUSIN18			-0.62	
CBI	Codon Bias Index	-1~1	0.07	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.55	[7]
RCA	Relative Codon Adaptation	≥ 0	1.1	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.32	[9]
B	Codon Usage Bias	0~2	0.4	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.46	[11]
gtAI	Genetic tRNA Adaptation Index	0~1	0.54	[2]
P2	P2 Index	0~1	0.54	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.4	[11]
GC	GC Content		0.42	
	GC Content at the First Position			

GC1	of Synonymous Codons		0.49	
GC2	GC Content at the Second Position of Synonymous Codons		0.38	
ENcp	Effective Number of Codon Pairs	20~61	21.84	[2]
CPS	Codon Pair Score	-1~1	0.03	[3] , [4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

Gene Name: Interferon omega-1

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

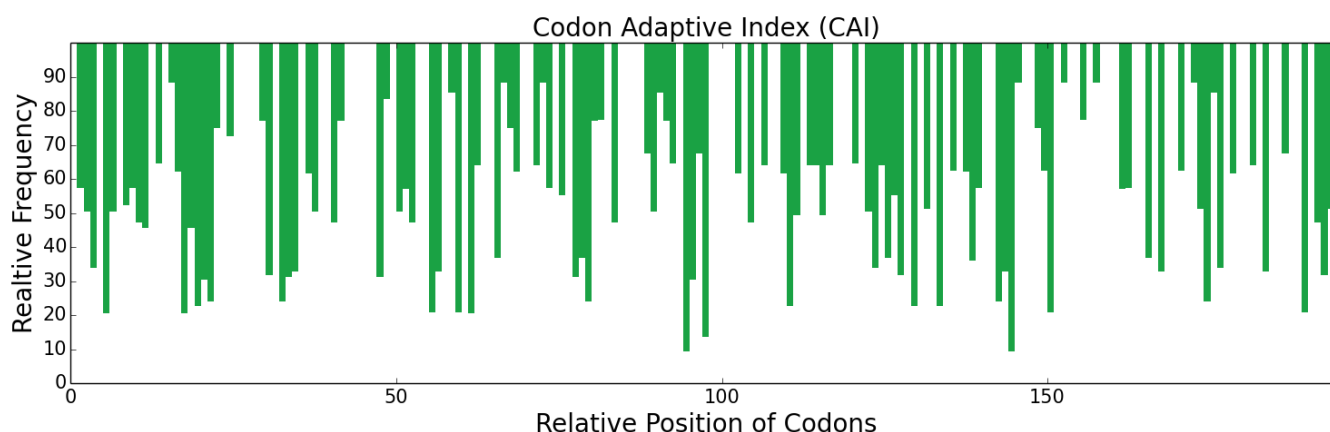
DNA Sequence Length: 585

DNA Sequence:

```
ATGGCCCTTCTATTTCCGCTTTTGGCAGCCCTGGTGATGACCTCTTATTCACCGGTGGGG
TCGCTCGGATGTGACTTGCCACAAAACACGGCTTGCTCAGCAGGAACACACTTGTTTTG
CTGCACCAAATGAGAAGAATTAGCCCTTTTCTTTGCCTGAAGGATCGAAGGGATTCCGA
TTTCCGCAGGAAATGGTAAAGGATCACAAATTCAGAAAAGCCCATGTCATGAGCGTACTC
CACGAGATGCTGCAACAAATTTTTTCCCTTTTCCACACCGAACGGTCGTCCGCGGCTTGG
AACATGACATTGCTGGATCAGTTGCATACAGGGTTACATCAGCAGTTACAGCATTGGA
ACCTGTCTTCTACAGGTAGTCGGCGAAGGGGAAAGTGCTGGGGCTATCTCTTACCCGCC
TTGACTCTCAGGCGGTATTTTCAAGGAATCCGAGTTTATTTGAAGGAGAAGAAATACTCT
GATTGCGCCTGGGAAGTAGTTAGGATGGAAATCATGAAAAGTCTCTTCTATCTACAAAC
ATGCAGGAAAGGTTGAGATCCAAGGATCGAGATCTGGGCAGTAGT
```

Protein Sequence:

```
MALLFPLLAALVMTSYSPVGSILGCDLPQNHGLLSRNTLVLLHQMRRI SPFLCLKDRRDFR
FPQEMVKGSQQLQKAHVMSVLHEMLQQIFSLFHTERS SAAWNMTLLDQLHTGLHQQLQHLE
TCLLQVVGESESAGAISSPALTLRRYFQGIRVYLKEKKYSDCAWEVVRMEIMKSLFLSTN
MQERLRSKDRDLGSS
```



Negative CIS Elements	Negative repeat Elements
1	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.2	[11]

ENC	Effective Number of Codons	20~61	57.12	[2 , 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.3	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.65	[5]
CDC	Codon Deviation Coefficient	0~1	0.15	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.55	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.18	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.13	[9 , 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.8	[11]
P	Codon Preference	≥ 1	1.18	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.29	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.62	[1]
CFD	Codon Frequency Distribution	0~1	0.1	[2]
FOP	Frequency of Optimal Codons	0~1	0.37	[3 , 4]
COUSIN59	Codon Usage Similarity Index	∞	1.18	[5]
COUSIN18			2.1	
CBI	Codon Bias Index	-1~1	0.08	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.75	[7]
RCA	Relative Codon Adaptation	≥ 0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.34	[9]
B	Codon Usage Bias	0~2	0.43	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.45	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.52	[2]
P2	P2 Index	0~1	0.44	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.54	[1]
GC	GC Content		0.48	
	GC Content at the First Position			

GC1	of Synonymous Codons		0.52	
GC2	GC Content at the Second Position of Synonymous Codons		0.37	
ENcp	Effective Number of Codon Pairs	20~61	20	[2]
CPS	Codon Pair Score	-1~1	0.04	[3] , [4]
Codon Volatility	Codon Volatility	0.5~1	0.75	[5]

Gene Name: GCSF

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

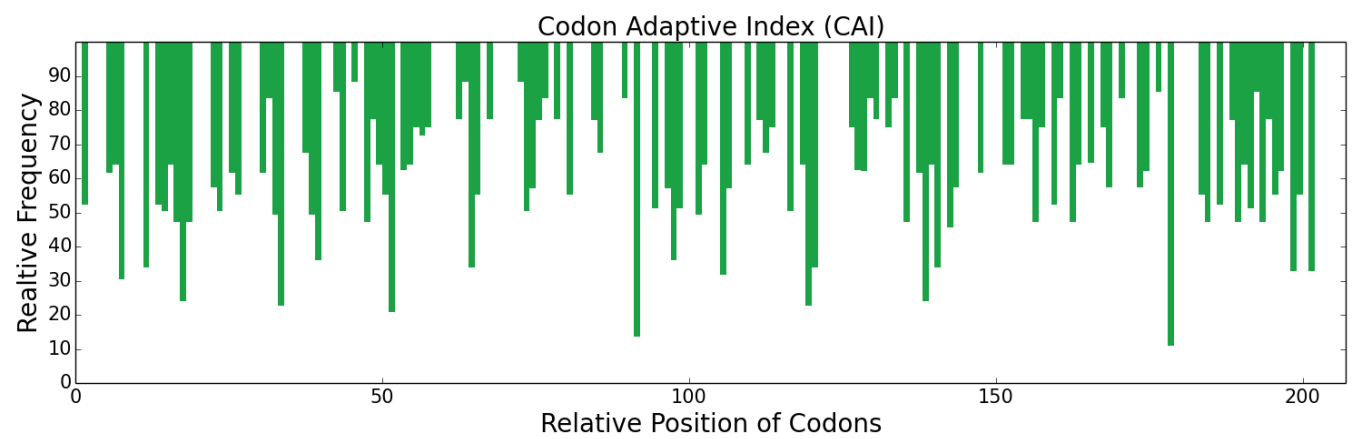
DNA Sequence Length: 621

DNA Sequence:

ATGGCAGGTCCAGCTACACAGTCGCCAATGAAGCTAATGGCACTTCAGCTGCTCCTGTGG
CATTCTGCCCTTTGGACAGTCCAAGAAGCTACACCTTTAGGGCCAGCTTCTTCCTTACCC
CAATCTTTTCCTTTTGAAATGTCTGGAGCAGGTCCGAAAGATCCAGGGAGACGGAGCTGCT
TTGCAAGAGAACTAGTCTCTGAGTGTGCTACTTACAACTTTGCCACCCTGAAGAGTTG
GTCTTGTTGGGTCACCTTGGGTATTCTTGGGCGCCATTGAGTTCTTGCCCCAGTCAA
GCTTTACAGTTGGCTGGCTGCTTGTCTCAGTTGCACTCCGATTGTTTCTTACCAGGGG
CTATTGCAAGCTTTGGAAGGAATCTCACCTGAGTTGGGACCTACTCTGGATACTCCAG
CTAGATGTGGCCGATTTTGCTACAACCTATTGGCAGCAGATGGAGGAGCTGGGAATGGCA
CCTGCTCTGCAGCCAACCAAGGAGCCATGCCTGCTTTTGCCTCAGCTTTCCAACGCAGA
GCTGGTGGTGTCTGGTTGCATCTCACCTGCAGAGTTTCCTGGAGGTCTCATACAGGGTC
TTGAGGCATTTGGCTCAACCA

Protein Sequence:

MAGPATQSPMKLMALQLLLWHSALWTVQEATPLGPASSLPQSFLKCLEQVRKIQGDGAA
LQEKLVSECATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCLSQLHSGFLFYQG
LLQALEGISPELGPTLDTLQLDVAADFATTIWQQMEELGMAPALQPTQGAMPAFASAFQRR
AGGVLVASHLQSFLEVSYRVLRLHLAQP



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference

RSCU	Relative Synonymous Codon Usage	0~6	1.43	[11]
ENC	Effective Number of Codons	20~61	50.51	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.52	[4]
DCBS	Directional Codon Bias Score	≥ 1	2	[5]
CDC	Codon Deviation Coefficient	0~1	0.27	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.44	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.2	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.19	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.77	[11]
P	Codon Preference	≥ 1	1.29	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.6	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.71	[11]
CFD	Codon Frequency Distribution	0~1	0.03	[2]
FOP	Frequency of Optimal Codons	0~1	0.42	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.68	[5]
COUSIN18			1.06	
CBI	Codon Bias Index	-1~1	0.16	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.48	[7]
RCA	Relative Codon Adaptation	≥ 0	1.19	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.32	[9]
B	Codon Usage Bias	0~2	0.4	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.48	[11]
gtAI	Genetic tRNA Adaptation Index	0~1	0.54	[2]
P2	P2 Index	0~1	0.73	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.49	[11]
GC	GC Content		0.52	

GC1	GC Content at the First Position of Synonymous Codons		0.61	
GC2	GC Content at the Second Position of Synonymous Codons		0.45	
ENcp	Effective Number of Codon Pairs	20~61	61	[2]
CPS	Codon Pair Score	-1~1	-0.02	[3] , [4]
Codon Volatility	Codon Volatility	0.5~1	0.73	[5]

Gene Name: Tumor necrosis factor

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

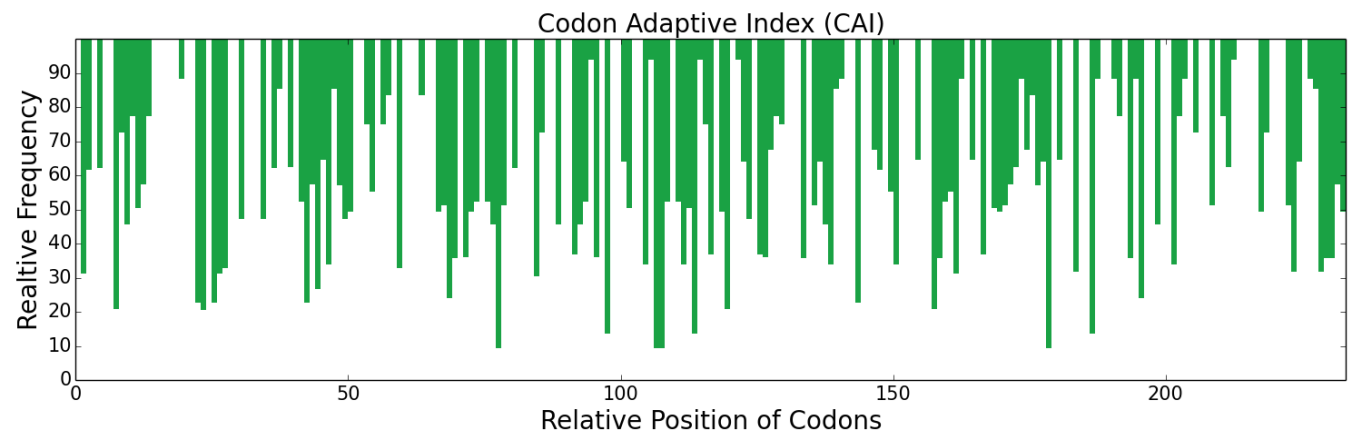
DNA Sequence Length: 699

DNA Sequence:

ATGAGCACAGAATCAATGATTTCGAGACGTGGAGCTTGCCGAGGAAGCTTTGCCAAAGAAA
ACTGGTGGGCGCAAGGGAGCAGGAGATGTCTGTTTTTGTCTCTGTTTTCATTCTTGATC
GTTGCAGGGGCCACGACCCTATTCTGCCTGTTACATTTTGGAGTCATTGGACCTCAAAGG
GAAGAATTTCTAGAGATTTAAGTCTCATATCTCCCTTAGCACAAAGCAGTGCGGAGTTCT
TCAAGAACTCCATCGGACAAGCCAGTGGCTCATGTAGTGGCAAATCCCCAAGCGGAAGGT
CAGCTTCAATGGCTAAATCGGCGGGCAAACGCACTACTTGCGAATGGAGTAGAATTACGA
GATAATCAGCTGGTTGTACCCTCCGAGGGATTGTACTTGATATACAGTCAGGTGCTATTC
AAAGGTCAAGGGTGTCCATCCACACATGTCCTATTGACTCATACCATTTCTCGAATAGCA
GTCAGCTATCAAACCAAGGTAAACCTTTTAAGTGCCATCAAATCCCTTGCCAGCGGGAA
ACCCCAAGGCGCTGAAGCGAAACCATGGTATGAGCCAATATATCTCGGTGGTGTGTTT
CAACTAGAGAAAGGTGACAGATTGAGTGCTGAGATCAATAGACCAGATTACTTAGACTTT
GCTGAAAGTGGCCAGGTTTATTTTCGGCATAATAGCCTTA

Protein Sequence:

MSTESMIRDVELAEELPKKTGGPQGSRRCLFLSLFSFLIVAGATTLFCLLHFGVIGPQR
EEFPRDLISPLAQAVRSSSRTPSDKPVAVVNPQAEGLQWLNRRANALLANGVELR
DNQLVVPSEGLYLIYSQVLFKQGCPSTHVLLTHTISRIAVSYQTKVNLLSAIKSPCQRE
TPEGAEAKPWYEPIYLGGVFQLEKGDRLSAEINRPDYLDFAESGQVYFGI IAL



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.16	[11]
ENC	Effective Number of Codons	20~61	59.74	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.17	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.43	[5]
CDC	Codon Deviation Coefficient	0~1	0.08	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.58	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.11	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.08	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.9	[11]
P	Codon Preference	≥ 1	1.08	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.2	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.63	[11]
CFD	Codon Frequency Distribution	0~1	0.08	[2]
FOP	Frequency of Optimal Codons	0~1	0.37	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	-0.01	[5]
COUSIN18			-0.29	
CBI	Codon Bias Index	-1~1	0.08	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.55	[7]
RCA	Relative Codon Adaptation	≥ 0	1.06	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.25	[9]
B	Codon Usage Bias	0~2	0.36	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.42	[11]
gtAI	Genetic tRNA Adaptation Index	0~1	0.51	[2]
P2	P2 Index	0~1	0.48	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
			0.4	

GC	GC Content	0~1	0.46	[1]
GC1 GC3	GC Content at the First Position GC Content at the Third Position of Synonymous Codons		0.56	
GC2	GC Content at the Second Position of Synonymous Codons		0.43	
ENcp	Effective Number of Codon Pairs	20~61	20	[2]
CPS	Codon Pair Score	-1~1	0.09	[3] , [4]
Codon Volatility	Codon Volatility	0.5~1	0.75	[5]

Gene Name: Alpha-1-antitrypsin

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

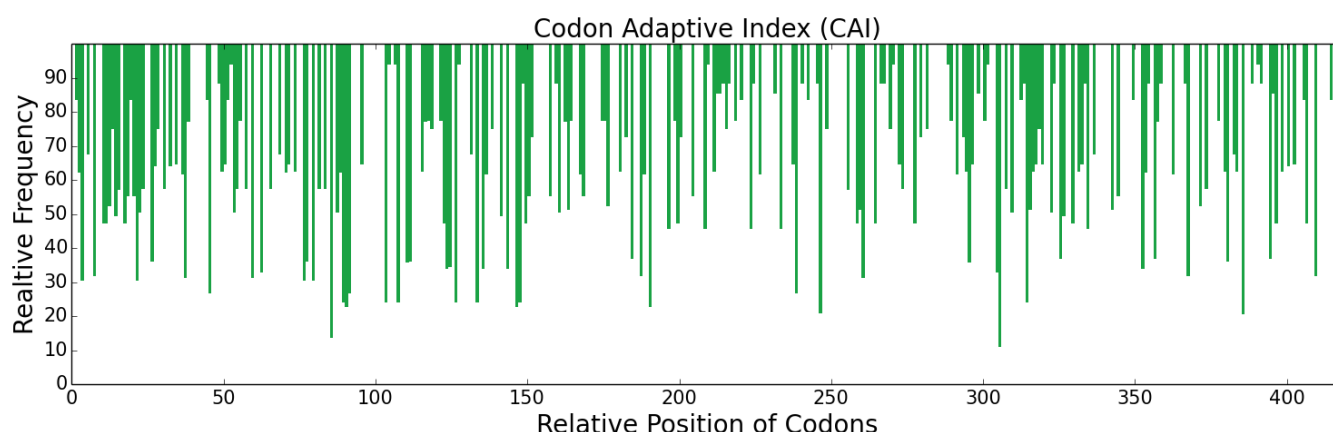
DNA Sequence Length: 1254

DNA Sequence:

```
ATGCCTTCATCGGTTTCTGCGGCATTTTGCTGCTGGCAGGATTATGCTGTCTGGTCCCT
GTCTCGCTTGCCGAAGATCCCCAGGGAGATGCCGCTCAGAAGACCGATACAAGCCACCAT
GATCAAGATCATCCTACGTTTAAACAAAATCACCCCTAATCTTGCCGAGTTTGCCTTTAGC
TTGTACAGGCAATTGGCCCATCAATCCAACCTCAACCAACATCTTTTTTTTCGCCCCGTTTCG
ATTGCCACTGCCTTTGCGATGCTTTCACTCGGGACGAAGGCTGATACCCATGATGAAATT
TTGGAAGGTCTCAATTTTAATCTCACTGAAATACCCGAAGCTCAAATCCACGAGGGATTT
CAAGAGCTGCTACGTACTCTCAATCAACCAGATTCCCAACTCCAATAACAACCTGGAAAC
GGTTTATTTCTATCTGAAGGGCTCAAACCTGGTCGACAAGTTTTTGGGAAGATGTCAAGAAA
CTTTACCACAGTGAGGCTTTTACAGTCAACTTTGGTGATACTGAGGAGGCAAAGAAGCAA
ATCAACGACTACGTAGAAAAGGGCACACAAGGGAAGATTGTTGATTTGGTGAAGGAGCTG
GACAGAGATACTGTCTTTGCTTTGGTGAATTACATCTTCTTCAAAGGAAAATGGGAGAGA
CCTTTTGAAGTGAAAGATACAGAAGAAGAAGATTTCCATGTGGATCAAGTTACCACGGTT
AAAGTTTCTATGATGAAACGATTGGGAATGTTTAAACATTCAACATTGCAAGAAGCTGAGT
AGCTGGGTTTTGCTGATGAAATATTTGGGAAATGCTACCGCCATTTTTTTTTCTGCCAGAC
GAAGGAAAGTTGCAACATTTGGAAAATGAGTTGACACATGACATCATAACCAAGTTCTTG
GAGAATGAAGATAGGCGCTCTGCCTCTCTTCATTTGCCTAAACTCAGTATCACCGGAACC
TACGATCTTAAATCTGTATTAGGTCAACTGGGTATCACCAAAGTGTTTCCAACGGTGCT
GATTTGAGTGGTGTCCTGAAGAAGCTCCTTTGAAGCTATCAAAGCTGTACACAAAGCT
GTTTTGACAATTGATGAAAAAGGCACTGAAGCTGCAGGTGCCATGTTTTTGGAGGCTATC
CCCATGTCCATCCCACCGGAAGTTAAATTTAATAAACCATTTGTATTCCTGATGATCGAA
CAGAACACCAAGTCTCCTCTGTTTATGGGCAAGGTTGTTAACCTACTCAAAAA
```

Protein Sequence:

```
MPSSVSWGILLLAGLCCCLVPVSLAEDPQGDAQKTDTSHHQDHPNFKITPNLAEFASF
LYRQLAHQSNSTNIFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTETPEAQIHEGF
QELLRTLNPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKKLYHSEAFVNFVDTEEAKKQ
INDYVEKGTQGKIVDLVKELDRDVFALVNYIFFKGKWERPFVEKDTSEEDFHVDQVTV
KVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNTAIFFLPDEGKLQHLENELTHDIIITKFL
ENEDRRSASLHLPKLSITGTYDLKSVLGLGITKVFNSGADLSGVTEEAPLKLSKAVHKA
VLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVNPQK
```



Negative CIS Elements	Negative repeat Elements
2	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.21	[11]
ENC	Effective Number of Codons	20~61	57.22	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.24	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.74	[51]
CDC	Codon Deviation Coefficient	0~1	0.16	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.47	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.15	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.12	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.85	[111]
P	Codon Preference	≥ 1	1.11	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.35	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.74	[11]
CFD	Codon Frequency Distribution	0~1	0.04	[21]
FOP	Frequency of Optimal Codons	0~1	0.48	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.74	[51]
COUSIN18			2.08	
CBI	Codon Bias Index	-1~1	0.21	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.39	[71]

RCA	Relative Codon Adaptation	≥ 0	1.13	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.2	[9]
B	Codon Usage Bias	$0 \sim 2$	0.29	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.51	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.59	[2]
P2	P2 Index	$0 \sim 1$	0.52	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.45	[1]
GC	GC Content		0.43	
GC1	GC Content at the First Position of Synonymous Codons		0.51	
GC2	GC Content at the Second Position of Synonymous Codons		0.33	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	27.78	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0.04	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.77	[5]

Gene Name: Tissue-type plasminogen activator

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

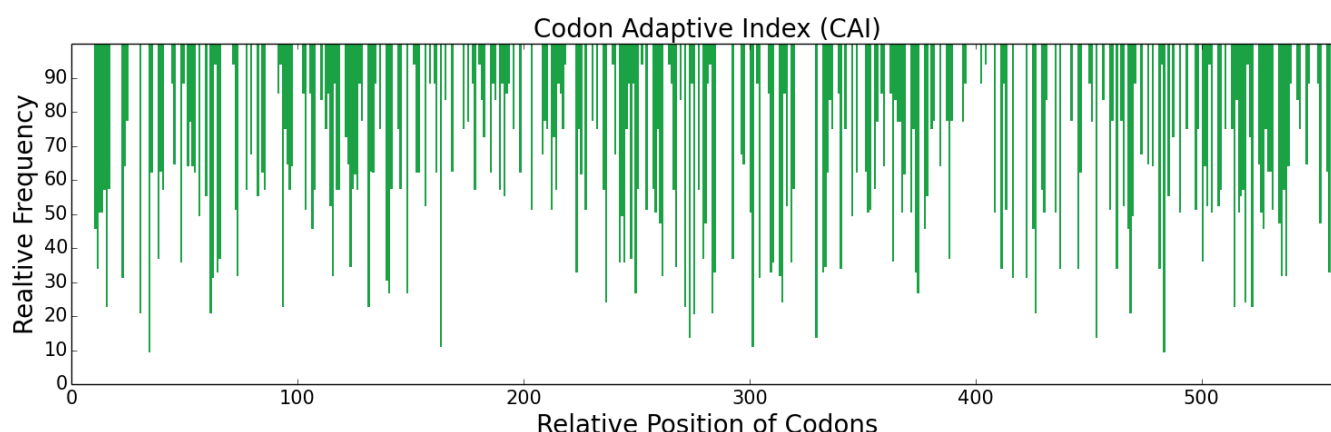
DNA Sequence Length: 1686

DNA Sequence:

```
ATGGATGCTATGAAGAGAGGTTTGTGTTGTGTGCTACTTCTTTGCGGGGCCGTTTTTGT  
TCTCCAAGCCAGGAGATTTCATGCTAGATTTTCAAGAGGTGCTCGGTCATACCAAGTAATC  
TGCAGAGATGAAAAAACCCAAATGATATATCAACAGCACCAGTCATGGTTAAGACCAGTC  
TTGCGAAGCAATAGGGTAGAATACTGTTGGTGTAAATAGTGGCAGAGCTCAATGCCATTCC  
GTTCCAGTCAAGTCATGCTCTGAACCAAGATGTTTCAATGGGGGAACCTGCCAGCAAGCT  
TTGTACTTCAGTGATTTTCGTGTGCCAATGTCTTGAAGGATTCGCAGGCAAAATGCTGCGAA  
ATTGACACCCGTGCCACATGCTATGAGGATCAAGGGATCTCATATAGAGGAACCTGGTCCG  
ACGGCCGAATCTGGAGCCGAATGTACGAACTGGAATTCATCAGCTTTGGCACAAAAACCA  
TATTACAGGTCGCAGACCTGATGCTATCAGATTGGGTTTGGGAAACCACAACTATTGCAGA  
AATCCTGACAGAGATTCAAAACCTTGGTGCTATGTCTTCAAAGCTGGAAAGTACTCATCT  
GAATTTTGTAGTACTCCAGCTTGTTCGAGGGAAACAGTGACTGCTATTTTCGGAAATGGT  
TCTGCTTACAGGGGAACACATAGTTTACTGAGTCTGGAGCTTCTTGCCCTCCCATGGAAT  
TCCATGATATTAATAGGAAAAGTATATACGGCCCAAAATCCAAGTGCTCAAGCCCTTGGA  
CTGGGCAAGCATAATTATTGCCGTAACCTGATGGGGATGCGAAACCGTGGTGCCATGTA  
CTGAAAAATCGAAGGTTGACTTGGGAATACTGTGATGTACCATCTTGTTCACCTGTGGT  
CTTCGCCAATATAGCCAACCACAATTCAGGATAAAAGGTTGGCTCTTCGCAGATATAGCC  
TCTCATCCATGGCAAGCTGCTATTTTTTGCAGCATAGGCGTTACCTGGAGAAAGATTC  
CTATGTGGAGGTATTTTAATTTTCATCTTGTGGATCCTTAGTGCTGCCACTGTTTCCAG  
GAAAGATTCCCCCTCACACCTTACAGTTATTCTTGAAGGACGTACAGAGTGGTCCCA  
GGAGAGGAAGAACAGAAGTTTGGAGTAGAGAAGTACATTGTTACAAAAGAAATTTGATGAT  
GATACTTATGATAATGATATTGCTCTTTTGCACATAAAAGTGATTCTAGCAGATGTGCT  
CAAGAAAGCTCTGTTGTGCGAACTGTTTGCTTCTCCAGCTGATCTTCAACTACCAGAT  
TGGACTGAGTGTGAACATACAGGTTACGGTAAACACGAAGCGTTGTCTCCTTTTACAGT  
GAGAGACTAAAGGAGGCACATGTGCGATTATATCCATCTTCCAGATGTACCTCTCAGCAT  
TTGCTAAATCGGACTGTCACTGACAACATGCTTTGTGCTGGAGATACTAGAAGTGGAGGT  
CCCCAGGCAAAATCTTCATGATGCATGCCAAGGAGATTCTGGAGGGCCTCTTGTCTGCCTC  
AATGACGGGAGAATGACCTTGTGGGAATCATCAGTTGGGGTCTGGGCTGCGGCCAGAAA  
GATGTTCTGAGTTTACACCAAAGTTACTAACTATCTGGATTGGATCAGGGATAACATG  
AGACCA
```

Protein Sequence:

```
MDAMKRGLLCVLLLCGAVFVSPSQEIHARFRRGARSYQVICRDEKTQMIYQQHQSWLRPV  
LRSNRVEYCWNSGRAQCHSVPVKSCSEPRCFNGGTCQQALYFSDFCVCPEGFAGKCCE  
IDTRATCYEDQGISYRGTWSTAESGAECTNWNSSALAQKPYSGRRPDAILGLGNHNYCR  
NPDRDSKPWCYVFKAGKYSSEFCSTPACSEGNSDCYFNGSAYRGTHSLTESGASCLPWN  
SMILIGKVYTAQNPSAQALGLGKHNCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCG  
LRQYSQPQFRIKGGLFADIASHWPQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQ  
ERFPPHHLTVILGRTYRVVPGEQKFEVEKYIVHKEFDDDTYDNDIALQLKSDSSRCA  
QESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQH  
LLNRTVTDNMLCAGDTRSGGPQANLHDACQGDGGPLVCLNDGRMTLVGIIISWGLGCGQK  
DVPGVYTKVTNYLDWIRDNMRP
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.24	[11]
ENC	Effective Number of Codons	20~61	50.48	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.18	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.54	[51]
CDC	Codon Deviation Coefficient	0~1	0.09	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.44	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.1	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.1	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.86	[111]
P	Codon Preference	≥ 1	1.09	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.44	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.72	[11]
CFD	Codon Frequency Distribution	0~1	0.05	[21]
FOP	Frequency of Optimal Codons	0~1	0.46	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.35	[51]
COUSIN18			0.08	
CBI	Codon Bias Index	-1~1	0.18	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.2	[71]

RCA	Relative Codon Adaptation	≥ 0	1.09	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.15	[9]
B	Codon Usage Bias	$0 \sim 2$	0.25	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.47	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.54	[2]
P2	P2 Index	$0 \sim 1$	0.58	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.36	[1]
GC	GC Content		0.45	
GC1	GC Content at the First Position of Synonymous Codons		0.5	
GC2	GC Content at the Second Position of Synonymous Codons		0.49	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	29.06	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0.04	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.76	[5]

Gene Name: Interleukin-2

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

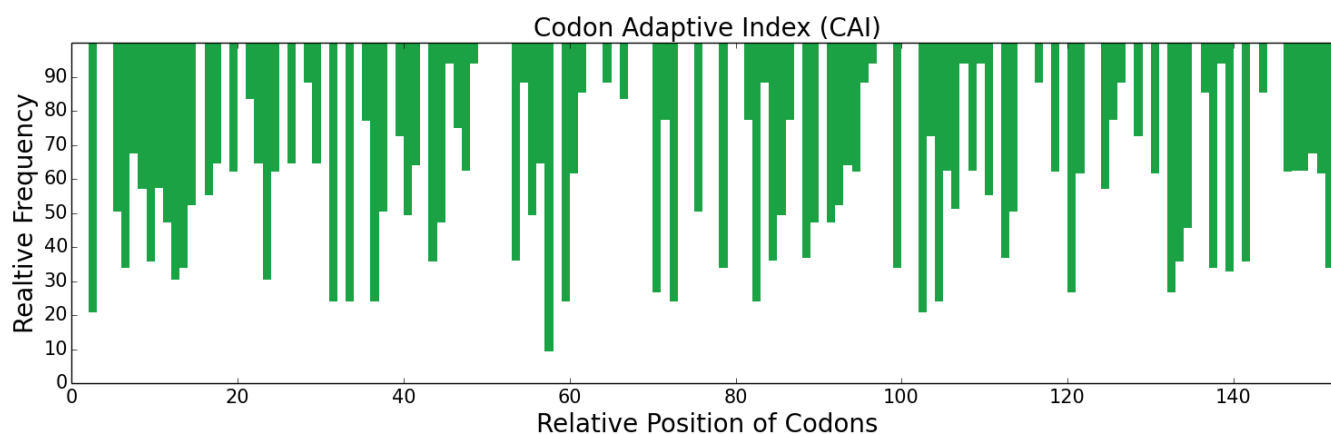
DNA Sequence Length: 459

DNA Sequence:

```
ATGTACCGAATGCAACTTCTATCCTGCATAGCCCTGTCGCTAGCATTGGTCACCAACTCA
GCTCCTACCTCGTCATCTACCAAGAAAACCCAACTCCAACCTCGAACACCTCCTTTTGGAC
TTACAGATGATACTGAATGGAATCAATAACTACAAGAACCCCAAATTAACCCGGATGCTC
ACATTCAAGTTTATATGCCTAAGAAGGCTACGGAGCTCAAGCATCTTCAATGTCTAGAA
GAAGAGCTCAAACCCCTTAGAGGAAGTACTGAACCTGGCACAGTCAAAAAATTTTCATCTA
AGACCACGAGACCTCATCAGTAATATCAATGTCATTGTACTTGAATTGAAAGGTTTCAGAA
ACGACATTTATGTGCGAGTATGCTGACGAAACAGCTACGATAGTGGAAATTCCTAAATAGG
TGGATAACTTTCTGTCAATCAATCATCTCCACACTAACA
```

Protein Sequence:

```
MYRMQLLSICIALSLALVTNSAPTSSSTKKTQLQLEHLLLDLQMLNGINNYKNPKLTRML
TFKFYMPKKATELKHLLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSE
TTFMCEYADETATIVEFLNRWITFCQSIISTLT
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.2	[11]
ENC	Effective Number of Codons	20~61	61	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.27	[41]

DCBS	Directional Codon Bias Score	≥ 1	1.62	[5]
CDC	Codon Deviation Coefficient	0~1	0.16	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.62	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.16	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.12	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.83	[11]
P	Codon Preference	≥ 1	1.13	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.23	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.63	[1]
CFD	Codon Frequency Distribution	0~1	0.08	[2]
FOP	Frequency of Optimal Codons	0~1	0.31	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	-0.22	[5]
COUSIN18			-0.19	
CBI	Codon Bias Index	-1~1	-0.04	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.7	[7]
RCA	Relative Codon Adaptation	≥ 0	1.05	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.35	[9]
B	Codon Usage Bias	0~2	0.45	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.42	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.52	[2]
P2	P2 Index	0~1	0.58	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.48	[1]
GC	GC Content		0.41	
GC1	GC Content at the First Position of Synonymous Codons		0.44	
GC2	GC Content at the Second Position of Synonymous Codons		0.32	

ENcp	Effective Number of Codon Pairs	20~61	20	[2]
CPS	Codon Pair Score	-1~1	0.08	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.75	[5]

Gene Name: Hepatitis B large envelope protein

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

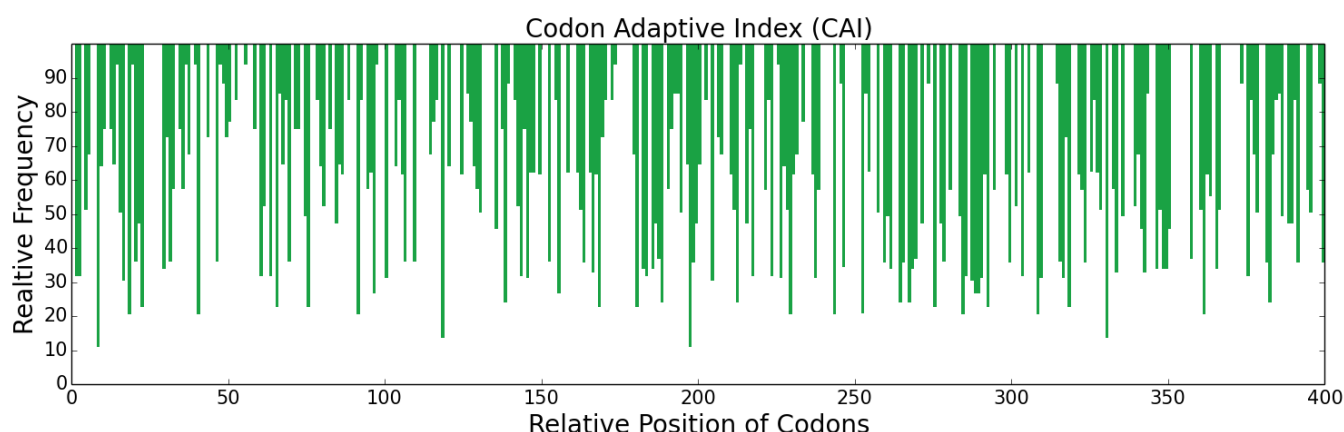
DNA Sequence Length: 1200

DNA Sequence:

```
ATGGGCGGCTGGAGTTCCAAGCCACGCCAGGGAATGGGAACCAATCTTTCGGTTCCGAAT
CCCCTGGGGTTTTTTCAGATCATCAACTAGACCCCGCCTTTGGAGCCAATTCCAACAAT
CCGGATTGGGACTTTAACCCCAATAAAGACCACTGGCCTGAAGCTAATCAAGTTGGAGCT
GGCGCATTTGGCCCAGGGTTACCCCTCCCCATGGAGGATTGTTAGGGTGGTCTCCTCAG
GCACAAGGAATTCTGACCACATTGCCTGCTGCTCCGCCTCCAGCCTCAACGAATAGACAA
AGCGGTAGACAGCCTACACCCATTTCTCCCCATTGAGAGATTCCCACCCTCAAGCGATG
CAGTGGAACTCTACAACCTTCCACCAGGCCCTTTTGGATCCAAGAGTGAGAGGACTCTAT
TTTCTGTCAGGCGGAAGCTCATCAGGTACAGTTAACCCCGTTCCTACGACTGCTTCACCA
ATTTCAAGTATATTTTCAAGGACAGGGGACCCTGCTCCTAATATGGAATCTACTACTTCC
GGGTTTCTAGGCCCACTACTGGTACTCCAAGCCGGATTCTTCCTTTTGACCCGCATACTG
ACCATTCTCAATCGTTGGACTCCTGGTGGACAAGTCTCAATTTTCTGGGAGGCGCTCCA
ACTTGCCCTGGCCAAAATAGCCAGAGTCCGACATCCAACCACTCTCCAACAAGCTGCCCC
CCAACTTGTCCGGGTATATCGTTGGATGTGTTTGAGACGATTCATCATTTTTCTTTTATA
TTACTATTGTGTCTCATATTTCTCCTAGTATTGCTGGATTATCAAGGGATGCTGCCCCTT
TGCCCATTGTTACCGGGCACTTCGACGACGAGCACAGGGCCATGCAGAACTTGTACAATA
CCAGCACAAAGGCACTTCAATGTTTCCGAGCTGTTGTTGTACTAAACCCAGCGACGGGAAC
TGTACATGCATACCAATCCCTTCAAGTTGGGCGTTTGCCAGGTTTTTATGGGAATGGGCA
TCCGTGAGGTTCTCTTGCTAAGTCTACTAGTGCCATTTGTTCAATGGTTTGTAGGTTTG
AGTCCGACAGTCTGGCTAAGTGTATTTGGATGATGTGGTATTGGGGCCCTTCCCTTTAC
AACATACTCTCCCCTTTCTTACCACTGCTGCCTATATTTTTTTGCCTTTGGGTTTATATA
```

Protein Sequence:

```
MGGWSSKPRQGMGTNLSVNPPLGFFPDHQLDPAFGANSNNPDWDFNPNKDHWPENQVGA
GAFGPGFTPPHGGLLGWSPQAQGILTTLPAAPPPASTNRQSGRQPTPI SPPLRDSHPQAM
QWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRTGDPAPNMESTTS
GFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTS NHSPTSCP
PTCPGYRWMCLRRFIIIFLIFLLLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCRTCTI
PAQGTSMFPPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFVGL
SPTVWLSAIWMMWYWGPSLYNILSPFLPLLPFIFFCLWVYI
```



Negative CIS Elements	Negative repeat Elements
2	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.09	[11]
ENC	Effective Number of Codons	20~61	54.49	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.25	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.7	[51]
CDC	Codon Deviation Coefficient	0~1	0.13	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.52	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.15	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.07	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.86	[111]
P	Codon Preference	≥ 1	1.08	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.22	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.63	[11]
CFD	Codon Frequency Distribution	0~1	0.08	[21]
FOP	Frequency of Optimal Codons	0~1	0.36	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.28	[51]
COUSIN18			-0.3	
CBI	Codon Bias Index	-1~1	0.08	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.53	[71]

RCA	Relative Codon Adaptation	≥ 0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.25	[9]
B	Codon Usage Bias	$0 \sim 2$	0.31	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.42	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.51	[2]
P2	P2 Index	$0 \sim 1$	0.41	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.43	[1]
GC	GC Content		0.49	
GC1	GC Content at the First Position of Synonymous Codons		0.48	
GC2	GC Content at the Second Position of Synonymous Codons		0.56	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	31.63	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0.08	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.74	[5]

Gene Name: HPV major capsid protein L1

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

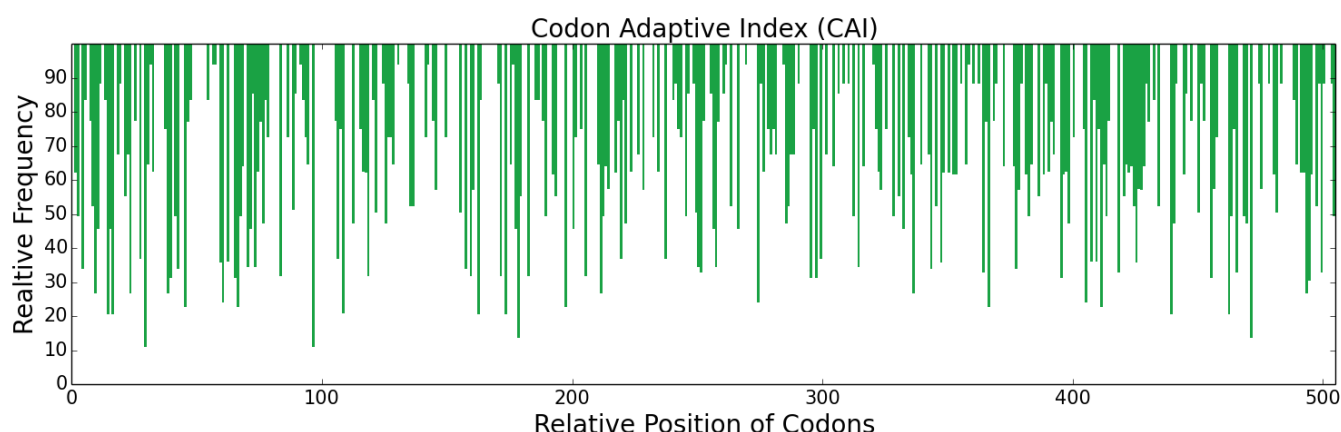
DNA Sequence Length: 1515

DNA Sequence:

```
ATGTCATTATGGCTACCTTCTGAGGCAACGGTGTATTTGCCTCCGGTGCCGGTTTCCAAA
GTTGTCTCCACGGATGAGTACGTAGCTCGCACCAATATCTACTACCATGCTGGAACGAGC
AGATTACTAGCTGTTGGGCACCCCTTACTTTCCAATTAAGAAGCCTAACATAATAAGATA
CTCGTTCCCAAGGTTAGCGGGTTACAGTACCGTGTGTTCGGTATCCACCTGCCTGACCCA
AACAGTTTGGCTTTCCAGACACTAGTTTCTACAATCCTGACACCCAACGCTTGTTTGG
GCTTGTGTTGGTGTGAGGTAGGACGAGGTCAACCACTGGGTGTTGGAATCTCAGGCCAT
CCTCTTTTGAACAACTGGACGACACCGAAAATGCTTCTGCTTATGCAGCAAACGCTGGT
GTTGACAATAGAGAGTGCATTTCTATGGACTACAAGCAAACCTCAACTTTGTCTAATTGGC
TGCAAGCCGCTATTGGTGAACATTGGGGTAAAGGCTCTCCGTGTACCAATGTGGCGGTC
AACCCAGGCGATTGTCTCCTTTGGAGTTAATTAACACAGTCATTCAAGATGGGGATATG
GTGGACACTGGATTTGGCGCTATGGATTTTACCACGTTACAGGCCAACAAAGTCAGAGGTA
CCTCTGGATATCTGTACTTCCATTTGCAAGTACCCAGACTACATCAAGATGGTATCTGAA
CCTTATGGAGACTCTTTATTCTTTTATCTTCGTAGGGAGCAAATGTTCTGTCGTCACCTG
TTCAATAGAGCAGGTGCTGTGGGTGAAAATGTTCCAGATGATCTCTATATCAAGGGATCC
GGATCCACTGCTAATCTGGCATCCTCCAACCTATTTTCCAACCTCCAAGCGGAAGCATGGTA
ACTTCCGATGCTCAGATTTTCAACAAACCATATTGGTTACAACGTGCTCAGGGTCATAAC
AATGGAATCTGCTGGGGAAACCAATTATTTGTCACTGTGGTTGACACAACGAGATCTACC
AACATGTCCCTATGTGCAGCTATATCAACTTCAGAAACAACATACAAAAACACCAATTTT
AAAGAATATTTGAGGCACGGGGAAGAGTATGATTTGCAGTTTATTTTTCAGCTATGCAAA
ATTACATTAACCGCTGATGTGCATGACATATATCCACTCCATGAACAGCACAAATCCTGGAA
GACTGGAACCTTTGGACTCCAACCCCTCCCGGAGGGACCTTAGAGGATACTTACAGGTTT
GTCACCTCACAGGCAATAGCCTGCCAGAAACACACTCCTCCAGCACCAAAGGAAGATCCG
CTGAAAAAGTACACATTCTGGGAGGTAAACCTTAAAGAGAAGTTTAGCGCCGACTTGGAT
CAATTTCCGTTAGGAAGGAAGTTTTTACTGCAAGCGGGTTTGAAAGCCAAGCCAAAATTT
ACACTTGGTAAAAGAAAGGCTACTCCTACCCTTCATCAACGTCGACAACTGCAAAAAGG
AAAAAGAGAAAATTA
```

Protein Sequence:

```
MSLWLPSEATVYLPPVPVSKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKI
LVPKVSGLQYRVFRIHLPDPNKFDPDTSFYNPDTQRLVWACVGVGVGRGQPLGVGISGH
PLLNKLDDTENASAYAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPTNVAV
NPGDCPPELINTVIQDGMVDTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSE
PYGDSLFFYLRRQMFVRHLFNAGAVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMV
TSDAQIFNKPWYLQRAQGHNNGICWGNQLFVTVVDTRSTNMSLCAAISTSETTYKNTNF
KEYLRHGEEDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDITYRF
VTSQAIACQKHTPPAPKEDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKF
TLGKRKATPTTSSTSTAKRKKRKL
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.12	[11]
ENC	Effective Number of Codons	20~61	55.76	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.16	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.45	[51]
CDC	Codon Deviation Coefficient	0~1	0.09	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.5	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.05	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.06	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.91	[111]
P	Codon Preference	≥ 1	1.07	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.24	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.71	[11]
CFD	Codon Frequency Distribution	0~1	0.05	[21]
FOP	Frequency of Optimal Codons	0~1	0.42	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.05	[51]
COUSIN18			1	
CBI	Codon Bias Index	-1~1	0.12	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.24	[71]

RCA	Relative Codon Adaptation	≥ 0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.14	[9]
B	Codon Usage Bias	$0 \sim 2$	0.21	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.49	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.56	[2]
P2	P2 Index	$0 \sim 1$	0.58	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.44	[1]
GC	GC Content		0.45	
GC1	GC Content at the First Position of Synonymous Codons		0.47	
GC2	GC Content at the Second Position of Synonymous Codons		0.42	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	29.5	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0.04	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.76	[5]

Gene Name: Covid-19 spike protein

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

DNA Sequence Length: 2994

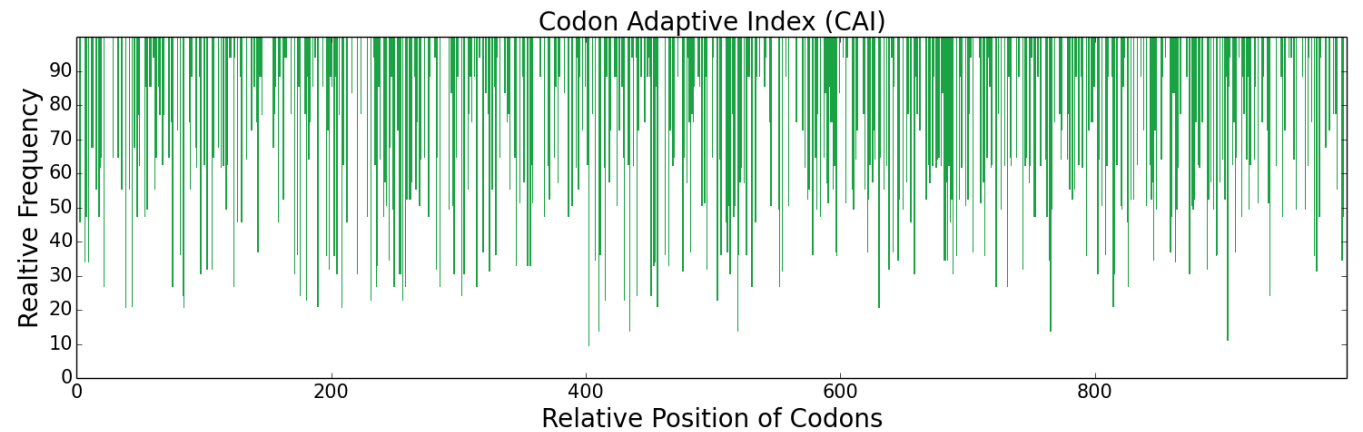
DNA Sequence:

```
ATGTTTGTGTTTTTGGTTCTACTGCCACTAGTTTCTCTCCCAATGTGTCAACCTGACAACC
AGAACGCAATTGCCACCAGCTTACACCAACTCTTTTACCAGAGGTGTCTACTACCCGGAT
AAGGTCTTTTCGATCTTCCGTTCTGCACTCAACTCAAGATCTGTTCTTACCATTCTTCTCT
AATGTCACCTGGTTCCACGCTATCCACGTTTCTGGTACCAACGGAACGAAGAGATTTGAC
AACCCCGTTCTCCCGTTTAACGATGGAGTCTATTTTGCTTCCACAGAAAAATCGAACATT
ATCAGAGGCTGGATTTTTGGCACCACCTTTGGATTCCAAGACACAATCATTGTTAATCGTT
AATAATGCTACGAATGTTGTGATTAAAGTGTGTGAATTCAGTTTTGTAACGACCCATTCT
TTGGGAGTATATTATCACAGAACAACAAGTCTTGGATGGAATCCGAGTTCAGAGTGTAT
TCTTCTGCAAATAATTGTACTTTTGAGTACGTTTCGCAACCCCTTCCTCATGGATTTGGAG
GGGAAACAGGGAACTTCAAGAATTTGCGAGAATTTGTTTTCAAGAACATAGACGGCTAT
TTCAAGATATATTTCGAAGCACACTCCGATCAACTTGGTGAGAGATTTGCCTCAAGGTTTT
TCGGCTTTGGAGCCATTGGTTGATCTGCCAATTGGGATTAATATCACGAGGTTCCAGACT
TTGCTGGCCCTTCATCGTTCTATTTAACGCCAGGTGACTCGTCGCTCTGGGTGGACGGCA
GGTGCAGCAGCCTACTATGTTCGGATACCTTCAGCCAAGAACCTTTTTGCTGAAATACAAC
GAAATGGCACCATTACGGATGCTGTTGATTGTGCTTTAGATCCTCTTTTCGGAAACCAAA
TGTAATCTCAAGTCGTTTACTGTGCGAGAAAGGAATTTATCAAACGTCTAATTTTAGAGTA
CAACCTACTGAGAGCATTGTTCAGATTCCCCAATATTACCAACTTGTGTCTTTTGGAGAG
GTCTTTAACGCTACCAGGTTTGTCTAGTGTATTATGCCTGGAACAGGAAGAGGATCAGTAAC
TGTGTTGCTGATTATTCTGTTCTGTACAACCTCAGCATCTTTTCTACCTTCAAGTGCTAT
GGTGTCTTCTCTACTAAGCTGAACGATCTTTGTTTTACAAATGTCTATGCTGACTCTTTC
GTTATCCGGGGTGATGAGGTTTCGTCAAATTCGCGCCGGTCAAACAGGGAAAAATTGCCGAC
TACAACATAAACTTCCAGACGATTTACCCGGGTGTGTTATCGCGTGGAATTCAAACAAT
CTCGACTCTAAAGTTGGTGAAACTATAATTATCTCTACAGGCTATTCCGAAAGTCTAAC
TTGAAGCCCTTTGAAAGGGACATTTCAACCGAAATTTATCAAGCTGGTAGCACCCCATGC
AATGGAGTAGAGGGATTCAACTGTTATTTCCCACTTCAAAGTTATGGCTTTCAACCAACC
AATGGTGTTGGGTACCAGCCCTACAGAGTTGTGGTACTTTTCGTTTGAGCTGCTTCATGCG
CCCGCCACAGTTTTCGGTCCCAAAAAATCTACGAACTTGGTGAAAAACAAATGTGTTAAC
TTCAATTTTAAACGGACTTACTGGTACTGGTGTTTTAAACGGAAAGCAACAAGAAATTTCTT
CCATTTCAACAATTTGGAAGAGATATTGCTGACACTACAGATGCAGTCAGAGATCCCCAA
ACCTTAGAGATTCTGGATATTACCCCTTGCAGTTTCGGAGGAGTCTCAGTAATAACTCCT
GGTACTAATACTAGTAATCAAGTTGCTGTCTTATACCAGGACGTTAACTGTACTGAGGTC
CCAGTAGCAATCCATGCCGACCAGTTGACTCCGACCTGGAGAGTCTACTCAACTGGCTCT
AATGTATTCCAAACTCGTGAGGTTGTTTAAATGGTGCCGAGCATGTGAACAACCTCGTAT
GAGTGTGACATTCCAATTGGTGCAGGTATCTGCGCTTCATATCAAACACAGACCAACTCA
CCTCGTCGTGCCCGTTTCAGTGGCATCGCAATCTATAATTGCATATACAATGTCTCTTGGT
GCAGAGAATTCTGTAGCTTACTCTAACAATAGTATTGCCATACCAACCAATTTACAATC
TCTGTTACGACTGAGATTTTACCAGTTTCAATGACGAAAACTCAGTTGATTGTACCATG
TATATTTGTGGCGATTCAACCGAATGCTCAAATTTGCTGCTGCAATATGGTTCAATTTGT
ACTCAGCTGAACCGTGCGTTAACTGGAATTGCTGTTGAGCAGGACAAGAACTCAGGAG
GTCTTTGCACAAGTCAAACAAATCTACAAAAACACCACCAATAAAGGACTTTGGAGGATTC
AACTTTTTCGCAGATTCTTCCAGATCCCTCAAAACCATCTAAACGATCGTTTATCGAAGAT
CTTTTATTCAATAAGGTGACGTTGGCAGATGCAGGTTTTATCAAGCAATATGGTGACTGT
```

CTTGGTGATATCGCAGCCCGTGACTTAATTTGTGCTCAGAAATTTAATGGTTTGACTGTA
CTGCCTCCTCTATTAAACAGATGAGATGATTGCTCAATATACTTCGGCTCTTTTAGCAGGA
ACTATCACATCTGGATGGACTTTTGGCGCAGGTGCCGCTTTACAAATACCCTTTGCCATG
CAGATGGCATATCGCTTCAATGGTATCGGAGTAACCCAAAATGTTCTGTATGAGAACCAG
AAATTAATCGCTAACCAGTTCAACAGTGCTATTGGAAAAATTCAGACAGTCTCTCTTCT
ACTGCTTCAGCTTTGGGTAAACTGCAAGACGTTGTTAATCAAAATGCTCAGGCTTTAAAC
ACTTTGGTTAAACAATTATCTTCAAATTTTGGAGCTATATCTAGCGTTCTGAATGATATT
TTGTCCAGATTGGACAAGGTTGAGGCTGAGGTCCAAATTGATCGTCTGATTACT

Protein Sequence:

MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVSNGTKRFDNPFVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIV
NNATNVVIKVCEFCNDPFLGVYHKNKSWMESEFRVSSANNCTFEYVSQPFLMDLE
GKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT
LLALHRSYLTGPDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIAD
YNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIQAGSTPC
NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVN
FNFNGLTGTGVLTESNKKFLPFQGFGRDIADTTDAVRDPQLEILDITPCSFGGVSVITP
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY
ECDIPIGAGICASYQTQTSNPRRARSVASQSIIAYTMSLGAENSVAYSNNNSIAIPTNFTI
SVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQE
VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLDAGFIKQYGDC
LGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAM
QMAYRFNGIGVTVQNVLYENQKLIANQFNSAIGKIQDLSLSTASALGKLQDVVNQNAQALN
TLVKQLSSNFGAISSVLNDILSRDLKVEAEVQIDRLIT



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.19	[11]
ENC	Effective Number of Codons	20~61	51.31	[2, 3]

RCBS	Relative Codon Bias Strength	≥ 0	0.16	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.5	[5]
CDC	Codon Deviation Coefficient	0~1	0.09	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.44	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.08	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.08	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.9	[11]
P	Codon Preference	≥ 1	1.06	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.37	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.73	[1]
CFD	Codon Frequency Distribution	0~1	0.04	[2]
FOP	Frequency of Optimal Codons	0~1	0.46	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.84	[5]
COUSIN18			0.57	
CBI	Codon Bias Index	-1~1	0.19	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.15	[7]
RCA	Relative Codon Adaptation	≥ 0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.12	[9]
B	Codon Usage Bias	0~2	0.18	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.5	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.58	[2]
P2	P2 Index	0~1	0.52	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.4	[1]
GC	GC Content		0.42	
GC1	GC Content at the First Position of Synonymous Codons		0.45	

GC2	GC Content at the Second Position of Synonymous Codons		0.41	
ENcp	Effective Number of Codon Pairs	20~61	39.14	[2]
CPS	Codon Pair Score	-1~1	0.01	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

Gene Name: C-Reactive Protein

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

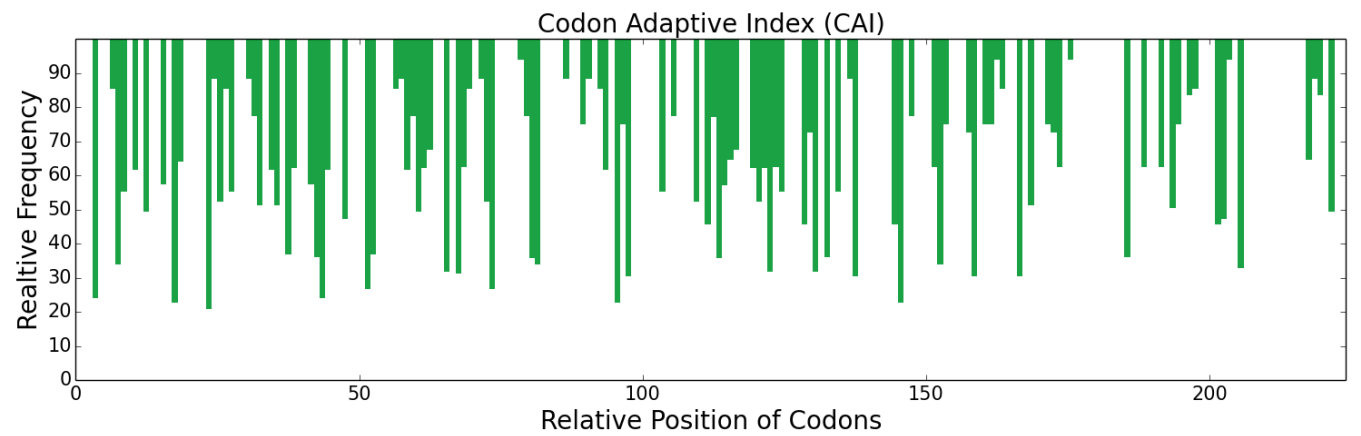
DNA Sequence Length: 672

DNA Sequence:

ATGGAAGCTCTTGTGTTTCCTAGTCTTGACATCTTTATCTCATGCCTTTGGGCAGACT
GATATGTCTCGAAAAGCATTTCGTCTTTCCAAAAGAGAGTGATACAAGTTACGTATCATTG
AAGCCCCCCTCACAAAGCCACTGAAGGCTTTTACGGTATGTTTGCATTTCTATACAGAG
TTATCATCCACTAGAGGCTACAGCATCTTCTCTTATGCAACGAAGAGACAAGATAATGAG
ATACTAATTTTTTGGTCTAAAGATATTGGATATTCTTTCACAGTTGGGGGATCGGAAATT
TTGTTTGAAGTCCCAGAGGTTACTGTTGCACCAAGTGCACATATGCACCTCCTGGGAATCA
GCATCAGGCATCGTCAATTTTGGGTGGACGGCAAGCCCAGAGTCAGAAAATCGTTGAAG
AAGGGTTACACTGTGGGGGCTGAGGCTTCTATTATCCTAGGACAAGAACAAGACTCGTTT
GGAGGAAATTTGCAAGGTTGCGAAAGTTTGGTTGGAGACATCGGTAATGTTAACATGTGG
GATTTTGTCTTGTCTCCCGATGAAATCAACACTATCTACCTTGGAGGTCCTTTCTCTCCA
AACGTGCTGAATTGGAGGGCTTTGAAGTACGAAGTTCAAGGTGAAGTTTTTACCAAACCT
CAATTATGGCCA

Protein Sequence:

MEKLLCFLVLTSLSHAFGQTDMSRKAFVFPKESDTSYVSLKAPLTKPLKAFTVCLHFYTE
LSSTRGYSIFSATKRQDNEILIFWSKDIGYSFTVGGSEILFEVPEVTVAPVHICTSWES
ASGIVEFWVDGKPRVRKSLKKGTVGAESIILGQEQDSFGGNFEGSQSLVGDIGNVNMW
DFVLSPDEINTIYLGPFSPNVLNWRALKYEVQGEVFTKPQLWP



Negative CIS Elements	Negative repeat Elements
1	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.25	[11]
ENC	Effective Number of Codons	20~61	52.79	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.25	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.64	[5]
CDC	Codon Deviation Coefficient	0~1	0.15	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.54	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.13	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.11	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.85	[11]
P	Codon Preference	≥ 1	1.11	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.34	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.74	[11]
CFD	Codon Frequency Distribution	0~1	0.04	[2]
FOP	Frequency of Optimal Codons	0~1	0.49	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.41	[5]
COUSIN18			1.47	
CBI	Codon Bias Index	-1~1	0.24	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.18	[7]
RCA	Relative Codon Adaptation	≥ 0	1.15	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.19	[9]
B	Codon Usage Bias	0~2	0.27	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.5	[11]
gtAI	Genetic tRNA Adaptation Index	0~1	0.59	[2]
P2	P2 Index	0~1	0.53	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
			0.42	

GC	GC Content	0~1	0.43	[1]
GC1 GC3	GC Content at the First Position GC Content at the Third Position of Synonymous Codons		0.45	
GC2	GC Content at the Second Position of Synonymous Codons		0.41	
ENcp	Effective Number of Codon Pairs	20~61	25.79	[2]
CPS	Codon Pair Score	-1~1	0.01	[3] , [4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

Gene Name: Lysosomal acid glucosylceramidase

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

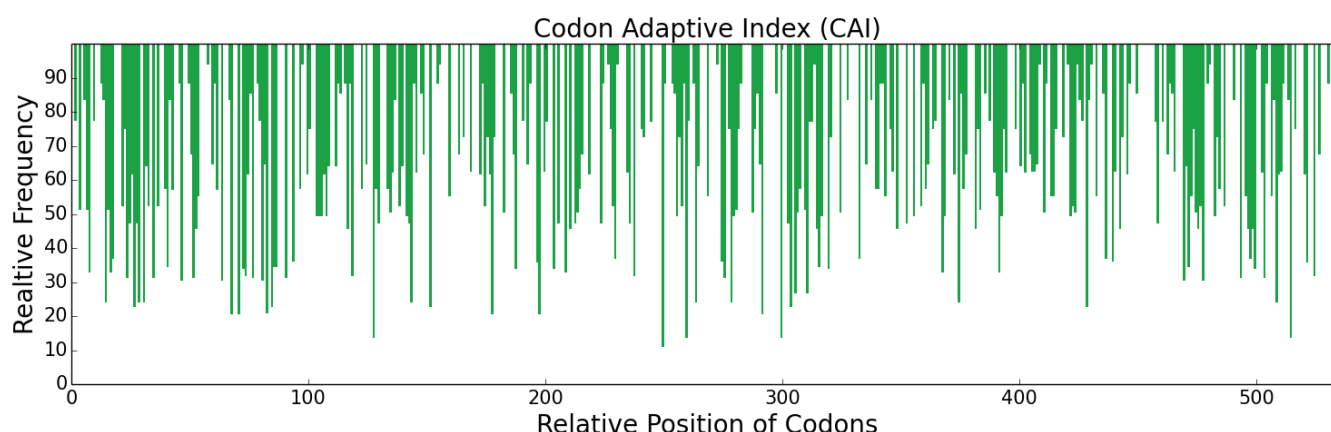
DNA Sequence Length: 1608

DNA Sequence:

```
ATGGAGTTTAGTTCTCCTAGTAGGGAAGAGTGTCCAAAACCTCTCAGTAGGGTATCTATT
ATGGCAGGAAGCCTGACAGGGCTGCTCTTGCTCCAGGCAGTTAGCTGGGCATCTGGTGCC
CGTCCTTGCAATTCCAAAATCGTTTGGTTATTCCAGCGTGGTCTGTGTTTGTAAATGCTACC
TATTGCGATTTCGTTTGATCCTCCGACTTTTCCGGCTCTAGGCACATTCAGCAGATATGAG
TCGACCCGATCTGGGCGTCGTATGGAATTGAGCATGGGTCCCATTCAAGCCAATCATACA
GGAAGTGGTTTATTATTAAACATTACAGCCAGAACAGAAATTCAAAAAGTGAAAGGCTTT
GGTGGTGCCATGACCGATGCTGCGGCCCTGAACATTTTGGCCCTTTCACCTCCAGCACAG
AACTTACTGCTCAAATCATACTTCTCCGAAGAAGGGATTGGTTATAATATTATTAGAGTC
CCAATGGCTTCCTGTGACTTTTCTATCAGAACTTACACATATGCAGACACACCGGACGAT
TTTCAACTTCATAACTTCTCCCTACCAGAAGAGGATACCAAATTGAAGATAACCGTTGATC
CACAGAGCTCTACAACTGGCTCAAAGGCCAGTGCTCTGCTTGCTTCCCATGGACATCT
CCAATTGGCTGAAAATAATGGAGCAGTAAATGGTAAGGGTTCAGTGAAGGGCCAACCA
GGAGACATTTACCACCAAACCTTGGGCTCGCTATTTTGTATAAATTCTTAGACGCATATGCG
GAGCATAAACTCCAGTTTGGGCTGTCACTGCTGAAAATGAACCCAGCGCTGGACTCTTA
AGTGGATATCCATTTCAATGTCTTGGATTACCCCCGGAACATCAAAGAGATTTTCATTGCG
AGAGATCTGGGGCCAACGCTTGCCAACAGTACGCACCACAATGTGCGTTTATTGATGCTA
GACGATCAAAGACTTTTGTTCCTCATTGGGCTAAGGTAGTTTGTACCGATCCTGAAGCC
GCCAAATATGTCCATGGAATCGCTGTGCATTGGTACCTGGATTTTTTAGCTCCAGCAAAA
GCCACCTTGGGAGAGACTCATAGGTTATTTCTTAACACAATGCTCTTCGCCTCCGAAGCT
TGTGTGGGAAGTAAGTTCTGGGAGCAATCAGTCAGGTAGGATCATGGGATAGAGGAATG
CAGTATTCACATTCCATCATACCAATTTGCTTTATCATGTCGTCGGATGGACTGACTGG
AATTTAGCACTTAATCCTGAGGGTGGGCCTAATTGGGTCAGAACTTCGTAGATTCTCCC
ATCATTGTGGACATTACAAAAGATACTTTCTACAAGCAACCAATGTTTTACCACCTGGGT
CACTTTTCCAAATTCATCCAGAAGGTTTCGACGCGTGTGCGACTTGTGGCATCGCAAAAA
AATGATTTAGACGCCGTTGCATTGATGCATCCTGATGGTAGCGCTGTCGTGGTAGTGCTA
AACAGATCAAGCAAAGATGTCCCTCTCACAATCAAAGATCCTGCGGTTGGATTTTGGAA
ACAATATCTCCAGGCTACTCCATTCATACTTATTTGTGGAGACGCCAG
```

Protein Sequence:

```
MEFSSPSREECPKPLSRVSIAGSLTGLLLLQAVSWASGARPCIPKSFYSSVVCVNCAT
YCDSFDPPTFPALGTFSTRYESTRSGRRMELSMGPIQANHTGTGLLLTLQPEQKFQKVKGF
GGAMTDAAALNILALSPPAQNLLLSYFSEEGIGYNIIRVPMASCDFSIRTYTYADTPDD
FQLHNFSLPEEDTKLKIPLIHRALQLAQRPVSLASPWTSPTWLKTNGAVNGKSLKGQP
GDIYHQTWARYFVKFLDAYAEHKLQFWAVTAENEPSAGLLSGYPFQCLGFTPEHQRFIA
RDLGPTLANSTHHNVRLMLDDQRLLLPHWAKVVLTDPEAAKYVHGIHVHWYLDLFLAPAK
ATLGETHRLFPNTMLFASEACVGSKFWEQSVRLGSDRGMQYSHSIITNLLYHVVGWTDW
NLALNPEGPNWVRNFVDSPIIVDITKDTFYKQPMFYHLGHFSKFIPEGSQVRGLVASQK
NDLDAVALMHPDGSAAVVVVLNRSSKDVPLTIKDPVGVFLETISPGYSIHTYLWRRQ
```



Negative CIS Elements	Negative repeat Elements
2	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.12	[11]
ENC	Effective Number of Codons	20~61	54.2	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.15	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.44	[51]
CDC	Codon Deviation Coefficient	0~1	0.08	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.49	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.09	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.07	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.88	[111]
P	Codon Preference	≥ 1	1.07	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.26	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.69	[11]
CFD	Codon Frequency Distribution	0~1	0.05	[21]
FOP	Frequency of Optimal Codons	0~1	0.4	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.02	[51]
COUSIN18			-0.19	
CBI	Codon Bias Index	-1~1	0.11	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.28	[71]

RCA	Relative Codon Adaptation	≥ 0	1.09	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.14	[9]
B	Codon Usage Bias	$0 \sim 2$	0.22	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.46	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.54	[2]
P2	P2 Index	$0 \sim 1$	0.53	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.42	[1]
GC	GC Content		0.46	
GC1	GC Content at the First Position of Synonymous Codons		0.51	
GC2	GC Content at the Second Position of Synonymous Codons		0.44	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	28.55	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0.04	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.75	[5]

Gene Name: Coagulation factor IX

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

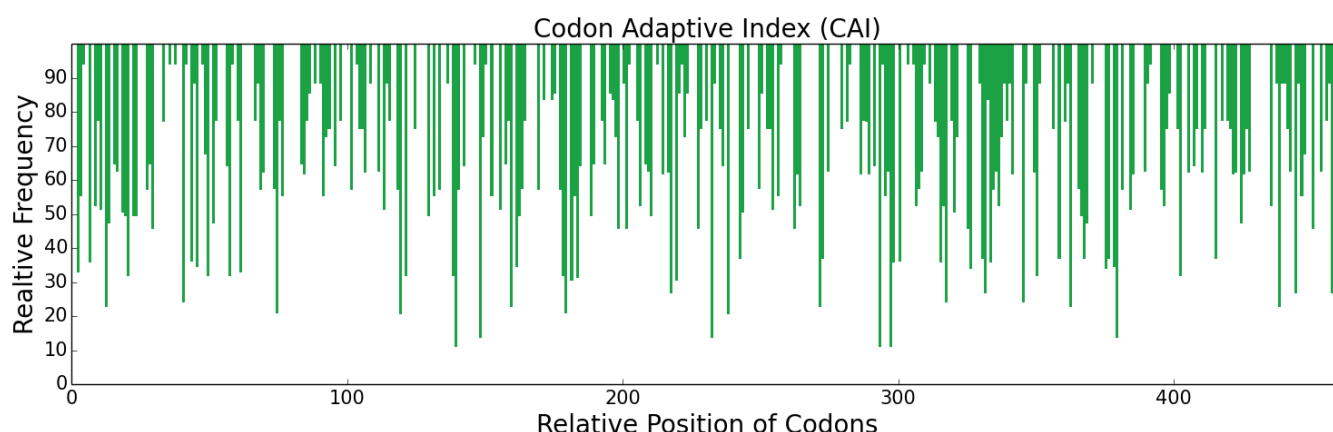
DNA Sequence Length: 1383

DNA Sequence:

```
ATGCAAAGGGTCAATATGATAATGGCAGAGAGTCCAGGGCTGATTACCATCTGTCTTTTA
GGCTACTTATTATCTGCTGAATGCACCGTGTTCCTGGATCACGAAAATGCTAATAAGATT
CTCAATAGACCCAAACGTTACAATTCCGGCAAGCTGGAGGAATTTGTTTCAGGGCAATTTG
GAGAGGGAATGTATGGAAGAGAAATGCTCATTGGAAGAGCCCGAGAGGTCTTTGAAAAC
ACTGAAAGAACCACAGAGTTCTGGAACAATATGTCGACGGAGATCAGTGTGAGTCTAAC
CCATGCTTGAATGGAGGATCATGTAAAGATGATATCAACAGTTATGAGTGTGGTGCCCG
TTTGGCTTTGAAGGAAAGAACTGTGAATTAGATGTCACTTGCAACATTAACGCGCCG
TGCGAACAGTTTTGTAAGAATTCTGCGGACAATAAGGTCGTTTGTAGTTGTACCGAGGGG
TACCGTTTAGCCGAGAACCAAAAGTCTTGCGAACCTGCTGTTCCCTTCCCATGCGGCCGA
GTTTCGGTCAGCCAGACTTCTAAGTTAACCAGAGCTGAGACCGTTTCCCTGACGTGGAT
TATGTGAATTCTACTGAGGCAGAAACCATCTTAGATAATATTACACAATCAACGCAATCG
TTCAATGACTTCACTAGAGTTGTGGGAGGTGAGGATGCGAAACCAGGACAGTTTCCGTGG
CAAGTTGTACTTAACGGAAGGTTGATGCCTTCTGTGGAGGAAGTATTGTCAATGAAAAG
TGGATTGTGACAGCAGCTCATTGTGTTGAAACTGGGGTAAAGATCACTGTTGTTGCTGGA
GAACACAATATTGAAGAAACAGAGCACACAGAACAGAAAGCGCAATGTCATCCGCATAATT
CCCCATCATAATTACAATGCAGCCATCAATAAGTATAACCACGACATAGCACTCTTGGAG
CTTGACGAACCATTGGTGCTAAACTCTTATGTAACGCCTATATGCATCGCAGACAAAGAG
TATACAAACATTTTTCTCAAATTTGGTTTCAGGCTATGTTTCTGGTTGGGGAAGAGTATTT
CACAAAGGGAGATCTGCCTTAGTACTGCAATATTTGAGAGTCCACTAGTAGATCGTGCG
ACTTGCTTGAGAAGTACAAAGTTTACTATCTATAATAACATGTTTTCGCGCAGGATTCCAT
GAAGGAGGCAGAGATTCATGTCAGGGAGATTCAGGAGGTCCACATGTAAGTGAAGTTGAG
GGAACATCATTTCTGACAGGAATCATTTCTTGGGGTGAAGAATGTGCAATGAAAGGGAAA
TATGGAATCTACACGAAAGTCTCCAGATACGTGAACTGGATCAAGGAGAAAACGAAACTA
ACA
```

Protein Sequence:

```
MQRVNMIMAESPLITICLLGYLLSAECTVFLDHENANKILNRPKRYNSGKLEEFVQGNL
ERECMEEKCSFEEAREVFENTERTEFWKQYVDGDQCESNPCLNGGSCKDDINSYECWCP
FGFEGKNCELDVTCNIKNRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVFPFCGR
VSVSQTSKLTRAETVFPDQVYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPW
QVVLNGKVDAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHEQKRNVIIRI
PHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIAADKEYTNIFLKFGSGYVSGWGRVF
HKGRSALVLQYLRVPLVDRATCLRSTKFTIYNMFCAGFHEGGRDSCQGDSSGGPHVTEVE
GTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKETKLT
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.13	[11]
ENC	Effective Number of Codons	20~61	55.24	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.16	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.47	[51]
CDC	Codon Deviation Coefficient	0~1	0.09	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.51	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.06	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.06	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.92	[111]
P	Codon Preference	≥ 1	1.05	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.23	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.72	[11]
CFD	Codon Frequency Distribution	0~1	0.05	[21]
FOP	Frequency of Optimal Codons	0~1	0.42	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.17	[51]
COUSIN18			0.2	
CBI	Codon Bias Index	-1~1	0.1	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.24	[71]

RCA	Relative Codon Adaptation	≥ 0	1.06	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.15	[9]
B	Codon Usage Bias	$0 \sim 2$	0.22	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.48	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.57	[2]
P2	P2 Index	$0 \sim 1$	0.44	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.41	[1]
GC	GC Content		0.43	
GC1	GC Content at the First Position of Synonymous Codons		0.47	
GC2	GC Content at the Second Position of Synonymous Codons		0.39	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	36.69	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0.05	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.77	[5]

Gene Name: Adalimumab light chain

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

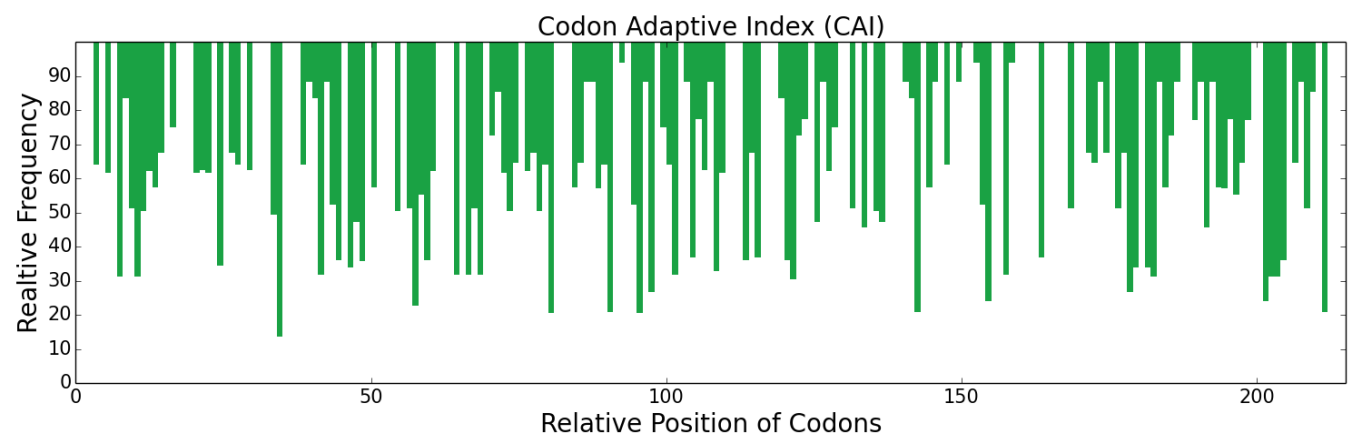
DNA Sequence Length: 645

DNA Sequence:

```
ATGGATATTCAGATGACACAAAGCCCTAGTAGCCTTTTCAGCCTCCGTTGGAGATAGAGTT
ACAATCACATGTCGTGCTTCCCAGGGTATCAGAACTACTTAGCGTGGTACCAACAGAAA
CCTGGCAAAGCACCCAAGCTACTGATATACGCCGCTTCTACTCTTCAAAGTGGGGTCCCC
TCAAGATTTTCTGGCTCTGGCAGTGGCACTGACTTCACACTTACCATTTTCATCCCTTCAG
CCGGAAGATGTTGCCACCTATTATTGCCAGCGATACAATAGAGCACCGTATACGTTTGGA
CAGGGCACTAAAGTAGAGATCAAAAGGACAGTTGCTGCTCCCTCCGTATTTATTTTCCT
CCCTCGGACGAGCAACTGAAATCAGGAACTGCTAGTGTTGTGTGTCTTCTGAACAACTTT
TATCCTCGAGAAGCCAAAGTTCAGTGGAAAGTTGATAATGCACTCCAATCTGGCAATTCT
CAAGAATCTGTAAGTGAACAAGATAGTAAGGATTCCACCTATTCCTTGAGTTCCACGCTA
ACTCTAAGCAAAGCCGACTATGAAAAGCACAAAGTGTATGCCTGCGAGGTCACCCACCAA
GGTCTCAGCAGCCCCGTTACCAAAAGTTTCAACCGAGGTGAATGT
```

Protein Sequence:

```
MDIQMTQSPSSLASVGDVRTITCRASQGIRNYLAWYQQKPGKAPKLLIYAASLTQSGVP
SRFSGSGSGTDFTLTISSLQPEDVATYYCQRYNRAPYTFGQGTKVEIKRTVAAPSVFIFP
PSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLISKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference

RSCU	Relative Synonymous Codon Usage	0~6	1.22	[11]
ENC	Effective Number of Codons	20~61	55.4	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.22	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.56	[5]
CDC	Codon Deviation Coefficient	0~1	0.13	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.55	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.15	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.11	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.87	[11]
P	Codon Preference	≥ 1	1.09	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.27	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.67	[11]
CFD	Codon Frequency Distribution	0~1	0.05	[2]
FOP	Frequency of Optimal Codons	0~1	0.36	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.03	[5]
COUSIN18			-0.05	
CBI	Codon Bias Index	-1~1	0.05	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.57	[7]
RCA	Relative Codon Adaptation	≥ 0	1.06	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.29	[9]
B	Codon Usage Bias	0~2	0.37	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.46	[11]
gtAI	Genetic tRNA Adaptation Index	0~1	0.54	[2]
P2	P2 Index	0~1	0.43	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.41	[11]
GC	GC Content		0.46	

GC1	GC Content at the First Position of Synonymous Codons		0.49	
GC2	GC Content at the Second Position of Synonymous Codons		0.48	
ENcp	Effective Number of Codon Pairs	20~61	20	[21]
CPS	Codon Pair Score	-1~1	0.07	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[51]

Gene Name: Adalimumab heavy chain

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

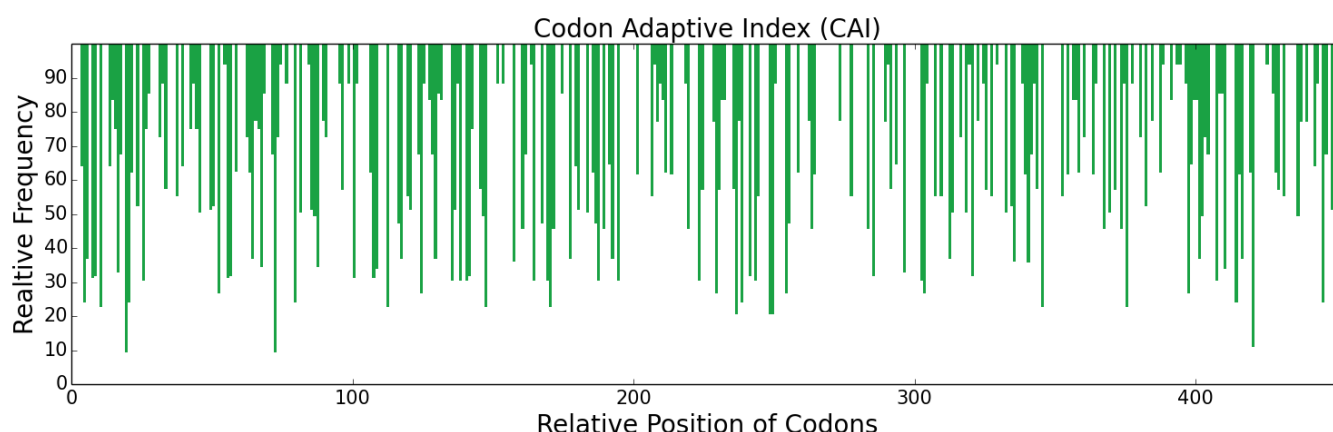
DNA Sequence Length: 1356

DNA Sequence:

```
ATGGAAGTTCAGCTCGTAGAAAGCGGCGGTGGGTTGGTTCAGCCTGGAAGGTCCTTGCGG
CTCTCATGTGCAGCTTCGGGATTCACCTTTTGATGACTATGCCATGCATTGGGTGAGACAG
GCTCCAGGAAAAGGACTTGAATGGGTAGTGCAATTACGTGGAATAGCGGCCATATCGAT
TACGCTGACTCAGTAGAGGGACGTTTCACTATTTCCCGGGACAATGCTAAAAACTCTCTC
TACCTTCAAATGAATAGTTTACGTGCTGAGGACACTGCTGTTTACTATTGCGCTAAAGTT
AGCTATTTGTCTACTGCTTCAAGCCTAGATTACTGGGGGCAAGGTACTCTGGTAACTGTC
AGTTCTGCTTCCACGAAAGGTCCTTCCGTATTCCCTTTGGCTCCATCGAGTAAATCGACT
TCGGGCGGAACTGCTGCCTTAGGGTGTGGTTAAAGATTATTTTCCAGAACCCGTTACT
GTGTCCTGGAATTCGGGTGCTCTGACTTCGGGGGTGCATACTTTCCAGCTGTATTGCAG
AGTTCTGGTCTTTACTCACTGTGCTGTGGTTACCGTACCATCGTCTTCTTTGGGTACT
CAAACATACATTTGTAACGTCAATCACAAACCTTCAAACACAAAGGTTGATAAGAAAGTG
GAACCAAAGTCGTGCGATAAGACTCACACGTGCCCTCCTTGTCAGCCCCGGAGCTCTTG
GGTGGCCCATCGGTCTTTTTGTTTCCGCCGAAACCAAAGGATACGCTGATGATTTCAAGA
ACTCCAGAGGTGACATGTGTTGTTGTTGATGTTTCTCATGAGGATCCAGAAGTCAAGTTT
AACTGGTACGTGGATGGCGTTGAAGTTCACAATGCCAAGACCAAGCCAAGGGAAGAACAA
TACAACTCGACGTATAGAGTTGTCTCTGTCTTGACTGTACTTCATCAAGACTGGCTTAAT
GGCAAGGAGTACAAATGCAAGGTCTCTAATAAGGCTCTTCCAGCACCCATTGAAAAACA
ATATCCAAAGCCAAGGGGCAACCAAGAGAACCACAAGTCTACACATTGCCTCCTTCAAGA
GACGAATTGACAAAAAACCAAGTGTCTCTTACTTGCTTGGTAAAAGGGTTTTATCCATCT
GACATTGCAGTTGAGTGGGAATCAAATGGTCAACCTGAAAATAATTACAAAACGACCCCT
CCTGTATTAGACTCCGATGGTTTCGTTCTTCTTATACTCTAAGCTCACAGTAGATAAGTCA
CGCTGGCAACAAGGTAATGTTTTCTCATGCTCTGTCATGCATGAAGCTTTACACAACCAC
TAACTCAGAAATCTCTCTCCTTGAGTCCAGGGAAA
```

Protein Sequence:

```
MEVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSAITWNSGHID
YADSVETRFTISRDNKNSLYLQMNSLRAEDTAVYYCAKVSYLSTASSLDYWGQGLTVTV
SSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQ
SSGLYSLSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKHTCPPCPAPELL
GGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ
YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR
DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKS
RWQQGNVVFSCSVMEALHNHYTQKSLSLSPGK
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.2	[11]
ENC	Effective Number of Codons	20~61	52.5	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.21	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.55	[51]
CDC	Codon Deviation Coefficient	0~1	0.13	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.47	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.1	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.09	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.88	[111]
P	Codon Preference	≥ 1	1.09	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.35	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.71	[11]
CFD	Codon Frequency Distribution	0~1	0.06	[21]
FOP	Frequency of Optimal Codons	0~1	0.47	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.61	[51]
COUSIN18			0.6	
CBI	Codon Bias Index	-1~1	0.22	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.21	[71]

RCA	Relative Codon Adaptation	≥ 0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.17	[9]
B	Codon Usage Bias	$0 \sim 2$	0.24	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.49	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.56	[2]
P2	P2 Index	$0 \sim 1$	0.6	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.42	[1]
GC	GC Content		0.46	
GC1	GC Content at the First Position of Synonymous Codons		0.5	
GC2	GC Content at the Second Position of Synonymous Codons		0.46	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	61	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0.01	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.75	[5]

Gene Name: Rituximab light chain

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

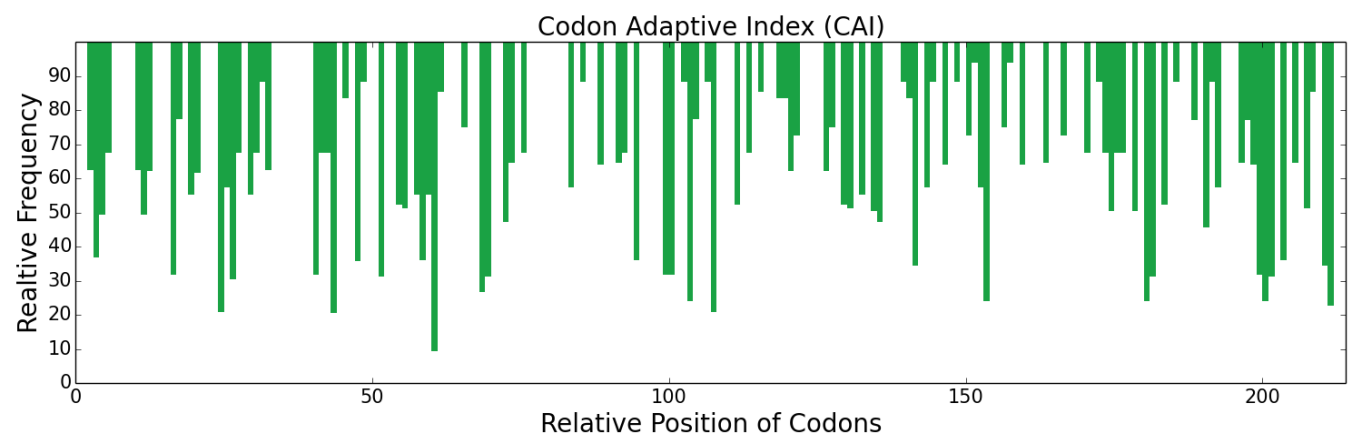
DNA Sequence Length: 642

DNA Sequence:

ATGCAAAATCGTATTATCCCAATCTCCAGCTATCTTATCAGCTTCTCCAGGCGAGAAGGTC
ACAATGACTTGTGCGAGCCTCGTCCTCTGTCTCCTATATCCATTGGTTTCAACAAAAGCCA
GGCTCCTCCCCGAAGCCTTGGATATATGCTACTAGCAACTTGGCAAGTGGTGTCCCCGTC
CGGTTCTCTGGTTCTGGATCTGGTACGAGCTACTCTCTGACCATTTCAGAGTTGAAGCT
GAAGATGCTGCCACTTATTACTGTGCAATGGACCTCCAACCCCCCACTTTTGGTGGC
GGCACTAAACTCGAGATTAAACGAAGTGTGCTGCACCATCCGTTTTCATTTTTCTCCTCT
TCAGACGAACAATTGAAGTCAGGAAGTGAAGTGTGCTGTCTTCTGAACAACTTTTAT
CCTCGTGAAGCCAAAGTTTCAAGTGGAAAGTTGACAATGCCCTCCAATCTGGAAATTCTCAG
GAATCTGTTACCGAACAAGACTCTAAGGATTCCACTTATTCCCTTTCTCCACTCTTACT
CTCAGCAAGGCAGATTATGAAAAGCACAAGGTGTATGCCTGTGAAGTTACCCACCAGGGC
CTCAGCTCTCCCGTTACCAAGAGTTTCAACCGTGGGGAATGT

Protein Sequence:

MQIVLSQSPAILSASPGEKVTMTCRASSSVSYIHWFQQKPGSSPKPWIYATSNLASGVPV
RFSGSGSGTSYSLTISRVEAEDAATYYCQQWTSNPPTFGGGKLEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLSTLT
LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference

RSCU	Relative Synonymous Codon Usage	0~6	1.35	[11]
ENC	Effective Number of Codons	20~61	54.55	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.3	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.8	[5]
CDC	Codon Deviation Coefficient	0~1	0.18	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.48	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.19	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.18	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.79	[11]
P	Codon Preference	≥ 1	1.14	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.47	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.72	[11]
CFD	Codon Frequency Distribution	0~1	0.05	[2]
FOP	Frequency of Optimal Codons	0~1	0.47	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.6	[5]
COUSIN18			1.59	
CBI	Codon Bias Index	-1~1	0.22	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.46	[7]
RCA	Relative Codon Adaptation	≥ 0	1.12	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.28	[9]
B	Codon Usage Bias	0~2	0.39	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.52	[11]
gtAI	Genetic tRNA Adaptation Index	0~1	0.59	[2]
P2	P2 Index	0~1	0.5	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.43	[11]
GC	GC Content		0.47	

GC1	GC Content at the First Position of Synonymous Codons		0.48	
GC2	GC Content at the Second Position of Synonymous Codons		0.5	
ENcp	Effective Number of Codon Pairs	20~61	22.99	[21]
CPS	Codon Pair Score	-1~1	0.02	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.75	[51]

Gene Name: Rituximab heavy chain

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: *Pichia pastoris* (budding yeasts)

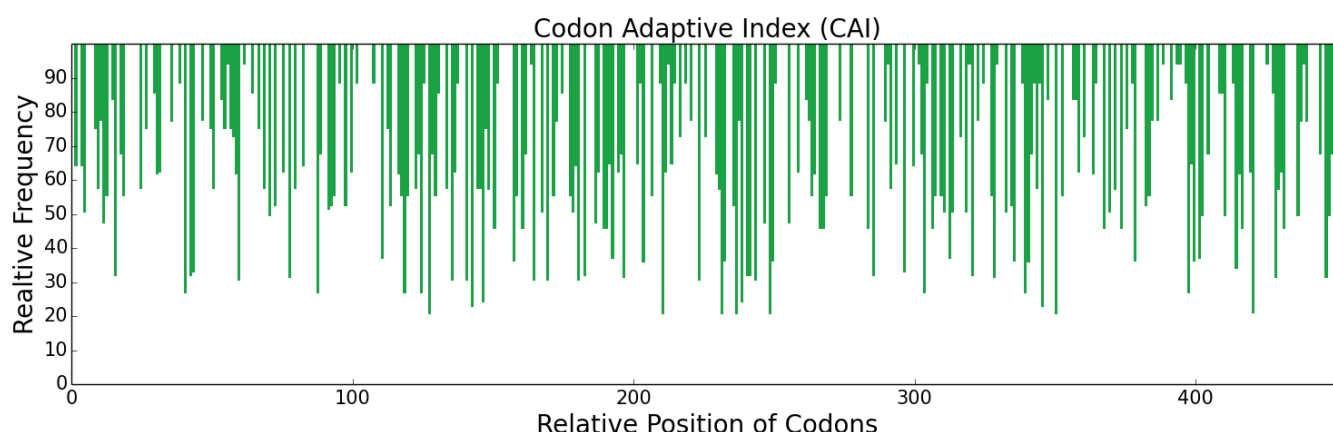
DNA Sequence Length: 1356

DNA Sequence:

```
ATGCAGGTTTCAGCTTCAACAACCAGGAGCCGAGCTGGTCAAGCCTGGCGCTTCCGTCAAG
ATGTCTTGTAAGGCCTCTGGATACACTTTTCACATCATACAACATGCACTGGGTAAACAA
ACGCCAGGCAGGGGTTTGGAGTGGATTGGAGCCATTTACCCTGGAAATGGAGACACATCG
TACAATCAAAAGTTCAAGGGAAAGGCCACTTTAACTGCAGATAAGTCATCTAGCACTGCC
TACATGCAGTTGTCTTCTTTGACGTCCGAAGATAGTGCAGTCTACTATTGTGCAAGATCA
ACTTATTACGGTGGTGATTGGTATTTTAACGTATGGGGAGCAGGTACTACAGTCACGGTC
TCTGCTGCCTCCACGAAAGGTCCGTCCGTCTTCCCATTGGCCCCATCGTCAAAATCTACT
TCGGGTGGGACTGCCGCCCTCGGATGCTTGGTGAAAGATTACTTTCCAGAACCCGTCACT
GTGTCCTGGAATTCGGGTGCTCTTACTTCGGGTGTCCACACTTTCCCAGCTGTCCTTCAG
TCGTCTGGCTTGTA CTCTGT CATCTGTGGTGACCGTACCATCATCCAGCTTGGGTACT
CAAACCTATATATGTAACGTCAACCATAAACCGTCAAATACCAAAGTTGACAAGAAAAGCT
GAGCCAAAGTCGTGTGACAAGACTCATACATGCCCCCCTGTCCAGCACCGGAGCTCTTG
GGCGGCCCATCGGTTTTTCTGTTTCCGCCCAAACCAAAGGATACTCTGATGATTTCAAGA
ACTCCTGAGGTACATGTGTGGTGGTGCATGTTTCTCATGAGGATCCAGAAGTCAAGTTT
AACTGGTACGTGGATGGCGTTGAAGTTCACAATGCCAAGACCAAGCCAAGGGAAGAACAG
TACAATTCCACGTATAGAGTGGTCTCTGTCTTACTGTACTTCATCAAGACTGGCTTAAT
GGCAAGGAGTACAAATGTAAGGTGAGCAATAAGGCTCTTCCAGCACCCATTGAAAAACG
ATATCCAAAGCCAAAGGGCAACCTAGAGAACCGCAAGTCTACACTTTGCCTCCTTCAAGA
GACGAATTGACAAAAAACCAAGTGTCTCTTACTTGCTTGGTGAAGGGATTTTATCCCTCT
GATATTGCAGTCGAGTGGGAGTCTAATGGTCAACCTGAAAATAATTACAAAACGACCCCC
CCAGTATTAGATTCCGATGGTTCTTTCTTCTTATACTCTAACTAACAGTGGATAAGTCA
CGATGGCAACAAGGTAATGTTTTTCAGCTGCTCAGTGATGCATGAAGCTTTACACAATCAC
TACACTCAAAGTCCTTGAGCTTATCCCCAGGTAAA
```

Protein Sequence:

```
MQVQLQQPGAELVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRGLEWIGAIYPNGDTS
YNQKFKGKATLTADKSSSTAYMQLSSLTSEDSAVYYCARSTYYGGDWYFNVWGAGTTTV
SAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQ
SSGLYSLSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKAEPKSCDKTHTCPPCPAPELL
GGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ
YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR
DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKS
RWQQGNVVFSCSVMHEALHNHYTQKSLSLSPGK
```



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.23	[11]
ENC	Effective Number of Codons	20~61	48.85	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.27	[41]
DCBS	Directional Codon Bias Score	≥ 1	1.64	[51]
CDC	Codon Deviation Coefficient	0~1	0.17	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.45	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.12	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.11	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.86	[111]
P	Codon Preference	≥ 1	1.12	[121]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.38	[131]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.72	[11]
CFD	Codon Frequency Distribution	0~1	0.04	[21]
FOP	Frequency of Optimal Codons	0~1	0.45	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.72	[51]
COUSIN18			0.63	
CBI	Codon Bias Index	-1~1	0.19	[61]
Dmean	Mean Dissimilarity-based Index	0~2	0.22	[71]

RCA	Relative Codon Adaptation	≥ 0	1.1	[8]
CUFS	Codon Usage Frequency Similarity	$0 \sim \sqrt{2}$	0.22	[9]
B	Codon Usage Bias	$0 \sim 2$	0.28	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	$0 \sim 1$	0.5	[1]
gtAI	Genetic tRNA Adaptation Index	$0 \sim 1$	0.57	[2]
P2	P2 Index	$0 \sim 1$	0.51	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	$0 \sim 1$	0.46	[1]
GC	GC Content		0.47	
GC1	GC Content at the First Position of Synonymous Codons		0.47	
GC2	GC Content at the Second Position of Synonymous Codons		0.46	
ENcp	Effective Number of Codon Pairs	$20 \sim 61$	28.46	[2]
CPS	Codon Pair Score	$-1 \sim 1$	0	[3] , [4]
Codon Volatility	Codon Volatility	$0.5 \sim 1$	0.75	[5]