GenRCA Rare Codon Analysis Report

Gene Name: YFP

Reference Source: Codon Usage Database - Kazusa

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

DNA Sequence Length: 582

DNA Sequence:

ATGTTCAAAGGAATCGTAGAAGGCATTGGGATCATTGAAAAGATCGACATTTACACTGAT
CTTGATAAATATGCTATTAGGTTTCCTGAAAATATGCTCAATGGTATCAAAAAGAGTCT
TCCATCATGTTTAATGGTTGTTTTCTGACTGTGACTTCCGTTAACAGTAATATAGTGTGG
TTTGATATTTTTGAAAAAGAAGCTAGGAAATTGGATACATTTAGAGAATACAAAGTTGGC
GACAGAGTAAACTTAGGAACATTCCCCAAATTTGGTGCTGCATCTGGTGGTCACATTCTA
AGTGCTAGAATTAGTTGTGTGGCCAGTATTATTGAGATCATCGAAAATGAAGATTATCAA
CAAATGTGGATTCAAATCCCTGAAAACTTTACGGAATTTCTGATTGATAAGGATTATATA
GCCGTCGATGGCATATCTCTAACCATTGACACAATCAAGAATAATCAATTTTTTATATCA
TTACCGTTGAAAATAGCACAAAACACAAACATGAAATGGAGGAAGAAGGGTGACAAAGTT
AACGTGGAACTGAGCAATAAAATTAACGCAAATCAATGCTGG

Protein Sequence:

MFKGIVEGIGIIEKIDIYTDLDKYAIRFPENMLNGIKKESSIMFNGCFLTVTSVNSNIVW
FDIFEKEARKLDTFREYKVGDRVNLGTFPKFGAASGGHILSARISCVASIIEIIENEDYQ
QMWIQIPENFTEFLIDKDYIAVDGISLTIDTIKNNQFFISLPLKIAQNTNMKWRKKGDKV
NVELSNKINANQCW
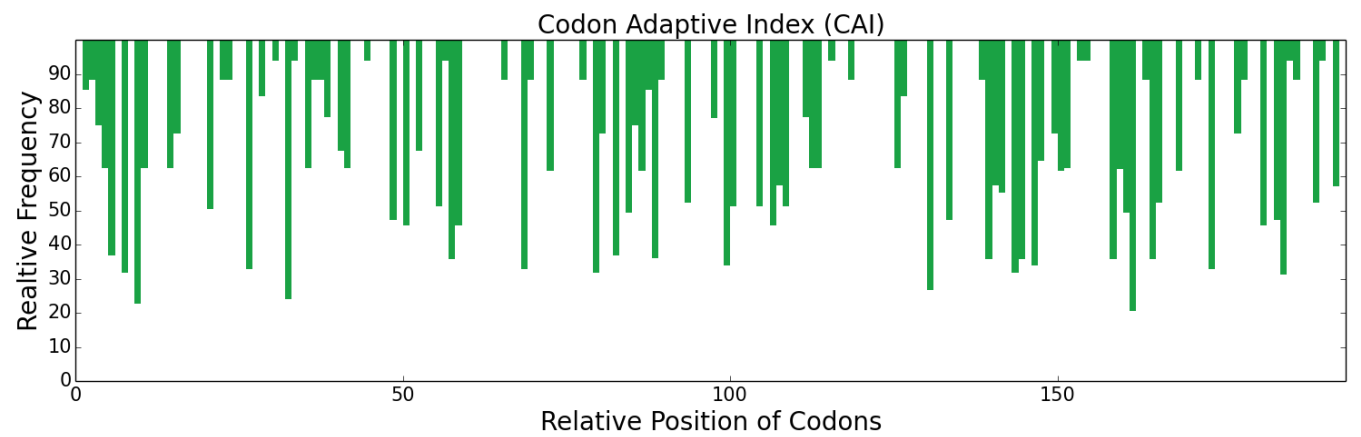
Codon Adaptive Index (CAI)

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	1.28	[1]
ENC	Effective Number of Codons	20~61	53.87	[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	0.1	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.52	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.26	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.18	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.77	[11]
P	Codon Preference	≥1	1.12	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.33	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.75	[1]
CFD	Codon Frequency Distribution	0~1	0.02	[2]
FOP	Frequency of Optimal Codons	0~1	0.48	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.57	[5]
COUSIN18			0.35	
CBI	Codon Bias Index	-1~1	0.18	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.48	[7]
RCA	Relative Codon Adaptation	≥0	1.09	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.34	[9]
B	Codon Usage Bias	0~2	0.35	[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	0.6	[2]
P2	P2 Index	0~1	0.43	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
	GC Content at the Third Position			

GC3	of Synonymous Codons	0~1	0.36	[1]
GC	GC Content		0.35	
GC1	GC Content at the First Position of Synonymous Codons		0.38	
GC2	GC Content at the Second Position of Synonymous Codons		0.3	
ENcp	Effective Number of Codon Pairs	20~61	20	[2]
CPS	Codon Pair Score	-1~1	0.03	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.79	[5]

Gene Name: Beta-galactosidase

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)
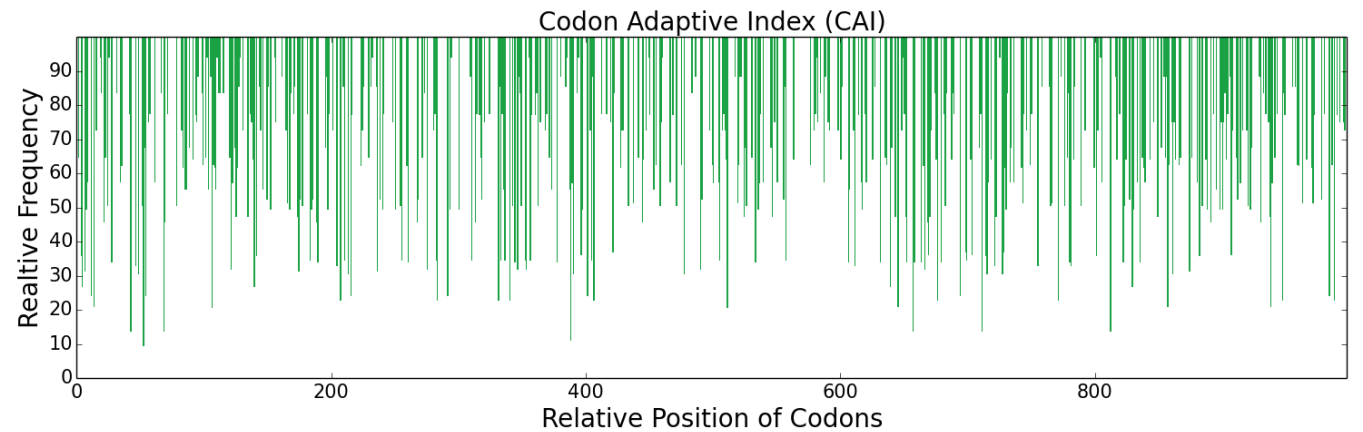
DNA Sequence Length: 2994

DNA Sequence:

```
ATGACCATGATAACGGATAGCTTAGCCGTTGTTCTCCAACGAAGAGACTGGGAAAATCCT
GGTGTGACCCAACTTAATAGACTAGCTGCTCATCCTCCATTTGCCTCATGGAGAAACTCT
GAAGAGGCGAGAACTGATAGGCCATCGCAACAACTTCGGTCCCTCAACGGAGAGTGGAGA
TTTGCCTGGTTTCCAGCTCCTGAAGCGGTGCCAGAGTCTTGGTTGGAATGTGATCTTCCA
GAAGCTGACACAGTTGTCGTCCCATCCAACTGGCAGATGCACGGATATGATGCTCCTATC
TACACCAATGTCACTTATCCGATCACAGTCAATCCTCCTTTTGTTCCTACTGAAAACCCA
ACCGGCTGCTACTCCCTGACATTCAATGTTGACGAATCTTGGCTGCAAGAGGGACAGACG
AGAATAATTTTCGACGGTGTCAACTCTGCATTCCATTTATGGTGTAATGGAAGATGGGTT
GGTTATGGTCAAGACAGTAGATTACCTTCTGAGTTCGATCTGAGCGCATTTCTTAGAGCT
GGTGAGAACCGTTTAGCAGTTATGGTGCTAAGATGGTCTGATGGTTCCTATTTAGAGGAT
CAAGACATGTGGAGGATGTCTGGGATTTTCCGTGATGTTTCGTTGCTCCACAAGCCAACT
ACTCAAATTTCAGACTTCCATGTTGCTACCAGATTCAATGATGATTTCAGCAGAGCAGTT
TTAGAGGCTGAAGTTCAAATGTGTGGAGAATTAAGAGATTACCTTCGTGTTACTGTTTCA
CTATGGCAAGGTGAAACTCAAGTGGCATCTGGTACCGCTCCTTTTGGCGGTGAAATTATT
GACGAGCGTGGGGGTTACGCTGATAGAGTTACTCTCAGATTAAATGTTGAAAACCCAAAG
TTATGGTCTGCTGAAATTCCAAACTTGTATCGTGCTGTTGTGGAGTTGCACACCGCAGAT
GGAACTTTGATTGAGGCTGAAGCTTGTGATGTTGGGTTCCGTGAGGTCCGTATTGAAAAC
GGGTTGCTTTTGCTAAACGGCAAACCTCTTTTGATTCGTGGCGTTAATCGTCACGAACAT
CACCCTCTTCATGGACAAGTTATGGACGAGCAAACCATGGTCCAAGATATTCTATTGATG
AAACAAAACAACTTCAATGCTGTCCGCTGCTCGCATTACCCTAATCATCCCTTATGGTAC
ACTCTCTGTGACCGTTATGGGTTGTACGTTGTTGATGAGGCTAACATTGAGACTCATGGA
ATGGTACCTATGAACAGATTGACAGACGACCCAAGATGGCTTCCAGCTATGAGTGAAAGA
GTTACCAGAATGGTGCAGAGAGACAGAAATCACCCATCTGTCATTATCTGGTCTCTTGGA
AATGAATCTGGTCATGGTGCCAACCACGATGCTCTTTACAGATGGATCAAGTCGGTTGAT
CCATCTAGACCTGTTCAATATGAAGGTGGTGGCGCAGATACTACTGCTACTGACATTATC
TGCCCAATGTACGCCCGTGTTGACGAGGACCAGCCGTTTCCAGCTGTTCCAAAATGGAGT
ATCAAAAAATGGCTGTCCCTTCCAGGTGAAACCAGACCACTAATTTTATGCGAGTACGCC
CATGCTATGGGTAACTCCCTGGGTGGATTTGCCAAGTACTGGCAAGCATTCCGTCAATAC
CCAAGATTGCAGGGTGGTTTTGTTTGGGATTGGGTTGATCAATCTTTGATCAAGTACGAC
GAGAATGGAAACCCTTGGTCTGCCTATGGTGGAGACTTTGGTGATACTCCAAACGACAGA
CAGTTCTGTATGAACGGTCTAGTCTTTGCTGACAGGACTCCACACCCAGCCTTAACTGAG
GCCAAGCATCAACAACAGTTTTTCCAATTTAGACTATCTGGTCAAACCATCGAAGTTACG
TCCGAATACCTGTTTCGACATTCTGACAATGAGCTGCTACATTGGATGGTTGCGCTAGAT
GGTAAGCCACTAGCTTCAGGCGAAGTGCCCCTGGACGTTGCTCCTCAGGGGAAGCAACTA
ATTGAGCTTCCTGAATTGCCACAGCCTGAGTCTGCTGGTCAACTCTGGTTGACTGTACGT
GTTGTTCAGCCCAACGCTACTGCTTGGTCCGAGGCGGGTCATATATCGGCCTGGCAGCAA
TGGAGGCTGGCTGAAAATTTGTCGGTAACATTACCTGCTGCCTCCCATGCCATTCCACAT
TTGACTACAAGTGAGATGGACTTTGTATCGAGTTGGGTAACAAGAGGTGGCAATTCAAC
AGACAATCTGGTTTCCTTAGTCAAATGTGGATTGGGGATAAAAGCAACTTTTGACCCCT
CTAAGGGATCAGTTCACTAGAGCTCCACTTGATAACGACATTGGTGTTTCTGAAGCTACA
AGAATAGACCCAAATGCCTGGGTTGAAAGATGGAAGGCGGCTGGTCATTATCAGGCTGAG
GCTGCTCTACTTCAGTGTACTGCAGATACGTTAGCTGATGCCGTTCTAATTACCACAGCC
```

CATGCTTGGCAGCATCAAGGAAAGACTCTGTTTATTTCCAGAAAAACCTATCGAATCGAT
GGATCGGGACAGATGGCTATCACCGTTGATGTTGAAGTTGCTAGCGATACCCCACATCCA
GCCAGAATAGGTCTTAACTGTCAGTTAGCTCAAGTGGCTGAGAGAGTCAACTGGTTAGGA
TTAGGACCTCAGGAGAACTATCCCGACAGATTGACCGCAGCTTGCTTTGACAGATGGGAC
CTTCCATTATCCGATATGTACACTCCATATGTGTTTCCTTCTGAGAACGGACTGCGATGC
GGTACCAGAGAGTTGAACTACGGGCCTCACCAATGGAGAGGTGATTTCCAATTCAACATC
TCAAGATACAGTCAACAACAGTTGATGGAAACAAGTCACAGACATTTGTTGCATGCAGAA
GAGGGTACTTGGCTCAACATCGATGGGTTTCACATGGGAATTGGTGGAGACGAT

Protein Sequence:

MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLRSLNGEWR
FAWFPAPEAVPESWLECDLPEADTVVVPSNWQMHGYDAPIYTNVTYPITVNPPFVPTENP
TGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLPSEFDLSAFLRA
GENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDVSLLHKPTTQISDFHVATRFNDDFSRAV
LEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGGYADRVTLRLNVENPK
LWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLLLLNGKPLLIRGVNRHEH
HPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLWYTLCDRYGLYVVDEANIETHG
MVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSLGNESGHGANHDALYRWIKSVD
PSRPVQYEGGGADTTATDIICPMYARVDEDQPFPAVPKWSIKKWLSLPGETRPLILCEYA
HAMGNSLGGFAKYWQAFRQYPRLQGGFVWDWVDQSLIKYDENGNPWSAYGGDFGDTPNDR
QFCMNGLVFADRTPHPALTEAKHQQQFFQFRLSGQTIEVTSEYLFRHSDNELLHWMVALD
GKPLASGEVPLDVAPQGKQLIELPELPQPESAGQLWLTVRVVQPNATAWSEAGHISAWQQ
WRLAENLSVTLPAASHAIPHLTTSEMDFCIELGNKRWQFNRQSGFLSQMWIGDKKQLLTP
LRDQFTRAPLDNDIGVSEATRIDPNAWVERWKAAGHYQAEAALLQCTADTLADAVLITTA
HAWQHQGKTLFISRKTYRIDGSGQMAITVDVEVASDTPHPARIGLNCQLAQVAERVNWLG
LGPQENYPDRLTAACFDRWDLPLSDMYTPYVFPSENGLRCGTRELNYGPHQWRGDFQFNI
SRYSQQQLMETSHRHLLHAEEGTWLNIDGFHMGIGGDD

Codon Adaptive Index (CAI)

	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	1.28	[1]
ENC	Effective Number of Codons	20~61	46.3	[2, 3]

RCBS	Relative Codon Bias Strength	≥0	0.25	[4]
DCBS	Directional Codon Bias Score	≥1	1.77	[5]
CDC	Codon Deviation Coefficient	0~1	0.13	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.38	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.11	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.11	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.82	[11]
P	Codon Preference	≥1	1.13	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.66	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.76	[1]
CFD	Codon Frequency Distribution	0~1	0.03	[2]
FOP	Frequency of Optimal Codons	0~1	0.52	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.55	[5]
COUSIN18			1.87	
CBI	Codon Bias Index	-1~1	0.29	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.09	[7]
RCA	Relative Codon Adaptation	≥0	1.13	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.12	[9]
B	Codon Usage Bias	0~2	0.2	[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	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.59	[2]
P2	P2 Index	0~1	0.66	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons		0.4	
GC	GC Content	0~1	0.46	[1]
GC1	GC Content at the First Position of Synonymous Codons		0.55	

GC2	GC Content at the Second Position of Synonymous Codons		0.43	
ENcp	Effective Number of Codon Pairs	20~61	35.62	[2]
CPS	Codon Pair Score	−1~1	−0.01	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

Gene Name: Cre

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

DNA Sequence Length: 1029

DNA Sequence:

```
ATGTCCAATCTTTTGACTGTCCATCAAAATCTTCCTGCTCTACCTGTTGATGCAACCTCT
GACGAAGTCAGGAAGAATCTAATGGATATGTTCAGAGACCGACAAGCATTTTCAGAGCAT
ACTTGGAAGATGCTGTTGAGCGTTTGTCGTTCATGGGCAGCGTGGTGCAAACTCAATAAC
CGTAAGTGGTTCCCCGCTGAACCTGAGGATGTGAGAGATTATCTATTGTATTTACAGGCT
CGTGGGCTGGCCGTTAAGACGATACAGCAACACTTAGGGCAGCTAAATATGTTGCATAGA
CGTTCCGGGCTACCTAGACCTTCAGATTCCAATGCAGTGTCTCTGGTCATGCGTCGGATA
CGTAAGGAGAATGTGGACGCCGGTGAACGAGCCAAGCAAGCTTTAGCGTTCGAAAGAACA
GACTTTGACCAAGTGAGAAGTCTCATGGAAAACTCAGATCGATGTCAAGATATTAGGAAC
CTGGCATTCTTGGGAATCGCTTACAACACTTTACTTCGTATCGCTGAAATTGCCAGGATC
AGGGTGAAAGATATTTCCAGGACTGATGGGGGCAGAATGTTGATTCATATTGGCCGAACA
AAAACGCTGGTTTCCACAGCAGGTGTAGAAAAGGCACTGTCGTTAGGAGTAACCAAACTC
GTTGAAAGATGGATATCCGTTTCTGGTGTTGCTGACGACCCTAACAACTATCTGTTTTGT
CGAGTCAGAAAAAATGGAGTGGCAGCTCCATCAGCTACTTCCCAGCTCAGTACCAGAGCT
CTGGAGGGTATCTTTGAGGCCACTCACAGGTTGATCTATGGTGCCAAGGATGATAGTGGT
CAAAGATATCTGGCCTGGTCTGGACACAGCGCCCGGGTTGGTGCTGCCCGTGACATGGCA
AGAGCGGGGGTTTCCATACCAGAAATTATGCAAGCTGGCGGCTGGACGAATGTTAATATT
GTCATGAACTATATACGGAATCTTGATTCAGAGACTGGTGCCATGGTAAGACTACTAGAA
GATGGAGAC
```

Protein Sequence:

```
MSNLLTVHQNLPALPVDATSDEVRKNLMDMFRDRQAFSEHTWKMLLSVCRSWAAWCKLNN
RKWFPAEPEDVRDYLLYLQARGLAVKTIQQHLGQLNMLHRRSGLPRPSDSNAVSLVMRRI
RKENVDAGERAKQALAFERTDFDQVRSLMENSDRCQDIRNLAFLGIAYNTLLRIAEIARI
RVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVERWISVSGVADDPNNYLFC
RVRKNGVAAPSATSQLSTRALEGIFEATHRLIYGAKDDSGQRYLAWSGHSARVGAARDMA
RAGVSIPEIMQAGGWTNVNIVMNYIRNLDSETGAMVRLLEDGD
```
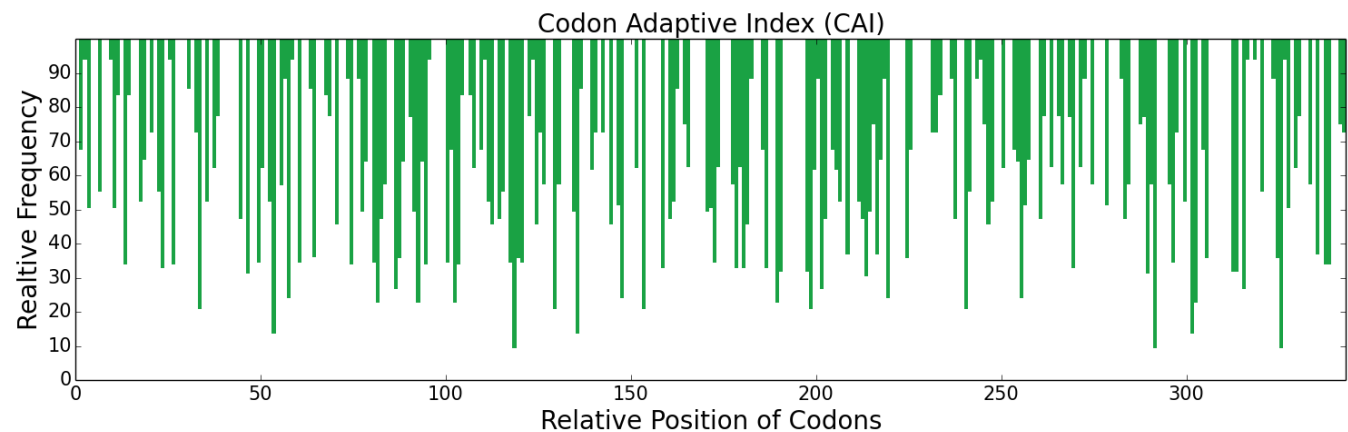
Codon Adaptive Index (CAI)

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	[1]
ENC	Effective Number of Codons	20~61	55.81	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.17	[4]
DCBS	Directional Codon Bias Score	≥1	1.54	[5]
CDC	Codon Deviation Coefficient	0~1	0.09	[6]
MILC	Measure Independent of Length and Composition	–1~1	–0.52	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.1	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.07	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.87	[11]
P	Codon Preference	≥1	1.08	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.25	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.66	[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	∞	0.19	[5]
COUSIN18			–0.07	
CBI	Codon Bias Index	–1~1	0.12	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.31	[7]
RCA	Relative Codon Adaptation	≥0	1.06	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.18	[9]
B	Codon Usage Bias	0~2	0.25	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.47	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.55	[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		0.43	
GC	GC Content		0.47	
GC1	GC Content at the First Position of Synonymous Codons	0~1	0.54	[1]
GC2	GC Content at the Second Position of Synonymous Codons		0.44	
ENcp	Effective Number of Codon Pairs	20~61	61	[2]
CPS	Codon Pair Score	-1~1	0.05	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.74	[5]

Gene Name: Luciferase

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

DNA Sequence Length: 1650

DNA Sequence:

```
ATGGAAGACGCCAAAAATATCAAGAAAGGTCCAGCCCCGTTTTATCCACTAGAGGATGGT
ACCGCCGGTGAACAACTCCATAAGGCCATGAAGCGTTATGCCCTGGTTCCTGGAACTATT
GCTTTCACTGATGCACATATCGAGGTCAACATTACATATGCTGAATATTTTGAGATGTCC
GTCAGATTAGCTGAAGCAATGAAGAGGTACGGTTTGAACACCAATCATCGTATTGTTGTC
TGTTCCGAAAACAGTTTACAATTTTTTATGCCGGTGTTGGGTGCATTATTTATTGGTGTT
GCTGTTGCTCCGGCTAACGATATTTACAATGAAAGAGAGTTGCTGAATAGTATGAACATC
TCACAGCCCACTGTCGTTTTTGTCTCAAAGAAAGGCTTACAAAAGATCTTAAATGTCCAG
AAGAAACTCCCCATTATCCAAAAAATCATCATCATGGATTCGAAAACTGATTACCAAGGG
TTCCAGTCAATGTATACGTTTGTCACTAGCCATCTGCCTCCGGGGTTTAATGAGTATGAC
TTTGTTCCAGAATCTTTTGACAGAGATAAGACAATTGCCCTGATTATGAATTCTTCAGGT
TCTACAGGATTGCCCAAGGGTGTTGCTTTGCCACATAGAACGGCTTGTGTCCGTTTCTCT
CACGCAAGAGATCCCATTTTTGGGAACCAGATAATACCAGACACTGCCATCTTATCAGTG
GTTCCCTTCCACCACGGATTTGGGATGTTTACAACCTTGGGGTACTTGATCTGTGGATTC
CGAGTTGTTCTCATGTATAGGTTCGAAGAGGAACTGTTTTTAAGATCGCTGCAAGACTAC
AAAATTCAAAGTGCGCTACTTGTTCCAACTTTATTTTCTTTTTTTGCCAAAAGTACGTTG
ATTGACAAGTATGACTTGTCCAATCTGCATGAAATCGCAAGTGGAGGTGCTCCGTTAAGT
AAAGAAGTTGGCGAAGCTGTAGCTAAGCGTTTCCATTTGCCTGGTATTCGTCAAGGTTAC
GGTTTAACGGAAACCACTAGTGCCATTTTAATAACACCAGAGGGCGACGACAAACCTGGA
GCCGTCGGTAAAGTTGTTCCCTTCTTTGAAGCCAAGGTAGTAGACTTGGACACTGGTAAA
ACATTGGGAGTGAATCAGAGAGGAGAGCTTTGTGTTCGTGGACCAATGATTATGTCCGGC
TACGTTAATAATCCTGAGGCTACTAATGCTTTAATTGACAAGATGGTTGGTTACATTCT
GGTGACATTGCTTATTGGGATGAAGATGAGCATTTCTTTATTGTCGATCGTCTTAAGTCC
TTGATTAAGTATAAGGGATATCAAGTGGCCCCAGCAGAATTGGAATCAATTTTGCTTCAA
CATCCTAATATATTTGACGCTGGTGTTGCTGGATTGCCAGATGATGATGCTGGCGAACTC
CCAGCAGCAGTGGTAGTCTTAGAACATGGAAAGACTATGACGGAGAAAGAGATTGTTGAC
TACGTCGCTTCGCAAGTTACCACTGCAAAGAAACTGAGAGGTGGGGTTGTCTTTGTCGAT
GAAGTTCCAAAAGGATTAACAGGTAAGCTTGATGCTAGAAAAATAAGAGAAATCTTAATT
AAAGCTAAAAAGGGTGGTAAATCCAAACTG
```

Protein Sequence:

```
MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNITYAEYFEMS
VRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAVAPANDIYNERELLNSMNI
SQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYD
FVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILSV
VPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRSLQDYKIQSALLVPTLFSFFAKSTL
IDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAILITPEGDDKPG
AVGKVVPFFEAKVVDLDTGKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDKDGWLHS
GDIAYWDEDEHFFIVDRLKSLIKYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL
PAAVVVLEHGKTMTEKEIVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILI
KAKKGGKSKL
```
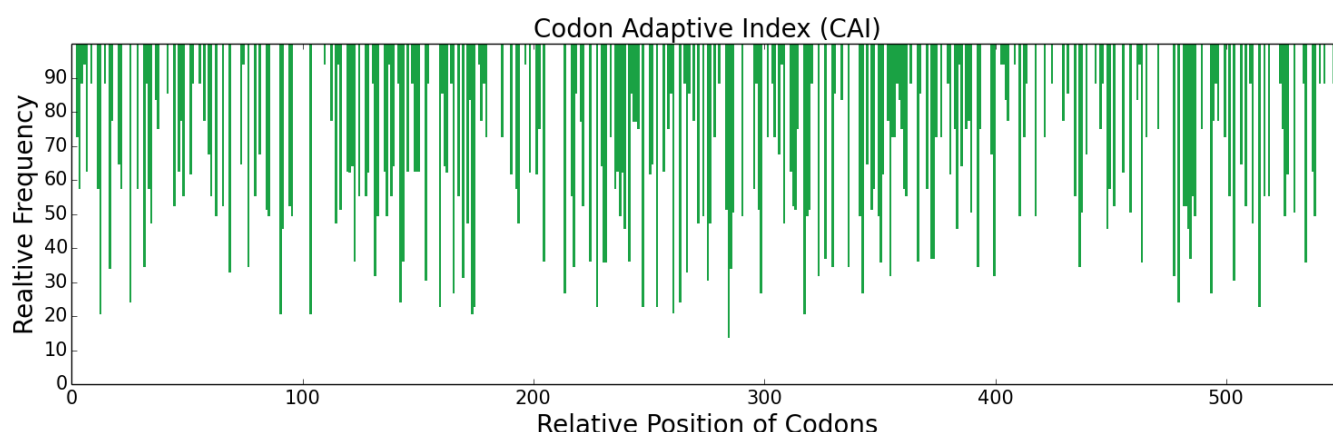
Codon Adaptive Index (CAI)

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	1.23	[1]
ENC	Effective Number of Codons	20~61	48.44	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.18	[4]
DCBS	Directional Codon Bias Score	≥1	1.55	[5]
CDC	Codon Deviation Coefficient	0~1	0.1	[6]
MILC	Measure Independent of Length and Composition	–1~1	–0.44	[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	1.09	[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	0.74	[1]
CFD	Codon Frequency Distribution	0~1	0.04	[2]
FOP	Frequency of Optimal Codons	0~1	0.48	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.67	[5]
COUSIN18			0.62	
CBI	Codon Bias Index	–1~1	0.22	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.2	[7]

RCA	Relative Codon Adaptation	≥0	1.1	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.19	[9]
B	Codon Usage Bias	0~2	0.22	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.51	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.6	[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	0.38	[1]
GC	GC Content		0.42	
GC1	GC Content at the First Position of Synonymous Codons		0.51	
GC2	GC Content at the Second Position of Synonymous Codons		0.36	
ENcp	Effective Number of Codon Pairs	20~61	35.07	[2]
CPS	Codon Pair Score	-1~1	0.01	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

Gene Name: `lacI transcription factor`

Reference Source: `Codon Usage Database - Kazusa`

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

DNA Sequence Length: `1080`

DNA Sequence:

```
ATGAAGCCAGTCACTCTCTATGACGTCGCCGAATATGCTGGAGTGTCCTATCAGACAGTG
TCCAGAGTGGTGAACCAAGCCTCGCATGTTAGCGCGAAGACGCGTGAGAAGGTTGAAGCT
GCCATGGCAGAGCTGAACTACATTCCAAACAGAGTTGCTCAACAACTAGCGGGTAAACAG
TCCCTTCTTATTGGTGTGGCAACATCTTCGCTGGCTTTGCATGCTCCTTCCCAGATTGTT
GCTGCCATTAAGAGCAGGGCTGACCAACTGGGTGCTTCGGTGGTAGTTTCCATGGTGGAG
AGATCCGGTGTTGAGGCCTGTAAGGCTGCGGTCCACAACCTGTTAGCTCAGAGAGTGTCA
GGGCTGATCATCAATTATCCTTTGGACGATCAGGATGCTATTGCAGTTGAGGCTGCCTGT
ACCAATGTTCCGGCTTTGTTCTTGGATGTCAGTGACCAGACTCCTATCAACTCAATCATC
TTTTCTCATGAAGACGGTACGAGGCTGGGAGTGGAACACCTGGTTGCACTAGGACACCAG
CAAATTGCTCTGCTAGCTGGCCCTCTGAGTTCCGTTAGTGCCAGACTGAGATTAGCTGGC
TGGCACAAGTACTTGACCAGAAATCAAATCCAGCCAATCGCCGAACGTGAGGGAGACTGG
TCCGCTATGTCCGGTTTCCAGCAAACTATGCAAATGCTTAACGAAGGAATCGTCCCCACT
GCAATGTTGGTTGCAAACGACCAAATGGCCCTGGGTGCCATGAGAGCTATCACTGAATCT
GGTCTCAGGGTTGGTGCTGATATCTCCGTCGTTGGCTACGATGACACCGAAGATTCTTCG
TGTTACATTCCTCCTCTCACCACTATCAAACAAGACTTTCGTTTATTAGGACAGACCTCG
GTTGACAGATTGTTGCAATTGAGTCAAGGTCAAGCTGTTAAGGGTAACCAACTTCTACCA
GTTTCCCTTGTTAAGAGAAAAACTACTCTTGCTCCCAATACTCAAACTGCTTCCCCAAGA
GCACTTGCTGATTCCCTAATGCAATTAGCCAGACAGGTTTCCCGTTTGGAATCCGGTCAA
```

Protein Sequence:

```
MKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYIPNRVAQQLAGKQ
SLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGVEACKAAVHNLLAQRVS
GLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSHEDGTRLGVEHLVALGHQ
QIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQTMQMLNEGIVPT
AMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPLTTIKQDFRLLGQTS
VDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSLMQLARQVSRLESGQ
```
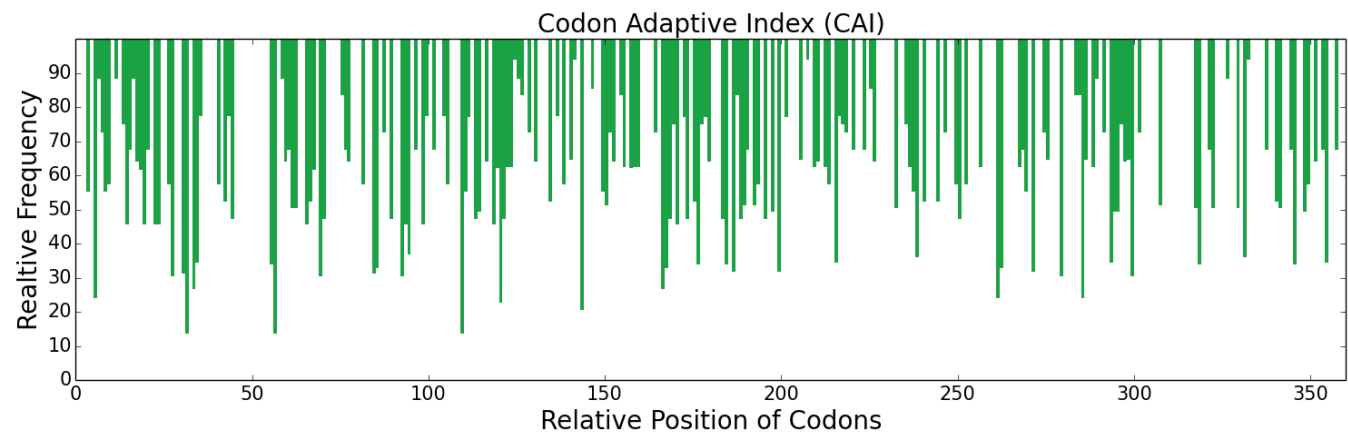

Codon Adaptive Index (CAI)

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	1.38	[1]
ENC	Effective Number of Codons	20~61	46.11	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.31	[4]
DCBS	Directional Codon Bias Score	≥1	1.89	[5]
CDC	Codon Deviation Coefficient	0~1	0.18	[6]
MILC	Measure Independent of Length and Composition	–1~1	–0.4	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.17	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.15	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.81	[11]
P	Codon Preference	≥1	1.2	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.59	[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	[1]
CFD	Codon Frequency Distribution	0~1	0.03	[2]
FOP	Frequency of Optimal Codons	0~1	0.47	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.42	[5]
COUSIN18			1.8	
CBI	Codon Bias Index	–1~1	0.23	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.21	[7]
RCA	Relative Codon Adaptation	≥0	1.12	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.23	[9]
B	Codon Usage Bias	0~2	0.28	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.52	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.58	[2]
P2	P2 Index	0~1	0.68	[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	0.47	[1]
GC	GC Content		0.49	
GC1	GC Content at the First Position of Synonymous Codons		0.59	
GC2	GC Content at the Second Position of Synonymous Codons		0.43	
ENcp	Effective Number of Codon Pairs	20~61	32.37	[2]
CPS	Codon Pair Score	-1~1	-0.03	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.74	[5]

Gene Name: `TetR transcription factor`

Reference Source: `Codon Usage Database - Kazusa`

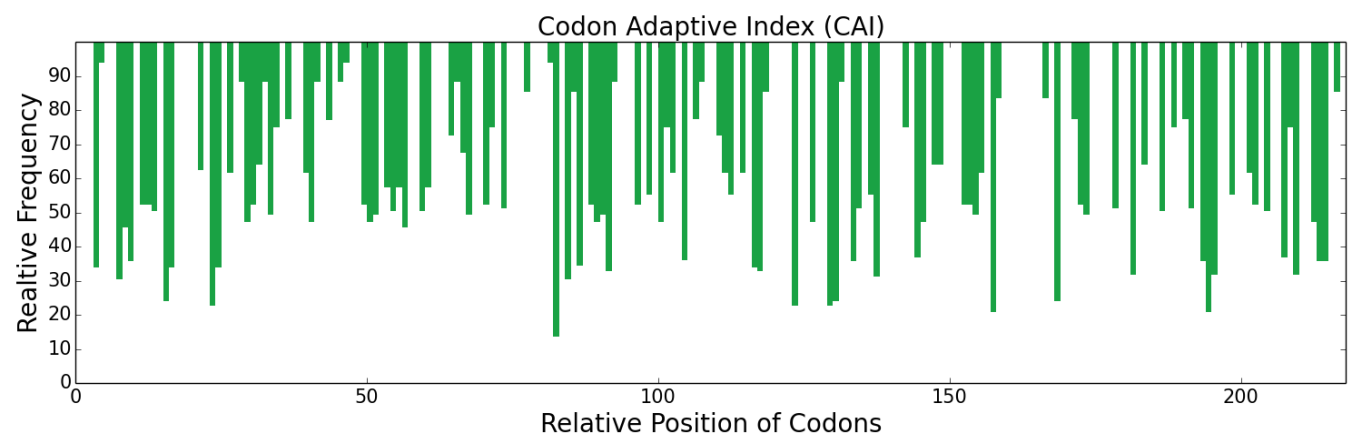**Expression Host Organism:** `Pichia pastoris (budding yeasts)`

DNA Sequence Length: `654`

DNA Sequence:

```
ATGGCTAGACTAAATAGAGAATCGGTGATAGATGCAGCACTTGAACTCCTAAACGAAACT
GGTATCGATGGGCTAACTACAAGAAAACTGGCACAGAAATTAGGAATTGAGCAACCAACA
CTGTATTGGCACGTTAAAAATAAGAGAGCACTGTTAGATGCCCTTGCCGTGGAAATTCTT
GCCAGACATCATGACTATTCCTTACCAGCTGCAGGAGAAAGTTGGCAATCTTTCTTGAGA
AACAATGCGATGTCGTTCCGTAGAGCACTGTTAAGGTATAGAGATGGTGCAAAGGTCCAT
CTGGGAACAAGACCCGATGAGAACAATACGACACAGTCGAAACACAACTAAGGTTCATG
ACTGAAAACGGGTTTTCTCTGAGAGATGGGCTCTATGCTATAAGTGCTGTCAGCCATTTT
ACTTTGGGAGCTGTACTGGAACAGCAGGAACATACTGCAGCATTAACAGATCGACCTGCT
GCTCCAGATGAAAACTTGCCTCCACTCTTGAGAGAGGCATTACAAATTATGGATAGTGAT
GATGGCGAACAGGCTTTTCTTCATGGATTGGAGAGTTTGATACGAGGCTTTGAAGTCCAA
TTGACAGCATTGCTTCAAATTGTAGGAGGCGATAAGCTGATAATACCATTCTGT
```

Protein Sequence:

```
MARLNRESVIDAALELLNETGIDGLTTRKLAQKLGIEQPTLYWHVKNKRALLDALAVEIL
ARHHDYSLPAAGESWQSFLRNNAMSFRRALLRYRDGAKVHLGTRPDEKQYDTVETQLRFM
TENGFSLRDGLYAISAVSHFTLGAVLEQQEHTAALTDRPAAPDENLPPLLREALQIMDSD
DGEQAFLHGLESLIRGFEVQLTALLQIVGGDKLIIPFC
```


Codon Adaptive Index (CAI)

Negative CIS Elements	Negative repeat Elements
0	0

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

Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.42	[1]
ENC	Effective Number of Codons	20~61	47.14	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.3	[4]
DCBS	Directional Codon Bias Score	≥1	1.75	[5]
CDC	Codon Deviation Coefficient	0~1	0.16	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.45	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.24	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.19	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.78	[11]
P	Codon Preference	≥1	1.17	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.53	[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	[1]
CFD	Codon Frequency Distribution	0~1	0.04	[2]
FOP	Frequency of Optimal Codons	0~1	0.45	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.06	[5]
COUSIN18			0.91	
CBI	Codon Bias Index	-1~1	0.2	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.46	[7]
RCA	Relative Codon Adaptation	≥0	1.11	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.31	[9]
B	Codon Usage Bias	0~2	0.42	[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	0.55	[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		0.33	
GC	GC Content		0.43	
		0~1		[1]

GC1	GC Content at the First Position of Synonymous Codons		0.57	
GC2	GC Content at the Second Position of Synonymous Codons		0.39	
ENcp	Effective Number of Codon Pairs	20~61	32.85	[2]
CPS	Codon Pair Score	-1~1	0.07	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.75	[5]

Gene Name: `araC transcription factor`

Reference Source: `Codon Usage Database - Kazusa`

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

DNA Sequence Length: `876`

DNA Sequence:

```
ATGGCCGAAGCTCAAAATGATCCACTGCTTCCAGGCTATTCTTTCAATGCTCATTTGGTT
GCCGGTTTAACCCCTATTGAAGCAAATGGATATCTGGATTTTTTCATAGACAGGCCATTG
GGGATGAAAGGTTATATTTTGAACCTGACCATCAGAGGTCAAGGGGTTGTGAAAAACCAA
GGCAGAGAATTCGTATGCCGTCCTGGTGACATCCTGCTGTTTCCTCCTGGAGAAATCCAT
CATTATGGGCGTCACCCTGAGGCCAGGGAATGGTACCACCAATGGGTGTATTTTAGACCC
AGAGCTTACTGGCATGAATGGCTGAATTGGCCATCAATATTTGCAAATACTGGCTTTTTC
CGACCAGACGAGGCTCACCAACCTCACTTCTCTGACCTATTTGGACAGATCATCAACGCT
GGACAAGGGGAGGGAAGATATTCTGAGTTATTGGCCATCAATTTGCTGGAGCAACTGCTG
CTACGGAGAATGGAAGCAATTAACGAATCATTGCATCCACCGATGGACAATAGAGTTCGG
GAAGCTTGCCAATATATATCCGATCATTTGGCTGACTCCAATTTTGATATAGCCTCAGTG
GCTCAACATGTTTGTCTTAGTCCCAGTAGACTATCTCATTTGTTCAGACAGCAGCTGGGT
ATCTCCGTACTGAGTTGGAGAGAGGATCAACGTATAAGTCAAGCAAAGTTGCTTTTGTCA
ACCACCCGTATGCCTATAGCTACGGTAGGAAGAAACGTTGGATTTGATGACCAATTATAT
TTCAGTCGTGTCTTTAAGAAATGCACTGGAGCCTCTCCTAGCGAGTTTAGAGCTGGCTGT
GAAGAGAAAGTGAATGATGTTGCCGTCAAGCTGTCC
```

Protein Sequence:

```
MAEAQNDPLLPGYSFNAHLVAGLTPIEANGYLDFFIDRPLGMKGYILNLTIRGQGVVKNQ
GREFVCRPGDILLFPPGEIHHYGRHPEAREWYHQWVYFRPRAYWHEWLNWPSIFANTGFF
RPDEAHQPHFSDLFGQIINAGQGEGRYSELLAINLLEQLLLRRMEAINESLHPPMDNRVR
EACQYISDHLADSNFDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRISQAKLLLS
TTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAGCEEKVNDVAVKLS
```
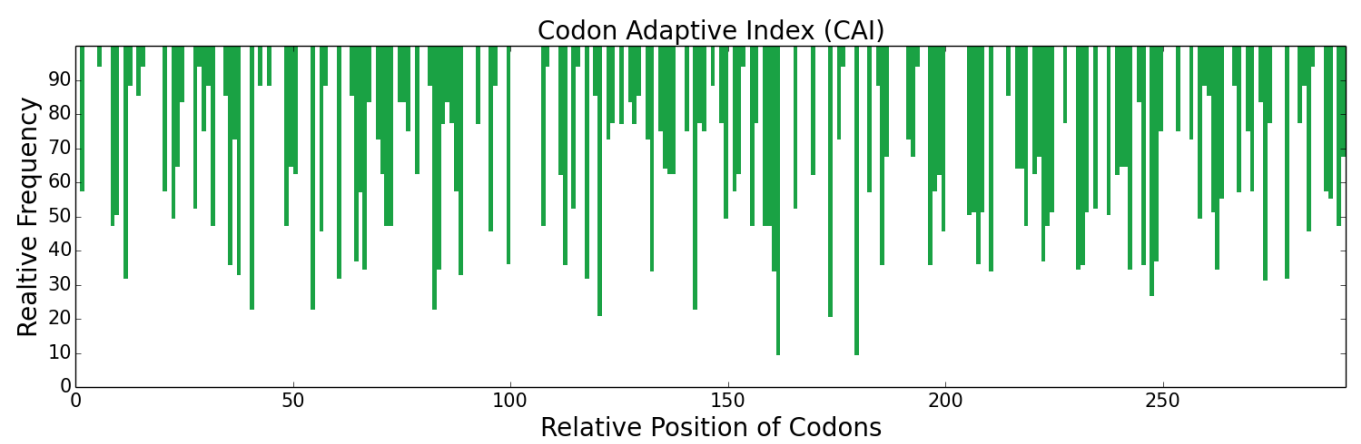
Codon Adaptive Index (CAI)

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	1.29	[1]
ENC	Effective Number of Codons	20~61	51.22	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.28	[4]
DCBS	Directional Codon Bias Score	≥1	1.71	[5]
CDC	Codon Deviation Coefficient	0~1	0.14	[6]
MILC	Measure Independent of Length and Composition	−1~1	−0.46	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.12	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.13	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.83	[11]
P	Codon Preference	≥1	1.18	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.45	[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	[1]
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	∞	0.07	[5]
COUSIN18			−0.37	
CBI	Codon Bias Index	−1~1	0.13	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.33	[7]
RCA	Relative Codon Adaptation	≥0	1.12	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.23	[9]
B	Codon Usage Bias	0~2	0.31	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.49	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.57	[2]
P2	P2 Index	0~1	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	0~1	0.41	[1]
GC	GC Content		0.45	
GC1	GC Content at the First Position of Synonymous Codons		0.54	
GC2	GC Content at the Second Position of Synonymous Codons		0.4	
ENcp	Effective Number of Codon Pairs	20~61	31.72	[2]
CPS	Codon Pair Score	-1~1	0.03	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

Gene Name: Regulatory protein GAL4

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)

DNA Sequence Length: 2643
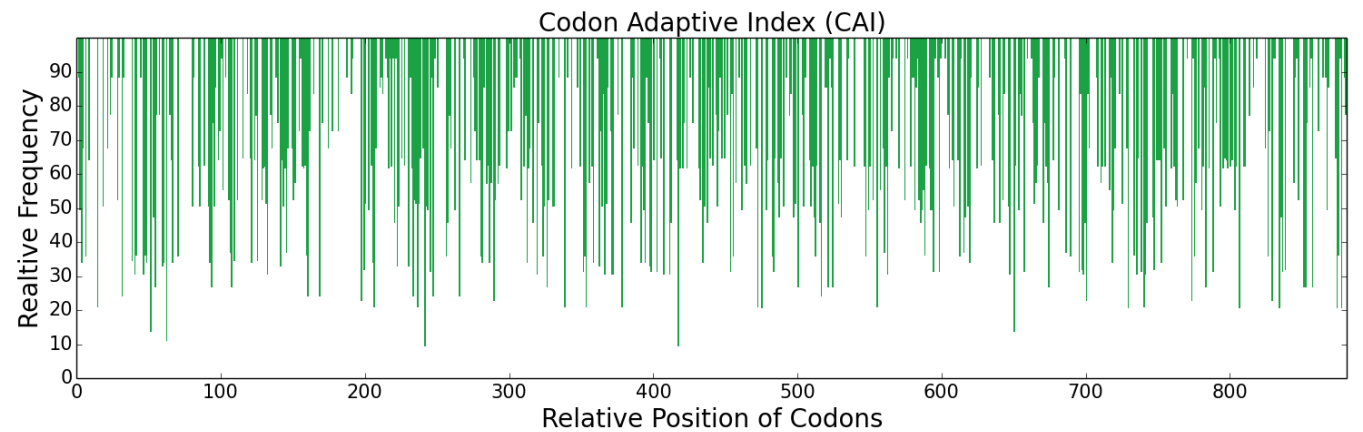
DNA Sequence:

```
ATGAAATTACTATCCTCTATAGAACAGGCTTGTGATATTTGTCGATTGAAGAAGCTTAAG
TGTTCCAAGGAGAAACCAAAGTGTGCAAAATGTCTCAAAAACAACTGGGAATGTCGTTAC
TCGCCCAAGACTAAAAGATCGCCCCTAACTAGAGCGCATCTGACGGAGGTTGAGTCTAGG
CTAGAACGCTTGGAGCAGCTATTTTTGTTGATATTTCCAAGAGAAGATTTGGATATGATT
CTTAAATGGATTCACTTCAAGATATCAAGGCTCTTCTAACGGGACTTTTCGTTCAGGAC
AATGTCAACAAAGATGCAGTAACGGATCGTTTGGCATCTGTTGAAACCGATATGCCTTTG
ACCCTAAGACAGCACCGTATTTCTGCAACATCAAGTTCGGAAGAGTCCTCTAACAAAGGA
CAAAGGCAGCTTACAGTATCCATTGATTCCGCAGCCCATCATGACAATTCAACAATCCCC
CTCGACTTTATGCCTAGAGATGCTCTCCATGGATTTGATTGGTCCGAAGAAGACGATATG
TCTGACGGTTTGCCATTTTTGAAAACTGATCCTAATAACAACGGTTTTTTTGGGGATGGC
AGTTTGTTATGTATCCTACGATCCATTGGTTTCAAACCTGAAAATTATACAAATTCAAAT
GTGAATAGGCTTCCAACCATGATCACTGATAGGTATACACTCGCAAGTCGAAGTACCACT
TCCCGGCTTTTACAAAGCTATCTCAATAACTTCCATCCATACTGTCCAATAGTGCACTCC
CCAACTTTAATGATGCTCTACAACAATCAGATTGAAATTGCCTCTAAAGACCAGTGGCAG
ATACTATTCAACTGCATCCTAGCCATTGGGGCATGGTGCATCGAAGGTGAATCTACAGAC
ATTGACGTTTTCTATTATCAAAATGCAAAGTCACACCTAACATCCAAGGTGTTTGAATCG
GGATCTATCATACTTGTTACGGCATTGCATCTTCTTTCTAGATATACTCAATGGCGACAA
AAAACTAACACATCTTACAACTTCCATTCATTTAGCATACGAATGGCCATTTCACTAGGT
TTGAATAGGGACCTTCCTTCGAGTTTCTCTGACTCGTCGATTTTGGAGCAAAGACGAAGA
ATTTGGTGGTCTGTGTATTCCTGGGAAATCCAACTATCATTACTATATGGAAGGAGCATT
CAGTTGAGCCAAAATACCATTTCGTTTCCATCTTCGGTGGATGATGTTCAGCGGACAACT
ACAGGACCAACAATTTATCATGGAATTATTGAAACAGCAAGACTACTTCAAGTGTTCACC
AAGATCTACGAGCTTGACAAAACCGTTACCGCTGAGAAGAGCCCTATATGTGCCAAGAAA
TGTTTAATGATCTGCAACGAAATCGAAGAAGTTTCCCGACAAGCTCCGAAGTTTTTACAG
ATGGATATAAGCACCACTGCCCTGACCAACCTTTTGAAAGAACACCCTTGGCTGAGTTTT
ACGAGATTTGAGCTTAAATGGAAGCAGCTTTCACTGATAATCTACGTGCTCAGAGATTTC
TTCACGAATTTCACGCAAAAGAAGAGTCAGCTGGAACAAGATCAGAATGATCATCAATCA
TACGAAGTTAAGAGATGTTCAATAATGTTATCAGATGCAGCTCAACGAACTGTCATGTCC
GTATCATCGTATATGGACAATCATAATGTTACACCATACTTTGCATGGAATTGTTCATAT
TATTTATTCAATGCAGTGTTAGTCCCCATCAAGACATTATTAAGCAATTCTAAAAGCAAT
GCTGAAAATAACGAGACAGCTCAGCTTTTGCAACAGATAAATACTGTACTGATGCTTCTA
AAAAAGTTGGCTACATTCAAGATCCAAACTTGTGAAAAGTATATTCAGGTGTTGGAAGAG
GTGTGTGCACCTTTTTTGCTTTCGCAATGTGCGATCCCATTACCTCACATTAGCTACAAT
AACTCTAATGGAAGTGCCATCAAAAACATTGTGGGATCAGCCACGATTGCTCAGTATCCA
ACTTTACCAGAAGAAATGTAAACAACATATCTGTTAAGTACGTTAGCCCTGGCTCGGTG
GGGCCTTCCCCAGTTCCATTGAAATCAGGTGCCTCATTTTCAGATTTGGTCAAACTATTA
TCCAACAGACCTCCAAGTAGAAACTCCCCGGTTACTATTCCCAGATCGACCCCAAGCCAT
CGATCGGTGACTCCATTTCTGGGCCAACAGCAGCAGCTACAATCCCTTGTTCCATTGACA
CCTTCAGCACTTTTTGGTGGTGCAAACTTTAATCAATCTGGGAATATAGCCGATTCCTCT
TTATCATTCACGTTTACAAACTCTAGCAATGGACCAAATCTTATTACAACCCAGACAAAC
TCACAGGCTTTATCACAACCGATTGCTTCATCAAACGTTCACGATAACTTCATGAATAAC
GAAATTACTGCTTCCAAGATAGACGATGGGAATAATTCTAAGCCGCTGAGCCCAGGCTGG
```

```
ACTGATCAAACTGCCTATAATGCATTTGGTATTACGACGGGAATGTTCAATACGACTACT
ATGGACGATGTTTATAACTATTTATTCGATGATGAAGATACCCCGCCCAATCCGAAGAAA
GAG
```

Protein Sequence:

```
MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVESR
LERLEQLFLLIFPREDLDMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDMPL
TLRQHRISATSSSEESSNKGQRQLTVSIDSAAHHDNSTIPLDFMPRDALHGFDWSEEDDM
SDGLPFLKTDPNNNGFFGDGSLLCILRSIGFKPENYTNSNVNRLPTMITDRYTLASRSTT
SRLLQSYLNNFHPYCPIVHSPTLMMLYNNQIEIASKDQWQILFNCILAIGAWCIEGESTD
IDVFYYQNAKSHLTSKVFESGSIILVTALHLLSRYTQWRQKTNTSYNFHSFSIRMAISLG
LNRDLPSSFSDSSILEQRRRIWWSVYSWEIQLSLLYGRSIQLSQNTISFPSSVDDVQRTT
TGPTIYHGIIETARLLQVFTKIYELDKTVTAEKSPICAKKCLMICNEIEEVSRQAPKFLQ
MDISTTALTNLLKEHPWLSFTRFELKWKQLSLIIYVLRDFFTNFTQKKSQLEQDQNDHQS
YEVKRCSIMLSDAAQRTVMSVSSYMDNHNVTPYFAWNCSYYLFNAVLVPIKTLLSNSKSN
AENNETAQLLQQINTVLMLLKKLATFKIQTCEKYIQVLEEVCAPFLLSQCAIPLPHISYN
NSNGSAIKNIVGSATIAQYPTLPEENVNNISVKYVSPGSVGPSPVPLKSGASFSDLVKLL
SNRPPSRNSPVTIPRSTPSHRSVTPFLGQQQQLQSLVPLTPSALFGGANFNQSGNIADSS
LSFTFTNSSNGPNLITTQTNSQALSQPIASSNVHDNFMNNEITASKIDDGNNSKPLSPGW
TDQTAYNAFGITTGMFNTTTMDDVYNYLFDDEDTPPNPKKE
```


Codon Adaptive Index (CAI)

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	1.1	[1]
ENC	Effective Number of Codons	20~61	53.75	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.13	[4]
DCBS	Directional Codon Bias Score	≥1	1.39	[5]
CDC	Codon Deviation Coefficient	0~1	0.07	[6]
MILC	Measure Independent of Length and Composition	−1~1	−0.47	[7]

ICDI	Intrinsic Codon Deviation Index	0~1	0.08	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.06	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.91	[11]
P	Codon Preference	≥1	1.04	[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	0.7	[1]
CFD	Codon Frequency Distribution	0~1	0.05	[2]
FOP	Frequency of Optimal Codons	0~1	0.42	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.41	[5]
COUSIN18			0.57	
CBI	Codon Bias Index	–1~1	0.13	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.25	[7]
RCA	Relative Codon Adaptation	≥0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.14	[9]
B	Codon Usage Bias	0~2	0.23	[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	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.54	[2]
P2	P2 Index	0~1	0.47	[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		0.41	
GC1	GC Content at the First Position of Synonymous Codons		0.43	
GC2	GC Content at the Second Position of Synonymous Codons		0.4	
ENcp	Effective Number of Codon Pairs	20~61	40.37	[2]
CPS	Codon Pair Score	–1~1	0.05	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

GenScript
Make Research Easy

Gene Name: Ribonuclease Barnase

Reference Source: Codon Usage Database - Kazusa

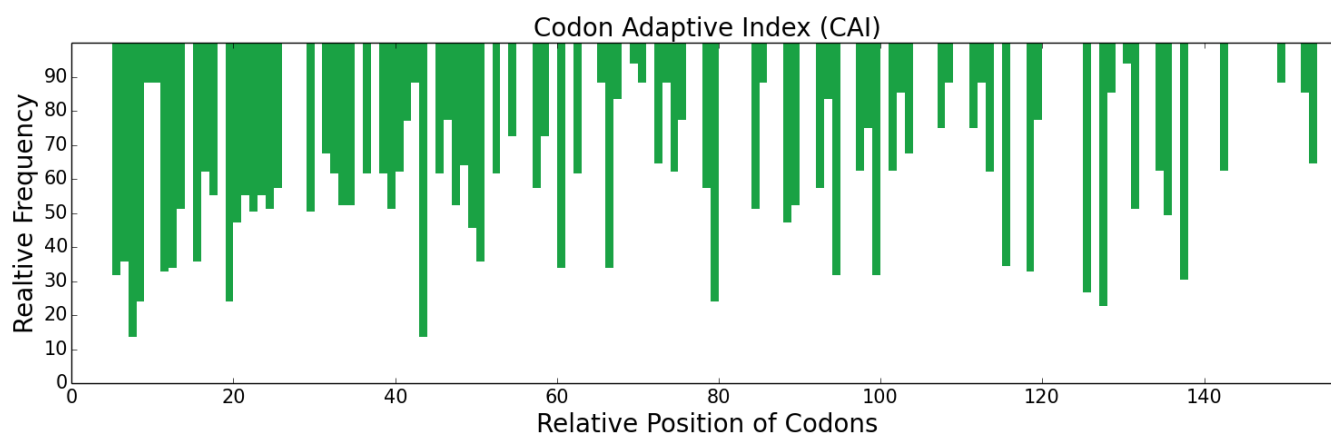Expression Host Organism: Pichia pastoris (budding yeasts)

DNA Sequence Length: 471

DNA Sequence:
```
ATGATGAAGATGGAAGGCATAGCGCTCAAAAAAAGGCTAAGTTGGATATCAGTCTGTCTC
CTGGTCCTTGTCAGTGCCGCTGGTATGCTTTTTTCCACAGCAGCAAAGACAGAAACAAGT
TCACACAAAGCGCATACAGAGGCACAGGTGATAAACACATTTGACGGTGTTGCCGACTAC
CTACAAACATACCATAAACTACCTGATAATTATATTACCAAATCAGAGGCTCAAGCCCTC
GGTTGGGTTGCTAGTAAAGGTAACCTGGCAGATGTTGCCCCTGGCAAGTCTATCGGAGGC
GATATCTTCTCCAACAGAGAAGGAAAATTGCCAGGAAATCAGGTCGTACTTGGAGGGAG
GCTGATATTAACTACACGTCTGGGTTCAGAAATAGTGATAGAATCTTATACTCGTCTGAT
TGGTTGATCTACAAGACTACTGATCATTATCAAACTTTCACCAAGATTAGA
```

Protein Sequence:
```
MMKMEGIALKKRLSWISVCLLVLVSAAGMLFSTAAKTETSSHKAHTEAQVINTFDGVADY
LQTYHKLPDNYITKSEAQALGWVASKGNLADVAPGKSIGGDIFSNREGKLPGKSGRTWRE
ADINYTSGFRNSDRILYSSDWLIYKTTDHYQTFTKIR
```

Codon Adaptive Index (CAI)

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	1.23	[1]
ENC	Effective Number of Codons	20~61	56.83	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.25	[4]

DCBS	Directional Codon Bias Score	≥1	1.57	[5]
CDC	Codon Deviation Coefficient	0~1	0.12	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.6	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.19	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.13	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.82	[11]
P	Codon Preference	≥1	1.13	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.26	[13]

Indices based on codon frequency in a reference set of genes				
Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.7	[1]
CFD	Codon Frequency Distribution	0~1	0.04	[2]
FOP	Frequency of Optimal Codons	0~1	0.42	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	1.06	[5]
COUSIN18			1.74	
CBI	Codon Bias Index	-1~1	0.13	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.42	[7]
RCA	Relative Codon Adaptation	≥0	1.08	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.27	[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	0.47	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.57	[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		0.41	
GC	GC Content		0.43	
GC1	GC Content at the First Position of Synonymous Codons	0~1	0.45	[1]
GC2	GC Content at the Second Position of Synonymous Codons		0.44	

ENcp	Effective Number of Codon Pairs	20~61	23.26	[2]
CPS	Codon Pair Score	−1~1	0.05	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

GenScript
Make Research Easy

Gene Name: `Cas9`

Reference Source: `Codon Usage Database - Kazusa`

Expression Host Organism: `Pichia pastoris (budding yeasts)`
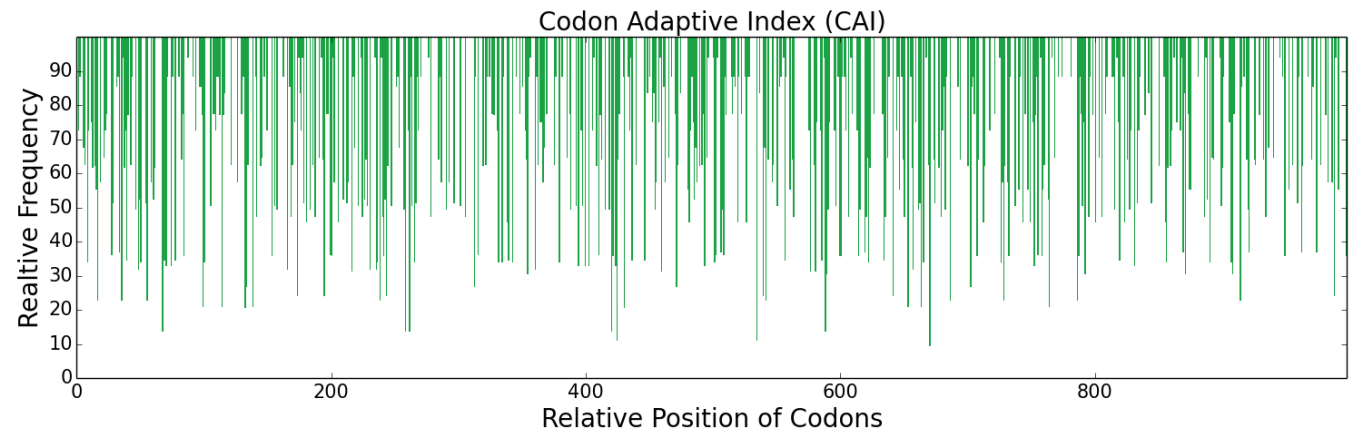
DNA Sequence Length: `2994`

DNA Sequence:

```
ATGGACAAAAAGTACTCCATCGGTCTAGACATTGGAACAAACTCAGTCGGGTGGGCCGTT
ATTACCGACGAGTACAAGGTTCCCAGTAAGAAGTTCAAAGTATTGGGGAATACAGACCGT
CACTCTATCAAAAAGAACTTAATTGGCGCACTATTGTTTGATAGTGGGGAAACTGCCGAA
GCAACAAGATTGAAGAGAACTGCGCGTCGTAGGTATACTAGAAGGAAAAACCGTATTTGT
TATTTGCAGGAGATATTTTCTAATGAAATGGCTAAAGTTGACGATTCTTTCTTCCACCGA
CTAGAAGAATCTTTTCTTGTTGAGGAGGACAAAAAACACGAACGACACCCTATTTTTGGT
AACATCGTTGATGAAGTTGCCTACCATGAGAAATATCCGACGATCTACCATTTGCGAAAG
AAACTGGTTGATTCAACCGATAAAGCTGACTTGAGATTGATATACCTTGCTTTAGCTCAT
ATGATTAAATTTAGAGGCCATTTCCTGATCGAAGGAGATCTCAATCCTGACAATAGTGAT
GTGGATAAGTTATTTATCCAACTGGTTCAAACCTACAATCAGCTCTTTGAGGAGAACCCC
ATAAATGCCTCTGGTGTGGATGCTAAGGCAATTTTGAGTGCCAGATTGAGCAAATCCAGA
AGACTTGAAAATCTGATTGCACAGCTTCCAGGCGAAAGAAGAATGGCCTATTTGGGAAT
CTGATAGCACTCTCATTGGGTCTTACTCCAAACTTCAAATCCAACTTTGATTTAGCGGAA
GACGCGAAACTTCAACTAAGTAAAGACACTTATGATGATGATTTGGATAATTTGCTGGCT
CAAATTGGTGATCAGTATGCCGATTGTTTTTAGCTGCCAAGAACTTGAGTGATGCTATT
TTGCTTTCTGATATTCTGAGAGTTAACACTGAAATTACGAAAGCTCCCTTGTCTGCTTCA
ATGATCAAAGATATGATGAGCACCACCAAGACCTAACTTTGCTAAAAGCTTTGGTGCGT
CAACAACTACCAGAAAAATACAAGGAGATTTTCTTTGACCAGTCGAAAAATGGTTACGCT
GGCTATATCGATGGAGGAGCCTCCCAAGAGGAATTTTACAAGTTTATCAAACCAATTCTA
GAAAAAATGGATGGTACCGAAGAGTTATTGGTTAAGCTTAACAGGGAAGACCTTTTGAGG
AAGCAAAGGACTTTTGACAATGGTTCAATTCCCCACCAGATTCATTTAGGTGAATTACAT
GCGATATTGAGGCGCCAGGAGGATTTTTATCCGTTTTTGAAAGACAATCGTGAAAAGATC
GAAAAGATTTTGACTTTTCGTATTCCTTACTATGTTGGACCTTTAGCCAGAGGAAATAGC
AGATTCGCCTGGATGACCAGAAAATCTGAAGAGACGATCACTCCTTGGAACTTTGAAGAA
GTCGTGGACAAAGGAGCATCCGCCCAATCATTTATCGAGAGGATGACCAATTTTGATAAG
AATCTACCCAATGAAAAGGTATTACCCAAACACTCTTTGTTGTATGAGTATTTCACTGTG
TATAACGAGTTGACTAAAGTGAAATATGTTACTGAAGGTATGCGCAAGCCTGCTTTTCTC
TCCGGGGAACAGAAGAAGGCCATCGTTGATCTTTTGTTCAAAACTAATCGTAAAGTTACT
GTCAAGCAGCTGAAGGAAGATTACTTTAAGAAGATTGAATGTTTTGACAGCGTTGAAATC
AGCGGAGTTGAAGACCGTTTTAATGCGTCGTTAGGAACTTACCATGACCTTTTGAAAATA
ATAAAGATAAAGACTTTCTGGATAATGAGGAAAACGAAGACATATTAGAAGATATTGTA
CTGACCCTAACATTGTTTGAGGATAGAGAAATGATTGAAGAGCGTCTGAAGACCTATGCT
CATCTCTTTGATGACAAAGTCATGAAACAGCTGAAGAGACGAAGATATACAGGCTGGGGA
AGATTAAGTCGAAAGCTAATTAATGGTATCCGGGATAAGCAGAGTGGTAAGACAATTCTG
GACTTCTTAAAATCTGATGGGTTTGCTAACAGAAACTTCATGCAGTTGATTCATGATGAT
TCATTGACGTTCAAAGAGGATATACAGAAGGCTCAAGTGTCAGGTCAAGGTGACTCTTTG
CATGAGCATATTGCTAATCTAGCCGGGTCACCAGCCATAAAAAAGGGTATTCTTCAAACT
GTCAAAGTTGTGGATGAGTTGGTCAAGGTGATGGGAAGGCACAAACCCGAAAATATAGTC
ATCGAAATGGCACGAGAAAATCAAACCACTCAAAAAGGTCAAAAAAACTCTAGAGAAAGA
ATGAAAAGAATTGAAGAAGGGATAAAAGAGCTTGGATCGCAAATTCTGAAAGAACACCCA
GTGGAAAACACCCAACTGCAAAACGAGAAATTGTACTTGTATTATTTACAAAATGGACGT
GATATGTATGTGGATCAGGAATTAGACATTAACAGGTTGAGTGACTATGATGTTGATCAC
```

```
ATTGTTCCTCAAAGTTTTTTGAAGGATGATTCAATTGATAACAAAGTGCTAACAAGATCA
GACAAGAATAGAGGAAAGTCTGACAATGTACCTTCGGAAGAGGTCGTCAAGAAGATGAAG
AACTATTGGAGACAATTGCTGAATGCAAAGCTAATTACCCAGAGAAAGTTTGATAATCTT
ACAAAGCTGAAAGAGGAGGACTATCGGAGTTGGATAAGGCTGGGTTCATTAAAAGACAG
TTAGTAGAAACTAGACAGATCACTAAGCACGTTGCTCAGATTCTGGATTCCAGAATGAAC
ACCAAGTACGATGAAATGATAAATTGATAAGAGAAGTCAAGGTTATCACTTTGAAGAGT
AAATTGGTATCAGATTTTAGAAAGATTTTCAGTTCTACAAGGTAAGAGAAATCAACAAC
TACCACCATGCCCATGATGCCTACCTCAATGCTGTCGTTGGTACTGCTTTGATA
```

Protein Sequence:

```
MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN
LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL
SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALI
```


Codon Adaptive Index (CAI)

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	1.18	[1]
ENC	Effective Number of Codons	20~61	50.33	[2, 3]

RCBS	Relative Codon Bias Strength	≥0	0.17	[4]
DCBS	Directional Codon Bias Score	≥1	1.52	[5]
CDC	Codon Deviation Coefficient	0~1	0.09	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.45	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.12	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.07	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.91	[11]
P	Codon Preference	≥1	1.06	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.36	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.75	[1]
CFD	Codon Frequency Distribution	0~1	0.04	[2]
FOP	Frequency of Optimal Codons	0~1	0.48	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.78	[5]
COUSIN18			0.88	
CBI	Codon Bias Index	-1~1	0.2	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.15	[7]
RCA	Relative Codon Adaptation	≥0	1.12	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.16	[9]
B	Codon Usage Bias	0~2	0.15	[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	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.6	[2]
P2	P2 Index	0~1	0.47	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons		0.41	
GC	GC Content	0~1	0.39	[1]
GC1	GC Content at the First Position of Synonymous Codons		0.47	

GC2	GC Content at the Second Position of Synonymous Codons		0.29	
ENcp	Effective Number of Codon Pairs	20~61	38.51	[2]
CPS	Codon Pair Score	-1~1	0.03	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.78	[5]

Gene Name: Toxin CcdB

Reference Source: Codon Usage Database - Kazusa

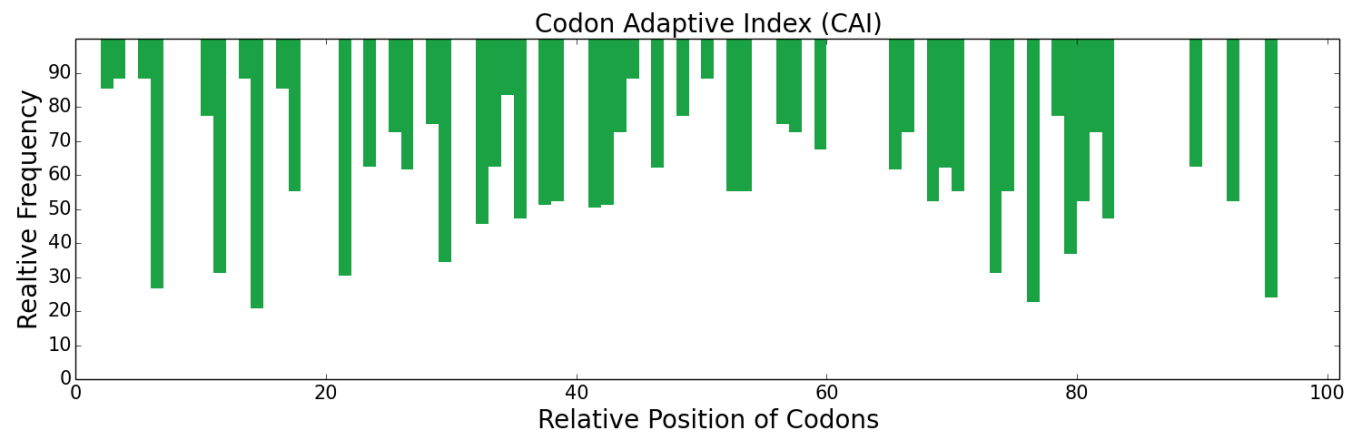Expression Host Organism: Pichia pastoris (budding yeasts)

DNA Sequence Length: 303

DNA Sequence:
ATGCAATTCAAAGTTTATACGTACAAGAGAGAGAGCAGATATCGATTGTTCGTCGATGTT
CAATCGGATATCATTGACACACCAGGACGTAGAATGGTGATCCCTCTGGCTAGTGCAAGA
TTGCTTAGTGACAAAGTTTCAAGAGAGTTGTATCCAGTCGTCCATATTGGAGACGAATCC
TGGAGAATGATGACTACAGACATGGCATCAGTCCCAGTTAGCGTCATTGGGGAAGAGGTA
GCAGACCTGTCTCATAGAGAAAACGATATCAAGAACGCAATTAACCTCATGTTTTGGGGT
ATT

Protein Sequence:
MQFKVYTYKRESRYRLFVDVQSDIIDTPGRRMVIPLASARLLSDKVSRELYPVVHIGDES
WRMMTTDMASVPVSVIGEEVADLSHRENDIKNAINLMFWGI

Codon Adaptive Index (CAI)

	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	1.6	[1]
ENC	Effective Number of Codons	20~61	43.46	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.55	[4]
DCBS	Directional Codon Bias Score	≥1	1.89	[5]
CDC	Codon Deviation Coefficient	0~1	0.22	[6]

MILC	Measure Independent of Length and Composition	-1~1	-0.54	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.33	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.28	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.64	[11]
P	Codon Preference	≥1	1.37	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.59	[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	∞	1.24	[5]
COUSIN18			2.32	
CBI	Codon Bias Index	-1~1	0.21	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.5	[7]
RCA	Relative Codon Adaptation	≥0	1.11	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.4	[9]
B	Codon Usage Bias	0~2	0.5	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.53	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.61	[2]
P2	P2 Index	0~1	0.59	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons		0.45	
GC	GC Content		0.43	
GC1	GC Content at the First Position of Synonymous Codons	0~1	0.48	[1]
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	-0.02	[3, 4]

Codon Volatility	Codon Volatility	0.5~1	0.76	[5]

Gene Name: T7 RNA Polymerase

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Pichia pastoris (budding yeasts)
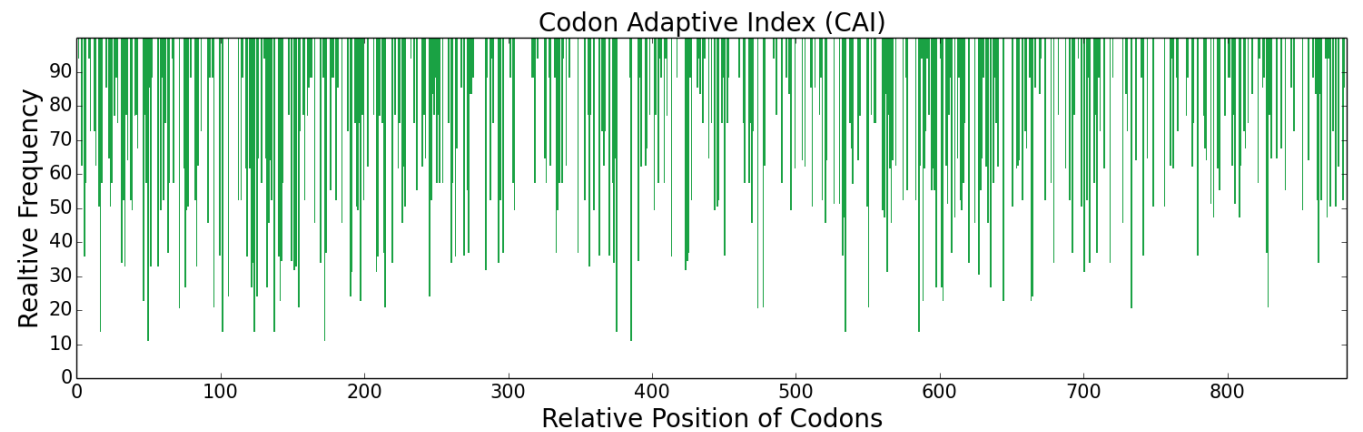
DNA Sequence Length: 2649

DNA Sequence:

```
ATGAATACTATCAACATAGCCAAGAATGACTTTTCTGACATCGAACTTGCGGCCATTCCA
TTCAACACCCTTGCCGATCACTATGGAGAAAGACTAGCAAGGGAGCAGTTGGCATTAGAA
CACGAGTCCTACGAAATGGGGGAGGCCCGCTTCAGGAAAATGTTTGAAAGGCAATTAAAA
GCAGGAGAAGTAGCCGATAATGCCGCTGCTAAGCCGTTGATTACAACGTTACTTCCAAAA
ATGATTGCAAGGATCAACGACTGGTTTGAAGAAGTGAAAGCTAAACGAGGTAAGAGACCC
ACTGCGTTTCAATTTCTCCAAGAAATTAAGCCAGAAGCAGTTGCATATATTACAATAAAG
ACAACGCTAGCGTGTCTCACCTCTGCCGATAATACCACGGTGCAGGCAGTTGCGTCTGCT
ATAGGGCGTGCCATTGAAGATGAGGCTCGTTTTGGCAGGATTCGAGACTTGGAGGCAAAG
CACTTCAAAAAAACGTGGAAGAACAACTAAACAAACGCGTAGGTCATGTCTATAAAAAG
GCATTCATGCAAGTGGTTGAAGCTGACATGCTCAGCAAGGGACTTTTAGGAGGGGAGGCA
TGGTCTTCATGGCATAAGGAGGATAGCATACACGTTGGAGTACGATGTATTGAAATGCTA
ATTGAGTCTACAGGAATGGTGTCACTTCATAGACAAAATGCTGGAGTTGTCGGTCAAGAT
TCAGAGACCATTGAACTCGCACCTGAGTATGCCGAGGCCATTGCCACTAGAGCTGGAGCT
CTAGCCGGTATATCCCCAATGTTCCAACCCTGTGTCGTACCTCCTAAACCATGGACTGGT
ATTACTGGTGGTGGCTATTGGGCAAATGGAAGAAGACCACTAGCATTGGTAAGAACTCAT
TCTAAAAAGGCCTTAATGAGATACGAAGATGTTTACATGCCAGAAGTTTATAAAGCCATT
AATATTGCTCAAAACACCGCCTGGAAAATCAACAAGAAAGTATTAGCCGTTGCCAATGTT
ATCACTAAATGGAAGCATTGTCCAGTAGAAGATATTCCAGCAATTGAGAGGGAGGAATTA
CCAATGAAGCCCGAAGACATCGACATGAACCCCGAAGCCCTAACCGCGTGGAAGAGAGCT
GCTGCTGCTGTTTATCGCAAGGATAAGGCTCGTAAATCAAGAAGAATCTCCTTGGAGTTT
ATGTTAGAGCAAGCCAATAAATTTGCCAATCACAAGGCTATATGGTTTCCATATAACATG
GATTGGAGAGGCCGTGTATATGCAGTTTCTATGTTCAATCCTCAAGGAAATGATATGACC
AAGGGATTGTTAACTCTTGCAAAGGGAAAACCCATTGGAAAAGAAGGTTACTACTGGTTG
AAAATTCATGGAGCCAACTGTGCCGGAGTGGACAAGGTTCCGTTTCCAGAACGAATCAAG
TTTATTGAAGAAATCATGAGAACATTATGGCCTGTGCTAAATCTCCTTTAGAAAACACA
TGGTGGGCTGAACAGGATTCACCATTTTGTTTCCTTGCTTTCTGTTTTGAGTATGCAGGT
GTGCAGCATCATGGTTTGAGTTACAACTGTAGTTTGCCCCTGGCGTTTGATGGATCCTGC
TCTGGTATTCAGCACTTTTCTGCTATGCTTCGAGATGAGGTTGGAGGAAGAGCTGTTAAC
TTACTGCCTAGCGAGACAGTGCAGGATATCTACGGTATTGTTGCAAAGAAAGTCAACGAA
ATTTTGCAAGCAGATGCGATCAATGGGACAGACAATGAGGTTGTCACAGTCACGGATGAA
AATACGGGGGAAATCAGTGAGAAAGTAAAGCTGGGTACCAAAGCATTAGCCGGACAATGG
CTAGCTTACGGTGTGACTAGATCGGTCACCAAAAGATCAGTGATGACGTTGGCCTATGGA
TCAAAGGAATTTGGGTTTAGACAACAAGTTCTTGAAGATACAATCCAGCCAGCAATTGAC
TCCGGTAAAGGGCTCATGTTCACTCAACCTAATCAAGCTGCAGGTTACATGGCCAAGCTA
ATTTGGGAGTCTGTTTCTGTTACAGTTGTTGCAGCTGTAGAGGCTATGAATTGGCTTAAG
AGCGCTGCAAAACTATTGGCTGCCGAGGTAAAAGACAAGAAGACAGGTGAAATTCTAAGA
AAAAGATGTGCTGTTCATTGGGTGACTCCTGACGGTTTTCCGGTTTGGCAGGAATACAAG
AAACCCATTCAAACCAGATTGAACCTTATGTTTTTGGGTCAATTTAGACTTCAACCAACT
ATCAATACAAACAAAGACTCTGAAATTGATGCTCACAAACAAGAATCAGGAATTGCTCCC
AACTTTGTTCACTCCCAGGATGGTAGTCATCTGAGAAAGACAGTCGTTTGGGCTCACGAA
AAATATGGTATCGAGAGTTTTGCTCTGATCCATGACTCCTTTGGAACTATTCCTGCTGAT
GCTGCCAATTTGTTCAAAGCTGTACGAGAGACCATGGTTGATACCTACGAATCCTGTGAT
```

```
GTCTTGGCTGATTTCTACGACCAATTTGCTGATCAATTACATGAATCTCAGTTGGATAAA
ATGCCTGCACTACCTGCAAAGGGTAATCTGAATCTTAGAGACATTCTTGAATCAGATTTT
GCATTCGCT
```

Protein Sequence:

```
MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLK
AGEVADNAAAKPLITTLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIK
TTLACLTSADNTTVQAVASAIGRAIEDEARFGRIRDLEAKHFKKNVEEQLNKRVGHVYKK
AFMQVVEADMLSKGLLGGEAWSSWHKEDSIHVGVRCIEMLIESTGMVSLHRQNAGVVGQD
SETIELAPEYAEAIATRAGALAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTH
SKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREEL
PMKPEDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIWFPYNM
DWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYYWLKIHGANCAGVDKVPFPERIK
FIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFDGSC
SGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAINGTDNEVVTVTDE
NTGEISEKVKLGTKALAGQWLAYGVTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAID
SGKGLMFTQPNQAAGYMAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILR
KRCAVHWVTPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP
NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETMVDTYESCD
VLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFA
```


Codon Adaptive Index (CAI)

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	1.15	[1]
ENC	Effective Number of Codons	20~61	50.3	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.16	[4]
DCBS	Directional Codon Bias Score	≥1	1.53	[5]
CDC	Codon Deviation Coefficient	0~1	0.1	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.44	[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	0.87	[11]
P	Codon Preference	≥1	1.06	[12]
MCB	Maximum-likelihood Codon Bias	≥0	0.36	[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	[1]
CFD	Codon Frequency Distribution	0~1	0.04	[2]
FOP	Frequency of Optimal Codons	0~1	0.48	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	0.86	[5]
COUSIN18			1.04	
CBI	Codon Bias Index	-1~1	0.2	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.13	[7]
RCA	Relative Codon Adaptation	≥0	1.1	[8]
CUFS	Codon Usage Frequency Similarity	0~√2	0.13	[9]
B	Codon Usage Bias	0~2	0.16	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.51	[1]
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	0.38	[1]
GC	GC Content		0.43	
GC1	GC Content at the First Position of Synonymous Codons		0.53	
GC2	GC Content at the Second Position of Synonymous Codons		0.39	
ENcp	Effective Number of Codon Pairs	20~61	35.63	[2]
CPS	Codon Pair Score	-1~1	0.03	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]