

# GenRCA Rare Codon Analysis Report

**Gene Name:** Herceptin light chain

**Reference Source:** Codon Usage Database - Kazusa

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

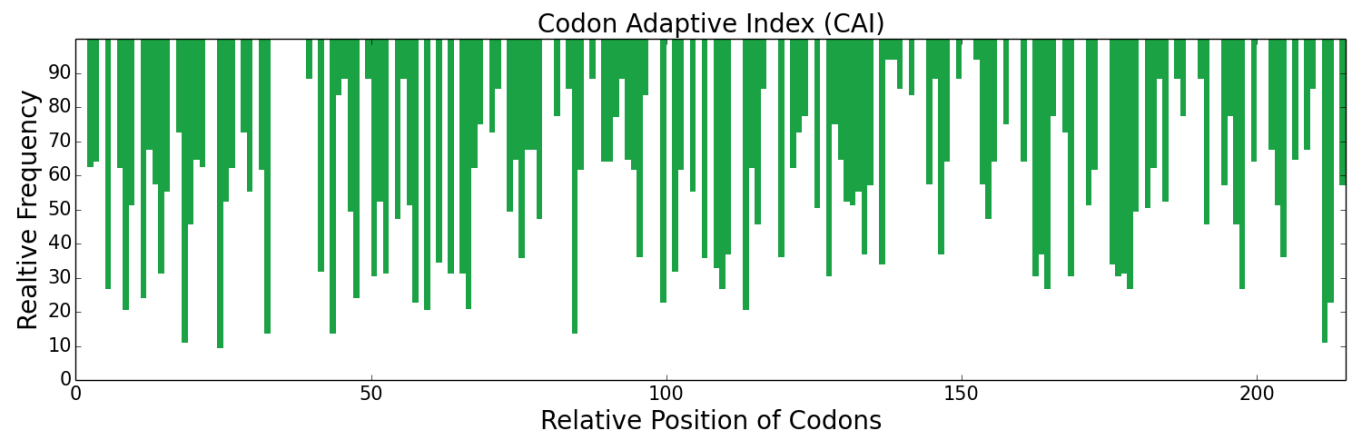
**DNA Sequence Length:** 645

**DNA Sequence:**

```
ATGGATATCCAGATGACGCAATCACCGAGTTCTCTCTCCGCCAGCGTCGGTGACCGCGTG
ACCATCACTTGTTCGGGCATCACAAAGACGTCAACACAGCGGTTGCTTGGTACCAACAAAAA
CCAGGCAAGGCGCCTAAATTACTCATTTATTTCGGCAAGCTTTCTGTATAGTGGGGTTCCG
TCTCGTTTTAGCGGTAGCCGATCAGGAAGTACTTCACTTTAACCATATCCTCCCTGCAA
CCAGAGGATTTTCGCGACATACTATTGTCAGCAGCACTATAACCACACCCCTACTTTTGGG
CAAGGCACAAAGGTTCGAAATAAAGAGGACGGTAGCTGCTCCGTCAGTGTTCATTTTTCCC
CCATCAGACGAGCAACTTAAGTCGGGAACCGCAAGTGTCGTATGCTTGCTAAATAATTC
TACCCTAGAGAAGCCAAAGTACAGTGGAAGTTGATAATGCCCTGCAGTCTGGAACTCT
CAGGAATCGGTAACGGAGCAAGACTCGAAGGATAGTACATACTCTCTATCGAGCACGTTA
ACTCTTTCAAAGCAGATTATGAGAAGCATAAAGTGTACGCTTGCGAGGTGACGCATCAG
GGTTTGTCCAGTCCCCTTACCAAGTCCTTCAACCGCGGGGAATGC
```

**Protein Sequence:**

```
MDIQMTQSPSSLSASVGDRTTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSVGP
SRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFP
PSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKSTYLSSTL
TLISKADYEKKHVKYACEVTHQGLSSPVTKSFNRGEC
```



| 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      | 0.99  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.13  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.41  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.09  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.68 | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0.01  | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0.01  | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.98  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1.01  | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                      | Range         | Value | Reference              |
|----------|----------------------------------|---------------|-------|------------------------|
| CAI      | Codon Adaptation Index           | 0~1           | 0.62  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution     | 0~1           | 0.09  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons      | 0~1           | 0.33  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | 0.03  | <a href="#">[51]</a>   |
| COUSIN18 |                                  |               | 0.07  |                        |
| CBI      | Codon Bias Index                 | -1~1          | 0     | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 0.73  | <a href="#">[71]</a>   |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1.02  | <a href="#">[81]</a>   |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.2   | <a href="#">[91]</a>   |
| B        | Codon Usage Bias                 | 0~2           | 0.3   | <a href="#">[101]</a>  |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference            |
|-------|-------------------------------|-------|-------|----------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.44  | <a href="#">[11]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.52  | <a href="#">[21]</a> |
| P2    | P2 Index                      | 0~1   | 0.52  | <a href="#">[31]</a> |

### 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   | 0.5  | <a href="#">[1]</a>    |
| GC                  | GC Content  |       | 0.49 |                        |
| GC1                 | GC Content at the First Position<br>of Synonymous Codons  |       | 0.48 |                        |
| GC2                 | GC Content at the Second Position<br>of Synonymous Codons |       | 0.47 |                        |
| ENcp                | Effective Number of Codon Pairs                           | 20~61 | 61   | <a href="#">[2]</a>    |
| CPS                 | Codon Pair Score  | -1~1  | 0.01 | <a href="#">[3, 4]</a> |
| Codon<br>Volatility | Codon Volatility  | 0.5~1 | 0.75 | <a href="#">[5]</a>    |

**Gene Name:** Herceptin heavy chain

**Reference Source:** Codon Usage Database - Kazusa

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

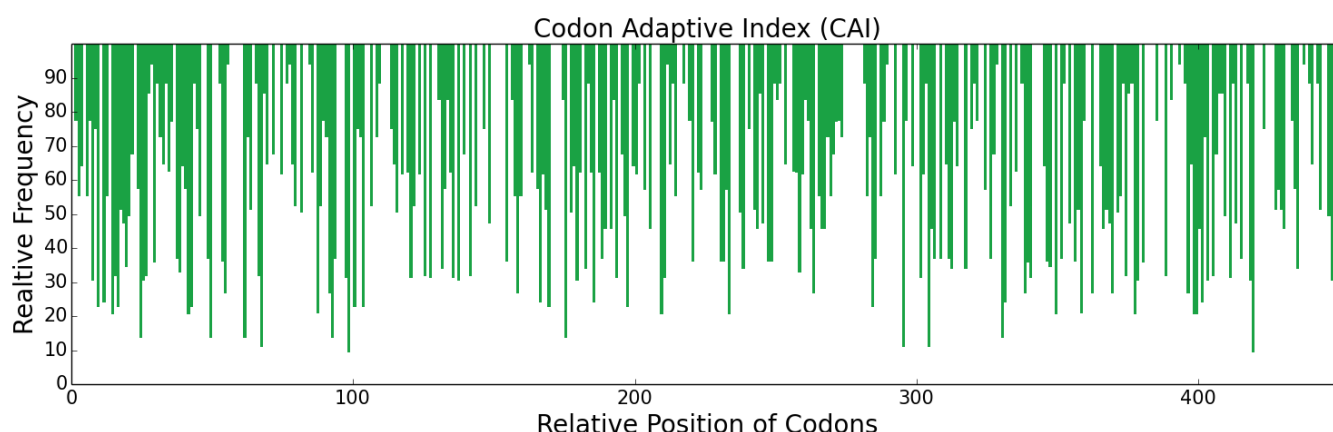
**DNA Sequence Length:** 1353

**DNA Sequence:**

```
ATGGAGGTCCAGTTGGTTCGAGTCGGGAGGGGGTCTCGTCCAACCGGGCGGGAGTCTGCGT
TTATCCTGTGCCGCGTCGGGCTTCAATATAAAAAGACACCTATATCCACTGGGTAAGGCAG
GCCCCGGGGAAGGATTAGAATGGGTAGCGAGAATTTATCCCACGAATGGTTACACTAGA
TACGCGGACAGTGTAAAGGCCGCTTCACCATTTCCGCTGATACATCTAAAAATACCGCA
TACCTTCAAATGAATTCATTGCGAGCAGAGGACACGGCGGTATACTACTGTAGCCGGTGG
GGGGGAGACGGGTTTTACGCAATGGACTATTGGGGTCAAGGAACCCCTGTTACAGTTTCA
AGCGCATCTACAAAGGGCCCAAGCGTTTTTCCTCTAGCCCCCTCAAGCAAGTCGACTTCC
GGTGGCACTGCAGCTTTGGGATGTCTGGTTAAGGATTACTTTCCCGAACCTGTCACGGTC
TCTTGGAATTCAGGTGCCCTCACAAGTGGGGTTCATACTTTTCCTGCGGTTCTTCAGTCG
TCAGGTCTATATTCAGTCTCTTCAGTAGTGACTGTGCCTAGCTCTTCCTTAGGGACTCAG
ACATATATTTGCAACGTGAACCATAAGCCGAGCAATACCAAAGTCGATAAGAAAGTTGAG
CCCAAGTCATGCGATAAGACTCACACATGTCCCCCTGCCCGGCTCCAGAACTTCTAGGT
GGACCAAGTGTGTTCTGTTTCCCCCAAAACCTAAAAGATACCTTGATGATCTCAAGGACA
CCTGAGGTGACGTGTGTCGTGGTGGACGTCTCCACGAGGACCCAGAAGTTAAGTTTAAC
TGGTATGTGCGACGGGGTAGAAGTCCACAATGCTAAGACAAAGCCACGCGAGGAACAGTAC
AACAGCACATATCGCGTGGTATCTGTATTGACCGTACTACACCAGGATTGGCTAAACGGA
AAAGAGTACAAGTGCAAGGTATCCAATAAGGCGCTCCAGCACCAATCGAAAAAACGATA
AGCAAGGCTAAGGGTCAGCCCCGTGAACCGCAAGTATATACTCTGCCACCCAGTCGAGAG
GAAATGACGAAGAACCAGGTGAGTCTGACGTGTCTTGTCAAAGGCTTCTATCCGTCGGAT
ATAGCTGTTGAATGGGAGTCTAACGGCCAACCTGAAAACAATTACAAAACGACCCCGCCG
GTGCTCGACTCGGATGGCTCCTTCTTCTTATACAGCAAACCTGACTGTAGATAAATCGCGG
TGGCAACAAGGAAACGTTTTTAGTTGCAGTGTGATGCATGAGGCCCTACATAATCATTAT
ACCCAAAAAAGTTTGTCTTTATCGCCAGGTAAA
```

**Protein Sequence:**

```
MEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTR
YADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGLVTVS
SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS
SGLYSLSVSVTVPSSSLGTQTYICNVNHKPSNTKVDKKEPKSCDKHTHTCPPCPAPELLG
GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY
NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRE
EMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSR
WQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
```



| 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      | 0.97  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.11  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.36  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.07  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.59 | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.97  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.62  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.09  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.33  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0.01  | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | 0.03  |                        |
| CBI      | Codon Bias Index               | -1~1     | 0.01  | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.65  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 1.01 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.18 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.28 | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range      | Value | Reference           |
|-------|-------------------------------|------------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | $0 \sim 1$ | 0.44  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.52  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.5   | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.51  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.49  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.51  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.46  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 40.9  | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.02  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.75  | <a href="#">[5]</a>                       |

**Gene Name:** Pegfilgrastim

**Reference Source:** Codon Usage Database - Kazusa

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

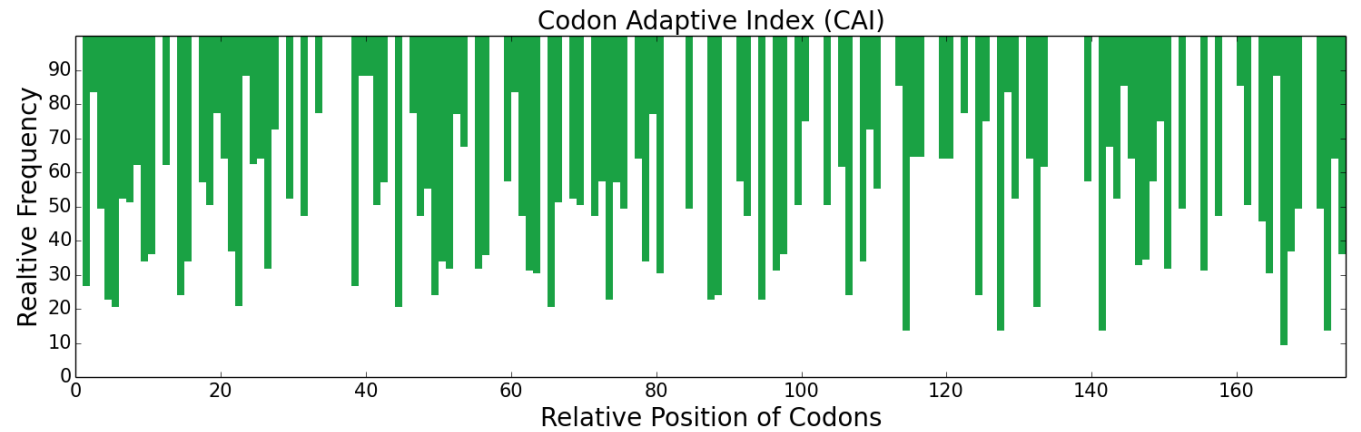
**DNA Sequence Length:** 525

**DNA Sequence:**

```
ATGACGCCTTTAGGGCCGGCAAGTTCACCTACCCCAATCATTTCTCCTAAAGTGCCTTGAG
CAGGTACGAAAAATCCAGGGCGACGGTGCAGCTCTGCAAGAGAAGTTGTGTGCTACGTAT
AAACTTTGCCATCCGGAAGAGCTGGTCCTCCTAGGCCACTCCTTGGGCATACCATGGGCC
CCTCTGAGCTCGTGTCCGAGTCAAGCACTTCAACTGGCCGGGTGCTTATCTCAGCTACAC
TCGGGTTTGTATTTTATACCAAGGGCTCTTGCAAGCCCTGGAAGGGATTAGCCCCGAAGTT
GGACCAACTCTTGATACACTCCAAGTAGACGTCGCTGATTTCGCGACCACCATTTGGCAG
CAGATGGAGGAAGTCCGAATGGCGCCTGCATTGCAGCCGACACAAGGTGCTATGCCAGCC
TTTGCGTCCGCATTCCAGAGGCGTGCCGGAGGCGTTTGTAGTTGCTAGCCATCTGCAATCT
TTCCTTGAAGTGTGCTATCGGGTATTAAGACATTTAGCGCAGCCC
```

**Protein Sequence:**

```
MTPLGPASSLPQSFLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWA
PLSSCPSQALQLAGCLSQLHSGFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQ
QMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRLHLAQP
```



| 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   | 0.99  | <a href="#">[11]</a>   |
| ENC  | Effective Number of Codons      | 20~61 | 61    | <a href="#">[2, 3]</a> |

|      |   |          |       |                         |
|------|---|----------|-------|-------------------------|
| RCBS | Relative Codon Bias Strength                  | $\geq 0$ | 0.2   | <a href="#">[4]</a>     |
| DCBS | Directional Codon Bias Score                  | $\geq 1$ | 1.54  | <a href="#">[5]</a>     |
| CDC  | Codon Deviation Coefficient                   | 0~1      | 0.11  | <a href="#">[6]</a>     |
| MILC | Measure Independent of Length and Composition | -1~1     | -0.71 | <a href="#">[7]</a>     |
| ICDI | Intrinsic Codon Deviation Index               | 0~1      | 0.02  | <a href="#">[8]</a>     |
| SCUO | Synonymous Codon Usage Order                  | 0~1      | 0.01  | <a href="#">[9, 10]</a> |
| Ew   | Weighted Sum of Relative Entropy              | 0~1      | 0.95  | <a href="#">[11]</a>    |
| P    | Codon Preference                              | $\geq 1$ | 1.03  | <a href="#">[12]</a>    |
| MCB  | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.02  | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                      | Range         | Value | Reference              |
|----------|----------------------------------|---------------|-------|------------------------|
| CAI      | Codon Adaptation Index           | 0~1           | 0.57  | <a href="#">[1]</a>    |
| CFD      | Codon Frequency Distribution     | 0~1           | 0.12  | <a href="#">[2]</a>    |
| FOP      | Frequency of Optimal Codons      | 0~1           | 0.3   | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | -0.04 | <a href="#">[5]</a>    |
| COUSIN18 |                                  |               | -0.2  |                        |
| CBI      | Codon Bias Index                 | -1~1          | -0    | <a href="#">[6]</a>    |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 1.1   | <a href="#">[7]</a>    |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1.04  | <a href="#">[8]</a>    |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.22  | <a href="#">[9]</a>    |
| B        | Codon Usage Bias                 | 0~2           | 0.35  | <a href="#">[10]</a>   |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference           |
|-------|-------------------------------|-------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.41  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.48  | <a href="#">[2]</a> |
| P2    | P2 Index                      | 0~1   | 0.47  | <a href="#">[3]</a> |

### 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   | 0.51  | <a href="#">[1]</a> |
| GC    | GC Content  |       | 0.54  |                     |
| GC1   | GC Content at the First Position of Synonymous Codons |       | 0.65  |                     |
|       |   |       |       |                     |



|                     |   |       |      |                        |
|---------------------|---|-------|------|------------------------|
| GC2                 | GC Content at the Second Position<br>of Synonymous Codons |       | 0.45 |                        |
| ENcp                | Effective Number of Codon Pairs                           | 20~61 | 61   | <a href="#">[2]</a>    |
| CPS                 | Codon Pair Score  | -1~1  | 0.01 | <a href="#">[3, 4]</a> |
| Codon<br>Volatility | Codon Volatility  | 0.5~1 | 0.74 | <a href="#">[5]</a>    |

**Gene Name:** Beta-glucuronidase

**Reference Source:** Codon Usage Database - Kazusa

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

**DNA Sequence Length:** 1953

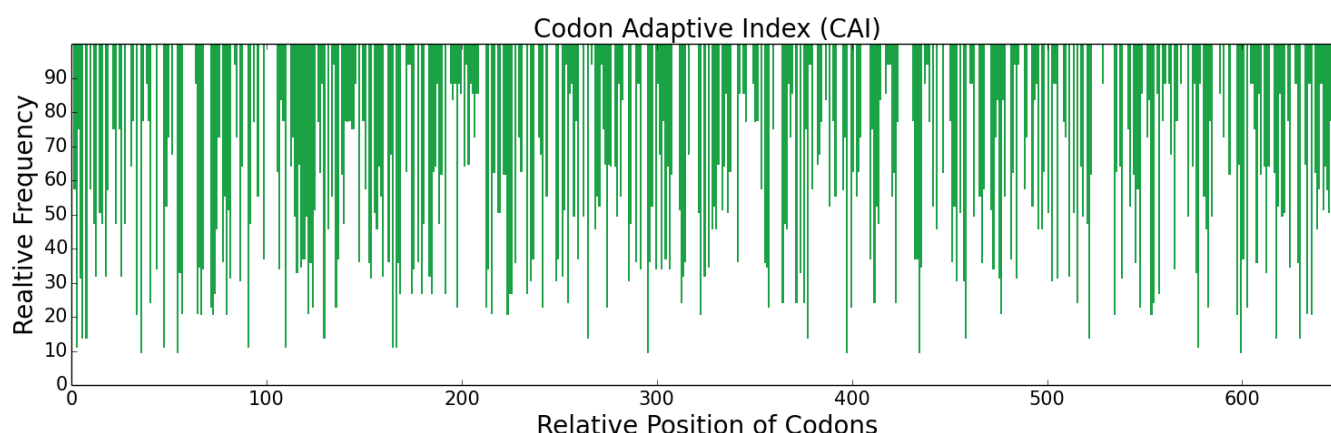
**DNA Sequence:**

```
ATGGCCCCCGGAAGCGCGGTTGCGTGGGCGCTCTGGGCCCACTTCTGTGGGGCTGCGCT
TTGGGACTGCAAGGAGGCATGCTGTACCCACAGGAGTCTCCGTCTCGGGAGTGTAAGAG
CTCGATGGTCTATGGTCTTTTCGCGCAGACTTTTCCGATAACCGGAGGCGAGGTTTTGAA
GAACAATGGTATCGACGTCCGCTATGGGAATCTGGGCCGACGGTGGACATGCCCCGTCCCG
AGTAGCTTTAATGACATTTTCGAGGATTGGCGCCTGAGACACTTTGTCTGGTTGGGTATGG
TACGAAAGAGAAGTTATCCTACCTGAGCGCTGGACTCAGGACTTAAGGACCCGTGTAGTA
TTACGAATAGGGAGTGCTCACTCATATGCGATTGTGTGGGTCAATGGGGTAGATACACTG
GAGCACGAGGGAGGATATTTGCCCTTTGAGGCCGATATAAGCAACCTTGTGCAGGTCGGC
CCATTGCCCTCCCGCTTGCGCATAACGATTGCTATCAATAATACGCTAACTCCCACTACG
CTGCCACCAGGCACGATCCAGTATCTGACAGATACGTCTAAGTATCCTAAAGGGTATTTT
GTTTCAAGATACCTATTTTCGACTTCTTCAACTACGCTGGGCTACAACGATCAGTTCTTCTT
TACACAACACCGACGACGTACATAGATGACATCACTGTTACGACTTCGGTAGAACAAAGAC
TCCGGGTTGGTCAACTACCAAATTTTCGGTAAAGGGCAGTAATCTCTTCAAATTAGAGGTA
AGATTGTTAGATGCGGAAAATAAGGTGGTCGCAAAATGGAACCGGGACCCAGGGTCAGTTA
AAAGTCCCTGGTGTTTCGCTGTGGTGGCCCTATCTAATGCATGAACGGCCCCGCATACCTA
TATTCCCTAGAGGTACAGCTAACAGCTCAAAGTCTCGGCCCCGTTTCCGATTTTTTAC
ACTCTTCCGGTTGGCATTTCGTACTGTGGCAGTGACCAAGAGTCAGTTCCTTATCAACGGT
AAGCCCTTCTACTTCCACGGTGTTAACAACACGAGGATGCCGATATACGTGGGAAGGGA
TTTGATTGGCCACTCCTGGTGAAAGACTTCAACCTCTTGAGGTGGCTCGGAGCGAATGCC
TTTAGAACCTCCCACTACCCCTTACGCAGAAGAGGTGATGCAAAATGTGCGACCGGTACGGG
ATCGTTGTCATCGACGAATGTCCAGGTGTAGGACTCGCACTGCCTCAATTTTTTCAATAAT
GTCTCACTCCACCATCATATGCAAGTTATGGAAGAGGTAGTACGGCGTGATAAAAATCAC
CCAGCAGTTGTGATGTGGTCAGTTGCTAACGAGCCCGCATCGCATCTTGAATCGGCGGGT
TATTATTTAAAGATGGTCATAGCCCATACTAAGAGTCTAGACCCTAGCCGACCTGTCACT
TTTGTATCAAATAGCAATTACGCTGCCGACAAGGGTGACCTTATGTGGATGTGATCTGT
CTTAACCTCGTATTATAGCTGGTACCACGACTACGGCCATCTTGAACCTATTCAGTTACAA
CTGGCGACACAATTTGAAAAGTGGTATAAGAAGTACCAAAAAGCCGATCATTTCAGAGCGAA
TACGGAGCAGAAACAATAGCCGGGTTCCATCAAGACCCTCCGCTCATGTTTCACGGAAGAG
TATCAAAAATCCTTGCTAGAGCAATATCATTTGGGTTTAGATCAGAAAAGGCGCAAATAC
GTCGTGGGGGAGTTAATTTGGAACCTTCGCTGACTTTATGACAGAACAATCTCCGACCCGG
GTATTGGGGAACAAAAAAGGAATATTCACAAGACAGAGGCAGCCAAAGTCAGCGGCATTTC
TTACTTAGAGAGCGTTACTGGAATAAGCGAACGAAACCCGATATCCGCATTTCAGTGGCC
AAAAGTCAGTGCCTTGAGAACTCGTTATTTACC
```

**Protein Sequence:**

```
MARGSAVAWAALGPLLWGCALGLQGMLYPQESPSRECKELDGLWSFRADFSNRRRGFE
EQWYRRPLWESGPTVDMVPVPSSFNDISQDWRLRHFGVWVYEREVILPERWTQDLRTRVV
LRIGSAHSYAIVWVNGVDLTLEHEGGYLPFEADISNLVQVGPLPSRLRITIAINNTLTPTT
LPPGTIQYLTDTSKYPKGYFVQNTYFDFFNAGLQRSVLLYTTPTTYIDDIIVTTSVEQD
SGLVNYQISVKGSNLFKLEVRLLDAENKVANGTGTQGQLKVPVSLWWPYLMHERPAYL
YSLEVQLTAQTSLGPFVSDFYTLFVGIRTVAVTKSQFLINGKPFYFHGVNKHEDADIRGKG
FDWPLLVKDFNLLRWLGANAFRTSHYPYAEVVMQCDRYGIVVIDECPGVGLALPQFFNN
```

VSLHHMQVMEEVVRRDKNHPAVVMWSVANEPASHLESAGYYLKMVIAHTKSLDPSRPVT  
 FVSNSNYAADKGAPYVDVICLSYYSWYHDYGHLELIQLQLATQFENWYKKYQKPIIQSE  
 YGAETIAGFHDPPPLMFTEEYQKSLLEQYHLGLDQKRRKYVVGELIWNFADFMTEQSPTR  
 VLGKKGIFTRQRPKSA AFLLRERYWKIANETRYPHSVAKSQCLENSLFT



| 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      | 0.96  | <a href="#">[1]</a>     |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.08  | <a href="#">[4]</a>     |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.25  | <a href="#">[5]</a>     |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.04  | <a href="#">[6]</a>     |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.56 | <a href="#">[7]</a>     |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[8]</a>     |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.95  | <a href="#">[11]</a>    |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[12]</a>    |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index | Description                  | Range | Value | Reference              |
|-------|------------------------------|-------|-------|------------------------|
| CAI   | Codon Adaptation Index       | 0~1   | 0.61  | <a href="#">[1]</a>    |
| CFD   | Codon Frequency Distribution | 0~1   | 0.11  | <a href="#">[2]</a>    |
| FOP   | Frequency of Optimal Codons  | 0~1   | 0.33  | <a href="#">[3, 4]</a> |

|          |                                  |               |       |                      |
|----------|----------------------------------|---------------|-------|----------------------|
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | -0.03 | <a href="#">[5]</a>  |
| COUSIN18 |                                  |               | -0.04 |                      |
| CBI      | Codon Bias Index                 | -1~1          | -0.01 | <a href="#">[6]</a>  |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 0.82  | <a href="#">[7]</a>  |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1.01  | <a href="#">[8]</a>  |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.19  | <a href="#">[9]</a>  |
| B        | Codon Usage Bias                 | 0~2           | 0.29  | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference           |
|-------|-------------------------------|-------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.44  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.52  | <a href="#">[2]</a> |
| P2    | P2 Index                      | 0~1   | 0.51  | <a href="#">[3]</a> |

### 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   | 0.52  | <a href="#">[1]</a>                       |
| GC               | GC Content   |       | 0.49  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |       | 0.55  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |       | 0.4   |   |
| ENcp             | Effective Number of Codon Pairs                        | 20~61 | 61    | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | -1~1  | 0.02  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | 0.5~1 | 0.75  | <a href="#">[5]</a>                       |

**Gene Name:** Prothrombin

**Reference Source:** Codon Usage Database - Kazusa

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

**DNA Sequence Length:** 1866

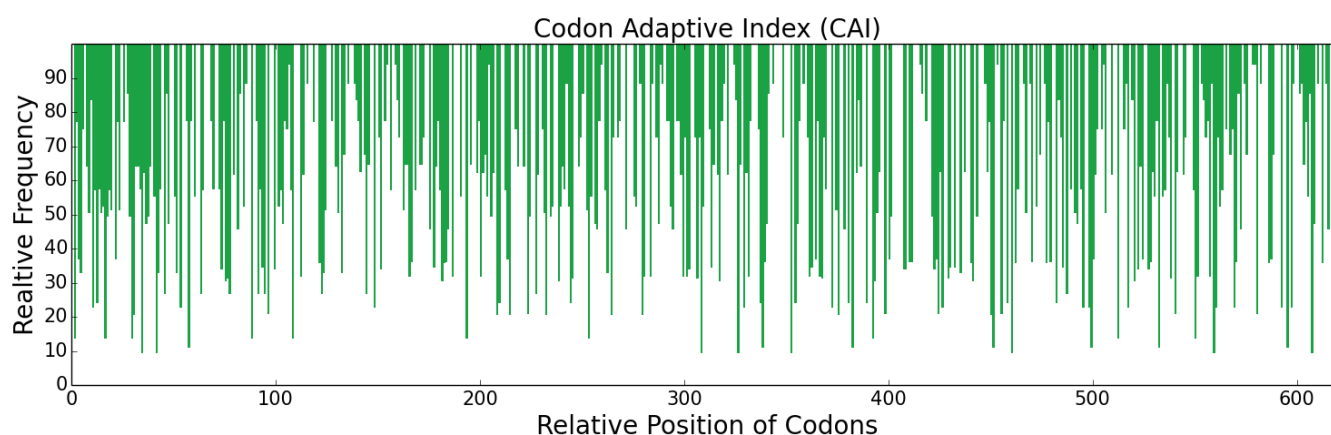
**DNA Sequence:**

```
ATGGCGCAGCTAAGGGGATTGCAGCTTCCTGGGTGCCTCGCCCTTGCAGCGTTATGCAGT
TTGGTACACAGTCAACACGTTTTCTTAGCGCCGCAGCAGGCCCGGTCAGTGTACAGAGA
GTCCGGAGGGCCAACACGTTCTTGAAGAAGTCAGGAAGGGGAACCTGGAGCGCGAGTGT
GTCGAAGAAACGTGCTCTTACGAAGAGGCCTTTGAAGCCCTAGAGTCGAGCACGGCTACA
GATGTGTTCTGGGCAAAATACACTGCGTGTGAGACGGCCCGTACGCCACGAGATAAGCTA
GCTGCATGCCTGGAGGGAAATTGCGCGGAAGGTTTGGGCACAACTATAGAGGTCACGTT
AACATAACGAGGAGTGGTATTGAGTGTGAGCTTTGGAGGTCCAGATATCCACATAAACCT
GAGATCAACTCCACGACCCATCCAGGGGCTGACCTACAAGAGAATTTTTCAGAAAATCCT
GACTCTAGTACCACCGGCCCTGGTGCTACACCACCGACCCAACCTGTGAGACGTCAGGAG
TGCTCGATACCCGTGTGTGGCCAAGATCAAGTCACTGTTGCGATGACCCCAAGATCAGAG
GGCTCATCCGTCAATTTATCACCACCGCTCGAACAATGCGTACCGGATAGAGGACAGCAA
TACCAGGGTCGATTAGCTGTTACGACACATGGACTTCCGTGTTTAGCATGGGGCTTCGGCA
CAGGCCAAAGCACTCAGCAAGCATCAGGACTTCAACAGTGCGGTCCAACCTGGTGGAGAAT
TTTTGCAGGAACCCGGACGGTGACGAGGAAGGTGTGTGGTGTTACGTGCGAGGTAAACCG
GGCGATTTTGGCTATTGTGACCTGAATTATTGTGAGGAGGCAGTGGAAGAGGAGACAGGC
GACGGCCTAGATGAAGACAGCGACCGGGCAATTGAAGGACGTACCGCTACATCGGAGTAT
CAAACATTTTTTAATCCTCGGACCTTTGGGTCAGGCGAGGCTGATTGTGGACTCCGCCCC
CTGTTTCGAAAAAAGTCTTTGGAAGACAAGACTGAACGGGAACTCCTGGAGTCTTATATT
GACGGCCGTATTGTAGAGGGCAGCGACGCCGAAATTGGGATGAGTCCGTGGCAAGTGATG
CTCTTTCGAAAGTCACCTCAGGAATTGCTCTGTGGTTCGCTCGCTTATCTCTGATCGATGG
GTATTAAGTCTGCTCATTGTCTACTATACCCCCCTGGGATAAGAAATTTCACTGAGAAC
GATTTACTAGTACGAATCGGGAAGCATAGCCGTACTCGTTACGAAAGGAACATCGAAAAG
ATATCGATGTTAGAAAAGATTTATATCCACCCGCGCTACAATTGGCGAGAGAACCTCGAT
CGGGATATAGCCTTGATGAAACTTAAGAAACCCGTTGCATTTTCCGATTATATACACCCC
GTTTGTCTCCCTGACCGTGAAACGGCTGCCTCTCTTCTGCAAGCCGGGTACAAGGGGCGC
GTAACAGGATGGGGAAATCTTAAGGAAACATGGACTGCGAACGTTGGAAAAGGGCAACCT
TCGGTTCTACAGGTAGTTAACCTACCCATCGTCGAGCGCCCAGTCTGCAAAGACAGCACT
CGAATCAGAATTACAGACAACATGTTTTGCGCGGGCTACAAACCTGACGAGGGCAAACGG
GGGGACGCATGCGAAGGAGATTCCGGAGGGGCCCTTCGTGATGAAATCCCCATTTAATAAT
CGATGGTATCAAATGGGTATAGTATCCTGGGGTGAAGGGTGTGATCGCGATGGGAAATAC
GGTTTCTATACCCACGTCTTCCGGCTGAAGAAATGGATACAAAAAGTGATTGATCAGTTC
GGTGAG
```

**Protein Sequence:**

```
MAHVRGLQLPGCLALAAALCSLVHSQHVFAPQQARSLQLQVRRANTFLEEVKGNLEREC
VEETCSYEEAFEALSSSTATDVFWAKYTACETARTPRDKLAACLEGNCAEGLGTNYRGHV
NITRSGIECQLWRSRYPHKPEINSTTHPGADLQENFCRNPDSSTGPPWCYTDDPTVRRQE
CSIPVCGQDQVTVAMTPRSEGSSVNLSPPLEQCVPDRGQQYQGR LAVTTHGLPCLAWASA
QAKALSKHQDFNSAVQLVENFCRNPDGDEEGVWCYVAGKPGDFGYCDLNYCEEAVEEETG
DGLDESDRAIEGRTATSEYQTFNPRFTGSGEADCLRPLFEKKSLEDKTERELLESYI
DGRIVEGSDAEIGMSPWQVMLFRKSPQELLCGASLISDRWVLTAAHCLLYPPWDKNFTEN
DLLVRIGKHSRTRYERNIEKISMLEKIYIHPRYNWRENLDRIALMKLKKPVAFSDYIHP
```

VCLPDRETAASLLQAGYKGRVTGWGNLKETWTANVGKGQPSVLQVNLPIVERPVCKDST  
 RIRITDNMFCAGYKPDEGKRGDACEGDSGGPFVMMKSPFNRRWYQMGIVSWGEGCDRDGKY  
 GFYTHVFRLLKKWIQKVIDQFGE



| 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      | 0.97  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.08  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.22  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.04  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.56 | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.96  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0     | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index | Description                  | Range | Value | Reference              |
|-------|------------------------------|-------|-------|------------------------|
| CAI   | Codon Adaptation Index       | 0~1   | 0.6   | <a href="#">[11]</a>   |
| CFD   | Codon Frequency Distribution | 0~1   | 0.11  | <a href="#">[21]</a>   |
| FOP   | Frequency of Optimal Codons  | 0~1   | 0.33  | <a href="#">[3, 4]</a> |
|       |                              |       | 0.02  |                        |

|          |                                  |               |       |                      |
|----------|----------------------------------|---------------|-------|----------------------|
| COUSIN18 | Codon Usage Similarity Index     | $\infty$      | 0.06  | <a href="#">[5]</a>  |
| COUBI59  | Codon Bias Index                 | -1~1          | -0.01 | <a href="#">[6]</a>  |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 0.99  | <a href="#">[7]</a>  |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1     | <a href="#">[8]</a>  |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.19  | <a href="#">[9]</a>  |
| B        | Codon Usage Bias                 | 0~2           | 0.3   | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference           |
|-------|-------------------------------|-------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.45  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.52  | <a href="#">[2]</a> |
| P2    | P2 Index                      | 0~1   | 0.49  | <a href="#">[3]</a> |

### 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   | 0.52  | <a href="#">[1]</a>    |
| GC               | GC Content   |       | 0.51  |                        |
| GC1              | GC Content at the First Position of Synonymous Codons  |       | 0.56  |                        |
| GC2              | GC Content at the Second Position of Synonymous Codons |       | 0.45  |                        |
| ENcp             | Effective Number of Codon Pairs                        | 20~61 | 61    | <a href="#">[2]</a>    |
| CPS              | Codon Pair Score                                       | -1~1  | 0.02  | <a href="#">[3, 4]</a> |
| Codon Volatility | Codon Volatility                                       | 0.5~1 | 0.76  | <a href="#">[5]</a>    |

**Gene Name:** Xylanase

**Reference Source:** Codon Usage Database - Kazusa

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

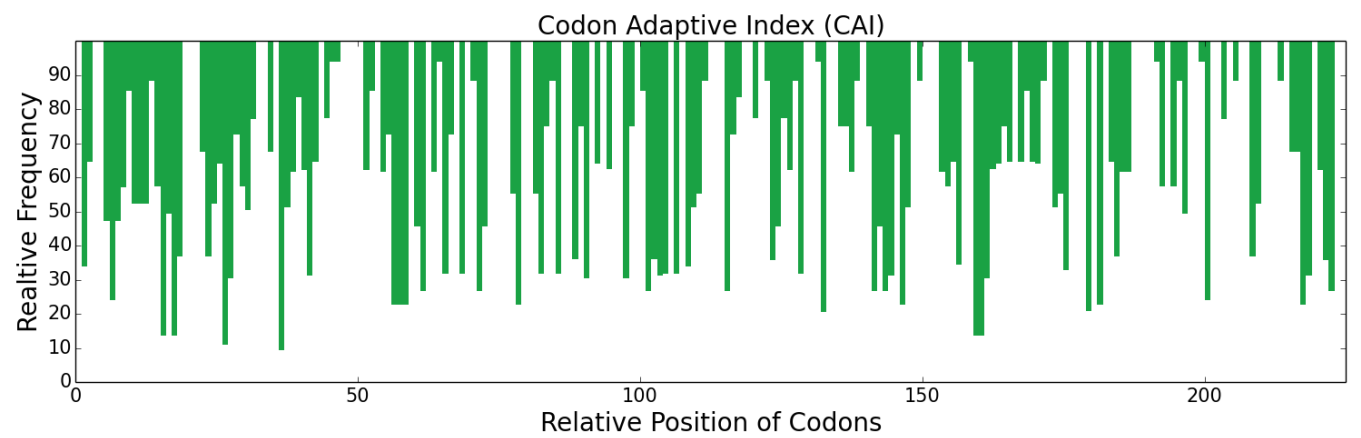
**DNA Sequence Length:** 675

**DNA Sequence:**

ATGCTAACCAAGAACCTGCTCCTGTGCTTCGCAGCAGCAAAAGCCGCGTTAGCGGTACCA  
CATGATTCCGTAGCACAGCGCTCGGACGCCCTTCACATGTTGTCCGAACGGAGTACACCT  
TCAAGCACCGGTGAGAATAATGGTTTTTACTACTCATTCTGGACAGACGGGGGGGGGAT  
GTGACGTACACAAATGGCGACGCTGGCGCTTATACGGTGGAATGGTCTAACGTCGGGAAC  
TTTGTGCGCGGAAAAGGCTGGAACCCCGGATCGGCTCAGGATATCACTTACTCGGGAAC  
TTCACGCCCAGCGGCAACGGCTACCTAAGTGTCTATGGTTGGACTACGGACCCTTTGATT  
GAGTACTATATAGTGGAGTCATATGGCGATTACAATCCGGGTTCTGGAGGAACATATAAG  
GGAACGGTGACGAGCGACGGGAGTGTATTATGATATTTACACAGCCACCCGTACTAATGCG  
GCGTCGATCCAGGGAACCGCTACCTTCACCCAGTATTGGAGTGTCTAGGCAAAACAAGCGA  
GTTGGGGGTACCGTAACAACATCTAACCATTTTAATGCCTGGGCCAAATTAGGTATGAAT  
CTCGGTACTCACAACATCAAATTGTAGCAACTGAAGGTTATCAATCCTCCGGGAGCTCT  
TCAATAACGGTTCAA

**Protein Sequence:**

MLTKNLLLCFAAAKAALAVPHDSVAQSRDALHMLSERSTPSSTGENNGFYYSFWDGGGD  
VTTYNGDAGAYTVEWSNVGNFVGKGWNPQSAQDITYSGTFTPSGNGYLSVYGWTTDPLI  
EYYIYESYGDYNPGSGGTYKGTVTSDSVYDIYTATRTNAASIQGTATFTQYWSVRQNK  
VGGTVTTSNHFNAWAKLGMNLGTHNYQIVATEGYQSSGSSSITVQ



| 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      | 0.98  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.17  | <a href="#">[4]</a>     |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.5   | <a href="#">[5]</a>     |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.11  | <a href="#">[6]</a>     |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.67 | <a href="#">[7]</a>     |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0.06  | <a href="#">[8]</a>     |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0.01  | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.95  | <a href="#">[11]</a>    |
| P     | Codon Preference                              | $\geq 1$ | 1.02  | <a href="#">[12]</a>    |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                      | Range         | Value | Reference              |
|----------|----------------------------------|---------------|-------|------------------------|
| CAI      | Codon Adaptation Index           | 0~1           | 0.61  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution     | 0~1           | 0.11  | <a href="#">[2]</a>    |
| FOP      | Frequency of Optimal Codons      | 0~1           | 0.31  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | -0.02 | <a href="#">[5]</a>    |
| COUSIN18 |                                  |               | -0.26 |                        |
| CBI      | Codon Bias Index                 | -1~1          | -0.02 | <a href="#">[6]</a>    |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 1.17  | <a href="#">[7]</a>    |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 0.99  | <a href="#">[8]</a>    |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.28  | <a href="#">[9]</a>    |
| B        | Codon Usage Bias                 | 0~2           | 0.32  | <a href="#">[10]</a>   |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference            |
|-------|-------------------------------|-------|-------|----------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.44  | <a href="#">[11]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.53  | <a href="#">[2]</a>  |
| P2    | P2 Index                      | 0~1   | 0.5   | <a href="#">[3]</a>  |

### Indices based on complex patterns of codon usage

| Index | Description | Range | Value | Reference |
|-------|-------------|-------|-------|-----------|
|       |             |       | 0.52  |           |

|                     |  |       |       |   |
|---------------------|--|-------|-------|---|
| GC                  | GC Content   | 0~1   | 0.51  | <a href="#">[1]</a>                       |
| GC1<br>GC3          | GC Content at the First Position<br>GC Content at the Third Position<br>of Synonymous Codons |       | 0.48  |   |
| GC2                 | GC Content at the Second Position<br>of Synonymous Codons                                    |       | 0.52  |   |
| ENcp                | Effective Number of Codon Pairs  | 20~61 | 23.55 | <a href="#">[2]</a>                       |
| CPS                 | Codon Pair Score   | -1~1  | 0.01  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon<br>Volatility | Codon Volatility   | 0.5~1 | 0.75  | <a href="#">[5]</a>                       |

**Gene Name:** Chymosin

**Reference Source:** Codon Usage Database - Kazusa

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

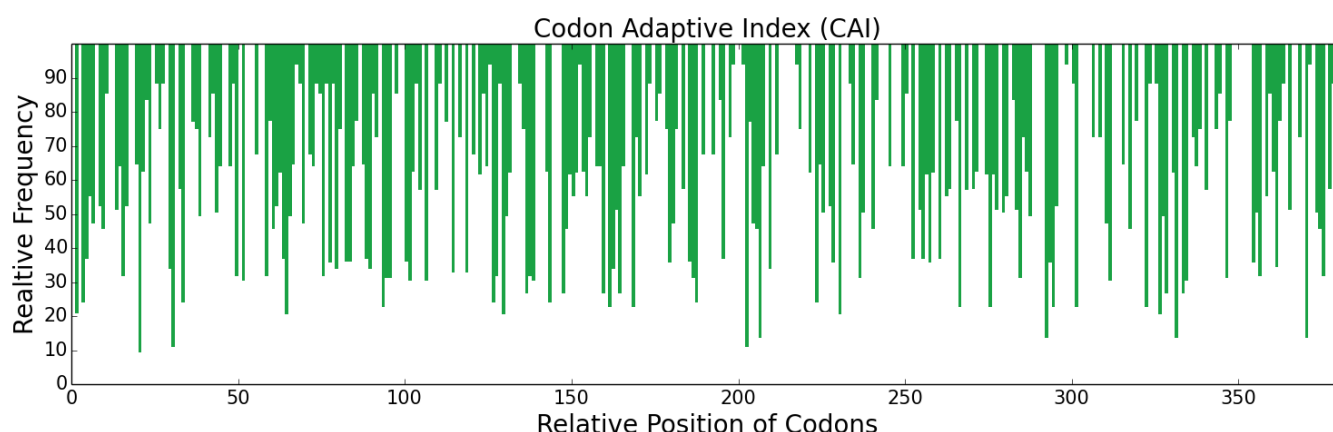
**DNA Sequence Length:** 1143

**DNA Sequence:**

```
ATGCGATGTCTCGTAGTCTGTTGGCAGTGTTCGCTTTGAGTCAGGGCGCAGAAATTACC
CGGATCCCTCTGTACAAAGGAAAATCTCTACGCAAGGCCCTCAAGGAACACGGATTATTG
GAAGACTTCCTTCAGAAGCAACAGTATGGCATTTCGTCTAAGTACTCCGGTTTTGGCGAG
GTGGCATCAGTACCGTTAACCAATTATCTGGATTCCCAGTATTTTCGGCAAAATATATCTA
GGAATCCCCCCCAGGAGTTTACCGTACTATTCGACACTGGGAGCAGCGATTTCTGGGTT
CCCTCGATCTATTGCAAGTCGAACGCTTGCAAAAACCAAAAGGTTTGACCCAAGGAAG
TCCTCTACATTCCAGAATCTCGGCAAACCGTTATCAATTCATTATGGAACGGGCTCGATG
CAAGGTATCCTCGGTTACGATACGGTGACAGTCTCAAATATCGTCGACATTCAGCAGACG
GTTGGGCTAAGTACGCAGGAACCAGGGGACGTCTTTACATATGCTGAGTTCGATGGAATA
CTGGGAATGGCCTACCCAGCCTCGCTTCCGAATACTCCATTCCTGTATTTGACAATATG
ATGAATCGCCACCTGGTGGCGCAGGATCTATTTCCGTTTACATGGATAGAAATGGACAA
GAATCAATGCTCACCTTGGTGCAATAGATCCGTCTTACTATACCGGTAGCCTTCATTGG
GTGCCTGTTACTGTTTCAGCAATACTGGCAGTTCACTGTAGATAGTGTAAACAATATCAGGT
GTAGTTGTGCGCTGTGAGGGGGGTTGCCAAGCCATCTTGGATACAGGGACAAGTAAGCTT
GTCGGTCCTAGTAGCGACATCTTAAACATTCAACAAGCGATAGGGGCAACTCAAATCAA
TATGGGGAATTTGATATTGACTGTGACAACCTGTTCGTACATGCCAACC GTTGTGTTTGAG
ATTAACGGGAAAATGTATCCGTTAACGCCATCAGCGTACACGTCGCAAGACCAGGGATT
TGCACCTCTGGATTCCAAAGCGAGAACCATTCTCAAAGTGGATACTTGGCGATGTCTTC
ATCCGTGAGTATTACAGTGTTTTTGACAGAGCGAATAACCTTGTGGGCTTGGCCAAAGCT
ATA
```

**Protein Sequence:**

```
MRCLVLLAVFALSQGAETRIPLYK GKSLRKALKEHGLLEDFLQKQQYGISSKYS GFGE
VASVPLTNYLDSQYFGKIYLGTPPQEFTVLFD TGSSDFWVPSIYCKSNACKNHQRF DPRK
SSTFQNLGKPLSIHYGTGSMQ GILGYDTVTVSNIVDIQQTVGLSTQEPGDVF TYAEFDGI
LGMAYPSLASEYSIPVFDNMMNRHLVAQDLFSVYMDRNGQESMLTLGAIDPSYYT GSLHW
VPVTVQQYWQFTVDSVTISGVVVACEGGCQAILDTGTSKLVGPSSDILNIQQAIGATQ NQ
YGEFDIDCDNLSYMP TVVFEINGKMYPLTPSAYTSQDQGFCTSGFQSENHSQKWILGDVF
IREYYSVFD RANNLVGLAKAI
```



| 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      | 0.97  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.11  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.36  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.06  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.6  | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.96  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.63  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.08  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.34  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | -0.03 | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | -0.06 |                        |
| CBI      | Codon Bias Index               | -1~1     | 0     | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.64  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 1.03 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.18 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.27 | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range      | Value | Reference           |
|-------|-------------------------------|------------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | $0 \sim 1$ | 0.45  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.54  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.49  | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.51  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.47  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.51  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.39  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 61    | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.03  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.76  | <a href="#">[5]</a>                       |

**Gene Name:** Alpha-amylase

**Reference Source:** Codon Usage Database - Kazusa

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

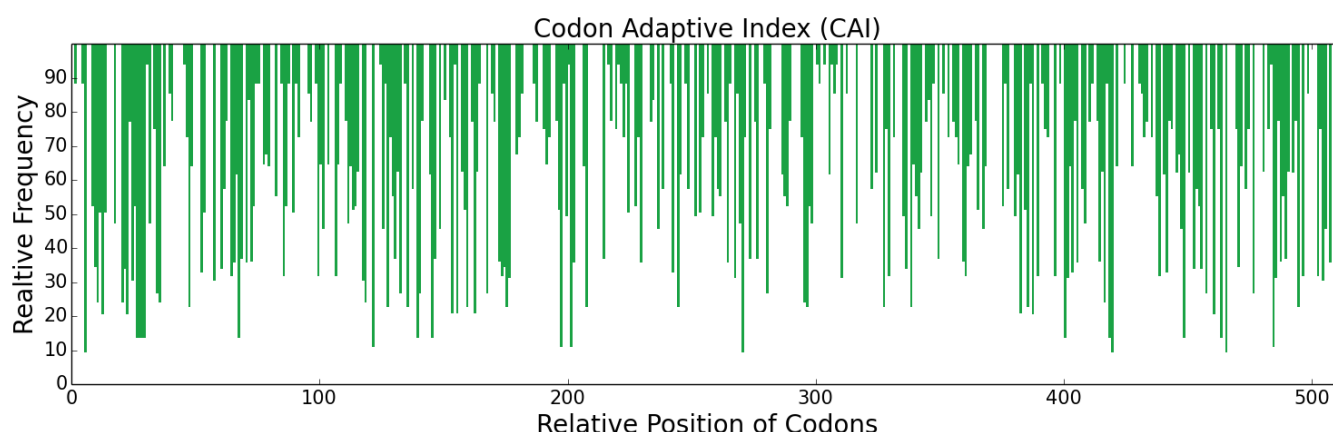
**DNA Sequence Length:** 1536

**DNA Sequence:**

```
ATGAAACAACAAAAACGGTTGTACGCACGTCTCCTTCCGCTTTTGTGTTGCTCTGATTTTT
CTCCTACCGCACTCGGCAGCGGCGGCGGAATCTGAAGGGAACGCTCATGCAGTACTTC
GAGTGGTACATGCCAAATGACGGGCAGCATTGGAAGAGGCTTCAAAACGATTTCGGCTTAC
CTAGCCGAGCATGGCATAACAGCGGTATGGATACCTCCCGCATATAAAGGTACCTCCAG
GCTGATGTGCGTTATGGCGCATATGATCTTTATGACTTGGGTGAATTCCACCAAAAAGGC
ACCGTGAGAACCAAGTACGGCACCAAGGTGAGCTGCAGAGTGCAATCAAGTCGCTCCAT
TCTCGCGATATTAATGTGTATGGGGACGTCGTAATCACGAACAAAGGGGGTGCCGATGCG
ACGGAGGATGTTACAGCGGTAGAAAGTGGATCCTGCTGACCGAAATCGAGTTATCAGTGGG
GAACACCGAATCAAAGCTTGGACGCATTTCCACTTTCCCGGCCGTGGGAGCACTTACTCC
GACTTCAAGTGGCATTGGTATCACTTTGATGGAACCGACTGGGATGAGAGTCGCAAATTA
AATCGCATATACAAGTTTCAGGGGAAGGCTTGGGATTGGGAAGTATCTAATGAGAACGGA
AATTATGACTATCTTATGTACGCAGACATAGATTACGATCACCTGATGTGGCTGCCGAA
ATTAAGAGGTGGGGGACATGGTATGCCAACGAATTACAACCTTGACGGTTTCAGATTAGAC
GCCGTCAAGCACATAAAATTTAGCTTCCTGCGGGACTGGGTAAACCACGTAAGAGAAAAA
ACGGGAAAGGAAATGTTTACAGTCGCAGAGTACTGGCAAAACGACCTCGGGGCACTGGAA
AATTATTTGAATAAGACAAATTTCAATCATAGCGTTTTTCGATGTTCCACTGCATTACCAA
TTTCATGCCGCTTCAACTCAAGGGGGAGGCTACGACATGGGTAAGTTACTAAACGGGACC
GTCGTGTCAAAGCACCCCTTTAAAGCTGTAACCTTCGTTGACAACCACGACACCCAACCC
GGCCAGTCCTTGGAGAGTACTGTGCAGACTTGGTTTAAGCCATTGGCATATGCCTTTATT
TTAACACGAGAAAGTGGGTATCCGTCCTGGCTTTTATGGAGACATGTACGGCACTAAAGGT
GCGAGCCAGAGGGAGATACCAGCCCTGAAGCACAAAATTGAGCCCATCCTCAAAGCGCGG
AAGCAGTACGCTTATGGTGCTCAGCATGATTATTCGACCACCATGACATTGTCGGCTGG
ACAAGGGAGGGAGATTTCATCCGTGGCGAACTCAGGTCTAGCCGCACTAATTACGGATGGA
CCGGGTGGAGCGAAGCGGATGTACGTTGGACGTCAGAACGCCGGAGAAACGTGGCATGAT
ATCACTGGAATCGCAGCGAGCCCGTCGTAATCAACTCAGAGGGGTGGGGCGAATTCCAT
GTTAACGGCGGATCGGTGTCTATATACGTCCAAAGA
```

**Protein Sequence:**

```
MKQQKRLYARLLPLLFLALIFLLPHSAAAAANLKGTLMQYFEWYMPNDGQHWKRLQND SAY
LAEHGITAVWIPPAYKGTSQADVGYGAYDLYDLGEFHQKGTVTRTKYGTKGELQSAIKSLH
SRDINVYGDVVIITNKGADATEDVTAVEVDPADRNRVISGEHRIKAWTHFHFPGRGSTYS
DFKWHWHYHFDGTDWDESRLNRIYKFQGKAWDWEVSNENGNIDYLMYADIDYDHPDVAAE
IKRWGTWYANELQLDGFRLDAVKHIKFSFLRDWVNHVREKTGKEMFTVAEYWQNDLGALE
NYLNKTNFNHVSFVDVPLHYQFHAASQTGGGYDMGKLLNGTVVSKHPLKAVTFVDNHDTQP
GQSLESTVQTFWKPLAYAFILTRSGYPSGFYGDYMTKGASQREIPALKHKIEPILKAR
KQYAYGAQHDFDHHDIVGWTREGDSSVANSGLAALITDGP GAKRMVGRQNAGETWHD
ITGNRSEPVVINSEGWGEFHVNGGSVSIYVQR
```



| 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      | 0.95  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.12  | <a href="#">[4]</a>     |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.33  | <a href="#">[5]</a>     |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.06  | <a href="#">[6]</a>     |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.58 | <a href="#">[7]</a>     |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[8]</a>     |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.95  | <a href="#">[11]</a>    |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[12]</a>    |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0     | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.63  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.1   | <a href="#">[2]</a>    |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.34  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0.02  | <a href="#">[5]</a>    |
| COUSIN18 |                                |          | 0.01  |                        |
| CBI      | Codon Bias Index               | -1~1     | -0.01 | <a href="#">[6]</a>    |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 1.05  | <a href="#">[7]</a>    |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 1.01 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.19 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.27 | <a href="#">[10]</a> |

### 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  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.55  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.5   | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.53  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.49  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.55  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.39  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 29.01 | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.01  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.77  | <a href="#">[5]</a>                       |



**Gene Name:** Cellulase

**Reference Source:** Codon Usage Database - Kazusa

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

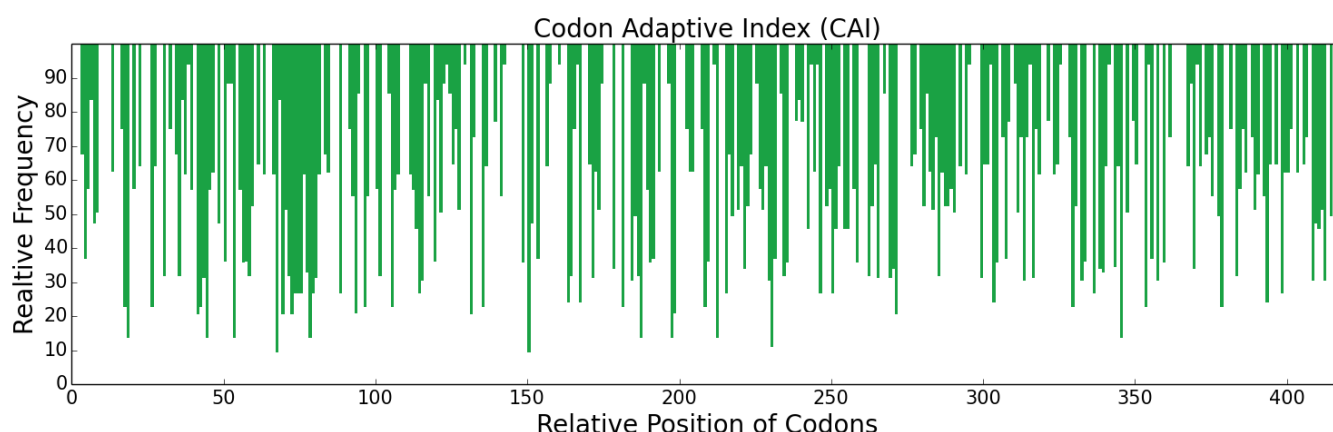
**DNA Sequence Length:** 1254

**DNA Sequence:**

```
ATGAACAAGTCCGTAGCCCCCTCTGCTTTTGGCTGCTTCTATCTTGTACGGAGGGGCGGTT
GCCCAACAGACTGTTTGGGGGCAGTGTGGTGGCATTGGATGGTCCGGCCCTACAAATTGC
GCTCCGGGGAGCGCGTGCTCAACTCTGAACCCCTATTATGCGCAATGCATACCCGGCGCA
ACTACCATTACAACCTTCTACACGGCCTCCGAGTGGCCCGACGACGACGACAAGGGCGACG
AGCACATCTTCCTCAACTCCACCAACGTCTTCTGGAGTCCGATTCGCTGGGGTCAACATT
GCCGGCTTTGATTTCCGGGTGCACAACTGATGGTACATGCGTGACGTCGAAAGTCTACCCC
CCTCTTAAAAATTTACCCGGAAGTAACAATTACCCGGACGGTATTGGGCAGATGCAACAC
TTTGTCAATGAAGATGGTATGACTATATTTCCGGCTGCCAGTAGGTTGGCAGTATTTGGTT
AATAACAACCTCGGCGGAAATCTCGATTCTACCAGCATCAGTAAATACGATCAACTAGTT
CAAGGGTGTTTGTGCGTTAGGCGCGTATTGCATAGTAGATATCCATAACTATGCGCGATGG
AACGGTGGAATCATCGGTCAAGGAGGGGCCACTAATGCGCAATTTACGTCCCTTATGGAGT
CAGCTAGCATCCAAGTATGCCAGTCAGTCGCGCGTATGGTTCGGCATAATGAACGAGCCT
CACGATGTGAATATCAATACGTGGGCAGCCACGGTGCAGGAAGTGGTGACTGCCATAAGA
AACGCTGGCGCAACCAGCCAATTCATTAGCCTACCGGGTAACGATTGGCAGTCCGCTGGA
GCATTCATCAGTGACGGCTCAGCAGCAGCCCTTTCTCAGGTTACAAATCCAGATGGTAGC
ACCACCAATCTCATATTTGACGTACACAAGTATCTTTGACTCGGACAATAGCGGAACACAT
GCTGAGTGTACAACCAATAACATTGACGGGGCATTTTCGCCCTTGGCTACGTGGCTAAGG
CAGAATAACCGTCAGGCGATTCTTACTGAGACCGGTGGTGGGAATGTACAATCGTGTATA
CAAGACATGTGTCAACAAATTCAGTATCTAAATCAGAACTCCGACGTCTACTTAGGGTAC
GTTGGATGGGGCGCCGGATCATTTGACAGTACATACGTCCTCACCGAAACCCCAACGTCA
TCAGGAAACTCATGGACCGACACTTCGCTGGTGAGTTCGTGTTTAGCTCGTAAA
```

**Protein Sequence:**

```
MNKSVAPLLAAASILYGGAVAQQTVWGQCGGIGWSGPTNCAPGSACSTLNPYYAQCIPGA
TTITTSTRPPSGPTTTTTRATSTSSSTPPTSSGVRFAGVNIAGFDGCTTDGTCVTSKVYP
PLKNFTGSNNYPDGIGQMQHFNEDGMTIFRLPVGWQYLVNNNLGGNLDSTSISKYDQLV
QGCLSLGAYCIVDIHNYARWNGGIIGQGGPTNAQFTSLWSQLASKYASQSRVWFGIMNEP
HDVNINTWAATVQEVVTAIRNAGATSQFISLPGNDWQSAGAFISDGSAAALSQVTNPDGS
TTNLI F DVHKYLDSDNSGTHAECTTNIDGAFSPLATWLRQNNRQAILTETGGGNVQSCI
QDMCQQIQYLNQNSDVYLG YVGW GAGSFDSTYVLTETPTSSGNSWTDTSLVSSCLARK
```



| 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      | 0.97  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.15  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.47  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.09  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.6  | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0.01  | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.96  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1.01  | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.6   | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.11  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.32  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | -0.16 | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | -0.3  |                        |
| CBI      | Codon Bias Index               | -1~1     | -0.01 | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.98  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 0.99 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.19 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.31 | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range      | Value | Reference           |
|-------|-------------------------------|------------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | $0 \sim 1$ | 0.43  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.51  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.49  | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.5   | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.5   |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.48  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.52  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 61    | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.02  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.75  | <a href="#">[5]</a>                       |

**Gene Name:** Lipase

**Reference Source:** Codon Usage Database - Kazusa

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

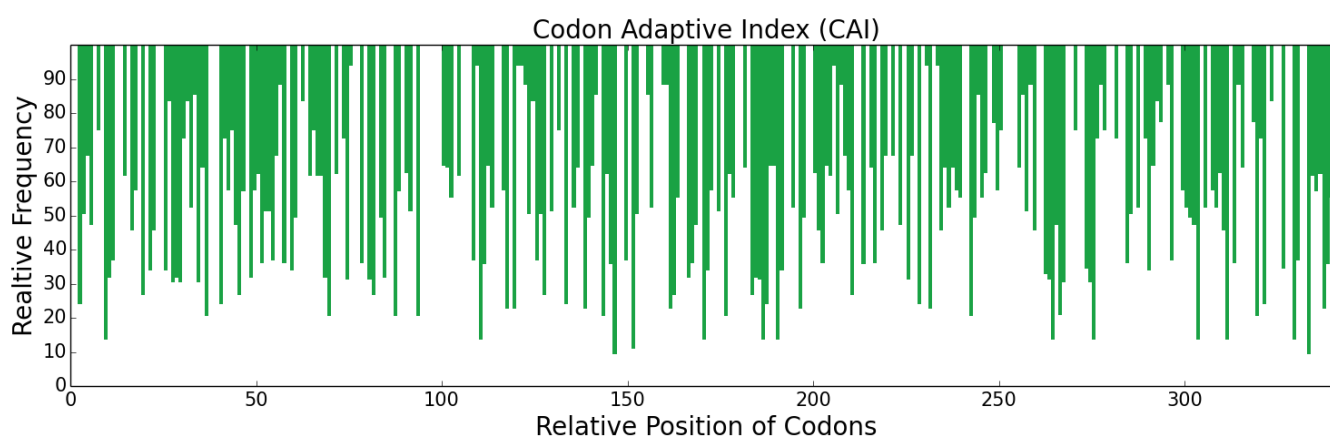
**DNA Sequence Length:** 1026

#### DNA Sequence:

```
ATGAAGCTCCTTTCCCTGACTGGAGTTGCGGGCGTATTGGCTACATGTGTGGCCGCTACG
CCACTAGTGAAGAGACTACCTTCGGGCTCGGACCCTGCATTCTCGCAGCCGAAGTCTGTT
CTCGACGCCGGACTGACGTGCCAAGGCGCCTCACCCAGTAGTGTATCCAAACCCATTCTA
TTAGTTCCCTGGTACAGGAACAACAGGCCCCGAATCATTTGACAGCAATTGGATTCCCTTG
AGCACGCAATTAGGCTACACTCCGTGCTGGATCAGTCCACCGCCATTTATGTTGAACGAT
ACCCAGGTCAACACAGAATACATGGTAAATGCGATAACCGCATTGTACGCCGGGTCTGGG
AATAATAAACTTCCTGTACTTACGTGGAGTCAAGGAGGTCTCGTTGCACAGTGGGGGTTA
ACCTTCTTTCCGTCAATACGGTCTAAGGTAGATCGCCTTATGGCTTTCGCACCAGATTAT
AAAGGGACGGTCTTGGCTGGCCCCCTGGATGCGCTAGCCGTTAGTGCTCCGTCAGTCTGG
CAACAGACTACGGGCAGCGCGCTACCACCGCGCTAAGAAACGCAGGTGGGTAACTCAA
ATCGTGCCCAACCAAAATCTTTATTCGCCACGGATGAAATAGTTCAGCCCCAAGTGTC
AACTCCCCACTGGATAGCTCCTACCTCTTTAATGGGAAGAATGTGCAGGCACAGGCCGTC
TGTGGTCCGTTATTTCGTCATCGATCACGCCGGATCTTTGACTTCTCAGTTCAGTTATGTG
GTTGGTAGGAGCGCGCTGCGATCGACTACTGGACAAGCTCGTTCGGCGGACTATGGAATT
ACTGACTGTAACCCCCCTTCCAGCAAACGACCTAACCCTGAGCAAAAAGTAGCTGCTGCC
GCATTACTGGCGCCAGCAGCTGCCGCAATCGTGGCGGGTCCCAAACAGAACTGTGAGCCG
GACCTCATGCCTTACGCTCGTCCATTTGCGGTAGGTAAGCGGACATGCTCAGGGATAGTC
ACGCCT
```

#### Protein Sequence:

```
MKLLSLTGAVGLATCVAATPLVKRLPSGSDPAFSQPKSVLDAGLTCQGASPSVSKPIL
LVPGTGTTGPQSFDSNWIPLSTQLGYTPCWISPPPFMLNDTQVNT EYMVNAIT ALYAGSG
NNKLPVLTWSQGGLVAQWGLTFFPSIRSKVDRLMAFAPDYKGTVLAGPLDALVSAPSVW
QQTTSALTTALRNAGGLTQIVPTTNLYSATDEIVQPQVSN SPLDSSYLFNGKNVQAQAV
CGPLFVIDHAGSLTSQFSYVVGRSALRSTTGQARSADYGITDCNPLPANDLTPEQKVAAA
ALLAPAAAAIVAGPKQNCEPDLMPYARPFVAVGKRTCSGIVTP
```



| 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   | 0.98  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61 | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | ≥0    | 0.11  | <a href="#">[4]</a>     |
| DCBS  | Directional Codon Bias Score                  | ≥1    | 1.33  | <a href="#">[5]</a>     |
| CDC   | Codon Deviation Coefficient                   | 0~1   | 0.05  | <a href="#">[6]</a>     |
| MILC  | Measure Independent of Length and Composition | -1~1  | -0.61 | <a href="#">[7]</a>     |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1   | 0.06  | <a href="#">[8]</a>     |
| SCUO  | Synonymous Codon Usage Order                  | 0~1   | 0.01  | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1   | 0.97  | <a href="#">[11]</a>    |
| P     | Codon Preference                              | ≥1    | 1.01  | <a href="#">[12]</a>    |
| MCB   | Maximum-likelihood Codon Bias                 | ≥0    | 0.01  | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                      | Range | Value | Reference              |
|----------|----------------------------------|-------|-------|------------------------|
| CAI      | Codon Adaptation Index           | 0~1   | 0.57  | <a href="#">[1]</a>    |
| CFD      | Codon Frequency Distribution     | 0~1   | 0.13  | <a href="#">[2]</a>    |
| FOP      | Frequency of Optimal Codons      | 0~1   | 0.3   | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index     | ∞     | 0.08  | <a href="#">[5]</a>    |
| COUSIN18 |                                  |       | -0.23 |                        |
| CBI      | Codon Bias Index                 | -1~1  | 0.01  | <a href="#">[6]</a>    |
| Dmean    | Mean Dissimilarity-based Index   | 0~2   | 0.93  | <a href="#">[7]</a>    |
| RCA      | Relative Codon Adaptation        | ≥0    | 0.99  | <a href="#">[8]</a>    |
| CUFS     | Codon Usage Frequency Similarity | 0~√2  | 0.24  | <a href="#">[9]</a>    |
| B        | Codon Usage Bias                 | 0~2   | 0.33  | <a href="#">[10]</a>   |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference           |
|-------|-------------------------------|-------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.41  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.49  | <a href="#">[2]</a> |
| P2    | P2 Index                      | 0~1   | 0.51  | <a href="#">[3]</a> |

## 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   | 0.51  | <a href="#">[1]</a>    |
| GC               | GC Content   |       | 0.54  |                        |
| GC1              | GC Content at the First Position of Synonymous Codons  |       | 0.56  |                        |
| GC2              | GC Content at the Second Position of Synonymous Codons |       | 0.53  |                        |
| ENcp             | Effective Number of Codon Pairs                        | 20~61 | 26.8  | <a href="#">[2]</a>    |
| CPS              | Codon Pair Score                                       | -1~1  | 0.04  | <a href="#">[3, 4]</a> |
| Codon Volatility | Codon Volatility                                       | 0.5~1 | 0.73  | <a href="#">[5]</a>    |

**Gene Name:** Phytase

**Reference Source:** Codon Usage Database - Kazusa

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

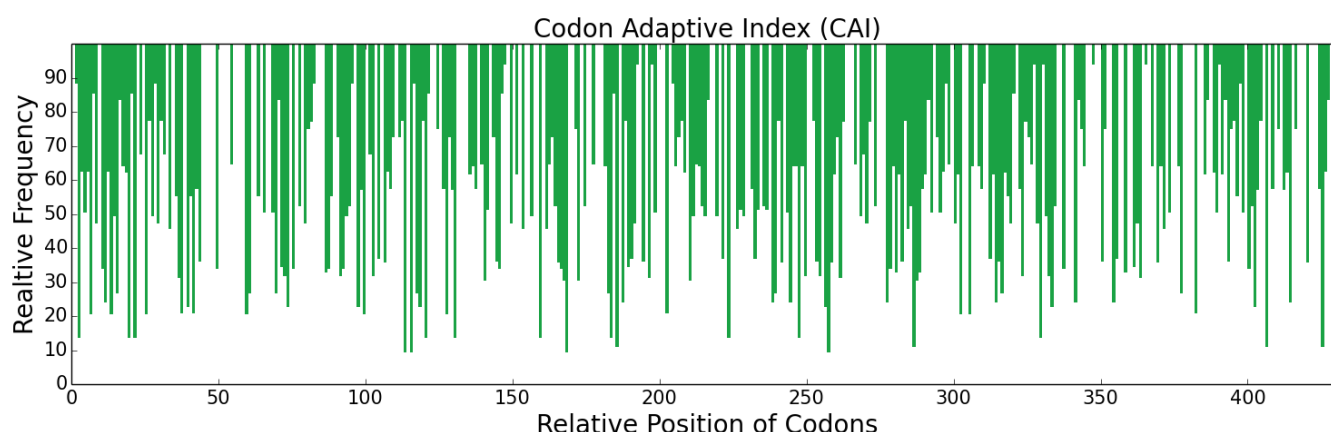
**DNA Sequence Length:** 1296

**DNA Sequence:**

```
ATGAAAGGATCCTTATCCCGTTCTGTCTCTACTCATCCCGTTAACGCCTCAGTCAGCG
TTCGCGCAATCCGAACCGGAGTTAAACTGGAGTCCGTTGTGATTGTCAGCCGACATGGG
GTCCGAGCCCCCACTAAGGCTACTCAACTAATGCAAGATGTTACCCAGATGCTTGGCCG
ACGTGGCCAGTCAAGCTTGGTTGGCTTACGCCTCGTGGCGGGGAACATAATTGCATACCTG
GGACACTATCAAAGACAAAGGCTAGTCGCTGACGGCCTATTAGCAAAAAAGGGTGCCCG
CAATCCGGCCAAGTAGCTATAATCGCCGACGTTGACGAGCGGACTCGGAAAACGGGGGAG
GCGTTTCGCTGCTGGATTGGCCCCGGACTGCGCGATTACTGTTTCATACACAGGCCGATACC
TCGAGTCCAGACCCCCCTATTCAATCCACTGAAGACAGGTGTGTGTCAATTAGATAACGCG
AACGTGACCGACGCAATACTATCGCGGGCTGGTGGATCGATTGCAGATTTTACCGGTCAT
AGACAGACGGCGTTCCGCGAACTCGAGCGTGTACTGAATTTTCCCCAAAGCAATCTTTGT
TTGAAGCGAGAAAAACAGGACGAGTCATGTTTCGTTAACCAGGCATTACCTTCTGAATTA
AAGGTATCTGCGGATAACGTGAGTTTAACTGGTGCCGTAAGTTTGGCAAGTATGCTCACG
GAGATATTTCTTCTCCAGCAGGCGCAGGGCATGCCAGAGCCCGGCTGGGGGCGGATAACA
GACAGCCACCAATGGAACACCTTGTTATCCCTGCACAACGCACAATTTTACCTCCTACAG
AGGACACCCGAGGTGGCAGCTCGAGGGCCACACCTCTTTTGGACCTTATCAAAACCGCT
CTGACACCGCATCCACCGCAGAAGCAGGCCTATGGTGTAACTACCTCCACGTCAGTCCTG
TTCATTGCCGGCCACGACACCAATCTGGCGAATTTAGGCGGGGCATTGGAATAAACTGG
ACTCTCCCTGGACAGCCAGATAATACTCCACCCGAGGTGAACTCGTATTTGAAAGGTGG
AGACGTCTGAGCGATAATTCTCAGTGGATACAGGTGTCTCTTGTTTTTCAGACGTTGCAA
CAAATGCGAGATAAGACACCTTTGTCACTTAATACACCTCCCGGAGAGGTCAAACCTACT
CTAGCAGGGTGCGAGGAACGCAACGCCCAAGGAATGTGCTCACTCGCTGGATTTACTCAA
ATAGTTAACGAAGCCCGCATCCCTGCTTGTAAGTCTC
```

**Protein Sequence:**

```
MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP
TWPVKLGWLTTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE
AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH
RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT
EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA
LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW
RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ
IVNEARIPACSL
```



| 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      | 0.97  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.08  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.31  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.05  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.59 | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.97  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.57  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.13  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.31  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0.01  | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | 0.06  |                        |
| CBI      | Codon Bias Index               | -1~1     | -0.01 | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 1.08  | <a href="#">[71]</a>   |



|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 1.01 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.19 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.32 | <a href="#">[10]</a> |

### 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  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.49  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.5   | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.51  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.52  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.6   |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.47  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 61    | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | -0    | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.74  | <a href="#">[5]</a>                       |

**Gene Name:** Superoxide Dismutase

**Reference Source:** Codon Usage Database - Kazusa

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

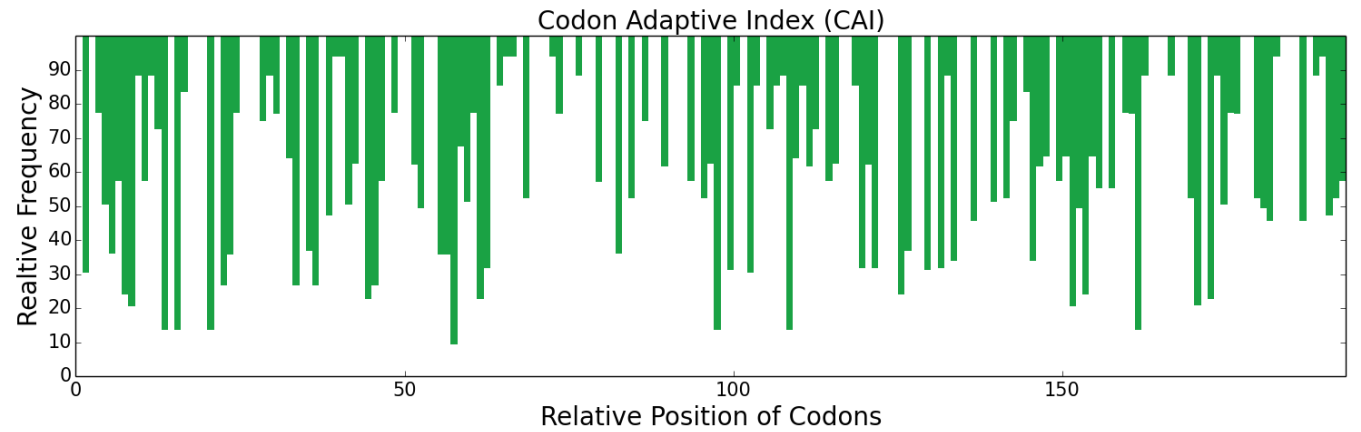
**DNA Sequence Length:** 579

**DNA Sequence:**

```
ATGTCGTTTGTAGCTTCCCGCCCTCCCGTATGCCAAAGACGCGTTGGCGCCTCATATTTCT
GCGGAAACGATAGAGTACCATTACGGAAAAACACCATCAGACGTACGTAACGAACCTGAAT
AATCTTATCAAGGGGACGGCCTTTGAGGGTAAGTCATTAGAAGAAATAATACGGTCCAGT
GAGGGGGGCGTTTTCAATAATGCTGCACAAGTTTGGGAATCACACTTTTTATTGGAACGTC
TTGGCTCCCAACGCAGGTGGAGAACCAACAGGTAAGGTTGCCGAAGCAATCGCGGCTAGC
TTCGGTTCGTTGCTGACTTCAAAGCGCAGTTCACAGACGCTGCCATCAAGAACTTCGGC
TCAGGCTGGACTTGGCTCGTAAAGAACAGCGATGGCAAAGTACTATTGTGTCTACTAGT
AACGCAGGAACCTCTAACAACCGATGCCACCCCGTTACTCACCGTCGATGTCTGGGAG
CACGCGTATTACATTGATTATAGAAACGCACGACCAGGGTATCTTGAGCACTTTTGGGCA
TTAGTGAATTGGGAATTTGTGGCTAAAAATCTGGCAGCC
```

**Protein Sequence:**

```
MSFELPALPYAKDALAPHISAETIEYHYGKHHQTYVTNLNLIKGTAFEGKSLEEIIRSS
EGGVFNNAQVWNHTFYWNCLAPNAGGEPTGKVAEIAAASFSGFADFKAQFTDAAIKNFG
SGWTWLVKNSDGKLAIVSTSNAGTPLTTDATPLLTVDVWEHAYYIDYRNARPGYLEHFWA
LVNWEFVAKNLAA
```



| 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   | 0.99  | <a href="#">[11]</a> |

|      |   |          |       |         |
|------|---|----------|-------|---------|
| ENC  | Effective Number of Codons                    | 20~61    | 61    | [2, 3]  |
| RCBS | Relative Codon Bias Strength                  | $\geq 0$ | 0.18  | [4]     |
| DCBS | Directional Codon Bias Score                  | $\geq 1$ | 1.53  | [5]     |
| CDC  | Codon Deviation Coefficient                   | 0~1      | 0.13  | [6]     |
| MILC | Measure Independent of Length and Composition | -1~1     | -0.69 | [7]     |
| ICDI | Intrinsic Codon Deviation Index               | 0~1      | 0.08  | [8]     |
| SCUO | Synonymous Codon Usage Order                  | 0~1      | 0.02  | [9, 10] |
| Ew   | Weighted Sum of Relative Entropy              | 0~1      | 0.94  | [11]    |
| P    | Codon Preference                              | $\geq 1$ | 1.03  | [12]    |
| MCB  | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.02  | [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           | 0.1   | [2]       |
| FOP      | Frequency of Optimal Codons      | 0~1           | 0.34  | [3, 4]    |
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | -0.18 | [5]       |
| COUSIN18 |                                  |               | -0.44 |           |
| CBI      | Codon Bias Index                 | -1~1          | -0    | [6]       |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 0.84  | [7]       |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1     | [8]       |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.29  | [9]       |
| B        | Codon Usage Bias                 | 0~2           | 0.3   | [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   | 0.5   | [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   | 0.51  | [1]       |
| GC    | GC Content  |       | 0.49  |           |
|       | GC Content at the First Position                      |       |       |           |

|                     |   |       |       |   |
|---------------------|---|-------|-------|---|
| GC1                 | of Synonymous Codons                                      |       | 0.52  |   |
| GC2                 | GC Content at the Second Position<br>of Synonymous Codons |       | 0.44  |   |
| ENcp                | Effective Number of Codon Pairs                           | 20~61 | 20.62 | <a href="#">[2]</a>                       |
| CPS                 | Codon Pair Score  | -1~1  | 0.03  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon<br>Volatility | Codon Volatility  | 0.5~1 | 0.76  | <a href="#">[5]</a>                       |

**Gene Name:** Alcohol Dehydrogenase

**Reference Source:** Codon Usage Database - Kazusa

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

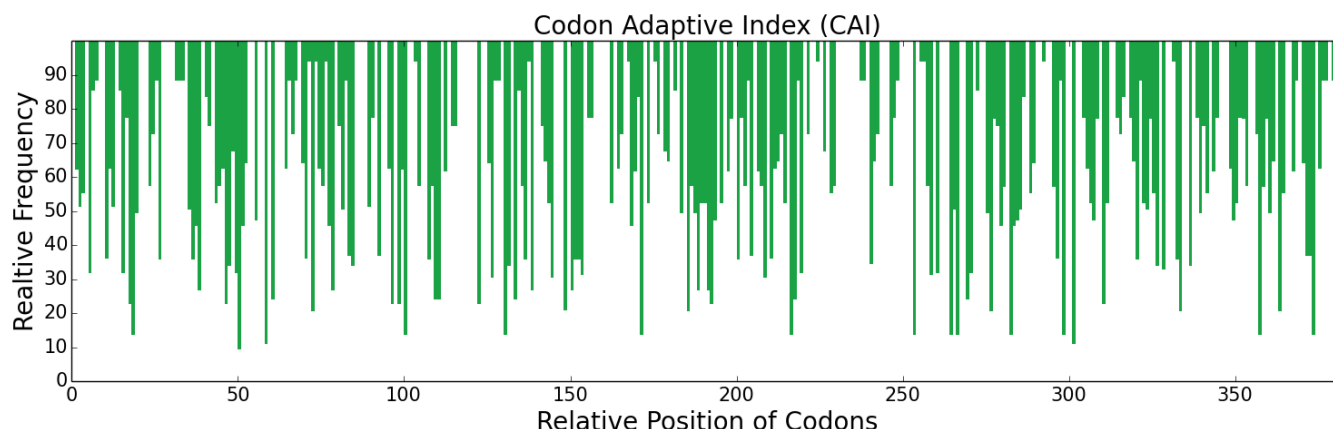
**DNA Sequence Length:** 1146

**DNA Sequence:**

```
ATGTCAAGTGTCACTGGCTTCTATATTCCACCCATCAGTTTTTTTCGGCGAGGGGGCGTTA
GAAGAAACTGCCGACTATATAAAGAACAAGGATTATAAAAAAGCTCTTATAGTGACGGAT
CCTGGAATTGCAGCCATCGGGCTATCCGGCCGGGTGCAGAAGATGCTGGAAGAACGCGAT
CTCAACGTTGCTATCTATGACAAAACCTCAGCCCCAATCCGAATATCGCCAATGTGACGGCT
GGACTTAAAGTACTAAAGGAACAAAACAGTGAGATTGTAGTTTCTATCGGGGGTGGGTCA
GCGCATGATAATGCCAAGGCTATAGCCCTCCTCGCTACAAACGGAGGAGAAATTGGTGAT
TACGAAGGGGTTAACCAGTCGAAAAAAGCTGCGCTACCACTCTTCGCCATAAATACGACT
GCTGGAACCGCATCGGAAATGACTCGATTACGATAATAAGCAACGAGGAGAAGAAGATT
AAGATGGCAATTATCGACAACAATGTGACACCTGCGGTTGCAGTTAATGACCCATCCACC
ATGTTTCGGTTTACCACCGGCCTTAACGGCAGCAACGGGGCTGGATGCATTGACACACTGT
ATAGAGGCCTATGTATCTACAGCCTCGAACCCCATCACCGACGCATGTGCGCTCAAAGGC
ATTGACTTGATTAATGAATCCTTGGTCGCCGCTTACAAGGATGGTAAGGATAAAAAAGCT
CGTACCGACATGTGTTACGCCGAGTATTTGGCTGGTATGGCGTTTAATAATGCCAGCTTG
GGCTACGTTTCATGCGCTTGCGCATCAACTCGGCGGTTTCTACCATTTACCGCACGGAGTG
TGCAACGCGGTGCTGCTTCCTCATGTCCAGGAAGCTAATATGCAATGCCCCAAAGCGAAG
AAGCGCTTGGGTGAGATCGCACTGCACTTTGGGGCATCTCAAGAGGACCCTGAAGAGACC
ATAAAAGCACTTCACGTCTTAAACAGGACTATGAATATACCGAGAAACCTAAAGGAGTTA
GGAGTCAAGACAGAGGATTTTGAAATCCTGGCAGAGCACGCCATGCATGACGCGTGCCAC
TTAACCAACCCGGTCCAATTTACAAAAGAACAGGTAGTAGCGATTATCAAAAAAGCTTAT
GAGTAC
```

**Protein Sequence:**

```
MSSVTGFYI P P I S F F G E G A L E E T A D Y I K N K D Y K K A L I V T D P G I A A I G L S G R V Q K M L E E R D
L N V A I Y D K T Q P N P N I A N V T A G L K V L K E Q N S E I V V S I G G G S A H D N A K A I A L L A T N G G E I G D
Y E G V N Q S K K A A L P L F A I N T T A G T A S E M T R F T I I S N E E K K I K M A I D N N V T P A V A V N D P S T
M F G L P P A L T A A T G L D A L T H C I E A Y V S T A S N P I T D A C A L K G I D L I N E S I V A A Y K D G K D K K A
R T D M C Y A E Y L A G M A F N N A S L G Y V H A L A H Q L G G F Y H L P H G V C N A V L L P H V Q E A N M Q C P K A K
K R L G E I A L H F G A S Q E D P E E T I K A L H V L N R T M N I P R N L K E L G V K T E D F E I L A E H A M H D A C H
L T N P V Q F T K E Q V V A I I K K A Y E Y
```



| 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      | 0.98  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.09  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.34  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.08  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.6  | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.97  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1.01  | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.63  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.1   | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.34  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0     | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | -0    |                        |
| CBI      | Codon Bias Index               | -1~1     | 0     | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.68  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 0.98 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.2  | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.28 | <a href="#">[10]</a> |

### 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  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.54  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.49  | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.51  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.48  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.54  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.38  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 61    | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.02  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.76  | <a href="#">[5]</a>                       |

**Gene Name:** Nattokinase

**Reference Source:** Codon Usage Database - Kazusa

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

**DNA Sequence Length:** 1143

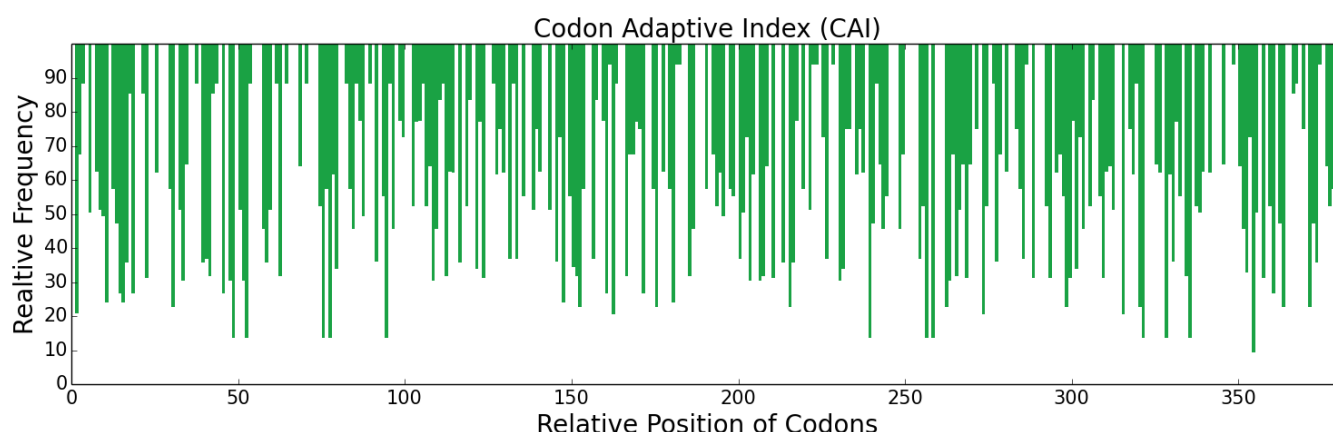
**DNA Sequence:**

```
ATGCGATCCAAAAAGCTTTGGATCAGTTTACTCTTTGCCCTGACGCTCATATTCACGATG
GCTTTTCAGCAACATGTCAGCTCAAGCTGCCGGGAAGAGTTCGACCGAAAAAGAAATACATA
GTAGGCTTCAAACAAACGATGTCGGCGATGAGTTTCGGCGAAAAAGAGGATGTGATAAGT
GAAAAAGGCGGTAAAGTTCAAAGCAGTTTAAATACGTTAACGCAGCGGCCGCGACACTA
GATGAAAAAGCCGTGAAAGAGTTAAAGAAAGATCCCTCTGTCGCGTATGTGGAAGAGGAC
CATATTGCACACGAGTATGCACAGTCGGTGCCTTATGGCATCTCACAAATAAAGGCACCT
GCTCTACACAGCCAAGGTTATACAGGATCAAACGTAAAAGTAGCTGTCATTGATAGTGA
ATCGATTCTAGTCATCCCGACCTCAACGTCCGTGGCGGGGCCTCTTTTGTACCTTCTGAG
ACGAATCCGTATCAAGATGGCTCCTCCACGGAACGCATGTTGCCGGGACTATCGCTGCC
CTCAATAATTCTATTGGCGTGTGGGTGTTGCCCCATCCGCATCATTATACGCCGTCAAG
GTACTTGACTCGACAGGTTTCGGGCCAGTACAGCTGGATTATAAACGGGATAGAGTGGGCC
ATTAGTAATAATATGGACGTAATTAATATGTCGCTAGGAGGACCAACAGGATCAACTGCG
CTGAAAACCGTGGTTCGATAAGGCTGTGTCTCTGGTATTGTTGTAGCAGCGGCTGCGGGT
AACGAAGGGTTCGTCCGGCAGTACCAGCACCGTTGGATAACCGGCAAAGTATCCCTCCACT
ATCGCTGTTGGAGCCGTAAATTCTAGCAACCAAAAGAGCAAGCTTTTCATCCGTTCGGGAGC
GAGCTAGACGTGATGGCACCTGGTGTGTCAGCATCCAGAGTACTTTGCCGGGTGGAACATAT
GGGGCGTACAACGGTACCTCAATGGCGACACCCACGTCGCTGGCGCGGCTGCACCTTATC
TTGTCAAAGCATCCAACCTGGACTAATGCTCAGGTGAGGGACCGGCTTGAAAGCACTGCA
ACGTACCTGGGGAACCTCTTCTATTACGAAAGGGGCTGATAAATGTTTCAGGCAGCCGCG
CAG
```

**Protein Sequence:**

```
MRSKKLWISLLFALTLIFTMAFSNMSAQAAGKSSTEEKYIVGFKQTMSAMSSAKKKDVIS
EKGGKVQKQFKYVNAAAATLDEKAVKELKKDPSVAYVEEDHIAHEYAQSVPYGISQIKAP
ALHSQGYTGSNVKVAVIDSGIDSSHPDLNVRGGASFVPSETNPYQDGSSHGTHVAGTIAA
LNNSIGVLGVAPSASLYAVKVLDTGSGQYSWIINGIEWAISNNMDVINMSLGGPTGSTA
LKTVVVDKAVSSGIVVAAAAGNEGSSGSTSTVGYPKYPSTIavgavNssNQRASFSSVGS
ELDVMAPGVSIQSTLPGGTYGAYNGTSMATPHVAGAAALILSKHPTWTNAQVRDRLESTA
TYLGNSFYYGKGLINVQAAAQ
```





| 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      | 0.97  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.13  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.4   | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.07  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.6  | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.96  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1.01  | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.6   | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.09  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.31  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0.04  | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | 0.06  |                        |
| CBI      | Codon Bias Index               | -1~1     | -0    | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.93  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 0.99 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.19 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.3  | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range      | Value | Reference           |
|-------|-------------------------------|------------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | $0 \sim 1$ | 0.44  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.53  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.5   | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.51  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.5   |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.52  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.48  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 37.41 | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.02  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.74  | <a href="#">[5]</a>                       |

**Gene Name:** Hyaluronidase-1

**Reference Source:** Codon Usage Database - Kazusa

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

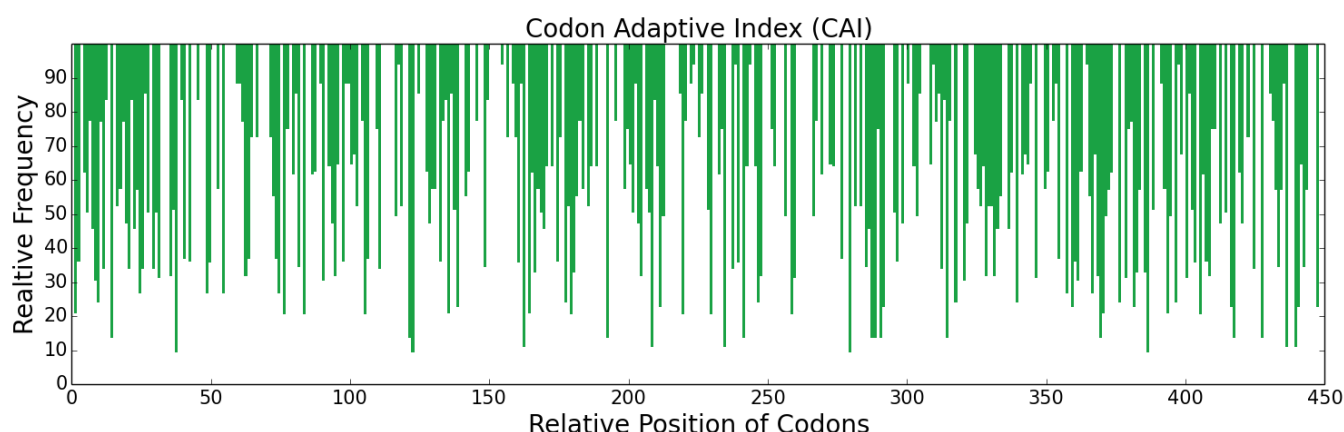
**DNA Sequence Length:** 1350

**DNA Sequence:**

```
ATGCGACCCCTTTTCACTTGAGGTGTCGCTCCACCTACCTTGGGCGATGGCAGCCCACCTG
CTACCTGTGTGCACGCTATTTCCTTAACCTACTTAGCATGACTCAAGGCAGTCGGGATCCT
GTAGTTCCCAACCAACCTTTTACTACGATATGGAACGCCAACACGGAATGGTGTATGAAA
AAACACGGCGTAGACGTTGACATTTCTATTTTTTGACGTCGTAACGAACCCGGGACAAACA
TTCCGTGGTCCGAACATGACAATCTTTTATTCGTCTCAGCTGGGCACCTACCCCTATTAT
ACCTCCGCAGGTGAGCCGGTATTTGGTGGACTACCACAAAACGCTTCTTTAAATGCACAT
TTGGCGCGGACTTTCCAAGATATCCTGGCCGCCATGCCCGAGCCTCGATTCAGTGGGTTG
GCTGTTCATCGATTGGGAGGCTTGGCGTCCTAGATGGGCTTTTAATTGGGACACTAAAGAC
ATATATCGCCAACGATCAAGGGCCCTTGTGCAGAAGCAGCATCCCGACTGGCTCGCACCG
AGGGTCGAGGCCGCTGCACAGGATCAGTTTGAAGGTGCGGCTGAAGAGTGGATGGCCGGA
ACCCCTAAACTGGGCCAAGCCCTTCGCCCTCAGGGGTATGGGGTTTTTACAACCTCCCG
GAGTGTATATAATTACGACTTCAAGAGTCCGAACACACAGGACGCTGTCCACTAAATATA
TGTGCGCAGAATGATCAGCTCGGCTGGTTGTGGGGACAGTCTAGAGCTTTATACCCGAGC
ATTTACTTGCCAGCTGCTTTAGAGGGTACAAAGAAGACCCAGATGTTTGTACAACATCGG
GTTGCAGAAGCATTTCGTGTGGCGGCGGGAGCGGGGGATCCAAAGCTTCCCGTTCTGCCA
TATATGCAGTTATTCTACGATATGACCAATCACTTCCCTACCTGCGGAGGAACTCGAACAT
TCGCTGGGTGAATCCGCCGCACAGGGCGCAGCAGGCGTGGTCTTGTGGGTGTCATGGCTC
TCTACATCCACCAAAGAAAGCTGTCAAGCCATCAAGGAGTATGTAGATACTACGTTGGGG
CCCTCGATCTTGAATGTCACGTCCGGCGCGCGATTATGCTCACAAGTTCTCTGTAGCGGA
CACGGGAGGTGCGCTAGGCGGCCAAGTTACCCAAAAGCCGATTAATTCTCAATTCCACT
AGCTTCAGTATAAAGCCGACACCCGGCGGAGGACCACTGACTCTTCAAGGGCGTTGTCA
CTGGAAGACAGACTAAGAATGGCGGTTGAATTCGAGTGCCGTTGCTATCGCGGTTGGCGC
GGGACCCGTTGCGAACAAATGGGGGATGTGG
```

**Protein Sequence:**

```
MRPFSLEVSLHLPWMAAHLLPVCTLFLNLLSMTQGSRDVPVVPNPFTTIWNANTEWCMK
KHGVDVDISIFDVVTNPGQTFRGPNTIFYSSQLGTYPPYTSAGEPVFGGLPQNASLNAH
LARTFQDILAAMPEPRFSGLAVIDWEAWRPRWAFNWDTKDIYRQRSRALVQKQHPDWLAP
RVEAAAQDQFEGAAEEWMAGTLKLGQALRPQGLWGFYNFPECYNYDFKSPNYTGRCPINI
CAQNDQLGWLWQSRALYPSIYLPAALEGTQKKTQMFVQHRVAEAFRVAAGAGDPKLPVLP
YMQLFYDMTNHFLPAEELEHSLGESAAQGAAGVVLWVSWLSTSTKESCQAIKEYVDTTLG
PSILNVTSGARLCSQVLCSGHRCARRPSYPKARLILNSTSFSIKPTPGGGPLTLQGALS
LEDRLRMAVEFECRCYRGWRGTRCEQWGMW
```



| 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      | 0.94  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.1   | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.33  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.05  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.59 | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.93  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.58  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.12  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.32  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0.1   | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | 0.14  |                        |
| CBI      | Codon Bias Index               | -1~1     | -0    | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.87  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 1.01 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.19 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.3  | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range      | Value | Reference           |
|-------|-------------------------------|------------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | $0 \sim 1$ | 0.43  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.51  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.5   | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.53  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.53  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.56  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.49  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 29.32 | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | -0.01 | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.75  | <a href="#">[5]</a>                       |

**Gene Name:** Laccase

**Reference Source:** Codon Usage Database - Kazusa

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

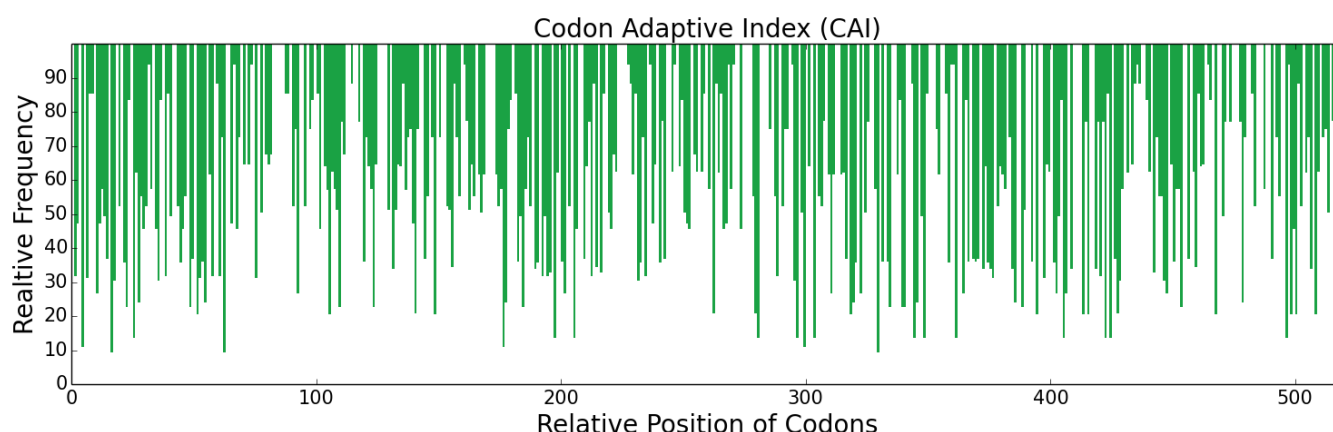
**DNA Sequence Length:** 1557

**DNA Sequence:**

```
ATGGGCCTGCAACGCTTTAGCTTCTTCGTTACGCTGGCCTTAGTAGCTCGGTCGTTGGCA
GCTATAGGGCCTGTTGCGTCACTCGTTCGTTGGCAAATGCCCCAGTGTCGCCTGATGGCTTC
TTAAGAGATGCAATAGTGGTCAACGGGGTAGTTCCGAGCCCCCTCATTACAGGCAAGAAA
GGCGACCGGTTTCAACTGAATGTGGACGATACCTTGACCAATCATAGCATGCTTAAGTCC
ACCTCCATTTCATTGGCATGGTTTCTTCCAAGCAGGAACGAACCTGGGCAGATGGACCTGCT
TTCGTGAACCAAGTGCCCCGATCGCCAGTGGGCACTCCTTTTGTATGATTTTCACGTTCCC
GACCAGGCCGGGACCTTTTGGTACCATAGTCATCTAAGTACCCAGTATTGCGACGGACTG
CGAGGACCATTGTAGTCTACGACCCGAAGGACCCACATGCAAGTCGTTATGACGTCGAT
AATGAGAGTACCGTCATTACACTTACAGATTGGTACCATACAGCAGCCCGCCTCGGACCT
AGATTTCCCTTAGGGGCCGACGCAACTCTAATAAACGGCTTAGGCAGGTCTGCGTCAACT
CCCACGGCTGCATTGGCGGTGATTAACGTACAGCACGGCAAACGTTACAGGTTACAGACTT
GTGTCCATCTCTTGTGATCCAAATTATACATTCTCGATAGACGGCCATAATCTGACCGTT
ATAGAGGTAGATGGTATCAATTCTCAGCCTCTTCTGGTGGATTCCATCCAAATCTTCGCT
GCCCAACGATATTCATTCGTGCTGAATGCCAATCAAACCTGTGGGTAACTACTGGGTCCGA
GCGAACCCAAACTTTGGAAGTGTGCGCTTTGCGAGGGAATTAATTCGGCGATTCTTCGC
TACCAGGGTGCGCCAGTTCGCAGAGCCAACAACGACACAACTACATCAGTAATTCGCTC
ATAGAAACGAACCTTCACCCATTGGCCCGGATGCCCGTTCCCGGGTCTCCAACACCTGGG
GGGGTTGATAAAGCGCTCAACTTAGCGTTCAACTTTAACGGAACAACTTTTTCATAAAT
AATGCGACTTTTACGCTCCCACTGTACCCGTTATTGCTACAGATACTAAGCGGTGCACAG
ACAGCCCCAAGACCTACTCCAGCTGGGAGTGTTTACCCCTTGCCGGCTCATAGCACCATC
GAAATAACGTTACCTGCGACGGCTCTAGCTCCAGGTGCTCCGCACCCGTTTCATCTACAC
GGCCACGCGTTTCGCGTTGTACGATCGGCCGGTTCAACTACCTATAATTATAACGATCCT
ATCTTTAGGGACGTCGTCGACGGGTACCCCGCCGCGGGGATAACGTAACATATCCGT
TTCCAGACCGATAATCCTGGTCCGTGGTTTTTACACTGTCACATTGATTTTCACCTCGAC
GCTGGTTTCGCAATTGTTTTTGCCGAAGATGTAGCTGACGTCAAGGCTGCGAATCCGGTG
CCGAAAGCATGGTCAGACCTATGTCCGATCTACGACGGACTTTCTGAGGCGAACCAA
```

**Protein Sequence:**

```
MGLQRFSSFFVTLALVARSLAAIGPVASLVVANAPVSPDGFLRDAIVVNGVVPSPPLITGKK
GDRFQLNVDDTLTNHSMKSTSIHWHGFFQAGTNWADGPAFVNQCPIASGHSFLYDFHVP
DQAGTFWYHSHLSTQYCDGLRGPFVVYDPKDPHASRYDVNDNESTVITLTDWYHTAARLGP
RFPLGADATLINGLGRSASTPTAALAVINVQH GKRYRFRFLVSI SCDPNYTFSIDGHNLTV
IEVDGINSQPLLVD SIQIFAAQRYSFVLNANQTVGNYWVRANPNFGTVGFAGGINSAILR
YQGAPVAEPTTTQTTSVIPLIETNLHPLARMPVPGSPTPGGVDKALNLAFNFGNTNFFIN
NATFTPTPTVPVLLQILSGAQTAQDLLPAGSVYPLPAHSTIEITLPATALAPGAPHPFHLH
GHAFVVRSA GSTTYNYNDPIFRDVVSTGTPAAGDNVTIRFQTDNPGPWFLHCHIDFHLH
AGFAIVFAEDVADVKAANPVPKAWSDLCP IYDGLSEANQ
```



| 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      | 0.98  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.1   | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.38  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.07  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.58 | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.98  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1.01  | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0     | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.59  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.12  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.32  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0.08  | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | 0.09  |                        |
| CBI      | Codon Bias Index               | -1~1     | -0    | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.85  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 0.96 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.19 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.31 | <a href="#">[10]</a> |

### 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  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.52  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.5   | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.5   | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.51  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.57  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.46  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 61    | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.01  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.74  | <a href="#">[5]</a>                       |



**Gene Name:** Renin

**Reference Source:** Codon Usage Database - Kazusa

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

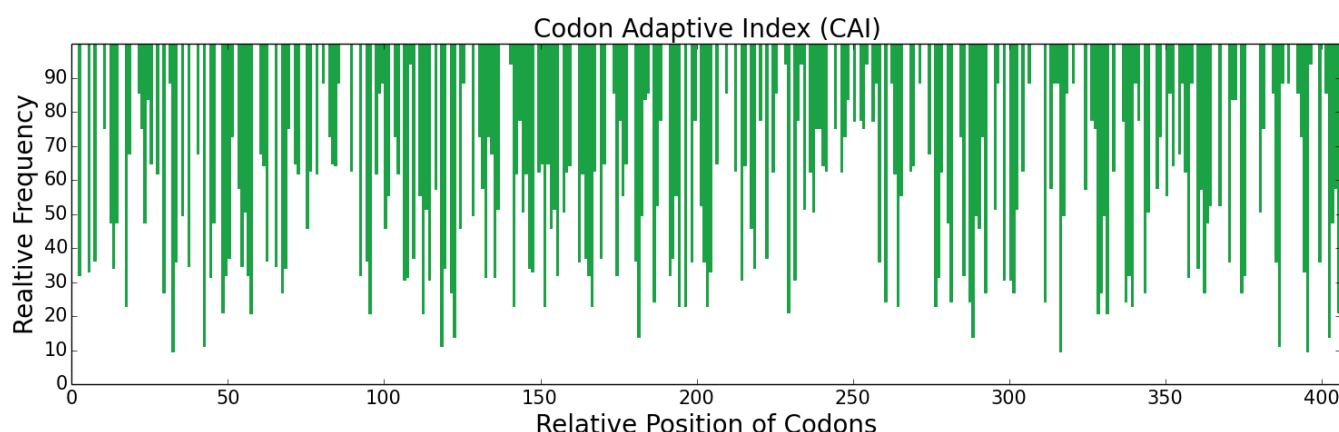
**DNA Sequence Length:** 1218

**DNA Sequence:**

```
ATGGATGGCTGGAGAAGGATGCCCAGATGGGGATTGCTGCTACTGTTGTGGGGGTCCTGT
ACTTTTCGGACTGCCTACCGATACAACCTACGTTTAAACGGATATTTTTAAAGCGTATGCCA
TCCATTTCGCGAAAAGCCTGAAGGAACGAGGCGTAGACATGGCCCCGTCTTGGCCCGGAATGG
TCCCAGCCCATGAAGCGTTTGACGCTAGGAAACACCACATCTTCTGTGATCTTGACAAAC
TATATGGACACCCAGTATTACGGTGAAATCGGTATTGGCACTCCCCCGCAAACATTCAAA
GTGGTCTTTGACACAGGTTTCGAGCAATGTATGGGTCCCGAGTTCGAAGTGCTCTCGCCTA
TACACGGCGTGTGTGTATCATAAGTTATTTGACGCCAGCGACTCCAGCAGTTACAAGCAT
AATGGGACAGAGCTTACACTAAGGTACTCAACCGGGACCGTGAGTGGCTTTCTTTTACAG
GATATTATAACAGTAGGCGGGATCACTGTAACCCAAATGTTTCGGCGAGGTCACCGAAATG
CCCGCGTTACCTTTTCATGCTCGCAGAGTTTGATGGCGTAGTCGGGATGGGGTTTATAGAG
CAAGCAATAGGGAGGGTTACCCCAATTTTCGATAACATCATTTTCGCAGGGTGTGCTAAAG
GAGGATGTATTTTCACTTACTACAATCGAGATTTCGGAGAATAGTCAATCACTTGGAGGA
CAGATCGTTTTTGGGAGGTTTCAGACCCCTCAACACTACGAGGGAAAATTTTCACTATATAAAC
CTCATTAACAGGGGTCTGGCAAAATCCAGATGAAAGGTGTTTCCGTTGGGAGCTCAACT
CTGCTCTGTGAAGACGGCTGTCTCGCGTTAGTGGACACGGGTGCTAGTTATATTTTCGGGT
TCGACGAGTTCTATCGAAAAATTGATGGAAGCTCTCGGTGCCAAAAAACGGTTATTCGAT
TATGTTGTTAAGTGCAACGAGGGACCGACGTTACCGGATATCTCTTTTACCTCGGCGGG
AAAGAGTACACGCTTACTTCTGCCGACTACGTCTTCCAGGAATCCTATTCAAGCAAAAAG
CTATGCACGCTGGCAATTCATGCAATGGATATACCTCCTCCAACGGGCCCAACTTGGGCT
CTTGGAGCTACTTTTCATACGCAAATTTTATACTGAATTCGACAGGCGGAATAACAGAATA
GGTTTCGCGCTGGCCCGA
```

**Protein Sequence:**

```
MDGWRRMPRWGLLLLLLWGSCTFGLPTDTTTTFKRIFLKRMPHSIRESLKERGVDMARLGPEW
SQPMKRLTLGNTTSSVILTNymDTQYYGEIGIGTPPQTFKVVFDTGSSNVWVPSSKCSRL
YTACVYHKLFDAASDSSSYKHNGTELTLYSTGTVSGFLSQDIITVGGITVTQMFGEVTEM
PALPFMLAEFDGVVGMGFIEQAIGRVTPIFDNIISQGVLEKEDVFSFYNNRDSSENSQSLGG
QIVLGGSDPQHYEGNFHYINLIKTVWQIQMKGVSVGSSTLLCEDGCLALVDTGASYISG
STSSIEKLMEALGAKKRLFDYVVKCNEGPTLPDISFHLGGKEYTLTSADYVFQESYSSKK
LCTLAIHAMDIPPTGPTWALGATFIRKFYTEFDRNRNRIGFALAR
```



| 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      | 0.95  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.09  | <a href="#">[41]</a>    |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.34  | <a href="#">[51]</a>    |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.05  | <a href="#">[61]</a>    |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.6  | <a href="#">[71]</a>    |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[81]</a>    |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.95  | <a href="#">[111]</a>   |
| P     | Codon Preference                              | $\geq 1$ | 1     | <a href="#">[121]</a>   |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[131]</a>   |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                    | Range    | Value | Reference              |
|----------|--------------------------------|----------|-------|------------------------|
| CAI      | Codon Adaptation Index         | 0~1      | 0.6   | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution   | 0~1      | 0.11  | <a href="#">[21]</a>   |
| FOP      | Frequency of Optimal Codons    | 0~1      | 0.32  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index   | $\infty$ | 0.01  | <a href="#">[51]</a>   |
| COUSIN18 |                                |          | 0.03  |                        |
| CBI      | Codon Bias Index               | -1~1     | -0    | <a href="#">[61]</a>   |
| Dmean    | Mean Dissimilarity-based Index | 0~2      | 0.71  | <a href="#">[71]</a>   |

|      |                                  |                   |      |                      |
|------|----------------------------------|-------------------|------|----------------------|
| RCA  | Relative Codon Adaptation        | $\geq 0$          | 1.02 | <a href="#">[8]</a>  |
| CUFS | Codon Usage Frequency Similarity | $0 \sim \sqrt{2}$ | 0.19 | <a href="#">[9]</a>  |
| B    | Codon Usage Bias                 | $0 \sim 2$        | 0.29 | <a href="#">[10]</a> |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range      | Value | Reference           |
|-------|-------------------------------|------------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | $0 \sim 1$ | 0.45  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | $0 \sim 1$ | 0.53  | <a href="#">[2]</a> |
| P2    | P2 Index                      | $0 \sim 1$ | 0.5   | <a href="#">[3]</a> |

### Indices based on complex patterns of codon usage

| Index            | Description  | Range        | Value | Reference                                 |
|------------------|--|--------------|-------|---|
| GC3              | GC Content at the Third Position of Synonymous Codons  | $0 \sim 1$   | 0.52  | <a href="#">[1]</a>                       |
| GC               | GC Content   |              | 0.48  |   |
| GC1              | GC Content at the First Position of Synonymous Codons  |              | 0.48  |   |
| GC2              | GC Content at the Second Position of Synonymous Codons |              | 0.44  |   |
| ENcp             | Effective Number of Codon Pairs                        | $20 \sim 61$ | 61    | <a href="#">[2]</a>                       |
| CPS              | Codon Pair Score                                       | $-1 \sim 1$  | 0.02  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon Volatility | Codon Volatility                                       | $0.5 \sim 1$ | 0.75  | <a href="#">[5]</a>                       |

**Gene Name:** Avidin

**Reference Source:** Codon Usage Database - Kazusa

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

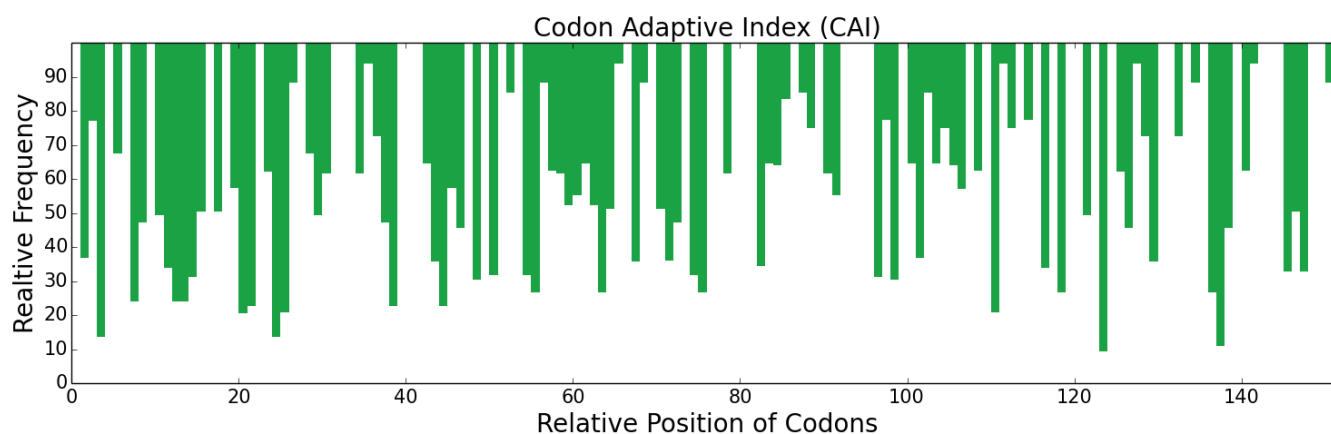
**DNA Sequence Length:** 456

#### DNA Sequence:

```
ATGGTACACGCGACTTCCCCACTCCTGTTGTTACTACTCCTCAGCCTTGCTCTTGTTGCC
CCGGGGTTGTTCAGCGCGAAAATGTTCCCTTAACAGGTAAGTGGACAAATGACCTGGGGTCT
AACATGACCATAGGGGCCGTGAACTCGAGAGGCGAATTCACTGGCACGTATATCACAGCA
GTCACCGCAACGAGTAATGAAATAAAAGAAAAGTCCCCCTGCATGGCACGCAAAACACAATT
AACAAAGCGTACCCAGCCTACTTTTCGGATTACAGTCAACTGGAAGTTTAGCGAGTCGACT
ACCGTATTCACCGGACAGTGCTTTATCGATCGAAATGGAAAGGAGGTTCTAAAGACGATG
TGGTTATTGCGGTCTTCAGTGAATGACATAGGTGATGACTGGAAAGCTACGCGCGTGGGT
ATCAATATTTTTTACTAGGCTTAGGACTCAAAAAGAG
```

#### Protein Sequence:

```
MVHATSPLLLLLLLLLSLALVAPGLSARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITA
VTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCfidrngkevlkTM
WLLRSSVNDIGDDWKATRVGINIFTRLRTQKE
```



| 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   | 0.98  | <a href="#">[11]</a>   |
| ENC  | Effective Number of Codons      | 20~61 | 61    | <a href="#">[2, 3]</a> |
| RCBS   | Relative Codon Bias Strength    | ≥0    | 0.13  | <a href="#">[41]</a>   |

|      |   |          |       |                         |
|------|---|----------|-------|-------------------------|
| DCBS | Directional Codon Bias Score                  | $\geq 1$ | 1.45  | <a href="#">[5]</a>     |
| CDC  | Codon Deviation Coefficient                   | 0~1      | 0.07  | <a href="#">[6]</a>     |
| MILC | Measure Independent of Length and Composition | -1~1     | -0.76 | <a href="#">[7]</a>     |
| ICDI | Intrinsic Codon Deviation Index               | 0~1      | 0.06  | <a href="#">[8]</a>     |
| SCUO | Synonymous Codon Usage Order                  | 0~1      | 0.01  | <a href="#">[9, 10]</a> |
| Ew   | Weighted Sum of Relative Entropy              | 0~1      | 0.94  | <a href="#">[11]</a>    |
| P    | Codon Preference                              | $\geq 1$ | 1.02  | <a href="#">[12]</a>    |
| MCB  | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.02  | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                      | Range         | Value | Reference              |
|----------|----------------------------------|---------------|-------|------------------------|
| CAI      | Codon Adaptation Index           | 0~1           | 0.59  | <a href="#">[1]</a>    |
| CFD      | Codon Frequency Distribution     | 0~1           | 0.12  | <a href="#">[2]</a>    |
| FOP      | Frequency of Optimal Codons      | 0~1           | 0.31  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | 0.04  | <a href="#">[5]</a>    |
| COUSIN18 |                                  |               | -0.79 |                        |
| CBI      | Codon Bias Index                 | -1~1          | -0    | <a href="#">[6]</a>    |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 0.99  | <a href="#">[7]</a>    |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1.01  | <a href="#">[8]</a>    |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.25  | <a href="#">[9]</a>    |
| B        | Codon Usage Bias                 | 0~2           | 0.31  | <a href="#">[10]</a>   |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference           |
|-------|-------------------------------|-------|-------|---------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.43  | <a href="#">[1]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.51  | <a href="#">[2]</a> |
| P2    | P2 Index                      | 0~1   | 0.49  | <a href="#">[3]</a> |

### 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   | 0.52  | <a href="#">[1]</a> |
| GC    | GC Content   |       | 0.47  |                     |
| GC1   | GC Content at the First Position of Synonymous Codons  |       | 0.43  |                     |
| GC2   | GC Content at the Second Position of Synonymous Codons |       | 0.47  |                     |

|                     |                                 |       |       |                        |
|---------------------|---------------------------------|-------|-------|------------------------|
| ENcp                | Effective Number of Codon Pairs | 20~61 | 20    | <a href="#">[2]</a>    |
| CPS                 | Codon Pair Score                | -1~1  | -0.01 | <a href="#">[3, 4]</a> |
| Codon<br>Volatility | Codon Volatility                | 0.5~1 | 0.74  | <a href="#">[5]</a>    |

**Gene Name:** GFP

**Reference Source:** Codon Usage Database - Kazusa

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

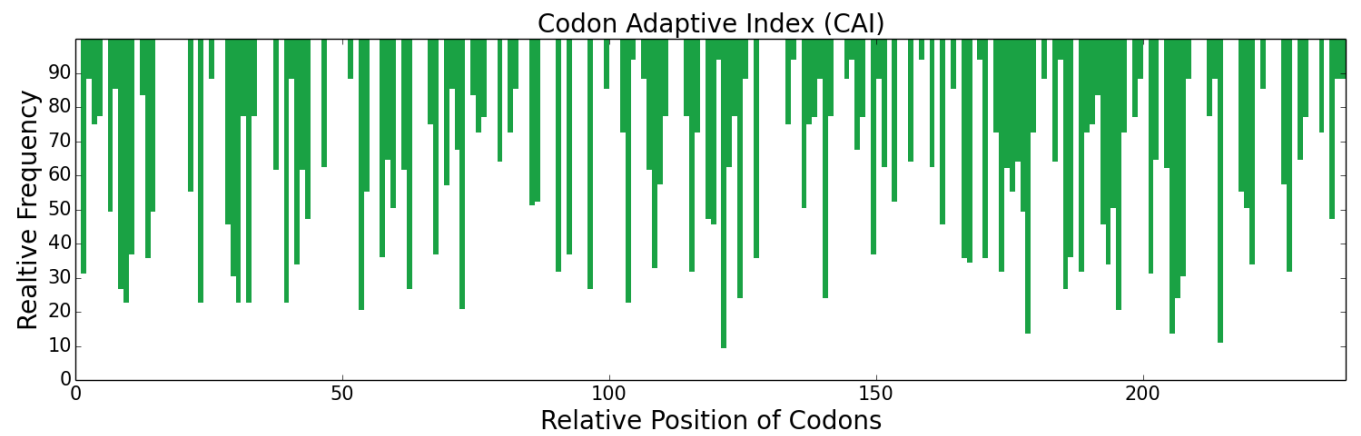
**DNA Sequence Length:** 714

**DNA Sequence:**

ATGAGCAAAGGAGAGGAATTATTCACGGGGGTAGTTCCTATATTAGTTGAATTGGATGGT  
GATGTCAACGGGCATAAATTTTCTGTGTCGGGGGAGGGGGAGGGTGATGCTACATACGGG  
AAACTAACACTGAAGTTTATCTGTACTACTGGTAAATTGCCGGTCCCATGGCCCACCCTT  
GTTACAACGTTTTCTTACGGAGTACAATGCTTCTCCCGATACCCTGACCACATGAAGCAG  
CATGACTTCTTTAAGAGTGCAATGCCAGAAGGCTACGTACAAGAAAGAACGATTTTTTTC  
AAGGATGACGGGAATTACAAAACAAGGGCCGAGGTTAAGTTTGAGGGCGACACTCTGGTG  
AATCGGATCGAGCTCAAAGGTATAGATTTTAAGGAAGATGGAAATATTCTTGACACAAA  
CTCGAGTACAACACTATAATTCCCACAACGTATATATCATGGCAGATAAGCAGAAGAATGGT  
ATCAAGGTGAACTTCAAGATACGTCATAATATAGAAGACGGCTCAGTCCAGTTAGCGGAC  
CATTATCAACAGAATACGCCCATTGGCGACGGACCTGTGCTACTTCCGGACAACCACTAT  
TTGAGCACCCAATCAGCGCTCTCGAAAGATCCAAACGAGAAACGCGATCATATGGTCCTT  
CTAGAATTGCTTACTGCTGCCGGCATTACCCACGGTATGGACGAACTGTATAAA

**Protein Sequence:**

MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGDATYGKLTCLKFICTTGKLPVPWPTL  
VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV  
NRIELKGIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD  
HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK



| 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      | 0.98  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.11  | <a href="#">[4]</a>     |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.37  | <a href="#">[5]</a>     |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.06  | <a href="#">[6]</a>     |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.67 | <a href="#">[7]</a>     |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0     | <a href="#">[8]</a>     |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0     | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.97  | <a href="#">[11]</a>    |
| P     | Codon Preference                              | $\geq 1$ | 1.01  | <a href="#">[12]</a>    |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.01  | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                      | Range         | Value | Reference              |
|----------|----------------------------------|---------------|-------|------------------------|
| CAI      | Codon Adaptation Index           | 0~1           | 0.67  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution     | 0~1           | 0.08  | <a href="#">[2]</a>    |
| FOP      | Frequency of Optimal Codons      | 0~1           | 0.37  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | 0.02  | <a href="#">[5]</a>    |
| COUSIN18 |                                  |               | 0.03  |                        |
| CBI      | Codon Bias Index                 | -1~1          | 0.02  | <a href="#">[6]</a>    |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 0.75  | <a href="#">[7]</a>    |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1.03  | <a href="#">[8]</a>    |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.18  | <a href="#">[9]</a>    |
| B        | Codon Usage Bias                 | 0~2           | 0.24  | <a href="#">[10]</a>   |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference            |
|-------|-------------------------------|-------|-------|----------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.48  | <a href="#">[11]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.57  | <a href="#">[2]</a>  |
| P2    | P2 Index                      | 0~1   | 0.51  | <a href="#">[3]</a>  |

### Indices based on complex patterns of codon usage

| Index | Description | Range | Value | Reference |
|-------|-------------|-------|-------|-----------|
|       |             |       | 0.5   |           |



|                     |  |       |       |   |
|---------------------|--|-------|-------|---|
| GC                  | GC Content   | 0~1   | 0.45  | <a href="#">[1]</a>                       |
| GC1<br>GC3          | GC Content at the First Position<br>GC Content at the Third Position<br>of Synonymous Codons |       | 0.53  |   |
| GC2                 | GC Content at the Second Position<br>of Synonymous Codons                                    |       | 0.32  |   |
| ENcp                | Effective Number of Codon Pairs  | 20~61 | 27.16 | <a href="#">[2]</a>                       |
| CPS                 | Codon Pair Score   | -1~1  | 0.02  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon<br>Volatility | Codon Volatility   | 0.5~1 | 0.77  | <a href="#">[5]</a>                       |

**Gene Name:** RFP

**Reference Source:** Codon Usage Database - Kazusa

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

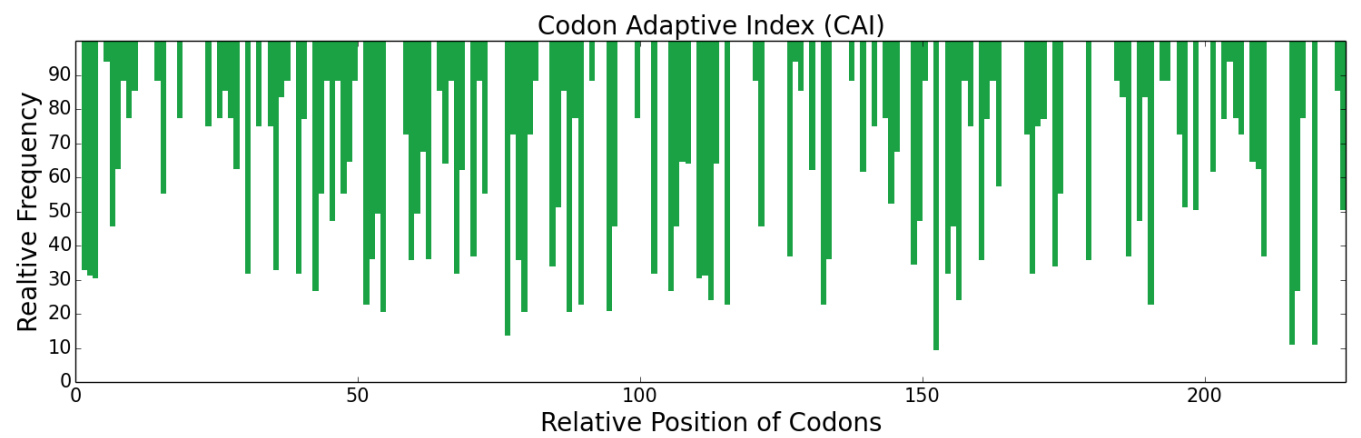
**DNA Sequence Length:** 675

**DNA Sequence:**

ATGAGGAGCTCGAAGAATGTGATCAAAGAGTTTCATGAGATTTAAAGTCAGAATGGAGGGT  
ACTGTTAACGGACATGAGTTTCGAGATCGAAGGCGAAGGAGAAGGAAGGCCTTATGAAGGC  
CACAAACACGGTCAAACCTGAAAGTCACCAAAGGTGGGCCCTTACCGTTTGCTTGGGACATA  
TTATCCCCCAATTCCAGTATGGCTCAAAGGTATATGTCAAGCATCCAGCGGACATACCG  
GACTATAAGAAGCTAAGTTTCCCGAGGGGTTTAAATGGGAACGAGTGATGAACTTTGAG  
GATGGTGGCGTTGTTACGGTGACCCAGGATTCGAGCCTCCAGGATGGGTGTTTTATTAC  
AAAGTGAAGTTTATTGGTGTAATTTCCCATCAGATGGGCCCGTTATGCAAAAAAAGACA  
ATGGGATGGGAGGCATCCACTGAACGTCTGTATCCACGGGATGGCGTGCTCAAAGGAGAA  
ATACACAAAGCCTTGAAGTTGAAGGACGGCGGACACTACCTAGTCGAATTTAAGTCTATA  
TACATGGCTAAGAAACCTGTACAACCTGCCTGGGTACTATTATGTTGACAGTAAGCTTGAT  
ATTACATCTCACAATGAGGACTACACCATCGTAGAACAATACGAACGCACGGAGGGTCGC  
CATCATTTGTTTCCTT

**Protein Sequence:**

MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI  
LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCFIY  
KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVCLKGEIHKALKLKDGGHYLVEFKSI  
YMAKKPVQLPGYYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL



| 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      | 0.98  | <a href="#">[11]</a>    |
| ENC   | Effective Number of Codons                    | 20~61    | 61    | <a href="#">[2, 3]</a>  |
| RCBS  | Relative Codon Bias Strength                  | $\geq 0$ | 0.16  | <a href="#">[4]</a>     |
| DCBS  | Directional Codon Bias Score                  | $\geq 1$ | 1.5   | <a href="#">[5]</a>     |
| CDC   | Codon Deviation Coefficient                   | 0~1      | 0.08  | <a href="#">[6]</a>     |
| MILC  | Measure Independent of Length and Composition | -1~1     | -0.67 | <a href="#">[7]</a>     |
| ICDI  | Intrinsic Codon Deviation Index               | 0~1      | 0.06  | <a href="#">[8]</a>     |
| SCUO  | Synonymous Codon Usage Order                  | 0~1      | 0.01  | <a href="#">[9, 10]</a> |
| Ew    | Weighted Sum of Relative Entropy              | 0~1      | 0.95  | <a href="#">[11]</a>    |
| P     | Codon Preference                              | $\geq 1$ | 1.01  | <a href="#">[12]</a>    |
| MCB   | Maximum-likelihood Codon Bias                 | $\geq 0$ | 0.02  | <a href="#">[13]</a>    |

### Indices based on codon frequency in a reference set of genes

| Index    | Description                      | Range         | Value | Reference              |
|----------|----------------------------------|---------------|-------|------------------------|
| CAI      | Codon Adaptation Index           | 0~1           | 0.66  | <a href="#">[11]</a>   |
| CFD      | Codon Frequency Distribution     | 0~1           | 0.08  | <a href="#">[2]</a>    |
| FOP      | Frequency of Optimal Codons      | 0~1           | 0.37  | <a href="#">[3, 4]</a> |
| COUSIN59 | Codon Usage Similarity Index     | $\infty$      | 0.15  | <a href="#">[5]</a>    |
| COUSIN18 |                                  |               | 0.31  |                        |
| CBI      | Codon Bias Index                 | -1~1          | 0.02  | <a href="#">[6]</a>    |
| Dmean    | Mean Dissimilarity-based Index   | 0~2           | 0.67  | <a href="#">[7]</a>    |
| RCA      | Relative Codon Adaptation        | $\geq 0$      | 1.05  | <a href="#">[8]</a>    |
| CUFS     | Codon Usage Frequency Similarity | 0~ $\sqrt{2}$ | 0.21  | <a href="#">[9]</a>    |
| B        | Codon Usage Bias                 | 0~2           | 0.25  | <a href="#">[10]</a>   |

### Indices based on adaptation to the tRNA levels and their supply

| Index | Description                   | Range | Value | Reference            |
|-------|-------------------------------|-------|-------|----------------------|
| tAI   | tRNA Adaptation Index         | 0~1   | 0.49  | <a href="#">[11]</a> |
| gtAI  | Genetic tRNA Adaptation Index | 0~1   | 0.57  | <a href="#">[2]</a>  |
| P2    | P2 Index                      | 0~1   | 0.48  | <a href="#">[3]</a>  |

### Indices based on complex patterns of codon usage

| Index | Description | Range | Value | Reference |
|-------|-------------|-------|-------|-----------|
|       |             |       | 0.52  |           |

|                     |  |       |       |   |
|---------------------|--|-------|-------|---|
| GC                  | GC Content   | 0~1   | 0.46  | <a href="#">[1]</a>                       |
| GC1<br>GC3          | GC Content at the First Position<br>GC Content at the Third Position<br>of Synonymous Codons |       | 0.52  |   |
| GC2                 | GC Content at the Second Position<br>of Synonymous Codons                                    |       | 0.33  |   |
| ENcp                | Effective Number of Codon Pairs  | 20~61 | 33.53 | <a href="#">[2]</a>                       |
| CPS                 | Codon Pair Score   | -1~1  | 0.04  | <a href="#">[3]</a> , <a href="#">[4]</a> |
| Codon<br>Volatility | Codon Volatility   | 0.5~1 | 0.77  | <a href="#">[5]</a>                       |