
GenRCA Rare Codon Analysis Report

Gene Name: TLR5M

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Danio rerio (zebrafish)

DNA Sequence Length: 2463

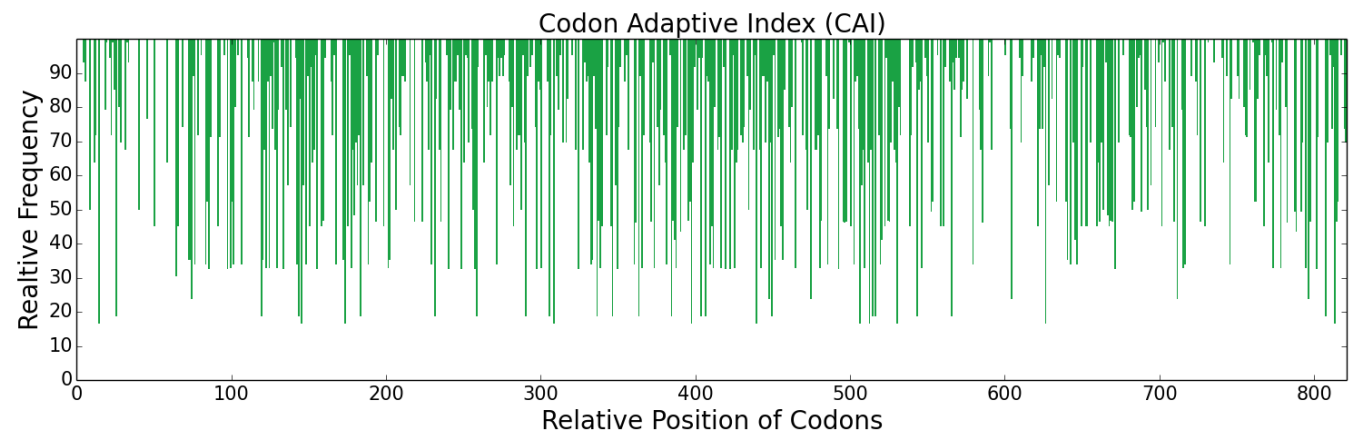
DNA Sequence:

```
ATGTGGACACTGGCCTTTTCACATGGTTTTTCATCGGTTTCATACCTACAGGGAATCGCATGC
TACCCATCATGCACTTTTATACGGCATTGTGGCTAATTGCGCCATGCAGCACCCTACTGG
GTTCTGCTCTGCCTCCCAACATCAGACCTCTACCTGGAGATGAACTACATCGGTGAG
ATCAACAGCAGCTCGCTCAGAAGATATGACCAGCTGCAACAAGTAGATCTTGGAAGTCAG
AAAGTGCAGCTTGTCTATAAGGAACAACGCTTTCCTCAGGCAGAGAAAAGTATAAAGTTG
GTCCTTGGCAACACCCAGCTTCAGCTGGAGCCAAGGGCTTTTGAGGACTGTTTAACCTTA
CAAAATTTGTTTTTGGATCATTGCAATTTGGCAGACTCTATACTGGCAGAAAGCTATCTG
CAGCCACTTTTATCCCTAGAAAAGCTTGATCTCTCTGGTAATAAAATAGTGAGACTCCGA
CCAGGACTGTTCTTTTCAAACTTACAAAGTTTACACAACCTAAACCTCAAAATTGAATCGT
ATTGAAAGTTTATGTGAAGAGGATCTTGTCTGGTTTCCAGGGGAAATACTTCACACTCCTG
AACTTGCAATCCAATTACGTTAGATATAGTGATGATTTTGACTGGGAAAGATGTGGGAAC
CCTTTTCAGAGGGATGGCCTTTAATACCCCTTGACTTATCCAGCAATGGGTTTCAGCGTGACC
ATAGCAAGAAATTTTTTTAAAGCAATACAGGGGTACCCCAATTGCTCATGTTATATTATCT
GGACACATGGGTAAAGGCTTTTTCATTTAACAACCTTCCAGATCCAGATAAAAAACACATTT
GAAGGCCTCATGAATAGTTCAGTTACCATTTTAGATCTGTCTAAAAACTATATATTTACT
TTGCAGAAGGCTTTTTTTAAGTCTCTAAAGGATGCAAAAATCATTGACATTTCCATGAAC
AAAATCAATCAGATAAACAGAAATGCCTTTGATGGTCTTCAAGATCATTACGATTGCTC
AACCTGTCTATCTAACCTCTTAGGAGAAATATATTCTCACACATTTACCAATCTGACAGAC
CTTCGGGTGTTAGATTTGTCTATACTACATTGGGGCCTTGGGATATGAAGCATTGAGT
GGTCTTCTCTAAATTACGAGCGTTGTATCTGACGGGAAATTCAGTGCAGAGTCTAGGTTCT
CCTGCACCATTAACAACTTAGATTTTCTTCTCTTGGGCGACAATAGGTTGACATCTATA
TACAACATAAAATAAATTGGGTAATAACAGTATTTATGTGGATGTTGCAGACAATAAACTA
ACAAATTTGGATGATATTTTGTAAATTTTAACCCAATTCAGGCATCTCCAACTTTCTTC
TATGGTGGCAACTTGATCAAGTGGTGCAGTCAGAATGTTACAGTACCTCACATAATAGT
TTGCGAGTGCTGGATCTTCATGACAGCTCCCTGCATATAATCTGGGGGCGGGGGGAGTGC
CTCAACCTTTTTTGATCATCTAGAAAATTTGCAGGGTCTAAATTTAAGCTTAACTCACTT
GCGGCTCTCCCACGAGGGATTTTTTAATGGTCTAGGCTCAATCACAGAGATCGACCTCTCA
TCTAACGCCTTAACCTTTTTTGCAGCCAGATATTTTCCCGGTCAGCCTGAAAAGACTCGAC
CTCTCTAACAACCTTTTTAGCCACTCCAGACCCAAGGACTTTTCAGTCCATCAGCTTCCTT
AGCCTGGCTGCAAATCGGTTCCACTGTGATTGCAATCTGGAGAGCTTCCTGAAGTGGCTG
AAAGTGACACATGTAAACTTCTTCAGCCCAATTGATGAGTTCAGATGTGAGTTTCAGCT
GCTCTCCATAAACATTCTCTACTGGAATACTCCACAAACGTCAAACCATGTGAGGACGTC
CAAGACCTTAAGATTGCGCTTTTCATCCTCTCTGCTCTCCTCATCATTGCTTTTCATCCTC
GGTGGGATTGTTTACGCCCCTCTCCGAGGGCACATATTCATTATCTACAAAAAGATCATC
AGTAGGGTTGTCAAGGGCCCAAAACCGACAACCTTATGTTGACGAAGTGCAGTATGACGCC
TTCTCTGCTTTAGCAACAATGACTATGGGTGGGTAGAGGCTGCATTGCTTAAGAAGCTG
GATAACCAGTTTTTCAAAGGGGAACATCCTCCACTGCTGTTTCGAGGCCAGAGACTTCCTG
CCAGGTGAGGATCACCTTTTCAACATCAGAGATTCCATCTGGGGCAGTAGGAAGACTGTG
TGCGTCGTCTCCAAAGAGTTTCTTCAAAGATGGTTGGTGCTTGGAGGCATTCTCATTGGCC
CAGGGCCGGATGCTGGAGGAGCCGACGAGCATCCCCGATTATGTTGGTGGTAGGGAAGGTG
```

TGTATAAGGCTGCTGAGATACTTAATTTACACCCATTCTCTAGGGGTGAGCAAAAGCCAT
ATT

Protein Sequence:

MWTLAFHMFVIGSYLQGIACYPSCCTLYGIVANCAMQHXYWVPALPPNITDLYLEMNYIGE
INSSSLRRYDQLQQVDLGSQKVQLVIRNNAFLRQRKLIKLVLGNTQLQLEPRAFAGLFNL
QNLFLDHCNLADSILAESYLQPLLSLEKLDLSGNKIVRLRPGLFFSKLTKFTQLNLKLNK
IESLCEEDLVGFQGYFTLLNLQSNYVRYSDDFDWERCNPFPRGMAFNTLDLSSNGFSVT
IARNFFKAIQGTPIAHVILSGHMGKGSFNNLPDPDKNTFEGLMNSSVTILDLSKNYIFT
LQKAFLSPLKDAKIIDISMNKINQINRNAFDGLQDHLRLLNLSSNLLGEIYSHTFTNLTD
LRVLDLSYNYIGALGYEAFSGLPKLRALYLTGNSLRVLGSPAPLPNLDLFLLLGDNRLTSI
YNINKLGNNSIYVDVADNKLTLNDDIFVILTQFRHLQTFYGGNLIKWCSQNVTVPHNNS
LRVLDLHDSSLHIIWGRGECNLNFDHLENLQGLNLSLNSLAALPRGIFNGLGSITEIDLS
SNALTFLQPDIFPVSLKRLDLSNNFLATPDPRTFQSISFLSLAANRFHCDCNLESFLKWL
KVTHVNFFSPIDEFRCEFPAAHLKHSLLEYSTNVKPCEDVQDLKIALFILSALLIIAFIL
GGIVYARLRGHIIFIYKKIISRNVKGPKPTTYVDEVQYDAFLCFSNNDYGWVEAALLKKL
DNQFSKGNILHCCFEARDFLPGEDHLSNIRDSIWGSRKTVCVVSKEFLKDGWCLEAFSLA
QGRMLEEPTSIPIMLVVGKVCIRLLRYLIYTHSLGVSKSHI



Negative CIS Elements	Negative repeat Elements
3	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.1	[11]
ENC	Effective Number of Codons	20~61	54.57	[2, 3]
RCBS	Relative Codon Bias Strength	≥0	0.16	[41]
DCBS	Directional Codon Bias Score	≥1	1.49	[51]
CDC	Codon Deviation Coefficient	0~1	0.08	[61]
MILC	Measure Independent of Length and Composition	-1~1	-0.48	[71]
ICDI	Intrinsic Codon Deviation Index	0~1	0.05	[81]
SCUO	Synonymous Codon Usage Order	0~1	0.05	[9, 10]

Ew	Weighted Sum of Relative Entropy	0~1	0.93	[11]
P	Codon Preference	≥ 1	1.07	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.2	[13]

Indices based on codon frequency in a reference set of genes

Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.71	[1]
CFD	Codon Frequency Distribution	0~1	0.05	[2]
FOP	Frequency of Optimal Codons	0~1	0.38	[3, 4]
COUSIN59	Codon Usage Similarity Index	∞	-1.16	[5]
COUSIN18			-2.86	
CBI	Codon Bias Index	-1~1	0.06	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.5	[7]
RCA	Relative Codon Adaptation	≥ 0	1.07	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.17	[9]
B	Codon Usage Bias	0~2	0.24	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.22	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.26	[2]
P2	P2 Index	0~1	0.56	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.49	[1]
GC	GC Content		0.43	
GC1	GC Content at the First Position of Synonymous Codons		0.47	
GC2	GC Content at the Second Position of Synonymous Codons		0.34	
ENcp	Effective Number of Codon Pairs	20~61	35.19	[2]
CPS	Codon Pair Score	-1~1	0.04	[3, 4]
Codon Volatility	Codon Volatility	0.5~1	0.77	[5]

Gene Name: TLR5S

Reference Source: Codon Usage Database - Kazusa

Expression Host Organism: Danio rerio (zebrafish)

DNA Sequence Length: 1932

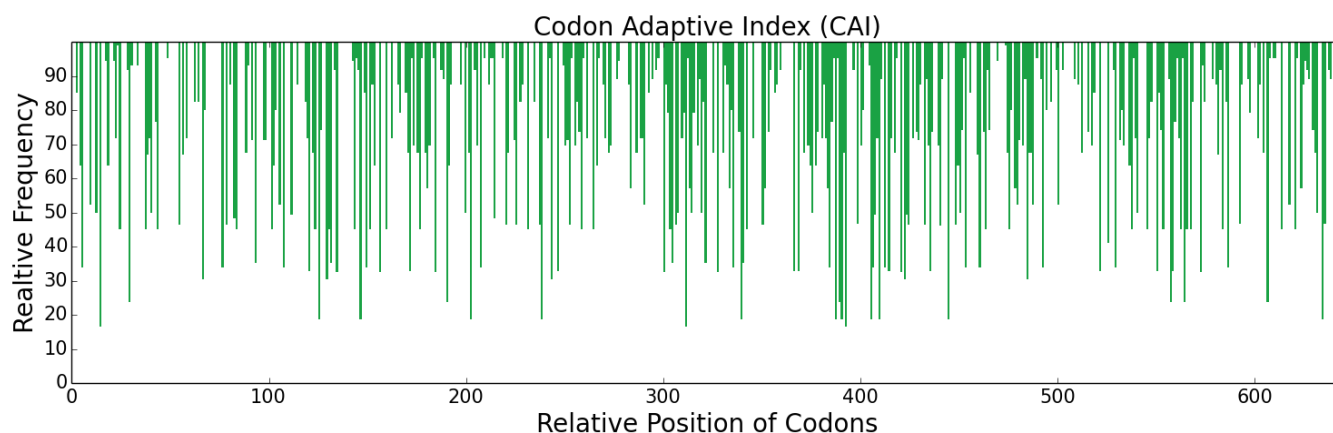
DNA Sequence:

```
ATGTGGACTCTGGGTCTTCAGGTGGCTGTCATCTGTGTTTTCTACAGGTGCCAGGTTGT
TTCCCATCATGCCTCATCGTGAACCTCTGTAGCCAACTGTGCCTACAAGAACCCTCCGCTCA
GTTCTCCCTCCCTCCTCCTCACATCACCCACCTGTACCTGGAGGGGAACCGCATCAGTGAG
ATCAACTCCAACCTCCCTGTCGGGCTGAAGGAGCTGCAGGAGCTGGACCTTGGAGGGCAG
TTTGTGCGTCTCATGATCAGAAACAATGCCTTCAGGGAGCAAAGACACCTGAGGAGGCTG
GTGCTCGGTGGCAACGTCCACCTTCAGCTGGAGCCGCAGGCTTTTGTGGGACTGTCCAGT
TTGCAGAATCTCTACTTATATCACTGTTTCGCTCCAACAGTCTATACTGGAGGAGGACTAC
CTGGAGCCACTCACCTCTTTAGAGACTCTTGACCTCTTTGGTAACAAGATAAAGAGACTC
CAGCCTTCAATGTTCTTTGCAAACATGACTAATTTGAAAAATCTGAATCTCAAACCTGAAT
GAAATTGACAAAATATGTGAGTCTGATCTGGTAGGTTTTTCAGGGAAAGCACTTTGAGGTT
CTGAATTTAGACTCTATTTTCTTAAGACCATGTTTAAAAAACGTGACTGGCAGAAATGT
GGGAATCCTTTTCAGGGGGATGTCCTTTTCAGACACTCGACCTGTCCAACAACGGGTTAAGC
GTGAGTAAATCGAAGCAGTTGTTTCACAGCCATTAGGGGGACCAAGATTTCCCATCTCAAA
CTGTCACTGGGACTCATGGGTAAAGGATTTTTCATTCAATATTTACCCTGATCCAGACAGC
AGCATGTTTGAAAGCCTGAATGACAGTTTCAGTCCACACTCTGGATCTGTCTAAAAACAAG
ATATTTGCACTCCAAGAGGGGGTTTTTCAGTGCACATAAAGAAGTTGCAATCATTGATGTT
TCCCAAAACAACGTGAATCAGATACACAGAAATGCCTTTGAAGGCCTTCAGGGACATTTA
CAAATGCTCAACCTGTCACACAACCTGCTGGGGGAAATCCATTCTGACACTTTTGCTTCT
CTGACAAACCTGCAGGTGTTGGACTTGCTTACAATCACATTGGTGTCTGGGTCATGAC
TCATTTAGTGAACCTCCCAAATTAAGTATTAAATCTAACAGGAACTCTCTGCGAGAC
ATTGGCTTCCCTGCCTTACTTCCGAGTTTAGATTACCTTCTGTTGAGTGACAATAAACTG
ATAGACTCGCCGGGGAGAAGTATCCATAGGTTTGTCTGGGGATATTTTGCATCTGGACATT
CGGGATAACAGATTAACAAACCTGGGGGGTGTTTATACCCTTGTGACTCTGCTGGACCGC
CTTCAGCATCTCTCTTATGGAGGAAACCAATCAAGTGGTGCAATCTCGGCAGAGAAGTC
AGGTTCAATTGATTGCAATAATGTCACAACCTGGATCTTCACGGCAGCTCCCTGCAGTCT
GTCTGGTCTCAGTGGACATGTCTGGATCTGTTTGACAATTTTCGGACATGTGATTACTCTG
AACTTGAGCCACAACGCGCTGCAGTCTCTTCTAGGGGCATTTTCAAGGGTCTCACCTCA
GTTGTGGAGATGGACCTCTCATCCAACAGCTTGACTTATCTCCAGCCTGATGTATTGCC
AAAAGTCTCAAAGTACTCGACCTCTCCAACAACCTTCATAGCCTCCCCTGACCCTGATGCT
TTTCGCTCTCTCAGCTCCCTTGACCTGAACATGAACCGATTTCACTGTGATGCAAACCTG
AAGAGTTTTCTGAATTGGGTAAACCAACACCACCGTGACACTCCTGACACCTGTCAAGGAG
CTCAAATGTGAATTTCCATCTGATTTCTATAATGTTTCCTCTGTTACGATTCTCTGATGAC
ATCACACAGCAG
```

Protein Sequence:

```
MWTLGLQVAVICVFLQVPGCFPSCLIVNSVANCAAYKNLRSVPPLPPHITHLYLEGNRISE
INSNSLSGLKELQELDLGGQFVRLMIRNNAFREQRHLRRLVLGGNVHLQLEPQAFVGLSS
LQNLVLYHCSLQQSILEEDYLEPLTSLETLDLFGNLIKRLQPSMFFANMTNLKILNLKLN
EIDKICESDLVGFQKGHFEVLNLDSIFLKTMTFKRDWQKCGNPFRGMSFQTLDSLNNGLS
VSKSKQLFTAIRGTKISHLKLSLGLMGKGFNFNIYPDPDSSMFESLNDSSVHTLDLSKNK
IFALQEGVFSALKEVAIIDVSQNNVNQIHRNAFEGLQGHQLMLNLSHNLLGEIHSDTFAS
LTNLQVLDLSYNHIGVLGHDSFSELPKLVNLNLTGNSLRDIGFPALLPSLDYLLLSDNKL
```

IDSPGRSIIHRFAGDILHLDIRDNRNLTLNKGGVYTLVTLLDRLQHLSYGGNPIKWCNLGREV
RFIDSNNVTTLDLHGSSLQSVWSQWTCLDLFDNFGHVITLNLSHNALQSLPRGIFKGLTS
VVEMDLSSNSLTYLQPDVLPKSLKVLDSLNNFIASPDPAFRSLSSLDLNMNRFHCDANL
KSFLNWTNTTVTLLTPVKELKCEFPSTDFYNVPLLRFSDDITQQ



Negative CIS Elements	Negative repeat Elements
0	0

Indices based on non-uniform usage of synonymous codon				
Index	Description	Range	Value	Reference
RSCU	Relative Synonymous Codon Usage	0~6	1.19	[11]
ENC	Effective Number of Codons	20~61	53.69	[2, 3]
RCBS	Relative Codon Bias Strength	≥ 0	0.22	[4]
DCBS	Directional Codon Bias Score	≥ 1	1.68	[5]
CDC	Codon Deviation Coefficient	0~1	0.14	[6]
MILC	Measure Independent of Length and Composition	-1~1	-0.45	[7]
ICDI	Intrinsic Codon Deviation Index	0~1	0.1	[8]
SCUO	Synonymous Codon Usage Order	0~1	0.09	[9, 10]
Ew	Weighted Sum of Relative Entropy	0~1	0.89	[11]
P	Codon Preference	≥ 1	1.1	[12]
MCB	Maximum-likelihood Codon Bias	≥ 0	0.39	[13]

Indices based on codon frequency in a reference set of genes				
Index	Description	Range	Value	Reference
CAI	Codon Adaptation Index	0~1	0.75	[11]
CFD	Codon Frequency Distribution	0~1	0.03	[2]
FOP	Frequency of Optimal Codons	0~1	0.44	[3, 4]

COUSIN59	Codon Usage Similarity Index	∞	3.14	[5]
COUSIN18			8.64	
CBI	Codon Bias Index	-1~1	0.16	[6]
Dmean	Mean Dissimilarity-based Index	0~2	0.28	[7]
RCA	Relative Codon Adaptation	≥ 0	1.12	[8]
CUFS	Codon Usage Frequency Similarity	0~ $\sqrt{2}$	0.15	[9]
B	Codon Usage Bias	0~2	0.2	[10]

Indices based on adaptation to the tRNA levels and their supply

Index	Description	Range	Value	Reference
tAI	tRNA Adaptation Index	0~1	0.25	[1]
gtAI	Genetic tRNA Adaptation Index	0~1	0.29	[2]
P2	P2 Index	0~1	0.54	[3]

Indices based on complex patterns of codon usage

Index	Description	Range	Value	Reference
GC3	GC Content at the Third Position of Synonymous Codons	0~1	0.55	[1]
GC	GC Content		0.46	
GC1	GC Content at the First Position of Synonymous Codons		0.51	
GC2	GC Content at the Second Position of Synonymous Codons		0.33	
ENcp	Effective Number of Codon Pairs	20~61	35.87	[2]
CPS	Codon Pair Score	-1~1	0.05	[3] , [4]
Codon Volatility	Codon Volatility	0.5~1	0.76	[5]