BIOBRICKS FOUNDATION PRESENTS

10K GENES PROJECT 2017-48

December 4, 2017

https://biobricks.org/10k-genes

Contributors:
Keoni Gandall
Stanley Qi
Zavier Henderson
Aaron Cravens
Adarsh Ambati

Editor:
Keoni Gandall
koeng101@gmail.com

Abstract

The 10K Genes Project is a BioBricks Foundation project aimed to synthesize and make freely available DNA in partnership with Twist Bioscience. Anyone around the world may request DNA to be synthesized through the 10K Genes Project as long as it is freely available to the scientific community. This document represents the week 48 of 2017 10K Genes Project order.

0 Introduction

This synthesis order is the first of our scheduled orders. Out of the 268kb synthesized this order, only about 6kb is for internal testing. This synthesis order will be placed with Twist at the end of day on Nov 22, 2017.

Contents

1 Gene Name: f1f1

Gene ID: BBF.2017.48.1

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TTGTTTAACT TTTCGAGACC TTAGGAGGTA AAaATATGTT CATTAGTACC GTGAGTCTTG
TACTGGCGCT TGCGACCAGT GCCATCTTCG CAGCTCCTAC AACCTTGTCG ACAAAAGATG
CATCTTCTGA CTTGCCGTTT TACTTCCCAA AACCCACCTA CCAGTCCATC CCATACGACC
ATTCTTCAAA CCTGGCGGCG CCCTCAAATG CCGATTCCGT CAAAATTGGT ACCGCGTATA
TTCGTGACAA ATTAAATTTG TCTCCAGATG AGTTCAAAAT CTCTAGCTCA TTTACAGATT CCGC-
CGGTGT GACGCATGTA TACGGTTCGC ACCTTGTTAA TGGAGCCAGC GTTGCGAACC ACCAGGC-
CGC CGTACACATC AAAGGAGGTC AAATTACCTC TTATTCTGCC TCATTCGGTA CCAAATCACA
TTTGACCAAA TCTGAATTAT CCATCGCGGA GCCCAAGGCG AAATTAACTT TTGAACAAGC
GAGTGGATTG GTCTCAAAGG CGAAGGGACT TCCCGTTTAT AGCGAGTTCA AACATGATTT
AGAATACGTA CAGCAAGCAG ACGGTACCCT GGTGTACGCG TATAAGTTTC AGTTGCGCGA
TAAATCGGTT AACTTCTGGT TCCAGGTTTG GGCCGACGCT AATACTGGGA AGCTGATCCA
AGCGGTAGAC TTCGCACATC GCGCAACGTA CAAAGCAATT AAACTGCCGG GTTTAACTCC
AGATGACGGG TTTTCTGATG TCAAGGACCC CGAATATAAG GCCAGCAGTC CAAATGGGTG
GACCGACGGA ACTACGACAG AGGGGAACAA CGTCTCGGCA CTGGACTTGC GTAAAGAGCC
TGCAGTCCCG GGTACGGGGG TCAATGGCGT CTTCGACTCT AAATTCAACC CCACGGCGGA
CATCACAGAT CCGGATAACA TGCAAGCAAC TGCTGTTAAC TTGTTCTACA TCACCAACAT
TATGCATGAC ATCACATACC AATATGGATT CAATGAAGCG GCTGGAAATT TTCAAAAAGA
TAACTTCGGC AAAGGGGGTG AAGGCAACGA CCCAGTCATT ATCTCTGTTC TGGAT

2 Gene Name: f1f2

Gene ID: BBF.2017.48.2

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TCATTATCTC TGTTCTGGAT GTCATTATCT CTGTTCTGGA TACGTCAGAC GAAGAAAACG
CTGATTTTTT TTCCCCTGCC GACGGCCAGC CGGGTCAGAT GCGCATGTTT CGCTTTGCAC
ACTTCAAACC GAACCGCAAC CCCGGTCTTG ACAATCAGGT GGCAATGCAT GAATATGGGC
ACGGAGTCTC CGGACGCTTA ACAGGCGGGC CCGCGGTGGT CGGCTGCCTT CGCGGCGCAG
AGGCCGGTGG TATGGGTGAA GGGTGGTCTG ATCTGTTTGC AATGATTGTT TCAGCAAAGG
AGTCACAAAA ATCCGATACT CCCATCGCCA TCGGGACGTA CATCCAGGGC GGTCCTGAAG
GTATTCGCAG TCACCCATAT ACTACTGACA TGAAAGTCAA CCCCTTGACA TATTCGACCT
TGAAAAAACG CCTGGAAGTA CATGACGCTG GCGAAGTGTG GGCCTCGATG CTGTGGGAAG
TTTATTGGAA CCTTGTGACG AAAAATGGCT TTTCGACCGA TATCTATGAT GCGAAATCGA
ATTCTGGAAA TACTATCACG ATGCAAAATA TGATCGGCGG CTTCATGCTG CAACCTTGTC
ACCCCACTTT GATCTCCGCT CGTGACGCCT TCATTGAGAG TGATGCAAAC CGTTATAAAG
GGGCGAACAA ATGTGAAATC TGGAAGGGCT TTGCAAAGCG CGGCCTGGGC GTGAAGGCTG
CAGACTTCAA CGACGATTTT AGTGTCCCGG CTGATTGCGA TACCGGGGCA TCGCCGCCTC
CCTCCAGCCG CGGTCCTCCT GCGACGACCG AAGGTGCAAC TACGACCGAG GGAGCAACCA
CGACGGAAGG TGCGACTACC ACGGAAGGTG CAACAACCAC TGAGGGGGCA ACCACCACGG
GGCGTGCGAC GACAACTGGG CGTGCTACCA CGACAGAGGG GGCGACGACT ACTGAAGGGG
CGACCACAAC CGAGCGCGCC ACCATCACTA CGAAACGTCC TACCACGACG GGGCGTGCAA

CAACCACGGC CCGCGGACCG AAACCTACGA GCGGCTGTAC GCGCATGGAT ATTTGTTGCA TTTACATTGG TAAGTCTTGC GATCAAAATA TTCTGTACAC TCGAGGGTC

3 Gene Name: f2f1

Gene ID: BBF.2017.48.3

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TTGTTTAACT TTTCGAGACC TTAGGAGGTA AAaATATGAT CCTTAAATCT TTGAGTAAGA TGTTCATCTC
CACTGTCAGT TTAGTGCTGG CGTTAGCCAC TTCAGCAATT TTTGCCGCCC CGACTACCCT
GAGTACAAAA GACGCATCAT CTGATTTACC TTTTTATTTT CCAAAACCGA CATATCAATC TATC-
CCATAT GATCATTCGT CCAACTTAGC GGCACCGAGT AATGCCGACT CTGTAAAGAT CG-
GTACTGCG TACATTCGCG ATAAGCTGAA CTTGTCTCCC GATGAATTCA AGATCTCGTC GAGTTT-
TACC GATTCTGCCG GAGTCACCCA CGTTTACGGT TCCCATCTGG TCAACGGCGC GTCCGTG-
GCG AACCATCAAG CGGCGGTGCA CATCAAAGGA GGACAGATCA CTTCTTATTC CGCTTCTTTC
GGAACAAAGT CGCACCTGAC GAAATCAGAA TTATCAATTG CCGAGCCCAA GGCGAAATTA
ACTTTCGAAC AGGCTTCCGG TTTAGTATCA AAAGCCAAAG GATTACCTGT CTACAGCGAA
TTCAAACATG ACTTAGAATA CGTCCAACAA GCGGACGGTA CCCTTGTGTA TGCTTATAAA
TTCCAATTAC GCGATAAGTC GGTAAATTTC TGGTTTCAGG TGTGGGCGGA TGCAAATACT
GGCAAGTTGA TCCAGGCCGT AGACTTCGCC CACCGTGCCA CATATAAGGC GATTAAGTTA
CCCGGGCTTA CCCCAGATGA CGGTTTCTCT GACGTAAAAG ACCCGGAATA TAAGGCTTCT
AGCCCTAACG GTTGGACTGA CGGCACAACT ACCGAGGGCA ACAATGTGTC TGCCCTGGAT
CTTCGTAAAG AGCCTGCGGT CCCAGGGACC GGTGTCAATG GGGTTTTTGA CAGTAAATTT
AACCCTACGG CTGACATTAC TGACCCGGAC AATATGCAGG CTACAGCAGT GAATCTTTTC
TATATCACGA ATATTATGCA TGATATCACG TACCAGTATG GGTTCAACGA AGCCGCGGGC

AACTTCCAAA AAGACAACTT TGGAAAAGGA GGTGAAGGAA ATGATCCAGT GATCATCAGC GTGTT

4 Gene Name: f2f2

Gene ID: BBF.2017.48.4

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

CCAGTGATCA TCAGCGTGTT TCCAGTGATC ATCAGCGTGT TAGATACTTC GGATGAGGAA
AACGCCGACT TCTTTTCTCC TGCCGATGGC CAACCCGGTC AGATGCGCAT GTTTCGTTTC
GCTCATTTTA AACCAAATCG CAATCCAGGG CTTGATAACC AAGTTGCAAT GCACGAGTAC
GGCCACGGCG TATCCGGGCG TTTGACGGGG GGACCAGCTG TTGTCGGTTG CTTACGCGGT
GCAGAGGCTG GTGGTATGGG TGAGGGGTGG AGCGATCTTT TCGCCATGAT CGTTTCAGCT
AAAGAAAGTC AAAAAAGCGA CACCCCGATT GCCATCGGAA CGTATATCCA AGGGGGACCG
GAGGGTATCC GTTCTCATCC CTACACCACG GATATGAAGG TGAACCCTCT TACCTATAGC
ACCTTAAAAA AGCGCCTGGA GGTCCACGAC GCCGGCGAGG TGTGGGCTAG TATGCTTTGG
GAAGTCTACT GGAATTTGGT GACGAAGAAT GGATTCAGTA CTGATATTTA CGACGCAAAG
TCAAACTCAG GTAACACGAT TACAATGCAG AATATGATTG GTGGGTTCAT GTTGCAACCT
TGCCACCCAA CTCTGATCAG TGCGCGTGAC GCTTTCATTG AATCAGACGC GAACCGCTAT
AAGGGCGCCA ACAAATGTGA GATTTGGAAG GGGTTTGCTA AACGTGGGCT TGGAGTTAAG
GCTGCGGACT TTAACGACGA CTTCTCCGTG CCTGCTGATT GTGACACAGG CGCGTCACCT
CCCCCTAGTT CACGTGGACC CCCTGCAACG ACAGAGGGCG CTACTACGAC GGAAGGCGCC
ACAACAACCG AAGGCGCGAC CACGACCGAG GGGGCAACCA CAACAGAGGG TGCCACCACT
ACCGGTCGCG CAACCACAAC TGGTCGTGCC ACGACAACCG AAGGTGCTAC CACCACCGAG
GGAGCCACCA CCACTGAACG CGCGACAATT ACCACAAAAC GCCCGACTAC AACTGGCCGC

GCCACTACTA CAGCCCGCGG ACCGAAACCC ACGTCGGGCT GCACCCGTAT GGATATCTGT
TGTATCTACA TCGGTAAATC ATGTGACCAA AACATTCTGT ACACTCGAGG GTC

5 Gene Name: p1

Gene ID: BBF.2017.48.5

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TTGTTTAACT TTTCGAGACC TTAGGAGGTA AAaATATGTC ATTCAATCAG CCTTGGGACT
GTTCACGTCA GTACGATGTC GGACAGGTGG TTGAATATCA GGGCCGTCAC TACGAGGTAG
TACAGGAGCA TCGCTCTCAA TCGGACTGGA ATCCTGAGTC CGCACCTTCG CTTTGGAAGG
AACATCATGG CCATCACGGG CATCATGGAC ATCAACAAGA GGGCCAACAT CACCAAAATA
ACGATAATGA CCGTTATGGA AACAATAATC AATACGGGAG CATCAATCCC TCTCAGAATT
ACAACCCCCC GGCTACAGGG GCGATGTACG GGCAAAACCA AAGTCAACCT GCCTACATGC
AGACGCCTGG GGGCGTCGGA GGAATGAACC AACAATCGAC ATATCAACAG CAAAGCACTT
ATCAACAACA ATCTGCTTAC CAGCAGCAAT CATCATACGA AAAGACGCAA GATTCCGAAA
AATCGGGCAT TGCGAGCATG CTGCCACAAA GTATGCAGGG CATGAGTAAT TTGGAACTGG
GCGGTATTGC CGCGGGTGGG ATTCTGGGAA TCGCCGGGCT TGCCGGGGCC GTGGCATACG
GTGTTGGACA ACATAAGCAG AATAAAAGCG AACATGCGTC TCGCTCTGCG TGGGAATCTG
AAGCCCGTAC CCGTGGCCAA CATTACATGC AGGCTCTTGG TTCTGGACAA CCGGTGCCTC
CGGTAGCTTG GGTGTTGACG GAAGGGTCAA GTATCCCCCA AAACGCGATT CAGGGCGGGA
ATGAGCGCGA TGGCTCTCCA ATCTACATTA CACGCACTTG GCACGAAAAT TCCATTCATG
TAGGGAAGAT GATTCGCTCC TGGAGCAAGG GTTGCAGCAT TAGTTATGGG GGAAAGGAAA
TTAGTAACAT TCCTACTTAT GAGATTCTTT GCGGTGATAT GCGCGCTATT CGTTGGAATG
ACGCAAGTGG TCCGTTGAAT CCTGCCTCAC ACCGCGGGTT GGTGGATGGG GGCCGTGAGG
CGAACGGACA GCCGATCTTC GTGGCGCAAG CGTTTTACGA AGGAGGTGTG CAAATTGGGA

AGGCTGCCAC AGGTAACCAT AATGGGTGTT ATATTGGGTA TGGTGGAAAG GAAATTTGCA TCTCCAACTA CAAAGTTTTA TGTCTTAATT AAATTCTGTA CACTCGAGGG TC

6 Gene Name: p2

Gene ID: BBF.2017.48.6

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TTGTTTAACT TTTCGAGACC TTAGGAGGTA AAaATATGAC GCAAACGGAC AGTGCCTTCC
ATTTATGCTT TTATAAAGAG ATGCGTCGTA TTTACCCACG TACCCGTGCA TTTCCCTGCA
CTGACATCGT TGATGGGTAT GAATACCGCA CATCTTTCGA AAACGGGGCA ATGGTCGTCT
CTCGTCAGAA AACCAACACA GCTGGTTCCT GTGAGATCGT GCTTGATTCT CGCTGGTTAG
GATGGGGTTC ACGCGGTAAG ACTGTGTCGA AAGCCGTCGT TAGTGATGAC CATAACAAGA
TTGCCTTCCT GCAGACCCGC CCCCAGGATG AACAGGGCTC GCTTGTTATT CGTACCTTGA
ACTCGAATGC GCCTGCAATG TTTACTTACA GTCAGTCCAT TCCTAACGTC TTTAACTTCG
TTTGGGCCAA TAACTCCAAA ACGATTTATT TCACTCGCTT GGATGACGTC TTGCGTTCTA
GCAAGGTTTA TCGTATCGAT GTAGACTCTC CCCATGAGGA AACCGTTGTA TATTCCGAAA
TGGACATTTT ATATTTTGTG GATCTTGCCC GTTCCAAGGA TAAGGACTCT CTGTTCTTGT
TGCATAACTA TGCCTCTACT GAGTTAAAGC TTTATCGTAC CTCGTTAAGT ACAGTGTTGA
CGGGGAAGGC CATTAATTTA GAATCTGATA CATCCATTGT TGCTGCCCCT GGATACGAAG
AAAGTATTGA TGATGTTGAA ATCTTTCCTA GCCACGCTGT GCTGACTATG AAACGCCAGG
GTCAGCCCTT CATTCGCAGC ATTGATCTTG TGAACATGCT TGCCAATGAC GTATTCCTTT
CCGACAGCTC AGGTGTAATT AGCCCTGAAC CGAACATTGA TCTGACAAGT AACACCTTTC
GCTTCTCGTA TTCGAGCCCC TTCGTAATTC AAGCCGTGTA TGAGTTAAAC TTGGTCCTGG
GGAAGCTGAT CAATTGGCAA GACATCCGCC CTTGCATGGA CACTAACAAT TTCGTTATTG
AAAAACAGAC CATTCCCTCG CGTGACGGAA CTTGTAATAT CCCTATTACG ATTTTGAAAA

ACAAAAACAT TTCTCATAGT CCTTCTAACC CCTGCGTTAT TCTTGTATAC GGCTCATACG GATACTGTAT TGAGACGAGT TTCCGCGTTG AACTGCTGCC CGCTCTTAGC CGTGGTGTGA CCGTCGTGTT CGCTCATGTC CGTGGCGGAG GAGAACTGGG AAATAGTTGG TATCACAAAG GGCGTGCCGA GAACAAGCTG TGCAGTTTCA CTGATCTTGA AGACGTAGCT GACATGTTAA TCTATAAGAA ATACACCAGC TCAGAACGCT TAGGCGGGAT TGGTACCTCA GCAGGAGGTA TGTTACTGGC GGGGACGGCC AATCGTCGCC CCGGAAAGTA TTTGTGAATT CTGTACACTC GAGGGTC

7 Gene Name: p3f1

Gene ID: BBF.2017.48.7

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TTGTTTAACT TTTCGAGACC TTAGGAGGTA AAaATATGCA CAAACTTTCA AACCCCGTCG
TGAGTCCCAA AGGTCATATG ACGGCATATA CCGTTAGCCA CTACAACACT TCTACGACCA
TTACAACCCG CTCGCTTTAT CTGTCACACC TGCCGCAAAG TGAATCAGAG CACCCCAATT
TATCGACACC TGAGACCGCC TTGGCCACAG GTCCCGTCTT GGTTGGCTCA GACGTGGGTT
CCGACCCCTT CTGGATGTCA GACTCCGTGC TGGCAGTATT ATGTAGCCGC GAGGGATTAA
ATCAGATCTG CGTCTCAAAC ATTACCAATG CGTTAGAGGG TACACCCGTT GAATTCAAGC
AAATCACACA GCTTGTCCAC GGTGTGACGA CAGCTAAATA CCACACTGCA TCAGGCTTAT
TAGCCTTCAC AAGCACAGTC GTTGTTGCTC CAAACCCGCG TAAGGTAAAG GGGAGTTCTG
CCATGGTGTA CGATAAATTA TTTGTCCGCC ATTGGGACAC CTATTTAGAT CCCACAGCTC
GCAGTCACCT GTTCGTCGCA CGCCTTACCC TTGACGCTTT ACATACGACG GTAGTAAAGG
AGGGTGACGT GGTCGACTTA TTTACTTCAC AACCAGATTT GGAGACACCT GTAGCTCCTT
TCGGCTCGAC TACGGACTAC GACTTTTCCC CGGATGGCAC AACGATCGTC TTCACTTCTC
GCGTCCCAGA GCGCAAAGCA GCCTGGAACA CCAACACCGA TGTTTACACC ATTTTGACCG
ATGGAACCCA ACCTCCTGTC AGCCTGAGCA AAAAAAACCT GGGGGCTGAT TCTCTGCCGA
CATTCTCCCC CTCTGGAGAT TTGATCGCTT GGCTGCAGAT GCGTGAGCCA GGCTATGAGT
CTGACAAAAA TGTAATCGTT GTATATAATC GCATTAAAGC AACCCAGCAC ATTCTTACAA AAAAG

8 Gene Name: p3f2

Gene ID: BBF.2017.48.8

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

AGCACATTCT TACAAAAAAG CAGCACATTC TTACAAAAAA GTGGGATCGC AGCCCGAGTT
CAGTGGCGTT TTTGAGCGAA ACCTTATTAG TGATTACGGC CCAGGATAAG GGCCACAATC
GCGTGTTCAC TGTTGATATC ACAACGGACA AAGTCGTTGA GCGCTATTCA GACTCTACCC
TGCATTCCAT TTCCGTCGTC CCCGGAACGA ATACGTTTAC TGCGCTGTCG AATTCCATGT
TATATCCCGA CAAATTAGTA GTCATGGACA TGGAGAAATG GACTGTTGCG TTCTTGACTC
ACTTCAACGA CGAACACTTA GCTTCACGCG AATTGGCTCT GTCTGAGGAA TTTTGGTTTA
AGGGGGCGAA GAATGAAAAA GTTCATGGGT GGTTGCTGAA ACCGCATTAT TTTAATAAAG
ATAAGAAATA CCCCCTGGCA TTTCTTATCC ATGGGGGTCC GGAGGGAGCT TGGAACGACG
ACTGGTCCTA CCGCTGGAAT CCACAGGTTT TCGCGGCACA AGGTTACTTC GTTGCTGTCA
TCAATTTCCA TGGAAGCACC GGCTTCGGAA GTGAGTTCAC TCGTAGCATT TTGGGTAATT
GGGGCGGCGC ACCGTTTGAA GACAATATGA AGGGTTTGAA GTATGTATTA GAGCAAAACC
CGCAAATCGA CCACCGTCGC GTAGCTGCCC TTGGCGCGAG CTATGGAGGA TTCATGATTA
ATTGGATTAA TGGGCATAGC AAAAAATTCG CATGCTTGGT GAACCACGAT GGAATCTTTG
ACCCTCGTTC GGCTTACTTT AGTACTGAGG AATTGTTTTT CAAGGAGAGC GAATTCAAAG
GGACGCCATA TGACCGTAAA GCTGCTAAAT TATACGAAAA GTATTCGCCA GCGCTGTATT
CCCATAAGTG GGTCACCCCC ACCTTGGTCA TCCACTCCGA ACGTGACTAC CGTTTGCCTT
ATCATTGAAT TCTGTACACT CGAGGGTC

9 Gene Name: c1f1

Gene ID: BBF.2017.48.9

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TTGTTTAACT TTTCGAGACC TTAGGAGGTA AAaATATGAC GCCTATTGCA GACTCATCTT
TAGACCCAAC CAGTGCCCGT CTGAGCAAGG ATCGTCTTTT TGCATCACTT AAGCCCTTAT
GCATTCAGTG CATGGATCCT ATTCTGTGCT GGAAATGGAA TCCCAGCTCT ACACGCCCAA
CTTCTTCACA CCCAGTGGCC ATGCTTCTTG CCGAGCTTGA TTCGCAGATT TCAGGACTGA
GCGTCACCGA TGCCGCTGAC CTTGTTGCCC CACTGCTGGC GTACATCAGT TTCCCCTTGG
TGGAAGCAAT TAAGTCACGC ATGGCAACGA GCACGACCAC CGAGTGCGTC TGCCGCTGTT
TGGCATTCTT ATTACAGACA AGTGCCCCAT CCGTGAAGTA CTTTCAACCC CCCCCTGCTT
TAGTGAGCCA ATTACTGCTT GAATTGCCAC TGATTATCTT TAGCCCATCA CTGATCAATA ATAAG-
CAAAC CTTGCCCCAA AACCCGCGTC AGTCTCTGGC CCTGACGGCC GTACGCCCGT CG-
GAGGATCT TAAACTTTGC GTTGTCAAAG CACTTAGCGC TCTGTTCGCC GGGCAGCAGG
CCATGCAATC CCGTCCGGAA TGCACAACAA GCATTGATTT CACGTCCGCT CAATTGCAAC
CAATTTTGGC ACATAACATG ATTCTGTTGC TTGACTTGTT GGCTACTGAG AAATCAATTG
AGCTGTGTGG GCATGCGCTT AACGCCCTGG CAGATCTGGT CCGTCTTATC TCGGTCTCGC
CTGACGCGTT AGCTCAATTC TTACCAGGAA TTGTCAGTGG TATTGTTAAG TACATTCTGA
TTGGTGACGA AAAGCGCCAT CATTCTCTTT TCGTCGGTTG CTTCGTCGTA CTGGAGTTGT
CGATCCAAGG GACCCTTTCC GACACGATGA GCACCTCCTT CATGGTAAAA CAGGACACTG
CCAAGGAGAT CCTTGAGCAG ATTGCATCAA AAGTAGACTT GAATCAATTG AAGAGCCTTA
CGTTACAAGA TGACTCAGCA TCCAGCAAGG AGGACAAATC GAAATCACAA CTTAAGCATA

TCGCCCCTCG TACAGCCAAC TGGTACAAAG TTAACAGCAA TAACATCTCG GCCATTTTAC CCAAATTAGC AAAGTTGCGC TTGCATCAGA ATCCCATCGT CCAGACAGCG TTGTTAGATT TTTCAACGCA ATTGTTTATC AGTTGCACGC AGTCTTTGTC AGGAACGGTG AGTTTTTTTG TGGACAATAT TATTATTTTA TGTTCTCTGT TGCAATCCGA GACGTTAAAC AACAAAAGTA AGTCTGGTAA TGATAGCCTG CACCTTGCAC AACACCAACT TAACACCTTG CAGACTACGC TTCAGTCCTC GATGGCAACA ACACGCGTGT TAACA

10 Gene Name: c1f2

Gene ID: BBF.2017.48.10

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

CAACAACACG CGTGTTAACA GCAACAACAC GCGTGTTAAC AGACCGCTTG TCGCAACTTC
TTGATCATCT TCCAACCCAG CTGACATCCA CGTCGGAGAC AATCCAATTG GAAACACTGT
TACTTATTAA CGGTTACATT CGCCTTTTAG AAATTAAGGC AGCCCCTATT ATTCAAGGGT
CCATTAATCA AGTCGTAAAA GCACTGCTTT CTGTGTTAGT TTTCGAAACA AATGGTGTAC
GTATGGTTGA GTCTCGCCAT CGCACTTTAG GAGCAGCTCT GACCCTTATG GAGAGTGTCT
CAGCAAGCAA CGCATTTGAC CAGGATATGT CTATTCCCGC TCCGCAACGC CATTTTCAGA
CGTTTTATGA CGATCGCATT TTAACCCAAA TCTGCCGCTT GTGCCGCCTG TTAGCGTACT
ACGGCGACGC CGAGATGTTG GTGGATGTTT TCAGCACGAA ATTAAACGTC GAACGTACGC
AAACCAGCGA AACGATCGAC CCGTTACACC AACAAGTGTC ACAGCAGCAT TTCTCGGCTG
CCTTATTTGT ACTGGGGCAA ATCGCCCTGG GTACATTGGG AGGGGGACTG AAGGACGACG
GAATCGACGT GGCTGAGTCC AATAGCACCA CACGTCTGTC CGCGGTTCGT CTGGCGCGCT
CCGTAATCGC GGAATTCCTG GACTTGTCTA TTTTACACAT GCCGACAAAT ATTAGCGAAA
AACGTCTTTT TTTAGAGCCA GAGGCCCCTA CACACTTGAT CAAATCCAAT CCAACTTACT
CTCGCAACTT ACAGTTACAG CCCACAGTTC GTGACTTTAA TCATGTGATC ATTTGTACAA
GCTTGGTGTT AGAGGGCCTT GGCCGCTTAG CTGACGTCCT GACTCCAGAT GAGGCGACAT
TGATCCTTAT TGATGGCTTG TACCCTATCC TTGAGAAGCT TGGGGATCGC AGCCATGCGG
TCTCCAATGC AGCAGGCTGG GCGCTGCAAC AATTTGCGCA GACGTGCTCT CGCTCCCCAT
TAGCGTTCCC GTCAGTTGTG CTTCCGTCAA TTGATTTTGA AAAAAAGGAT AACGAGACGC

TGCCTGCACA AATTGGAACG GTATGTAAGT TAGTGCTGAG TCACATCGAT TACCTGGTGG ACTCGGTGTC TCAGCGCATC CGCCACATCC GCCATTTCCC TATGGCTCCT AAAGTATTGA CAGCCGCCAT TCGTGTGACC GGAGTCGAGA TCGTTGGGCC TTTACTGACG GACTGCGTAG ACCTGCTTAT CGACGTATTG GACGAATTGG GCACATCTGG ACGCGGAGGG GACCTTCTGT TACAGGAGTT AGATAGTGCA AACCATATTA AGTTCAATAA CGATTCTAAA ATTACTCATC GCCATGCCAT CGGTTTGTTG GATAGCGACG ATCCAGGGCT G

11 Gene Name: c1f3

Gene ID: BBF.2017.48.11

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

ATAGCGACGA TCCAGGGCTG GATAGCGACG ATCCAGGGCT GGTCGGTGAG GTCATCAGCG TTTTATACGC ACTTATCGAA GTCATGGCAA ACGCACCGGA GTCTAAGCGT GCAGCTCAGG ACGTGCTGGC CAAACGTAGT GGAAATTTGC TGGGGAATGT CAATACATTC CGTGCTCTTA ACGATAAGAG TAAATTAATC CAGGATACTG ATATGCACAA AGTGGATGAT ATTTTAGGGA TCCAACATTG TTCAGCCGAA ATTCGCGCAT ACTTTATCGA GCGCCGTAAA CCAGTAGACG ATGATGCGGA CAAAAACGAT AGCTTATCGG CAAAGGAATT CTTCGAGCAG CGCCTGAAAG AAGCCAAACC AAATCGCGCC TCTATTGGTA ACGAGATGGC CGACGAATTG AGCGAACCTA CCCAAAAGCC CGATGCCGAC CCCATCCCTC CCACCCCATT CGAGAGCTTA TCGCATCAGA TCTTGTCTAA CAGCGTGCAT TTTTTAAGTT CCGACAGTCC CCTTATTCGC GTGCGTGTTT TACGTATGCT GCGTATGGGC ATCTCCTTGC TGTCAAACCG CCCTAGTGAT CTTAATCCCT TAATCCATAT TATTTGGCCT AAGCTTGTGT CTCGCGTCAA GGACACCCTT CACTACGTAG CGTTAGAAGC TGTTTTGGTA ATTGGTGCGG TGGTCAACGT AAGCCCCGAA TTTGTCCGTA AGCGCGTATC TGACGACATC ATTCCAATTT ACGCACAGCT GTTCAAAAGT TTGCAGACAA AAGCGTCGAG TGCTATTATG GGGAAGTCTA AAGCAAACGA GAGCGTCGCC AACGCTCGTT CTTCAGTCAA ACGCGACGTT GTGTCCCTTC TGCATCCGAC TTCTATCACC TCGGCTCGTA ATCATATCTT CGAAAGCATC GATACCACGA CTGTGGAGAC TCGCTTGTTT ATGGCGTGCC TTAGCACATT GGAGTCTATG TTTGATTCAA TGATGTTACT GCGTCACGAC ATCGATGTTC TGGTAACAGC ATTAATTGGA TTGTTGAACT TTCGCGTATA CTCCGTCGAT GTGCGCAACG CAGTGGTTGA

CCTTATTGGA CGCTTCATTC GTGGGGAAAA TAAAATGCCG GGCAACAACG GGGGCAATGA
CTGGGTTTGG GCTGTTGTTT GGATCACGGC AGGGGCTTCT CAGCTTGAAT TAACCGGGGG
CTGCCAGGAC AAAGATGTGT TACTGAAGTC CTTGTCAGTT ACGAATGGTA CCCGTGAAGC
GTTCGCGCCT CCAGGCACGT ATTCTGAATT TGATTCTGAA CGTTTGATCC AGGTGCTTGA
TATGTGCGGA GAGCCGGGTG TCTCAGTAGT TCTGAATGAT AAGTGGGCGG AATGCGACCT
TGCTTTCGGA GGACGTTGGA CGCAATCTAC AATTCTGTAC ACTCGAGGGT C

12 Gene Name: c2f1

Gene ID: BBF.2017.48.12

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TTGTTTAACT TTTCGAGACC TTAGGAGGTA AAaATATGGT TGAATCGCGC CATCGCACTC
TGGGGGCTGC ACTGACTTTG ATGGAGTCCG TTAGTGCGAG TAACGCATTT GATCAGGACA
TGTCCATCCC AGCGCCGCAG CGCCACTTTC AAACATTTTA CGACGACCGC ATTCTGACTC
AAATTTGTCG TCTGTGCCGT CTGCTGGCCT ATTACGGCGA CGCGGAAATG CTGGTTGATG
TTTTTAGCAC TAAATTGAAC GTGGAGCGTA CGCAAACGAG TGAGACGATT GATCCTCTGC
ATCAGCAGGT TTCACAACAA CACTTCAGTG CTGCGCTGTT CGTTTTAGGC CAGATTGCAC
TGGGGACCTT AGGGGGCGGC TTGAAGGACG ATGGAATCGA TGTCGCGGAG TCTAACTCGA
CCACCCGCTT ATCCGCCGTA CGTCTGGCAC GCTCTGTAAT CGCGGAGTTC CTGGATCTGT
CTATCTTACA TATGCCGACG AATATTTCGG AGAAACGCTT ATTCCTTGAG CCCGAGGCAC
CTACTCACCT TATTAAATCT AATCCGACCT ATTCCCGCAA TTTACAGCTT CAACCCACTG
TACGTGATTT CAACCACGTT ATCATTTGCA CATCCTTAGT CCTTGAAGGT TTAGGGCGTT
TAGCAGACGT GCTTACTCCG GATGAGGCCA CATTGATTCT TATCGATGGT TTGTACCCAA
TTTTAGAAAA GCTGGGTGAT CGTAGCCATG CAGTTAGCAA CGCCGCGGGT TGGGCTTTAC
AACAGTTCGC TCAGACCTGT AGTCGTTCAC CACTTGCATT TCCCAGCGTT GTACTTCCAT
CGATTGATTT TGAGAAGAAG GATAATGAGA CCCTGCCAGC CCAAATTGGA ACGGTCTGTA
AACTTGTATT GTCTCACATC GACTATCTGG TAGATTCGGT TTCGCAACGT ATTCGCCACA
TCCGTCATTT TCCGATGGCC CCGAAAGTAT TGACAGCGGC TATTCGCGTT ACTGGAGTAG

AGATCGTAGG TCCATTGCTT ACGGATTGCG TTGATTTACT TATTGACGTA TTAGATGAAT TGGGGACTTC AGGAC

13 Gene Name: c2f2

Gene ID: BBF.2017.48.13

Submitted by Adarsh Ambati on 2017.11.20. Contact at adarsh.ambati1@gmail.com.

Description

NA

Motivation

NA

References

NA

Sequence

TGAATTGGGG ACTTCAGGAC ATGAATTGGG GACTTCAGGA CGCGGTGGTG ACCTTCTGCT GCAAGAACTG GACTCCGCTA ATCACATTAA GTTCAATAAT GATTCAAAAA TCACACACCG CCATGCGATT GGGCTGTTGG ACAGTGATGA CCCTGGTTTA GTGGGTGAAG TAATCAGTGT GTTGTACGCT TTGATCGAGG TAATGGCAAA CGCACCGGAA AGCAAGCGTG CAGCCCAAGA TGTGCTGGCC AAGCGTAGTG GAAACCTGCT GGGCAACGTG AATACGTTTC GCGCTCTGAA TGACAAAAGC AAGCTGATTC AGGATACGGA CATGCACAAG GTGGACGACA TCTTAGGCAT TCAACACTGC AGTGCGGAAA TTCGCGCATA TTTTATCGAA CGTCGTAAGC CGGTTGACGA TGATGCAGAC AAAAACGACT CGCTTTCTGC TAAGGAGTTC TTCGAGCAGC GCTTAAAGGA GGCGAAGCCA AACCGCGCTT CAATCGGTAA CGAAATGGCA GACGAATTAT CTGAACCTAC TCAGAAACCA GACGCAGATC CGATTCCGCC AACTCCATTC GAATCGCTGA GCCACCAAAT TCTGTCGAAC AGTGTTCACT TTCTGTCCAG TGATAGCCCG CTGATTCGTG TTCGTGTTCT GCGTATGCTT CGTATGGGAA TCAGTTTACT TTCGAATCGC CCCTCAGATC TTAATCCACT TATTCACATT ATCTGGCCAA AGTTAGTCTC ACGTGTTAAA GACACTCTGC ATTACGTGGC ACTGGAGGCT GTGCTGGTTA TCGGCGCTGT TGTGAACGTG AGTCCCGAGT TCGTTCGCAA ACGCGTGTCC GACGACATCA TCCCCATTTA TGCACAGCTT TTCAAGTCTC TGCAGACGAA GGCTTCAAGT GCCATTATGG GCAAGTCCAA GGCCAACGAA TCCGTCGCGA ATGCTCGCTC GTCCGTAAAG CGTGACGTAG TTTCGCTTCT GCACCCGACA AGTATTACTT CAGCACGTAA

TCACATCTTC GAATCTATTG ATACCACCAC TGTGGAAACG CGCCTTTTTA TGGCGTGTTT
GTCCACACTT GAGAACGCCC GCAAGATTCT GTACACTCGA GGGTC

14 Gene Name: ACYX2008404

Gene ID: BBF.2017.48.14

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgC TGTTAGCGTT TCTTGTCTTA GTATTCTTCT CAGGGGGCTC CTGGGCCGCG AACAACAGTC TAAATGGGGA TTTCTTGCAA TGCATTAAAA AAAACGAGTA TTCTAGCATA CCGATTCCTA TCTTTACACC GGATAACTCT TCATTCACGA CGATATTTCG TTCATCCGCC AGGAACCTTC GTTTTCTTAC TCCAAACTCT ACACAAACTC CACAGTTTAT AATTACACCA ACGCATGAGA GCCATGTTCA ATCCGCTGTA GTGTGTTCCC AAAAGCATGG ATTTGATCTA AAAGTGCGTA GTGGCGGGCA CGACGTGGAG GGGCTGTCTT ATGTGTCCGA TACGCCGTAT GTCCTAGTTG ATCTAATAAA CTTCAGAGAC ATTACTGTGG ACCTGAAAGA AAAGACTGCA TGGATAGAGG CTGGCGCGAG CTTAGGAGAA GTTTACTACC AGGCGGCCAA CAAGTCAAAT AACACGTTAG GTTTCCCGGC TGGTTTTTGC CCTACTGTAG GAGTTGCGGG GCATATTTCC GGAGGCGGCT TCGGCGCACT AGTGCGTAAA TATGGACTTG CGTCAGACCA AGTCATTGAC GCGCGTATCG TAACTGTTGA TGGTAAGATT TACACAAAAG AGACGATGGG AAAAGACCTA TATTGGGCAA TTAGGGGGGG AGGCGCTAAC AACTTCGGGG TCCTACTTAG CTGGAAGGTA AAACTAGTTC CCGTAACGCC AGTGGTTACT GTTGCTACCA TCAGTCGTAC ACTAGAACAA GGCGCAACGG ATCTAGTTCA CAAATGGCAA TTTGTCGCAG ATAGACTTCA CGAGGACGTC TACATCGGCT TGACATTCAG CGTTGCCAAC TCCTCAAGGG CCGGTGGGAA GACTGTTCTT GCACAGTTCG CTTTCCTGTT CCTTGGCGGT AGCGACAGAC TGCTAGAGCT GATGGAGGAA TCTTTTCCCG AGCTTGGTCT GAAGCGTAAT GATACGACCG AGATGAAGTG GGTCGAATCC

CACGTTTACT TCTACGCGAG AGGCAGACCA ATTGAGTTAT TGTGGGATAG GGATCATGCC ACTAAGAGTT TTCTGAAAAT TAAAGCGGAC TACGTCAGGG AACCAATCAG TAAGTTCGGG CTGGAAGCCA TCTGGAGAAG GTTTGTCGGT GGAGACTCTC CGGCCATGTT ATGGACCCCC TTCGGTGGGA GGATGAACGA TATTAGCGAA TTTGAAACAC CCTACCCGCA TAGAGCAGGC AACATATACA ACATTATGTA TGTAGGCAAT TGGATGAATG AAACTGAGTC AGAGAAACAG ATCGATTGGA TGCGTCGTTT CTATAATAGT ATGGCACGTT ACGTGTCTAA AAATCCTCGT TCTGCGTATG TAAACTACAA AGATCTAGAC CTTGGCGTAA ACAGTAATAA TGTGTCTGAA GCTATGAGAT ATGTACAAGC CAGGTCATGG GGAACTAAAT ACTTTAAACA TAACTTTGAG CGTCTAGTGA AAGTGAAAAG TATGGTTGAT CCAAGTAACT TTTTCAAAAA CAAACAGAGT ATTCCCCCGG TAGGCACATG Gtaagagagg gcgcgccact tctaaataag cgaattGATC GAATCTCCTC TTAGCCGATT CGATTATCGC ATTCGGGACT GAT

15 Gene Name: YDEH2011686

Gene ID: BBF.2017.48.15

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgA AGACTGCTTG GATACAGGCC GGAGCAAGTT TGGGCGAGGT CTACTACCAG GCGGCGAACA AGAGCAATAA CACCTTGGGT TTCCCGGCGG GGTTTTGCCC AACGGTCGGA GTCGCGGGTC ACATTTCAGG TGGGGGTTTC GGCGCGTTGG TCAGGAAGTA CGGACTGGCC TCAGACCAAG TAATCGACGC ACGTATTGTA ACGGTCGATG GCAAAATTTA CACTAAGGAA ACTATGGGAA AAGATCTTTA CTGGGCTATT AGGGGCGGTG GCGCGAATAA TTTCGGGGTC CTGCTAAGTT GGAAGGTCAA GCTAGTGCCT GTTACCCCTG TCGTAACAGT CGCGACTATT TCTAGAACCT TGGAACAAGG AGCGACGGAT CTTGTTCATA AGTGGCAGTT TGTGGCCGAT AGATTACACG AGGATGTGTA TATCGGTCTA ACGTTTTCCG TAGCAAACAG TAGTAGAGCC GGGGGCAAGA CGGTAAGCGT CCAGTTCGCG TTTCTATTTT TGGGTGGGTC CGACCGTCTA CTGGAGCTTA TGGAAGAAAG TTTCCCTGAG CTGGGCCTAA AGAGAAACGA GACCACTGAA ATGAAATGGG TTGAATCACA CGTCTACTTT TATGCGCGTG GGCGTCCTAT CGAACTGTTA TGGGACAGGG ATCACGCCAC GAAGTCTTTT TTGAAGATCA AGGCGGACTA CGTAAGAGAG CCTATATCAA AGTCAGGTCT GGAGGCTATA TGGAGGCGTT TCGTTGGTGG CGACTCACCC GCAATGTTAT GGACACCCTT CGGGGGTAGA ATGAATGAAA TTAGCGAGTT TGAGACCCCG TACCCTCACA GGGCGGGGAA CATTTACAAC ATTATGTATG TAGGTAATTG GATGAACGAG ACTGAATCCG AGAAGCAAAT TGACTGGATG CGTCGTTTTT ACAACTCTAT GGCGAGATAT GTGAGTAAAA ATCCGAGATC AGCCTACATC AATTATAAAG ACCTAGACCT TGGCGTAAAC

AGGAATAATG TTAGCGAAGC AGTAGGCTAT GTACAAGCTA GAAGCTGGGG GAGAAAGTAC
TTCAAGAGCA ATTTTGAGAG GTTAGTGAAA GTCAAGTCAA TGGTGGACCC AGGTAATTTC
TTCAAGAACA AGCAATCAAT CCCTCCAGTT AGCACCTGGt aagagagggc gcgccacttc taaataagcg
aattGATCGA ATCTCCTCTT AGCCGATTCG ATTATCGCAT TCGGGACTGA T

16 Gene Name: JSVC2033447

Gene ID: BBF.2017.48.16

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgA TTACACCCAC GCACGAATCT CATGTGCAGT CTGCCGTGTT CTGCTGCAAG AAGCACGGGT TCGACTTGAA AGTCAGAAGC GGAGGCCATG ATGTGGAAGG ACTATCCTAC GTTTCCGACA CCCCGTACGT CCTGGTAGAT TTAATAAACT TCAGAGATAT AACTGTAGAT CTAAAAGAGA AAACCGCCTG GATACAAGCA GGAGCTTCAT TGGGAGAAGT GTATTACCAG AGTGCTAACA AGAGTAATAA CACATTGGGG TTCCCTGCGG GATTCTGCCC GACCGTTGGT GTTGCGGGGC ATATCTCTGG TGGGGGATTC GGTGCCTTAG TGAGGAAATA CGGCCTTGCC AGCGATCAAG TTATCGATGC TAGAATAGTA ACTGTTGATG GCAAGATCTA TACACGTGAA ACTATGGGTA AAGACTTGTA TTGGGCCATT AGAGGAGGGG GCGCAAACAA TTTCGGCGTG CTGTTGTCCT GGAAGGTGAA ACTGGTCCCG GTTACTCCCG TTGTCACTGT TGCAACAATC TCTAGAGAAC AAGGCGCGAC AGATTTAGTA CATAAGTGGC AATTTGTTGC AGACAGACTT CACGAAGACG TGTATATCGG CCTTACTTTT AGCGTTGCTA ACAGCAGTCG TGCTGGTGGG AAGACGGTAA GTGTACAATT TGCCTTCCTA TTCTTAGGTG GCTCTGACAG GCTATTGGAG TTAATGGAAG AAAGCTTTCC TGAGTTGGGC TTAAAACGTA ATGAAACCAC CGAGATGAAG TGGGTGGAGA GTCATGTCTA CTTTTACGCT AGAGGAAGAC CTATAGAACT ACTATGGGAC AGAGACAACG CGACTAAGAG TTTCCTTAAG ATTAAGGCAG ATTATGTCAG GCAGCCCATA TCTAAGTCAG GGCTTGAAGC GATTTGGCGT AGATTTGTTG GCGGGGACTC ACCAGCAATG CTGTGGACGC CATTCGGAGG GCGTATGAAT GAGATCAGCG AATTTGAAAC ACCCTATCCA

CACAGGGCTG GGAACATTTA TAACATAATG TATGTAGGAA ATTGGCTGAA TGAAACTGAA TCTGAGAAGC AGATTGATTG GATGAGGCGT TTCTACAACT CAATGGCAAG ATACGTCAGT AAAAATCCTA GATCTGCTTA CATAAACTAT AAGGATCTAG ATTTGGGGGT AAATAGTAAT AATGTCTCAG AGGCCGTCCG TTACGTACAA GCGAGAAGTT GGGGTACCAA GTACTTTAAA CACAACTTTG AACGTCTAGT TAAAGTTAAG AGCATGGTCG ATCCCTCCAA TTTCTTCAAG AATAAGCAAT CCATACCCCC TATTTCTACC TGGtaagaga gggcgcgcca cttctaaata agcgaattGA TCGAATCTCC TCTTAGCCGA TTCGATTATC GCATTCGGGA CTGAT

17 Gene Name: RRID2024547

Gene ID: BBF.2017.48.17

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgA TAAGCCAAAA CCAGAACTTC TCAATACCGA TACCCATTTA TACGCCGAAT AACAGCAGTT TTACGACGTT ATTCCGTAGC ACTGCGCGTA ACTTGAGATT TCTGTCACCC AATAGCACTA AGCAGCCTCA GTTCATTATA ACACCAACCC ACGAGTCCCA TGTTCAGTCT ACTGTAATCT GTAGCAGGAA GCATGGTTTC GACCTTATCG TTAGATCTGG GGGGCATGAC GTCGAGGGCT TATCCTACGT TTCAGATGTG CCCTATGTTA TGGTAGATCT GGTGGACTTT CGTAATATTA GTGTAGACCT AAAGGATAAT ACAGCCTGGA TTCAGGCAGG TGCCAGTCTG GGAGAAGTCT ATTACCAGGT CGCCAACAAG TCTAAGACGC TGGGGTTTCC TGCGGGGTTC TGTCCCACGG TAGGAGTTGG AGGGCATATA TCAGGAGGAG GTGTTGGAGC GCTGGTGCGT AAATACGGGC TATCATCCGA CCAGGTCATC GACGCTAACA TTGTCACCGT CGAGGGTAAG ATACTAAATA AAGAAACGAT GGGGAAGGAC TTGTACTGGG CGATTAGGGG CGGAGGAGCG TCAAACTTCG GAGTAGTTCT ATCATGGAAA GTGAAACTTG TGCCTGTGAC CCCGATTGTC ACGGTGGCGA CCATCGGAAG AACGCTGGAG CAGAACGCAA CAGACCTTGT GCATAAGTGG CAATTTATTG CAAACCGTCT GCATGAGGAT GTTTACATTG GGGTGACCTT CACGGTAGGA AACTCCTCCC GTGTAGGTGA GAAAACAGTA TTGGCCCAAT TCAGCTTTAT GTTTTTGGGG GGAGCGGATA GGCTTTTGGA GCTAATGGAT GAGAGCTTTC CTGAACTGGG GTTAAAGAGG AATGATACTA TCGAAATGAG TTGGGTTGAG AGCCACGTGT ACTTCTATGC AAGAGGTAGG CCAATTGAGT TACTGTGGGA TCGTGACCAT GCAACAAAAA TATTCTTGAA GGTCAAAGCA

GATTACGTCC GTGAACCTAT TAGCAAGGCT GGCCTGGAAG CCATATGGAG GCGTTTTTTA GGAGGTGATA GCGCTGCAAT GTTATGGACG CCTTTCGGTG CGAGAATGAA CGAAATTTCC GAATTTGAAA CCCCCTACCC TCACAGAGCT GGCAATATTT ACAACATTAT GTATGTGGGC AATTGGCTGA ACGAGACCGA CTCAGAAAAG CAGATTGACT GGATGAGACG TTTCTATACA TCTATGGGTC GTTATGTTAG TAAAAACCCG AGGTCCGCAT ATTTAAATTA TAAAGACCTG GACTTGGGAA CCAATTCAGA TGACAACGTG TCCGAATGGA TGAGGTACGT AAAAGCCAGG AGCTGGGGAA GGAAATATTT TAAATCAAAC TTCGAGAGAC TAGTAAAAGT TAAATCCGAA GTAGACCCCC ATAATTTTTT TAAGAACAAG CAGTCTATCC CACCGATTAG TAGCTGGCGT AAAAATtaag agagggcgcg ccacttctaa ataagcgaat tGATCGAATC TCCTCTTAGC CGATTCGATT ATCGCATTCG GGACTGAT

18 Gene Name: MIKW2005234

Gene ID: BBF.2017.48.18

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgA AGACTGCTTG GATACAGGCC GGAGCAAGTT TGGGCGAGGT CTACTACCAG GCGGCGAACA AGAGCAATAA CACCTTGGGT TTCCCGGCGG GGTTTTGCCC AACGGTCGGA GTCGCGGGTC ACATTTCAGG TGGGGGTTTC GGCGCGTTGG TCAGGAAGTA CGGACTGGCC TCAGACCAAG TAATCGACGC ACGTATTGTA ACGGTCGATG GCAAAATTTA CACTAAGGAA ACTATGGGAA AAGATCTTTA CTGGGCTATT AGGGGCGGTG GCGCGAATAA TTTCGGGGTC CTGCTAAGTT GGAAGGTCAA GCTAGTGCCT GTTACCCCTG TCGTAACAGT CGCGACTATT TCTAGAACCT TGGAACAAGG AGCGACGGAT CTTGTTCATA AGTGGCAGTT TGTGGCCGAT AGATTACACG AGGATGTGTA TATCGGTCTA ACGTTTTCCG TAGCAAACAG TAGTAGAGCC GGGGGCAAGA CGGTAAGCGT CCAGTTCGCG TTTCTATTTT TGGGTGGGTC CGACCGTCTA CTGGAGCTTA TGGAAGAAAG TTTCCCTGAG CTGGGCCTAA AGAGAAACGA GACCACTGAA ATGAAATGGG TTGAATCACA CGTCTACTTT TATGCGCGTG GGCGTCCTAT CGAACTGTTA TGGGACAGGG ATCACGCCAC GAAGTCTTTT TTGAAGATCA AGGCGGACTA CGTAAGAGAG CCTATATCAA AGTCAGGTCT GGAGGCTATA TGGAGGCGTT TCGTTGGTGG CGACTCACCC GCAATGTTAT GGACACCCTT CGGGGGTAGA ATGAATGAAA TTAGCGAGTT TGAGACCCCG TACCCTCACA GGGCGGGGAA CATTTACAAC ATTATGTATG TAGGTAATTG GATGAACGAG ACTGAATCCG AGAAGCAAAT TGACTGGATG CGTCGTTTTT ACAACTCTAT GGCGAGATAT GTGAGTAAAA ATCCGAGATC AGCCTACATC AATTATAAAG ACCTAGACCT TGGCGTAAAC

AGGAATAATG TTAGCGAAGC AGTAGGCTAT GTACAAGCTA GAAGCTGGGG GAGAAAGTAC
TTCAAGAGCA ATTTTGAGAG GTTAGTGAAA GTCAAGTCAA TGGTGGACCC AGGTAATTTC
TTCAAGAACA AGCAATCAAT CCCTCCAGTT AGCACCTGGt aagagagggc gcgccacttc taaataagcg
aattGATCGA ATCTCCTCTT AGCCGATTCG ATTATCGCAT TCGGGACTGA T

19 Gene Name: NJKC2059279

Gene ID: BBF.2017.48.19

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgA TTATCACTCC CAGTAATGAG TCTCACATAC AAGCCTCCGT CATATGTGCT CGTAAGCATA GGTTCGATTT GAAGGTAAGA AGCGGCGGGC ATGATGTCGA GGGTCTGAGT TACGCTTCAG ATGTCCCCTT TATCCTGGTT GACTTGATAG GTTTTCGTAA CATTACAATC GACGTACAAA ACAAGACCGC TTGGGTTCAA GCGGGAGCCA GTTTGGGGGA GGTTTATTAT GCCGTCGGTA ATAAGAGTAA AACCCTGGGA TTTCCTGCCG GGTTCTGTCC GACAGTCGGG GTAGGTGGCC ACATTTCCGG AGGGGGGTTT GGTGCGTTGG TTAGAAAGTA CGGCTTAGCT TCAGATCAGG TGATCGACGC CATCTTCATC AACGCTGAGG GTAAGATATT CGATAGAGAA ACAATGGGCA AAGATCTTTT TTGGGCGATT CGTGGAGGTG GAGCCTCTTC CTTCGGGGTG GTGGTAGCGT GGAAGGTTAA ATTGGTAGAG GTAACTCCCA TTGTGACGGT GGCTACGGTT AATAGACCGC TTGAACAGGG TGCTACCGAC TTGGTTCATA GGTGGCAGTT CGTAGCGGAC AGGTTGGATG ATGATATATA CATCGGCCTA ACAATGGTGG TGGCAAACGG CTCTCGTCCG GGTCAAAAAA CTGTCATAGC GCAATTTTCA TTTATGTTTC TTGGCCAAAC GGACAGACTT CTGGCTCTAA TGGAAGAGTC ATTCCCGGAA CTGGGCCTTA AGCGTTCAGA CTGTATCGAA ATGTCCTGGG TCCAATCTCA CATTTATTTT TACGCTCCGA ATAGAACGAT AGAATTCCTT TGGGATAGAG ACCATACGAC CAAGAGCTTT CTAAAGATCA AAAGCGATTA TGTTAAAGAA CCGATTTCCA AGGCAGGGCT TGAAGACATT TGGAAGTTTT ATATGACAGG AGATAGCCCA GCCATGTTAT GGACGCCGAT GGGCGCACGT ATGAATGAAA TCAGTGAGTT CGAGTCACCC

TATCCTCACC GTGCCGGGAT CATCTTTAAT GTGATGTATG CGGGCAGTTG GGTACAGGGC TCCGATAGTG AGAAACAATT AGACTGGATG CGTAAGTTTC ATGATCATAT GACCCCTTAC GTTTCCAAAA ATCCGAGGTT GGCATATTTG AACTATAAGG ACTTGGACCT AGGTGAAAAT GACAAAAATT CTAGTGATTT GAAGGCGTAT TTTAAAGCAC GTAATTCTTG GGGCCGTGAG TATTTTAAAG GAAATTTTGA AAGATTAGTT AAGGTTAAAT CTATGGTGGA TCCTACAAAC TTTTTTAAGA ACAAGCAAAG CATCTCACCG ATCtaagaga gggcgcgcca cttctaaata agcgaattGA TCGAATCTCC TCTTAGCCGA TTCGATTATC GCATTCGGGA CTGAT

20 Gene Name: FNXH2130139

Gene ID: BBF.2017.48.20

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgG AAGTGTACTA CCAAGCAGCA AATAAATCTA ATAATACGTT AGGCTTCCCT GCTGGCTTTT GTCCCACAGT GGGTGTGGCA GGACACATCA GTGGAGGAGG CTTCGGGGCT CTAGTGCGTA AGTATGGGTT AGCGAGCGAC CAAGTAATCG ACGCGAGAAT TGTTACGGTC GATGGGAAGA TCTACACAAG GGAAACCATG GGAAAGGACC TATACTGGGC GATACGTGGG GGTGGTGCTA ACAACTTCGG GGTGCTGTTA TCATGGAAGG TAAAACTAGT CCCGGTAACG CCTGTTGTTA CGGTAGCGAC GATTAGCCGT ACACTAGAAC AGGGTGCTAC GGACCTGGTT CACAAGTGGC AATTCGTGGC CGATAGGTTA CACGAAGACG TTTACATTGG GTTGACATTT AGTGTAGCTA ACAGCAGTAG GGCTGGAGGG AAGACCGTGT CCGTTCAGTT CGCCTTTCTT TTTCTGGGGG GCTCTGATAG GCTGTTAGAG TTGATGGAAG AGTCCTTTCC CGAACTTGGT CTAAAGCGTA ATGAGACTAC TGAGATGAAG TGGGTCGAGA GTCATGTGTA CTTTTACGCC AGAGGACGTC CTATCGAGCT ATTATGGGAT AGAGATCACG CTACTAAGTC ATTTCTGAAA ATTAAGGCGG ATTATGTACG TGAGCCCATA TCAAAATCCG GGCTTGAAGC CATTTGGCGT AGGTTCGTAG GGGGGGATTC CCCGGCCATG CTTTGGACCC CGTTCGGAGG CAGGATGAAT GAGATCAGTG AATTCGAGAC CCCATATCCA CACAGAGCGG GGAATATTTA TAATATTATG TATGTTGGGA ATTGGATGAA CGAGACAGAA TCCGAAAAAC AAATTGACTG GATGAGACGT TTCTATAATT CCATGGCCAG GTACGTCTCA AAAAACCCTA GGTCTGCGTA CATCAATTAC AAAGACCTTG ATCTAGGGGT TAACCGTAAC AATGTGAGTG AAGCGGTGGG CTATGTCCAA

GCCAGATCTT GGGGTAGAAA GTACTTTAAA TCCAATTTCG AGAGGCTGGT GAAAGTTAAG TCCATGGTAG ATCCAGGTAA CTTCTTTAAA AATAAACAAA GCATCCCTCC GGTAAGCACA TGGtaagaga gggcgcgcca cttctaaata agcgaattGA TCGAATCTCC TCTTAGCCGA TTCGATTATC GCATTCGGGA CTGAT

21 Gene Name: XMVD2010505

Gene ID: BBF.2017.48.21

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgA TTCTGTCATT GCTTTATTCA CTATTGCTGC TAAGTATCTC ATGGCAGACA AGCTCTGCCT CCTCTTTGGC TCCCGAAAAC TTAGTACATG ACGAGGATTT CGTCCAGTGT ATACACCTTT ACAAACCTCC TAACAGTATA CCGATCTTCA CTATCAATCA TTCAAGCTAC TCTCCCTTAC TTCAGAGCAG TGCGCGTAAT CTGCGTTTCC TGAGCCCGAC ATTCCCAAAG CCCCAATTCA TCGTGACACC CTTACATGAG TCTCACATAC AGGCGACGGT AGTTTGTTGT CGTAAACACG GACTAGATAT CAAGATTCGT AGTGGAGGTC ATGATGTGGA GGGAGAGTCC TTTATTAGTG ACACACCTTT TGTTCTGGTT GATCTAATAA ACTTCAGATC CATACATGTA GACGTAGAAA ATAAAACCGC ATGGGTACAA GCTGGTGCTA ACTTAGGAGA GATCTATTAT GAGATAGGTG AAAAATCAAA GACATTAGGG TTTCCTGCGG GTTACTGTCC AACAGTGGGC GTGGGTGGCC ACCTGTCAGG AGGTGGCTTC GGCGCATTAG TCCGTAAGTA CGGGCTGGCG GCGGATCAGG TAATCGACGC TCGTATCGTA AATGCAAATG GAGAAATCTT GGATAAAGAA ACGATGGGGG AGGATCTGTT TTGGGCCATC CGTGGTGGCG GTGCGAGTTC ATTCGGTGTC ATTTTATCAT GGCAAATTAA TCTGGTCCCA GTGACCCCGA TAGTGACTGT CGCTACAATC AGTAAAACAC TGGAGCAAGG CGCGATAGAT CTAGCTCATA TGTGGCAGTT TATCGCAGAT AGGTTGCATG AAGACTTATA TCTGGGTCTT ACATTTACAA AAGTAAACGG GACAGCGTTA GCCGTCTTTA GCGTAATGTT TCTGGGCGGA ATCGATAAGG TCCTACCCTT GATGAAGGAG AGTTTCCCGG AGTTGGGAAT GAAGAGAGAA GACTGCTTGG AAATGTCTTG GGTCGAGTCA

CACTTATATT TCTACGCTAG AGGTCAGCCG CAGGATTACT TGCTAAACAG GAATGACCCA
CTGTCCAAGG TCTTCCTTAA AGCCAAGTCA GACCACGTTA AGGAACCTAT ACCGAAAAAG
GGTCTTAAAG GACTTTGGAG GAGACTTGAG CAAGAAGACA TGCCTGTTCT TATGTTTACT
CCGTTCGGTG GAAAAATGAA TGAGATTTCT GAAAGTGAAA CGCCGTTTCC TCATCGTGTT
GGAAATATCT ATAACATATG TTATTTGACG AACTGGTTCC AGCCAAACGA AAGTGAAAAG
CACATAGATT GGTCTAAGCG TCTTTACAAA TACATGACTC CGTACGTGAG TAAGGACCCC
CGTACGGCGT ATATAAACTA CAAAGACCTA GACCTTGGTC AGTACAAGAA CGGTAAAGTG
TCCCATTCCA AGGCGAGGGT CTGGGGTAGA AAGTACTTTA AGAATAACTT CGAGAGATTA
GTTAAGGTGA AATCCTTAGT TGATCCTGGG AACTTCTTCA AGAATAAACA ATCTATTCCG
CCCGTGAAAA GTtaagagag ggcgcgccac ttctaaataa gcgaattGAT CGAATCTCCT CTTAGCCGAT
TCGATTATCG CATTCGGGAC TGAT

22 Gene Name: ERXG2008196

Gene ID: BBF.2017.48.22

Submitted by Aaron Cravens on 2017.11.20. Contact at acp1an0@stanford.edu.

Description

Specific and non-specific small molecule oxidase enzymes derived from plants.

Motivation

NA

References

NA

Sequence

ATCAGTCCCG AATGCGATAA TCGAATCGGC TAAGAGGAGA TTCGATCgga attcgatatc aaagcttgga gatctaaaag aaatacatgT TGGTTTCCTT ATTGTTATTT TCTTCTCTTA CATTGTCTTT TGCGACAAAG TCCTCCTCCA TCCACGGGGA CTTTATTCAG TGTCTAAATT CCCAGAAGCA CCTGCATTCC AGCTCCATAG ACTCTAACCA GTTTGTGGTT CCCGTACTAA CTCCTACCTC TTCCAACTTT ACGAACCTGT TCAGATCCTC TGCTAGGAAT TTAAGATATC TGAGTCCAAA CAGCACCAAG CCGCAAATAA TCATCACCCC GACGCATGAA TCCCATATTC AAGCGGCTGT CATCTGCTCC AAGAAGCATG GATTCGATTT GAAGATAAGG AGTGGAGGGC ACGATGTAGA AGGTATCTCC TACGTCTCTG ATACTCCGTT CGTCTTGGTA GACCTGATTT CATTCCGTAA TATTAGCGTA GATGTCAAAT CCAAAACTGC CTGGATACAA GGCGGAGCTT CTCTGGGAGA GGTATATTAT GCCGTAGCAA ATAAGAGCAA TACCTTAGGC TTTCCTGCTG GTTTTTGTCC AACCGTGGGG GTCGGCGGGC ACATCTCAGG GGGTGGATTT GGAGCACTTG TGAGAAAGTA CGGTCTTGCC TCTGACCAAG TTATAGATGC TCGTTTTATA AATGCTGACG GAAAGATCCT GGACAAAAGG ACGATGGGTA AGGAATTGTT CTGGGCTATT AGGGGTGGGG GAGCGTCCAA CTTCGGAGTC GTCGTGAGCT GGAAAGTCAA GTTAGTCGAT GTGACACCGA TAGTAACTGT AGCTACGGTA GACCGTGCCC TGGAACAAGG AGCAACCGAT TTAGTTCATC GTTGGCAGTT TATAGCGGAC AGGTTACACG AAGATTTGTA TATAGGGCTT ACCATGACGG TGAGGAATGG ATCTAGGGCG GGGGAAAAGA CTGTGATCGC TTCTTTTAGT TTCATGTTTC TGGGTAGGGT GGACAGGTTG TTGGAGCTTA TGGAAGAGTC TTTTCCTGAG CTTGGTTTGA AACGTAATGA CTGCACTGAA

ATGAGCTGGG TTGAGTCCCA TGTGTACTTC TTCGCGAGGG GACGTACTAT CGACCTTTTG TGGGACCGTG ATCACGCGAC TAAATCTTTT CTGAAAGTCA AATCAGATTA CGTCACTGAG CCAATTAGCA AAGCTGGGCT AGAGTCAATC TGGAAAAGAT ATATGACTGG TGACTCACCA GCAATGCTGT GGACTCCTAT GGGCGCTAGG ATGAATGAAA TCAGCGGGTA CGAAACTCCG TATCCTCACA GAGCCGGCAA CATTTATAAT ATTATGTATG TGAATAACTG GCAACAAGAG TCCGACACCA AGAAACAAAT CGATTGGATG AGAAAGTTCC ATAGCTACAT GACCCCGTTT GTATCAAAAA ACCCCAGGTC TGCCTATTTG AACTACAAGG ATCTAGATAT CGGTGTGAAC AGCGAGGATG GCAAGTTAAC TTACTCTAAG GCTAAAATAT GGGGTTCAAA GTACTTTGGC TCTAATTATG AAAGACTTGT GAGAGTGAAG AGCATGGTGG ATCCCGACAA TTTCTTTAAA AATATTCAAA GTATTCCTCC GGTGAGACAT TGGtaagaga gggcgcgcca cttctaaata agcgaattGA TCGAATCTCC TCTTAGCCGA TTCGATTATC GCATTCGGGA CTGAT

23 Gene Name: ctrB-yeast-1

Gene ID: BBF.2017.48.23

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Beta Carotene and Astaxanthin synthesis genes optimized for yeast and rice.

Motivation

NA

References

NA

Sequence

gaagacataA TGAATAACCC ATCTCTGTTG AATCATGCCG TGGAAACAAT GGCAGTGGGA AG-
TAAGTCTT TCGCAACCGC CAGCAAATTG TTTGACGCAA AAACTAGACG TAGTGTCTTG AT-
GTTGTACG CGTGGTGCAG ACATTGTGAC GATGTGATAG ACGACCAAAC ACTGGGATTT CAAGC-
TAGAC AGCCGGCCTT GCAAACGCCG GAACAACGTC TGATGCAGTT AGAAATGAAA ACCA-
GACAGG CATACGCGGG TTCACAAATG CACGAGCCTG CTTTCGCGGC TTTCCAAGAA GTAGC-
CATGG CACATGATAT CGCCCCCGCT TACGCTTTCG ACCACTTGGA GGGCTTTGCA ATG-
GACGTGA GAGAAGCACA GTACAGCCAG CTAGATGATA CTTTGCGTTA CTGTTATCAT GTCGCAGGGG
TCGTTGGCCT AATGATGGCG CAAATAATGG GCGTCCGTGA CAACGCAACG CTTGATCGTG
CGTGCGATTT GGGACTTGCT TTCCAACTAA CTAATATTGC GAGGGACATA GTGGATGACG
CTCATGCTGG TAGGTGCTAT CTGCCAGCTT CTTGGTTAGA GCACGAGGGC CTTAATAAGG
AGAATTACGC GGCCCCCGAA AACAGACAGG CGTTATCCAG AATAGCAAGG CGTCTAGTAC
AGGAAGCCGA ACCTTACTAC CTATCTGCGA CAGCCGGCTT AGCGGGGCTA CCTTTGCGTT
CAGCATGGGC AATTGCGACG GCGAAACAAG TTTACAGAAA AATTGGGGTA AAGGTCGAGC
AGGCGGGGCA ACAAGCTTGG GATCAACGTC AATCTACCAC CACACCAGAA AAATTGACCC
TTCTTTTAGC CGCGTCCGGC CAAGCTTTGA CAAGTAGAAT GAGAGCACAT CCTCCTCGTC
CTGCCCACTT ATGGCAAAGA CCTCTGTGAa tcgtcttc

24 Gene Name: crtI-rice-1

Gene ID: BBF.2017.48.24

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Beta Carotene and Astaxanthin synthesis genes optimized for yeast and rice.

Motivation

NA

References

NA

Sequence

gaagacataA TGAAGAAAAC TGTTGTGATT GGCGCCGGGT TTGGCGGTCT GGCGCTTGCA ATTAGGCTTC AAGCGGCGGG GATCCCTACG GTCTTGTTGG AGCAGAGGGA TAAACCTGGG GGGAGAGCGT ATGTGTGGCA CGACCAGGGG TTTACTTTTG ATGCAGGACC AACGGTCATT ACCGATCCGA CGGCACTGGA GGCACTGTTT ACCCTGGCTG GCAGGCGGAT GGAGGACTAT GTTCGCCTCC TGCCCGTCAA GCCGTTTTAT CGCCTGTGCT GGGAGAGTGG TAAAACTCTT GATTACGCAA ACGACTCTGC GGAACTGGAG GCACAAATCA CTCAATTCAA TCCCCGGGAC GTGGAGGGGT ATCGGCGGTT CCTTGCATAT TCGCAGGCAG TTTTTCAGGA GGGATATCTG CGCCTGGGAT CTGTGCCTTT CCTGTCATTC AGGGACATGC TTAGGGCTGG CCCACAACTT CTGAAATTGC AAGCGTGGCA AAGCGTGTAC CAATCCGTTA GCCGCTTTAT TGAAGATGAA CACCTTCGCC AAGCGTTTTC CTTCCACAGC CTGTTGGTCG GGGGCAATCC TTTTACCACA TCGAGTATCT ACACACTCAT TCATGCGTTG GAGAGGGAAT GGGGCGTTTG GTTCCCTGAA GGAGGAACCG GTGCTCTCGT CAACGGCATG GTGAAGTTGT TCACCGACCT GGGAGGCGAG ATTGAGTTGA ACGCCCGCGT GGAGGAGCTG GTTGTCGCGG ACAATAGGGT TTCTCAAGTG AGACTCGCTG ATGGACGGAT TTTCGATACT GACGCCGTGG CGTCTAACGC TGATGTTGTC AATACTTACA AAAAACTTCT CGGGCATCAT CCTGTTGGTC AAAAGAGGGC TGCCGCACTG GAAAGAAAAT CTATGAGTAA CAGTTTGTTC GTCCTGTATT TCGGATTGAA TCAACCGCAT TCACAGTTGG CACACCATAC CATCTGCTTC GGTCCTAGAT ACCGCGAATT GATTGATGAA ATTTTCACAG GGTCTGCGCT CGCTGACGAC TTCTCATTGT ATTTGCATTC CCCGTGCGTC

ACGGACCCTA GTCTGGCACC ACCAGGCTGT GCCAGCTTCT ACGTCCTGGC TCCAGTCCCC
CACCTTGGTA ATGCTCCGTT GGATTGGGCG CAAGAAGGGC CGAAGCTGCG GGATAGAATT
TTTGATTATT TGGAGGAGAG GTATATGCCA GGCCTTCGGA GCCAGCTCGT GACTCAGCGC
ATATTTACGC CTGCTGACTT TCACGACACC TTGGATGCGC ATCTCGGAAG TGCCTTTTCA
ATAGAGCCTC TGCTGACACA GAGCGCCTGG TTCAGACCAC ATAATAGGGA CAGTGACATT
GCCAACCTCT ACCTTGTGGG CGCTGGCACA CACCCGGGTG CTGGCATTCC CGGGGTGGTG
GCGTCAGCCA AGGCCACCGC TTCGCTGATG ATTGAGGATC TCCAGTGAat cgtcttc

25 Gene Name: crtI-yeast-1

Gene ID: BBF.2017.48.25

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Beta Carotene and Astaxanthin synthesis genes optimized for yeast and rice.

Motivation

NA

References

NA

Sequence

gaagacataA TGAAAAAGAC AGTGGTGATA GGGGCAGGGT TCGGCGGTTT GGCACTGGCG ATTCGTTTGC AGGCAGCTGG GATCCCCACC GTTTTGTTAG AGCAACGTGA TAAACCGGGC GGTCGTGCAT ATGTATGGCA CGACCAGGGT TTTACGTTTG ACGCCGGGCC GACTGTGATC ACGGACCCAA CAGCACTTGA AGCTCTATTC ACGCTTGCAG GAAGGAGGAT GGAGGATTAC GTTAGACTGC TACCAGTCAA GCCCTTTTAT AGACTGTGTT GGGAGTCCGG CAAGACGTTG GACTACGCAA ATGATAGTGC AGAGCTGGAA GCACAAATCA CACAGTTCAA CCCCCGCGAC GTAGAAGGTT ACAGGCGTTT TTTAGCCTAT AGCCAAGCGG TTTTCCAAGA AGGGTATCTA CGTTTGGGCA GTGTGCCGTT CTTGTCCTTT AGAGATATGT TGAGAGCTGG CCCTCAGTTA CTAAAATTAC AGGCATGGCA AAGTGTCTAT CAGAGTGTCT CCCGTTTCAT CGAGGACGAG CATCTGCGTC AAGCTTTCTC CTTCCACAGC CTTCTTGTCG GTGGTAACCC GTTCACAACT TCCTCAATCT ATACGTTAAT CCATGCTCTT GAAAGGGAAT GGGGAGTATG GTTTCCGGAG GGAGGCACGG GCGCCCTTGT CAATGGAATG GTAAAGTTGT TCACAGACTT AGGCGGAGAG ATTGAACTTA ATGCAAGAGT CGAGGAGCTA GTGGTGGCTG ACAACCGTGT CTCTCAAGTT CGTTTAGCGG ATGGTAGGAT TTTTGATACC GATGCAGTAG CTTCCAATGC GGACGTCGTA AATACGTACA AGAAACTGCT GGGACATCAC CCTGTTGGAC AAAAGCGTGC TGCCGCTCTT GAGAGAAAAT CTATGTCAAA TAGCTTGTTT GTGCTGTACT TCGGCCTGAA CCAGCCGCAT AGTCAGCTGG CGCATCATAC TATTTGTTTC GGTCCAAGGT ACAGGGAACT GATAGACGAG ATCTTCACAG GTTCCGCACT AGCAGACGAT TTCTCACTTT ATCTGCATTC ACCATGCGTT

ACAGACCCGT CTTTAGCGCC TCCGGGATGT GCGTCATTTT ATGTCCTGGC ACCGGTCCCG
CACCTTGGAA ACGCGCCGTT AGATTGGGCC CAAGAGGGCC CCAAACTGCG TGATCGTATC
TTTGATTATT TAGAAGAGAG ATACATGCCA GGCCTACGTT CTCAGCTTGT CACTCAAAGG
ATATTCACAC CCGCAGATTT CCACGACACC CTGGACGCGC ACCTTGGTTC CGCTTTCTCA
ATCGAGCCGC TACTGACCCA GTCCGCTTGG TTTCGTCCAC ACAATAGAGA TTCCGACATT
GCGAATTTGT ATTTAGTAGG AGCAGGAACA CATCCAGGCG CGGGCATACC AGGCGTAGTG
GCTTCAGCAA AGGCTACGGC GTCATTGATG ATCGAGGATT TACAATGAat cgtcttc

26 Gene Name: crtO-rice-1

Gene ID: BBF.2017.48.26

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Beta Carotene and Astaxanthin synthesis genes optimized for yeast and rice.

Motivation

NA

References

NA

Sequence

gaagacataA TGCCGACAAC TCAGCTCTAC CCTGCCCCTC AACATGATGC CTCGGCACGC TG-
GTACTTGA CTAACCCCTA CAGGGGGGTC GCAATTGGGT TGACTATACT TGTCGGGTGG
TTCGGGCTGC TTGCTTTCCT CCTGAATTAC ACCATAGATT GGACCAATCC ACTCACCTAC
TTGCTTGTGC TTCTGCAGAC GCATTTGTAT ACCGGAATGT TCATCACTGC ACACGACGCC
ATACACGGGG TTGTCGCCCC CACGAATCGC AGGCTTAACG ACGCTATTGG GTGGGTGGCT
ACAACACTGT TCGCGTTTAA TAACTATGCG AAACTCACTA AAGGACATCA TCTGCACCAC
AAGCATGCCG CTACAGATCA CGATCCTGAT TTTCATAGAG GAAATGCGAA TATCCTTTTG
TGGTTCTTGG CGTTCGCCCG GTCTTATGTT ACAATCTGGC AAGTGCTTCT GATGGCTCTT
ACATATAATG TGTTGAAAAT TTGGTTCCCA ATGGAAAATC TGATTTTGTA CTGGATGGTG
CCCGCAATTT TGAGCACGTT CCAACTGTTC TTCTTTGGAA CTTACCTCCC ACATCGGGGG
GATCACGAGA ACCCACCACA CAACGCTCGT TCTCAATCTC TTAATCATGT GTGGGCCTTC
GTTTCATGTT ACTTCTTTGG GTATCATTTT GAGCATCACG ATCAGCCATA TCTCCCTTGG
TGGAAACTGG CCGCTGCGAG GGAATCCAAT GGTCTTTGAa tcgtcttc

27 Gene Name: crtO-yeast-1

Gene ID: BBF.2017.48.27

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Beta Carotene and Astaxanthin synthesis genes optimized for yeast and rice.

Motivation

NA

References

NA

Sequence

gaagacataA TGCCCACTAC CCAGCTGTAT CCCGCACCCC AGCATGACGC CTCCGCTCGT TGGTATCTTA CAAACCCGTA CAGGGGTGTC GCAATCGGCC TTACAATATT GGTAGGATGG TTCGGCTTAC TAGCCTTTTT GCTAAACTAC ACCATCGACT GGACTAACCC ACTTACATAC TTACTGGTAC TACTGCAGAC GCACCTTTAC ACCGGTATGT TCATCACCGC TCATGACGCA ATACATGGCG TTGTAGCACC AACGAACAGG AGATTAAATG ACGCGATTGG GTGGGTCGCC ACCACACTGT TTGCTTTCAA CAACTACGCA AAGCTGACTA AAGGTCATCA TCTTCACCAT AAGCATGCAG CTACAGATCA TGACCCGGAC TTCCACCGTG GTAACGCAAA TATACTTTTG TGGTTTCTGG CTTTTGCTAG AAGCTATGTG ACCATTTGGC AAGTTCTATT GATGGCGCTT ACCTACAACG TGCTTAAGAT TTGGTTCCCG ATGGAGAACC TTATATTGTA TTGGATGGTC CCTGCTATTT TGTCAACCTT CCAGTTGTTC TTTTTTGGCA CGTACTTGCC CCATAGGGGT GACCATGAAA ACCCGCCGCA TAATGCACGT TCACAGAGCC TAAACCATGT ATGGGCCTTT GTCTCCTGTT ATTTTTTCGG TTACCACTTT GAGCATCACG ACCAACCATA CCTTCCTTGG TGGAAATTAG CTGCGGCAAG GGAATCCAAT GGCCTGTGAa tcgtcttc

28 Gene Name: crtB-b-rice-1

Gene ID: BBF.2017.48.28

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Beta Carotene and Astaxanthin synthesis genes optimized for yeast and rice.

Motivation

NA

References

NA

Sequence

gaagacataA TGTCCTTCGC GGCGACAACA ATGCTCAGTA AACTCCAGTC CATTTCAGTT AAGGCCAGGC GGATCGAACT CGCTAGAGAT ATAACCAGGC CGAAAGTGTG CCTTCACGCT CGGCGCTCTC TGGTCAGACT TCGGGTTGCG GCTCCACAAA CTGAAGAAGC GGTTGGCACA GTGCAAGCTG CTGGAGTTGG AGATGAACAT TCCGCGGACG TGGCCCTCCA ACAACTCGAC CGCGCGATAG CGGAGAGGAG GGCTAGGAGG AAAAGGGAGC AACTGTCGTA TCAAGCAGCA GCCATAGCAG CAAGCATCGG TGTTTCAGGA ATTGCGATTT TTGCGACTTA CCTCCGGTTC GCTATGCATA TGACTGTTGG CGGAGCTGTC CCATGGGGCG AGGTGGCGGG GACTCTCTTG CTGGTCGTTG GTGGAGCTCT TGGTATGGAG ATGTATGCAC GGTACGCACA CAAGGCAATT TGGCACGAAA GCCCACTGGG GTGGCTGCTC CATAAGTCCC ATCATACTCC GCGGACTGGG CCCTTCGAGG CGAATGATCT GTTCGCAATC ATCAACGGCC TTCCCGCCAT GCTTCTTTGC ACATTCGGGT TCTGGCTGCC AAACGTGCTG GGTGCTGCGT GTTTTGGAGC AGGACTTGGG ATCACCCTTT ACGGCATGGC TTACATGTTC GTCCATGACG GCCTCGTTCA CCGCCGGTTC CCAACCGGCC CAATTGCTGG CCTCCCGTAC ATGAAACGCC TCACCGTGGC TCACCAGCTT CATCACAGCG GAAAGTACGG TGGGGCGCCC TGGGGAATGT TCCTCGGTCC CCAGGAGCTT CAGCACATTC CCGGTGCGGC GGAGGAAGTT GAGAGGCTCG TGTTGGAGTT GGATTGGAGC AAGAGGTGAa tcgtcttc

29 Gene Name: crtB-b-yeast-1

Gene ID: BBF.2017.48.29

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Beta Carotene and Astaxanthin synthesis genes optimized for yeast and rice.

Motivation

NA

References

NA

Sequence

gaagacataA TGAGTTTTGC CGCTACCACG ATGTTGTCAA AGTTGCAGAG CATCAGCGTA AAGGCAAGGC GTATTGAATT GGCTCGTGAC ATCACTAGGC CGAAAGTATG TCTACACGCA AGGCGTTCTC TTGTAAGACT GCGTGTAGCG GCACCGCAAA CGGAAGAGGC AGTTGGAACT GTTCAGGCTG CGGGTGTCGG TGATGAGCAT AGTGCAGACG TCGCATTGCA ACAGTTAGAC AGGGCAATAG CTGAGCGTAG AGCACGTAGA AAAAGAGAAC AGTTAAGCTA CCAAGCCGCT GCCATAGCTG CATCAATCGG GGTATCCGGG ATCGCGATCT TTGCCACGTA TTTGAGATTT GCAATGCACA TGACTGTAGG TGGAGCTGTC CCCTGGGGGG AAGTCGCTGG AACTCTGCTG CTGGTAGTGG GGGGAGCCCT AGGGATGGAG ATGTACGCTA GGTACGCCCA CAAAGCTATC TGGCACGAAT CTCCTCTTGG GTGGCTGCTA CACAAATCAC ATCATACCCC ACGTACTGGC CCATTTGAAG CCAACGATCT TTTTGCCATC ATAAACGGGT TGCCAGCAAT GCTGCTATGC ACCTTTGGGT TTTGGCTACC AAACGTTCTG GGAGCGGCAT GCTTCGGAGC TGGGTTAGGA ATCACACTTT ATGGTATGGC ATATATGTTT GTACATGATG GACTGGTTCA TAGAAGGTTT CCGACGGGAC CCATAGCAGG CTTGCCTTAC ATGAAAAGGT TGACCGTTGC TCACCAATTA CATCATAGCG GTAAATACGG AGGGGCGCCC TGGGGGATGT TCCTAGGTCC ACAAGAGCTT CAACATATCC CTGGGGCTGC CGAAGAGGTA GAGAGGTTGG TCTTGGAGTT GGATTGGTCC AAACGTTGAa tcgtcttc

30 Gene Name: Mcherry-Zebrafish-1

Gene ID: BBF.2017.48.30

Submitted by Zavier Henderson on 2017.11.20. Contact at xthechallengerx@gmail.com.

Description

Mcherry codon optimized for zebrafish.

Motivation

NA

References

NA

Sequence

gaagacataA TGGTTAGTAA AGGAGAGGAA GATAACATGG CGATCATAAA GGAATTTATG CGATTCAAAG
TTCATATGGA GGGAAGCGTT AATGGGCACG AATTTGAAAT CGAAGGGGAG GGAGAGGGAC
GACCATATGA GGGAACGCAA ACCGCAAAAC TTAAGGTCAC GAAGGGCGGC CCCCTTCCTT
TCGCATGGGA TATTTTGTCC CCGCAGTTCA TGTACGGCAG CAAAGCCTAT GTAAAGCACC
CCGCCGACAT CCCGGATTAT TTGAAACTCT CTTTTCCAGA GGGGTTCAAA TGGGAAAGAG
TCATGAACTT TGAGGACGGG GGAGTAGTGA CTGTTACACA GGACAGTTCC CTTCAAGACG
GAGAATTCAT ATATAAAGTA AAACTCCGGG GGACGAACTT CCCTAGTGAC GGGCCTGTCA
TGCAGAAGAA AACGATGGGT TGGGAAGCGT CAAGTGAGCG TATGTACCCC GAGGATGGCG
CCTTGAAAGG AGAGATTAAG CAACGCCTTA AACTTAAAGA CGGCGGTCAT TACGACGCGG
AGGTAAAGAC CACATATAAA GCAAAGAAAC CTGTGCAACT TCCTGGTGCA TATAACGTGA
ACATTAAACT CGACATAACC TCACACAACG AGGACTACAC GATTGTTGAA CAGTATGAGC
GGGCGGAAGG GCGCCACTCA ACAGGCGGGA TGGACGAGTT GTATAAGTGA atcgtcttc

31 Gene Name: gene-15693

Gene ID: BBF.2017.48.31

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AAAAGAGGAA AGTCGAGGAT CCGAAGAAGA AACGGAAGGT GGGTTCCGGT
TCTATGCCTT CAGCTCAACG GTGCATCTGG GAGTGGAAGA GGGATATCTT CGTGACCAAG
AATCCGACGC TCCGGGAGTC CGTGGATGAA CTTAGCTTGC CAGGGACCAG GCGCATCGTA
CAGGGATGGA TCGACCAGCA AGCCCAATAC CCGGAAGATG GGTCAGCAGA CGAATATAGC
TTTTATGCCG AAGAGTGCTA CCCAACCTCT CATGACCGGC GAGCGTTCTT CCACCGCTTC
ATTGCCGAGG CCAGACCGCA TATCGGCTAC AAGCTGGTTG CGCAGTTGGC AGAAGCAGGG
TTCTTGAGAA CCATTTGGAC GACCAACTTT GACGGACTGG TTAGCAGAGC GTGCACAGCG
GCTAACGTCG TGTGCGTGGA AGTGGGCATG GACACACCCC ACAGGGCCTC ACGACCGCAA
GGGGATGACG AAGTCAGACT GGTGTCCCTC CACGGTGACT TTAGGTATGA TCTGCTGAAG
AACACCGCCA ATGAGCTGCG CGAGCAGGAT TTGGCCCTTA GGGAGGAACT GCTGCACGAA
CTCAAAGACT ACGACCTGGT GGTCATCGGA TATTCAGGGC GGGACGACAG CCTTATGCAA
GTGCTCTCTG CTGCCTACAG CGACCGCGCA TCTTGTAGGC TCTACTGGTG CGGGTTTGGC
GCGGAACCAG CACCGGAAGT GAGGCACCTT ATTAAGAGCA TCGACCCAGC CCGAGAGAGC
GCGTTCTACG TGGATACCGC CGGATTTGAC GACGTAATGA GCAGGCTTGC ACTCAGGCGA
CTGAGCGGTG AAAGCCTCGA AAGGGCCCAG AAGCTCATAG AAAGCGTCAC CCCGGTTGCT
GGCAAAAAGA TGGCCTTTAG TGTTCCACCA TTGGCCCCTA GCGCCTTGGT GAAGGGTAAT
GCCTACCGAT TGACCTGTCC GGCAAACGTC TTGAAACTTG ATATCGAACT TCCCGAGCAC

GGTTCCTGGC GCGATTGGCT GTCCGAACGA ATGACTCCAG AAAGGGGGCA GGCCGTTGTG
TTCGAGAAGG GAGCACTGGT TTTGGCCGAC ATGGCGGTTA CCGCTAAAGT TTTCGATGGA
TTTCTTAGGG TGAGCCCGAC ACGGGTGGAG ATAAGTGACG AGAACATTAT CGCTGACGGC
CGGATCGCCA GTCTTTACCG ACGAGCTCTC GTGAGCAGTG CCGCAAAAGC GCTCCAGATC
CAAACCGACC ACAGGAGGAG GATATGGGAG CCCGTGCACT ATGATACAAG GCAACTCGAC
GATGTGACGT ACCGCGTGCA TCGAGCTGTC TCCCTGACGA TAGTAGGGAT AGAGGGAGTG
CCCCATGTGG TGCTGATGCC AGAGGTCGTC GCATCTACGT TGGCGGGCGA CCTTGCGCCG
GTTGACAGTC AAAAGACTCT CCGCAATGCC ATTTACGGGT TCCAACATAA CGATAAGTTT
GATGCCGACC TCAGCTATTG GACCCACCGC CTTGTTGAGA AGGAGCTGGC TTCCAGCGGC
GAGGGCGTTT TCGTATTGAG CAAAGTGCCA CTTTATGCGG GCCTGGCACA AAAAGGTAAA
GCTCCTCTCC CACACAGGTT TGCACGCCAC GCTAAACAGC ATGGAATTAT TGTGCCCGAC
GCACCGCTTG TTTTCAGCGC CAAGGTTGGC TCTGGAGAGG TACGAAACCC CAATCCGCTG
CATGGGCTGG TGCAAAACCG GCCATGGGAC CACTCTCTTA CGGCGTCTGG TTTGTGTCCG
AGTACAGATG CTAGCGTGAT CTGCCCCGCA GACGCTGCTC CGAGGTTTGA GAGATTCCTC
CAATCTATGC AGGAGGTAGC AAGACCAAGC CAGAGCGAGA GGGACTATTT GCATGATTTT
CCCGGCTTCC CTGCGGCCTT TGGACTGCCA CTCCGAATGC CCGTGAGAGG GGACGCAAAC
TGGATTACCA TCGACGACGG AGTGAGCACC GATGCCCTGA CAGGGGCTAA GCAACTGGCG
CACCGAGTGT GCCAAGCACT CGACCACCTC CGCAGAGCAA GGCCCTCTGA CACGGCGATC
GTGTTCGTTC CCAGGAGATG GGAACCATAT AAGGTAGTGG ACACGCAGCA CGAAAGATTC
AATTTCCACG ATTACATTAA GGCCTACGCG GCCAGGCACA GTCAGAGCAC GCAGTTCGTC
AGAGAAGAAA CCATCCAAAG CCAATACGTG TGTAGGGTCC GGTGGTGGTT GAGTTTGGCA
CTGTATGTTA AGGCTATGCG GACCCCCTGG CGGCTGGATG CGCTTGATGA GAATACGGCT
TTTGTTGGTA TAGGGTACTC CCTGGACGCA GAGGCAGGGA GGGGCAACCA TGTACTGCTC
GGCTGCAGCC ACCTGTATTC TGCGAGGGGT GAGGGATTGC AGTTTAGGCT GGGCCGAATC
GAGAATCCCG TGGTGCGAGG AAGGAACCCC TTCATGAGCG AGGACGACGC AAGGAGGACC
GGAGACACCA TCCGGCAGCT TTTCTACGAT AGCAAAATGC ATATTCCGAC AAGGGTGGTG
ATACACAAGA GGACAAGGTT CACTGACGAG GAGCAGAGGG GGTTGGTACA AGGATTGGAC
GGTGTGAGGA ATATCGAGCT GATAGAGATC AACCAGGAAG AGAGCTTGCG ATATCTCAGC
AGCCAGATGA AGGACGGCAG ATTTGAGATC GACAAGTTCC CCCTGTTCAG GGGTACCACA
ATAGTTGAGT CAGATGACAC TGCATTGCTG TGGGTGCATG GAGCCACACC CAGCGCCGTG
AACAAGTACT GGAGGTACTA CCAGGGGAAG CGCCGCATTC CGGCGCCATT GAGGATTCGA
AGGTTCCTCG GGCAAAGCGA CGTAGTGCAG ATCGCGACCG AGATCTTGGG ACTGTCTAAA
ATGAACTGGA ATACGCTTGA CTACTATTCA AGGATGCCTG CGACTCTGGA TTCTGCAGGC
AGTATTGCCA AGTTCGGGTC ATATCTTGAT GGGTTTACGA GCGCACCCTA TGATTACAGA
CTTCTGATCT AGTGA

32 Gene Name: gene-20107

Gene ID: BBF.2017.48.32

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAAA AGAAAAGGAA GGTAGAGGAC CCCAAGAAAA AGCGCAAAGT AGGGAGCGGT
AGCATGAACT ATACCGCTGC TAACACAGCG AACTTCCCGA TATTTCTGAG CGAAATAAGC
TTTCTCACAA CCAATAACAT TTGCTTGAAC TGTTTCAAGC TTAACTACCA GGTAACGAGG
AAGATCGGTA ACCGATTTTC ATGGCAGTTC AGCAGGAAAT TCCCCGACGT TGTAGTGATA
TTCGAGGACA ACTGCTTCTG GGTCCTGGCA AAGGACGAGA AGTTCTTCCC CTCACCACAA
CAGTGGAAGG AAGCACTTAG CGATATCCAG GAGGTTCTTA GAGAGGACAT CGGGGACCAC
TACTACAGCA TCTATTGGCT TAAAGACTTT CAAATAAAGG CCCTGGTGAC CGCCCAACTG
GCGGTGAGGA TACTCAAGAT TTTCGGCAAA TTTAGCTACC CAATCGTCTT TCCCAAGGAT
AGCCAGATAT CAGAAAATCA AGTACAGGTC AGGCGCGAAG TTGACTTTTG GGCCGAGATC
ATCAATGACA CCAACCCCGC AATCTGTCTG ACCGTGGATA GTAGCATTGT GTACAGTGGC
GACCTTGAAC AGTTTTACGA AAACCACCCC TACAGGCAAG ACGCCGCTAA GCTGCTGGTG
GGACTGAAGG TAAAGACCAT CGAAACCAAT GGCACCGCGA AGATCATACG GATCGCCGGT
ACCATAGGCG AGCGCAGAGA GGACTTGCTG AAGAAGGCCA CAGGCTCAAT GTCACGACGG
AAACTGGAGG AAGCCCATCT CGAACAACCC GTCGTCGCAG TCCAGTTCGG AAAGAACCCC
CAGGAGTACA TATACCCGCT TGCGGCCCTT AAACCTAGCG TGACCGACGA AGATGAGAGC
CTCTTCCAGG TCAACCACGG AGACTTGTTG AAGGAAACCA AGATCCTGTA TGCGGAGAGG
CAGGAGCTTC TGAAGCTGTA CAAGCAGGAG GCCCAGAAAA CCCTGAACAA CTTTGGGTTC

CAGTTGAGGG AGAGGTCCAT CAATTCTCAG GAATATCCTG AGGTGTTTTG GACTCCCAGC
ATCAGCCTGG AGCAAACCCC AATCTTGTTT GGCAAGGGGG AGCGAGGTGA AAAAAGAGAG
ATTTTGAAGG GCCTGAGCAA AGGCGGAGTG TACAAAAGGC ACAGGGAATA CGTGGACACA
GCTCGCAAAA TTCGCCTGGC CATACTTAAG CCCGCTAACC TCCGCGTGGG CGACTTTCGG
GAGCAACTTG AGAAGCGATT GAAGCTTTAT AAGTTTGAGA CAATTCTGCC ACCGGAGAAC
CAAATTAACT TCAGTGTCGA AGGCGAAGGT TCCGAAAAGA GGGCCCGATT GGAAGAAGCG
GTCGACAGAC TCATAAGGGG GGAGATCCCC GTAGACATTG CACTGGTGTT CCTCCCGCAG
AGCGATAGGA ATGCAGACAA CACCGAGGAG GGAAGCCTTT ACAGTTGGAT CAAGAGAAAA
TTCCTCGATA GGGGCGTGAT TACACAGATG ATTTATGAGA AAACGCTTAA CAATAAGTCA
CAGTACAACA ACATCCTGAA CCAGGTGGTG CCGGGGATTC TTGCGAAGCT GGGAAATCTG
CCATACGTTC TTGCAGAGCC GCTTGAGATA GCCGACTACT TCATAGGCCT GGATGTGGGG
CGGATGCCAA AGAAGAATCT TCCGGGGAGC CTCAACGTGT GCGCGTCTGT CAGGCTCTAT
GGCAAGCAAG GCGAGTTCGT GCGCTGCCGC GTCGAGGACA GCTTGACCGA GGGCGAAGAG
ATTCCCCAGC GGATCCTGGA AAATTGCCTG CCCCAAGCAG AACTTAAAAA CCAAACTGTC
CTTATCTACA GAGATGGTAA ATTCCAGGGA AAGGAGGTGG ATAACCTTTT GGCTAGGGCT
CGCGCAATCA ATGCCAAGTT CATACTGGTT GAGTGCTACA AGACCGGTAT CCCCCGACTG
TATAACTTCG AGCAAAAACA GATCAACGCA CCCTCCAAGG GGCTGGCACT CGCGTTGAGC
AACCGAGAGG TGATCTTGAT TACGAGCCAA GTGAGCGAGA AGATAGGCGT TCCTCGGCCA
CTTAGACTCA AAGTGAATGA GCTGGGTGAA CAGGTGAACC TGAAGCAGCT GGTCGATACC
ACTCTTAAAC TCACGCTGCT CCACTATGGA TCTCTGAAAG ACCCACGGCT GCCTATTCCC
CTGTACGGTG CCGACATCAT AGCCTATCGG CGGCTGCAAG GAATCTACCC ATCCCTTCTC
GAGGATGATT GTCAGTTCTG GCTGTAGTGA

33 Gene Name: gene-4461

Gene ID: BBF.2017.48.33

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AAAAGCGCAA GGTAGAGGAC CCTAAAAAGA AGCGGAAAGT TGGCAGCGGG
TCAATGAACA CGCCTTTGAC GCATTACGTG CTCACCGAGT GGGAATCCGA TACAAATACT
AATGTATTGC ACATCCACCT GTACACCCTC CCCGTTAGGA ACGTGTTCGA GCAGCACAAG
GAGAACGGTA ACGCATGTTT CGATCTTCGC AAGCTGAATA GGAGTCTGAT CATCGACTTC
TACGACCAAT ATATCGTGAG CTGGCAGCCT ATAGAAAACT GGGGCGAGTA CACCTTCACC
CAGCACGAAT ACCGCAGTAT AAACCCAACA ATACTGGCCG AGAGGGCCAT CCTCGAACGA
CTCCTCTTGC GGACAATCGA AAGCGTCCAG CCCAAGAAGG AGATCGCAGC TGGTTCCCGC
AAGTTTACCT GGCTGAAGGC AGAGAAGGTC GTGGAGAACA TTAGCATCCA CAGGGTAATC
CAGTGCGACG TAACCGTGGA CTACGCCGGC AAGATCTCTG TGGGCTTTGA CCTCAATCAC
AGCTATAGGA CAAATGAGAG CGTGTACGAC CTCATGAAGT CTAACGCCAT CTTTAAGGGC
GACCGCGTGA TAGACATTTA CAATAATCTG CACTACGAGT TTGTAGAGAT TTCCAACTCC
ACAATAAATG ACTCCATCCC CGAGCTCAAC CAAAGTGTCG TCAACTACTT TACGAAGGAG
CGAAAGCAAG CATGGAAAGT GGATAAGCTG GAACAGAGCA TGCCAGTCGT GTACCTCAAG
GCATTCAACG GCAGTAGGAT TGCATACGCG CCTGCCATGC TCCAAAAAGA GCTGACCTTT
GAGAGTCTCC CGACCAACGT AGTACGGCAG ACGTCAGAAA TATTCAAGCA AAATGCCAAT
CAGAAAATCA AGACCTTGCT GGATGAAATC CAAAAGATTC TTGCCCGCAC CGACAAGATC
AAATTCAACA AGCAGAAGCT GTTGGTTCAG CAGGCCGGCT ACGAGATACT TGAACTGTCC

AACCCAAACC TCCAGTTTGG GAAGAACGTT ACTCAGACGC AACTGAAGTA TGGACTGGAT
AAAGGCGGAG TTGTGGCCTC CAAGCCGCTC AGCATCAATC TTCTGGTCTA CCCGGAACTT
ATAGACACCA AGCTCGATGT GATCAACGAT TTCAATGACA AACTGAACGC TTTGTCCCAC
AAATGGGGCG TGCCCCTGAG TATCCTGAAG AAGTCTGGAG CGTACCGCAA CAGACCCATT
GATTTCACTA ACCCCCACCA GCTCGCGATT CTGTTGAAGG AACTGACCAA GAACCTTTTC
CAGGAACTCA CGCTTGTGAT AATACCGGAA AAGATCAGCG GCATGTGGTA CGATCTGGTT
AAAAAGGAAT TTGGCGGCAA TAGCAGTGTT CCGACGCAAT TTATCACCAT CGAGACACTT
CAGAAGGCAA ACGACTATAT TCTGGGGAAT CTGCTCCTTG GCCTCTATAG CAAGTCCGGC
ATCCAACCAT GGATTCTTAA TAGCCCCCTT AGCTCCGACT GCTTCATCGG TCTGGACGTA
TCACATGAGG CGGGTCGCCA CAGCACCGGG ATAGTCCAAG TCGTAGGAAA GGACGGGCGC
GTGTTGTCAT CCAAGGCGAA TACGAGCAAT GAAGCCGGCG AGAAGATCCG CCACGAAACC
ATGTGCCAAA TAGTGTATAG CGCCATCGAC CAGTACCAGC AACACTACAA CGAGAGGCCT
AAGCACGTGA CCTTCCACCG CGACGGTTTT TGCAGGGAGG ATCTGCTGTC ACTCGACGAG
GTGATGAACT CCCTGGATGT CCAGTACGAC ATGGTGGAGA TCATCAAGAA AACCAATCGG
CGAATGGCAC TGACCGTCGG CAAACAAGGA TGGGAAACCA AGCCAGGACT GTGCTACCTG
AAGGACGAGA GCGCCTATCT GATCGCCACC AATCCGCACC CGAGGGTGGG CACCGCGCAA
CCCATCAAGA TTATCAAGAA GAAGGGGAGC CTCCCTATCG AGGCCATTAT ACAGGACATC
TACCACCTGA GCTTCATGCA TATCGGCTCA CTGCTTAAGT GCCGACTCCC CATCACAACT
TATTACGCCG ATCTGTCTAG CACCTTCTTT AACCGCCAAT GGCTTCCGAT CGATAGTGGC
GAGGCCCTTC ACTTCGTGTA GTGA

34 Gene Name: gene-8853

Gene ID: BBF.2017.48.34

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAGCGGAA AGTTGAGGAC CCCAAGAAAA AGCGCAAGGT GGGCAGCGGC
TCCATGCTTA TCTGGCAATT CAAGAGAATG CTCTACTGCC AGGCCAACAA CATCAAAGAG
GAAAAATTCA AAGACCTGGA GAGCGAGCGA AATCAAAACA CTATCCAGAG CTATTTTGAC
CTGAAGGGCG GCTATCCGGA AAGATATAGC CAGGAGGAAT ACTCCGCTTA TTTCGAGCAT
TGCTTCCCGA AGTCTATCAA CCGGAAGTAT TTCATGCAGA AAATAGTAGA GGGCCGAAAT
CCGAGCATAG GTCACAAGTG TTTGGGTGCC CTGTTCGACT GCAAAAAGGT AAACCACATC
TGGACAACCA ACTTCGACGA GCTCATCGAG AATGGGATTA AAAGCGTCAA CAATGCCAGC
AGCTTCGAGG TCATTAGTAT CGACAATCAG AGGCAGCTGG CCAACCTCAA CAACTACCCA
AGGGTGGTAA AACTTCACGG CGACTACAGG TACGACAAGC TCCAAAATAC CGTTGACGAA
CTGCAGACGC TGGAGAAGGA CCTCCATAAG TACTTCGCCG ATGTGCAAAG CAAGACCGGC
TTGATTGTGA TAGGCTACGG CGGAAACGAC CAGAGCATCA TGTCCGCCTT TGAAAAGACT
TTGGAGGCCG ACAACCCGTT CCCGTTTGGG CTTTACTGGT GCGTGAGGAC GGGCCAGAAA
ACCAACAAGA AGGTAATCGA ATTCATAGAG AAGGTTCACC AGAAGAACAA GGAAAAGCTT
GCTGCGTTCA TCGAAATCGA CTCTTTTGAC GATTTTCTTT ATGAGCTGTA TAAGACGAAC
AACCTTGCCA ACGATCACAT TGAAAATATC GCCAAAAGCC GCTTCGAAAA AGGAAGGCT
TTTACAGCCC CCCAGATCGG CACCTCCTTT ACGCCTATAA AGCTTAACGC CATAAAGGCC
AAGACTTACC CGAAAAGCAT CTATTCCTTT AAAACTGACC TCAAGGGGGG CAAGGATGAC

TGGGATAAAC TCAGGGAAAT CATTAAGGAC CAACCGGTGA GCGCGGCTCT GACCAATGAA
AACACGGTCG CCTTCGCAAG TGTCAACGAC ATCAAGAAAC TCTTCTCACA CACACTGAAG
TCAGAGATCA CCACCGTGGA CATAGATGAC AAGTTGATCT ATCGGCAGGA GTCTTTCTAC
CTGGGCATGC TTTACGATCT GATAGAGCAC AACCTCCTGA AGAAGTTCAA GTTGGAGAAA
GTGCCCAACA ATAGGCTCCG CAAGTATTAT AGCAAAAACT ACAAGCTGAA TACCGAGGAG
CTTCAGAAGT CCAAGATCAA GACCAGCCTG TCCGTCTACG AAGCGTTCGA GATTCAAATA
GAATTCCACA ATAAAGAGCT GTTCCTCATT ATCCTTCCGT CCATCCACAT AGACGACAAA
GCCGGGCTGA GCCGATTTGA GAAACAGGAG ATAGCCAATA AGATCATAAG CAAAAGGTGG
AACCGCATGG TTAACAACCA GCTTAGGTTC TGGCTGGGGC TCCTTAAGAA CGATAACACT
AACATAGAGT TCAGCATCGA CAGTTTCAAG ATTGATTTGG AAGAAAAGTT CTCCGGCGTC
GGGAGCTTTA CATCCTCTTA CTACATCTTT AAGGGCGCGT TTATTTCCAA CGAACCCAAG
CTTAGCTTCC ATATCTCCGA CAGCAATTAC AAAACAGTGC ACCCCCTGAA AGGCCTCAAG
AACTTCGGTC CACTGGATTA CTCATTTGAA AGCAAACAGA CCAATCAGCA GGCTATTAAA
CTTGGTATAA TCACTCCGAT CAGCGGCATG CAACGGATAC TCAAACACCT GAACGAACTT
AATAACGAGA TCCGCGCAGC TACGGAAAAG GAGTACCTGA CCGATTATTA CCCCTTTAGC
AACATCTACA AGAGATACCT TGACATCCCG CAGAATAAGG ATAGTAAATT CTTGGAACTC
GTGAATGAAG CCGAAGTGAA CAAACTGAAC CACCTCGAGT TTTATGACTT CCTCAAACGC
AAAATTGATT ACTTCTATAC AATTAGGGGC GAGTTCGACG TGCTTGTGTT GTATTTTCCC
AAAGGCTGGA CTAAGTTCCG CGAGCTGAAA AATGACAGTG TCTACTTTGA TCTGCACGAC
TCCATCAAGC TGTACTGTGC TAAGAAGAAT ATCAAGATCC AATTCGTGGA AGATAAGAGT
ATAGACTACC TCGACCCGGC CAAGGTTAAA TGGTGGTTGA GCCTCGGCTT GTATGTCAAA
GCGAACGGGC TGCCCTGGCG GAACGTGGTC GTAAACGAAA GCACCGCGTT TGTCGGGCTC
GACTTCGCGG TCCAGCGAAT AAACAACAGT AACAAGTACG TGCTGGGTAG CTCACAGATC
TTCGACAGCT CCGGACAAGG ACTCAGGTTT CTGTTGCAGC CCATCGAACA CCCTGTGTTT
ATCGGTAAAA ACCCCTTCAT GAGCAAGGAA GATGCGCGAC GGATGATTCT TAAATTGAAG
GAAGCGTATT TTAGGATTGA CGGTAACTCC AAGCTGGAAA AACTGGTGGT GCACAAAGTA
CTGCATTACA CAAATGATGA GATGACCGGC ATTTCCGAGG CGCTGGAAGG TATTGAGAAC
ATTGAGCTTC TGCAAATACA GAAGTATAGT AAGTGGAGGG CAATTAGAGG GGACATCGAT
CGGTATACGG GAAAGGTTAA GACCGACCCG CACAATTTCC CGATCCAACG GGGGACAGTG
ATCCAGCTCG ACGACTTCTC TTTCCTTCTG TGGACACATG GAAGTGTACA GGAGGACGAC
GTGGCTGGTA GGCACATGAA TTACTACCAG GGTAAGCGCG GGATTCCCGC ACCACTTCTC
ATACGGAGGT TTCGCGGCAC CGATCCGATT GAAATGACCG TGCGAGACAT CCTGTCACTC
ACCAAGATGA ACTGGAACGG AGGCGAACTT TACAAGACTC TGCCGGTGAC CCTGGATTTC
TCTAAACGGC TTTCTAAGTA TGCGAAGCAG GCAGAAACCC TCCAGGCAAT ACCCTACGAC
TTTCGGTTCT TCATGTAGTG A

35 Gene Name: gene-12442

Gene ID: BBF.2017.48.35

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGCGAAA GGTAGAGGAC CCCAAAAAGA AACGCAAAGT GGGCTCCGGA
AGCCTGAAGC TGAACCACTT CCCCCTTAAT CCCGACCTCC CCCTGTACAT CACAGAATAT
GCCCACCGGA ACCCGCGAGC GTTGCTCGGA TTCGTTAGGG GCCAAGGTTT CTGGGCGCAA
CAGGTCGGAG AACAGGTACA AGTGTACCAC GGTAGACCGC AGCCCACGTT CAGGGGAGTT
CAGGTGATCA GCCATACCAG GTTGGACCCC GACCATCCGG CTTTTGACCA AGGCGTTTTG
AGCCTCATCC GACAAGCACT GGTGAGGGCG GGATACGTGC TGACCTACAG GGAGAGGATG
GCTATTCATC CCAGACTGGA GAGGGTTGTG CTCAGACCCC CGGACCGGCA CCCAGCAGAG
TTGACCGTCC ATGCACATCT TCGATGGGAA TGGGAGCTTG AAAGGCACAG CGGACAACGC
TGGCTGGTTC TTCGACCCGG CAGGCGACAT CTGAGCGCCC TTCCATGGCC CGCAGAAGCA
GTACAAATGT GGTCCGCCGC ACTTCCGGCC ACATGCCAGA AGCTGCACGC CCTTTGTCTG
GACCGAGGCC AACAGATGGC CCTTTTGCGG CAAGAGGACG GCTGGCACTT CGCCAATCCC
GGTGCTGCCA CTCAAGGAAG GTGGCACCTG TCCTTTAGCC CCCAGGCCCT TCACGAGCTG
GGACTGGCAC AGGCTGCGCA CCATGCGGCT GCATTTAGGT GGGACGAGGT ACAGCGACTC
GTGCAACTGA CTGACCTGTG GAAGCCCTTC GTGACCTCTC TGGAGCCCCT TGAGGTAGCT
GCCCCCATCA TTGCCGGGAA AAGGCTGAGG TTTGGACGGG GTCTTGGCAG GGATGTCACG
GAGGTGCACA AGCGAGGTAT CCTGGAACCA CCCCCACTGC CCGTGCGACT GGCTGTCGTG
TCTCCCCATC TTCCTGATGA GCACGCGAAC GCCCAGTTGA GGCGGGAGTT GCTTGCTCAC

CTCCTCCCGC GACACCAAGT ACTGAGATCA GCGGAGAGCC GGCAAGGCCT CCACGAGCAC
CTGAGGAGGC AAGATCAGGA CGATACCCTG TATACCTTTT GGTCAGGCGG CGAGTACAGG
AAGCTGGGCT TGCCCCCCTT CGATCTCGCA CGAGGCCTGC ACACCTACGA CCCAGCTAGC
GGCCAGCTGC AACAACCGGC TGCCCTGGCA CCAGCACCCG CGCAGGCCAC GCAAGCGGGT
AGGCAGCTGA TAGCCCTGGT GGTGTTGCCC GACGACCTGA CGCGGTCTGT CCGGGACACC
CTGTTTCAGC AGCTCCAGCA GTTGGGCCTT AGGTGTCTGT TTAGTGTGAG CAGGACCCTG
CTGCACCGAC CACGCACAGA GTATATGGCA TGGGTAAACA TGGCCGTCAA GTTGGCTAGG
ACTGCAGGGG CCGTGCCTTG GGACCTGGCA GATCTGCCCG GTGTCACCGA GCAGACGTTT
TTCGTAGGCG TTGATCTGGG GCATGACCAC ACCCACCAAC AGTCCCTCCC GGCCTTCACC
CTGCACGACC ATAGGGGACG CCCTCTTCAA AGCTGGACGC CTCCCCGACG CACCAATAAT
GAGAGGCTGT CATTGGCCGA GCTTAAGAAG GGGTTGCATA GGCTTCTTGC ACGCAGGAGC
GTGGACCAAG TGATCGTGCA TCGGGACGGC CGATTCCTTG CTGGCGAGGT GGACGACTTC
ACTCTGGCGT TGCATGATCT CGGCATCCCG CAGTTTAGCT TGTTGGCAAT CAAGAAAAGC
AACCACAGCG TGGCGGTGCA AGCAGAGGAA GGATCCGTGC TTAGCCTGGA CGAACGACGA
TGCCTTCTTG TTACTAATAC CCAAGCCGCG CTTCCGCGGC CCACGGAGTT GGAACTGGTC
CATAGCGACA GGCTTAGTTT GGCGACCCTG ACCGAACAAG TATTCTGGCT GACCCGCGTG
TTCATGAACA ACGCGCAGCA TGCGGGCAGC GATCCAGCCA CCATCGAATG GGCCAACGGC
ATAGCCAGGA CTGGACAGCG AGTGCCCCTG GCCGGGTGGC GGCTGTAGTG A

36 Gene Name: gene-12924

Gene ID: BBF.2017.48.36

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AGAAGCGAAA AGTGGAGGAC CCTAAGAAGA AAAGAAAGGT GGGCTCAGGG
AGCATGGAGG CGTACATAAC GGAGATGGTG TCCAGGGAGA GGGCCAACGA GCTGGAGGTT
TACGTGTACG TGTTTCCACG GAAGCAATCC GACAACAACT ACGAGGGTGT GTATCACATA
ATGAGGGCGT GGCAACGGGC TAATGATCTG CCTCTGGCGT ATAATCAACA TACGATCATG
GCATTTTCCC CCGTGAGGCA TATGTGTGGC TACACGCCGA TGGAAACGCA GAAACGCCAT
ATTAACATTG ACTCCCCATT CGAGAGAGCC CTGCTGGAGC GACTGATAAA GAACAGCCTG
ATTTTTACAG CCGAGCGCCA TTTGCATGCC AAGCGGGTAG GCCATGCGCT TCGGCTGAAC
CAGGTGCAGC AAATCAGACA GGTGATCATC TATGAGGCCA TCGAGCTCTA TGTAAATATC
ATTGAGAATA GAATAAGCAT CGGCTTTCAC CTCACCCACC AGTTCGAGTA CGTATACACT
CTCCAGAGCA TGATAGAACA GGGAAAAACA ATCAGACCTG GAATGCGCGT CGTGCATTCT
AACGGAAGGC AGCATTATAC CTACACCGTG GAGAACGTAG CAACATATGG GGTGACCGAC
AGATGCCCGC TGCTGCAGAC CAGCATTTAC CAATACTACG TCGAAAAAGG CGCGCAGCAC
ATTTTGCGCA CCTTCACCCG ATCCACCAGG GTGATCCACG TAAGAACGAA AGAGCAGAGG
TTGAGCTACG CGGCGACACT CCTGAAACCG CTGTGTACTT TTGAAACCAT GCAACCCCAG
GACGTGCTCA ATGTCAGCAA GTGCATCAAA CTTAGCGCGA GCAAACGAAT GAAATGTACT
TACAGGTGGA TTCAGCAACT CCGGGCACAG TACCGACACC TGACCTTTGC GCCGAACCCC
TTCACGATCG CCCAGAATGG CTATAAACTT GATCAGCTCA GCACCCCCAA GGTGCACTTC

CACAGAGACT ACGCCACCGT CGTGAGCGGA ATGAAAACCG GCAAGCTTTA CAAAGGCGGT
AATATCAAGA TCAGCGTTCT CTTCGACGAG GACTTTTACT TGAAACACCA CATCACCAAG
AAGGACATAT ATCAATTCAT TGCAGTCCTG CAGAAAATCG CTATCGCACA AGGCGTGAAC
ATGACCATAA GCACGAGCAC CAAGTCCATT ACGGGCAAGT TCACGGACGA CTTTTTCCAC
CACTTCACCG AGGAGGTCGA AGCACTGCAG CCCATCTTCG CGCAAACCAC AGTTCTGGCA
TTCATTACCA GTACCCACCT GAGCAACAAG AAAACCAGGA GTTACCAGCT GCTGAAACAG
TACTTCGGCG GCAAGTGGGA CATTGCCTCT CAAGTCATCA CGGAAAAGAC GATTGAGGCG
TTCCAAAAAA TCTTGCACAA GCACGGCCTG AAGAATTTCT ACCCCAATGA CGAACAGCAC
TGTCTCCGCG TGATCGATGT CCTCAAGAAT GAGAGCTTCT ACTACACGGT CATGAACATC
CTCTTGGGAG TATATGTGAA AAGCGGCATC CAGCCCTGGA TCCTTGCTAA TACAACCCAC
TCAGACTGCT TCATCGGCAT CGACGTTAGC CACGAGAACG GAAACTCTGC GGCTGGGATG
ATGAATGTTA TCGGCAGCCA GGGCCACCTT ATCCAACAGG CGCCCCTGAA CGGCATATTG
GCGGGAGAAA AGATTGACGA CACCCTGCTC GCAAACTTGC TTAAACAAAT GATTAAGGCA
TACCACACCC AGTTCCAGCG CTTTCCCAAG CATATAACAA TCCACAGGGA CGGCTTTTGG
AGAGAACACA CTGCACTGGT CGAGAAGATC ATGAGCCACT ATGAGATTAC CTACGACATC
GTCGAGATCA TCAAAAAGCC TAATAGGAGG ATGGCTTTCT TCAACAGCGT GGACAACACC
TTTAGCACCA GGCAGGGGAC AGTGTACCAA CGGGGCAACG AAGCCTTTCT GTGCGCCACT
AACCCTCAGC AGAAAGTGGG CATGGCACAA CCAATCAAAA TACATCAGGT GACCAAGACC
CTGCCCTTCT CACACATCAT AGAAGATGTC TACAACCTCA GCTTCCTTCA TATTCACGCT
ATGAATAAGA TGCGACTGCC GGCCACCATA CATTATGCCG ACCTGTCTGC CACCGCTTAC
CAGAGGGGCC AAGTGATGCC CAGGAGCGGT AACCAGACAA ATCTGCCTTT CGTGTAGTGA

37 Gene Name: gene-13895

Gene ID: BBF.2017.48.37

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAAA AGAAACGCAA GGTAGAGGAT CCCAAGAAGA AAGGAAGGT GGGGAGCGGG
AGCGTTCACG CATTGCTCGC TCTGCTCGCG AACCGAGCCG GTGGAAGGAC CGCCAGAATG
GGAGACAGCT TGCTCACGTG GAGCCCTCCT GAGTCTCTGC TGCTTGAAGG GACCCTGAGC
TGGCGCGGCA ACACCTACAC ATACCGGCTT CGCCCACTGG CGAGAAGGGT GCTCAACCCT
AGGAATCCCA GTGAGCGCGA CGCCTTGTCC GCGTTGGCGC GACGACTCCT CCGAGAAGTG
CTTGAGCAAT TCAGGCGCGA GGGGTTTTGG GTTGAAGGTT GGGCCTTTTA CAGGAAGGAG
CACGCACGGG GTCCCGGGTG GCGCGTGCTG AAAGGTGCGG CGCTGGATCT GTGGGTTTCA
GCCGAGGGGG CCATGGTATT GGAGGTGGAT CCGACTTATC GAATCCTGTG TGACATGACA
CTCGAGGCGT GGCTTGCACA GGGACATCCA CCCCCGAAAC GCGTCAAGAA CGCGTACAAC
GACAGGACAT GGGAACTCCT GGGTCTGGGT GAGGAGGACC CGCAAGGCAT TCTTTTGCCA
GGCGGGCTGA ACCTCGTCGA GTACCACGCT AGTAAGGGCA GAATCAGGGA CGGCGGGTGG
GGTCGGGTTG CGTGGGTGGC AAATCCTAAA GACGCCAAAG AGAAGATCCC GCATTTGACG
AGCTTGTTGA TCCCCGTCTT GACCCTGGAA GATCTGCATG AAGAGGGGGG CTCTAACTTG
GCCCTCTCCA TCCCGTGGAA TCAAAGGCAA GAGGAAACCC TTAAAGTGGC CCTGTCCGTG
GCTCGCCGAC TCGGCGTCGA ACACCCCAAG CCCGTCGAGG CCAAAGCCTG GAGGATGAGG
ATGCCAGAGC TTCGCGCACG ACGCAGGGTG GGTAAGCCAG CGGACGCCCT TAGAGTGGGG
CTGTACCGGG CTCAAGAGAC TACCCTCGCA CTGCTTCGGC TCGATGGCGG CAGAGGATGG

CCTGACTTTC TGCTTAAAGC ATTGGAGAAC GCTTTTAGGG CCAGCCAGGC TAGGCTTCAT
GTTAGGGAAA TCCACGCGGA TCCTAGCCAG CCCCTTGCAT TTAGAGAAGC CTTGGAAGAA
GCGAAAGAAG CTGGTGTGCA GGCTGTCCTC GTACTCACGC CCCCACTGAG TTGGGAGGAG
CGACACCGCT TGAAAGCACT GTTCCTCAAA GAAGGACTCC CAAGTCAACT TCTGAACGTC
CCCATACAGA GGGAGGAAAG GCATCGGTTG GAAAACGCCC TGCTCGGGCT CCTGGCGAAA
GCGGGGCTCC AAGTAGTCGC CCTTGAGGGC GCATACCCTG CTGATTTGAC AGTTGGATTT
GATGCCGGAG GCCGCAAGTC CTTTAGGTTC GGAGGTGCCG CATGTGCTGT CGGCTCCGAC
GGAGGTCACT TGCTGTGGAG TCTGCCGGAA GCCCAAGCGG GCGAACGGAT ACCAGGCGAA
GTAGTTTGGG ACCTGTTGGA GGAGGCGTTG CTGGTGTTTA AGAGAAAAAG AGGGCGGTTG
CCCAGCCGGG TGCTTCTGCT GAGGGATGGC AGGCTTCCCA AGGACGAGTT CACCCTGGCA
CTTGCAAAGC TGAGGCAGCT CGGCATTGGC TTCGACCTCG TGTCCGTAAG GAAGAGTGGA
GGCGGAAGGA TTTATCCGAC CCGGGGAAGA TTGCTTGACG GCCTTCTGGT GCCCGTTGAA
GAGAGGACTT TTTTGCTCCT GACGGTGCAT AGGGAGTTCA GAGGCACCCC ACGGCCCCTC
AAATTGGTAC ACGAAGAAGG TGAGACACCT CTGGAGGCTC TCGCAGAGCA GATCTACCAC
CTGACGAGGC TGTATCCTGC ATCAGGTTTC GCATTTCCCA GACTGCCCGC ACCCCTGCAC
TTGGCAGATA GGCTCGTGAA AGAGGTGGGC CGATTGGGCG TGAGGCATCT CAAGGAAGTA
GACAGGGAAA AGCTGTTCTT TGTATAGTGA

38 Gene Name: gene-14034

Gene ID: BBF.2017.48.38

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAGCGAAA AGTGGAGGAC CCAAAAAAGA AAAGGAAGGT GGGTAGCGGC
AGCATGAACG CCGTGACCGT GGGCAGCACC CCAAGCGCCC AGGTACTCGT CGGTGTTCAG
CCATACGACG AAACCACCCT GGAGAGCCTG AGAAGTAAAC ACCGCGGAGA CTATCTCTTT
AAAAGAGGGG GAGAGAACGG CGATAGCATA CTTGCTGTGG CCCTGAAACC GAGTCTGCCG
GTCATCGGAG CAACCGAGGA GGATGTAATT CTTGCCGAGA GCCCATGGTT GTTGGCTCCA
CTTGCCTTGG AGACTTTGCT GCAATGCTTC GTGAGGCTTC AAAGGCCCAT CCTGAAAGCT
AGGCATCCCC TGAGAGTGCT CTCACAAAAA CCGGCAAATC TTTTCCCAGC CGATGCGGGG
GTCCCCCAGT GGCTGCAGAG GAGACTGGTG CTGGAATTCG ACACGCGCAC TGTTAGGGAC
AGGTCAGACG CTGCCTCTGT CGTGCTGGCA TGTGGCGTGA GGACTCGGAA TTTGATTGAT
GCCGACTGCG CGACACTGAT AGCAGCCGGT GTCCCCCTTG TGAATCGATA CGTGGTGACG
AGGCACCCTG CGGATGATCC CCGAGTGCAG GGCTATTTGA GGCTCGCCGG GAGGGTGACC
AGGATAGATG GCCCCAACCT GTACTTGGAG GATCATGGGG ATGGAGCAGC TGTGATCAAG
GCCTCCATGG CCTATCTGGA GCCCAGGAGG GAGAACGTGA TTTGGTGTGC CCACCATTTG
CTGGGGAGAA ATGCGGATAG AGTACTGGCG GAAGCGGATA ACGCAGCCGC AAAGCACTTG
AGCGGTCCCG AACGATTGGC CGTAGTGAAA AAGACTTTCG ACTACCTTAG GAGCCAGAAC
ATCGAGCTTG CGCCTGGAGT GCCCCTCACT CTGGGTAACG TTGTGGGGAA TGACAAGGGT
TCTTGGATCT TCCGGACGGA AACTCTGCCC AAGCCCCACC TGGTGTTCGA CCCGAGCGGG

ACCCGGATCG ATAGGTGGAA TGAGAGGGGA TTGGACGCTC ACGGGCCCTA TGATCAAAGG
ACCTTCACCC CTAAACAACT GAGGATTGCC GTCATATGTC AACTGCCCTA CGAAGGCCAG
GTCGATGCGT TCCTGGCAAA ATTTCTCGAC GGCCTTCCAG ACGTTAAGAC CGGCTACGGG
GACCGGGCCA GGGCGCCTTA TGCCAAGGGG TTCATCAGGA GGTACGGTCT GGAGAAGCCC
AAGGTGAGCA CCTTCGCAAC AAAAGGCGCT ACTGCTAAGG ACTATGCCGC TGCATGTAGG
GCGGCTGTGG AGGACGCAAC CGCAAGCGGC TTCGAGTGGA ATCTGGCTAT CGTGCAGATC
GACAAGGATT TCAAGGAGCT GAGTGACGTG GAGAATCCCT ACTTCACCAC CAAGGCCCTG
CTGCTGAAGC ATCGGGTGCC CGTCCAAGAG GTGACGCTGG AAACGATGAG GTTGGCAGAC
GAACAGCTGG TGTACGTGTT GAACAACATG AGCGTAGCCA CCTACGCCAA AGTGGGCGGT
ACTCCCTGGC TCTTGAAAGC GCAACCAACC GTGGCCCATG AGTTGGTAGT TGGAATCGGA
AGCCAGACTT TTAGTGCCTC AAGGCTGGGT GAGAAAGAGA GGGTTGTAGG CCTTACCACC
GTGTTCTCCT CCGACGGGAA ATATCTGCTG GACGACCGGA CTAGCGCCGT TGATTACGAC
AACTATAGCG AGGAGCTGTT TAAGAGCTTG TCCCGGTCAA TAGAATCAGT AAGGATCGCC
GATAACTGGC GAAGTACGGA CAGTGTCAGG CTGATTTTCC ATGTTTTCAA GCAGATGGCG
GACGAGGAAG CCGACGCGGT TGACAAGTTG GTGCAAAAGC TGGGTTTGGC ACAGGTTAAG
TTCGCGTTTC TGCACATCGT GGATGACCAC CCATTCGCCC TGTTTGACGA GAAGAACATA
GGTACAAAGA CATGGGGTGG GATATTCAAG GGCGTCTTGG CACCGGAAAG GGGCCTCGCG
GTAAACCTCT CTGGGGCCGA AACCCTGTTG TGCTTCACAG GCGGCAGGGA ACTGAAACAG
GCGAAGGATG GCCTGCCCGT GCCTAGTCTG CTGCGACTGC ACCACAGGAG TACGTTCAGG
GACATGACCT ACCTGACGGG GCAAGCCTTC AACTTCAGCT GTCACACCTG GCGCATGTTC
ACACCCGCTC CTGTTCCCAT CACAATACAT TACAGCGAGC TGATGGCGCG ACTCCTTACG
GGCCTCAGGC ACGTCCCGGA TTGGGATCCA GACACAATGC TGACCCCCAT CAGTCGAACC
CGGTGGTTCC TGTAGTGA

39 Gene Name: gene-14606

Gene ID: BBF.2017.48.39

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAAAGGAA GGTAGAGGAT CCAAAGAAAA AGCGGAAGGT TGGAAGTGGA
AGCCTCCCCA TCGTCCTGAA CGCCTTCCCA CTTAAAGTAC CCGAACTGGA GCTGGAAGTT
AGGCAAATAC CGTACGATAA AGAAACGCTT GACGGCCTCA GGGCTGCGCA CAAGGCCACC
CACGCTTTCC GCAGGCAGGG CGACAACATA CTGATTTTTT CCGGTGATGG CACATTTCCC
GCGTCTGGGA CGCCTCAAAC TATTGCACTG AAGGACAATT TCGGCGTGTT CTACAGCCTC
GTGAAGGATG GTCTTATCCG CCACCTTGCG GGGCTCGGGA GGAATCCCAG CGGGTTCAAC
CCCATAGAGT TGGTGTCCGC AAAACCCGAG GACAATCTGC TGGTCCCCAT ACTCGGGGAT
GCGTATCCTT TTAAGGTGTG CGCGAAATAC AGCATTGACA CCAGAACCGT GCTGGGGCAC
CCATGTCTGG TGATCGATTG CACGACCAGG AGGGTGTTGA AGGAAAATGG CTTGTTCTTT
TTGAACGCTG GGTTCGACCT CGCGGGAAGG TACGTGGTGA CGGAGCAAGA TGACGGGTAC
AGGAAATTGC TCGGCAGCGT GAGCGGCTGT AAGGGTGAAA CGCTGTACGT GACTAGGCCC
GATGGCCAAG TGGTGCAGGC CGAGGCTAAA AACGTGTACC TGGAGGCATC CCGCACAAAT
TTCGACGACT ATATTCTGCA CACCCACAGG GCTCAGAAGG ACGCGATCGT TGAACGAATC
AGACAGTCCG TTTCCGTGTT TAATGGGGGC GAAAATAAGA AAGCCCGAAT CGACACGCTG
AAGAAGTATA TCCAGTCCAA AACCATTCCC TTGATCGACG GCACCAGGAT TGAGATCCAA
GATTCCCCTA ACATACAGAA AGACTGCGGC CAGATGCAAA AACCGGTATT CGTCTTTAAC
GACAACGGCG AGGCGGACTG GGCGGAGAAG GGGCTGACCC AATCTGGGCC GTACACCAAG

AGGACCTTCG ACAGGAATGA CCCCTCCATT TGCGTGATCT GCGCCCAACA TGACAAGGGA
CGCGTTGAGC AGTTCGTCAG GAAGTTGCTT AAGGGCATTC CAAACTCCAA ATACTTCAGC
AACGGACTCG AGGGGAAGTT TACCCTGGGC ACTAGCAGGG TAGAAGTGTT CGCGACCGCT
ACTGACAGCG TAGACGCCTA CAAGAACGCT ATTGAAGCCG CAATACGGAA GAAGGCCGAC
GACGGCGGAA GGTGGGACCT GGCCCTGGTT CAAGTGAGGC AGAGCTTTAA GAAGTTGAAA
GTGACCGAGA ACCCCTACTA CCTTGGCAAA AGTCTGTTCT TCCTCCACCA GGTGCCCGTC
CAGGACTTTA CCATTGAGCT GTTGGCTCAG TCCGACTACT CCCTCGGCTA CTCTCTGAAT
AACATGGCCC TTGCATGCTA CGCGAAGATG GGCGGTGTGC CCTGGCTGCT TAAATCTTCA
CCCACCCTCA GCCATGAGCT TGTGATAGGC ATCGGCTCCG CCAACATCGG CCAGGAGAGA
GGAGCTGATA ATCAGAGAAT TATGGGCATC ACCACTGTGT TCAGCGGCGA CGGCAGCTAT
ATCGTGAGCA ATACATCTAA GGCTGTTGTC CCCGAAGCTT ACTGCGAGGC CCTTACCGCC
GTACTTGGCG AAACCATCGA AAAGATTCAG AAGAGGATGA ACTGGCAGAA GGGCGATACC
ATCAGATTGA TCTTCCACGC TCAGGTCAAG AAATTCAACA AGGAGGAAAT CGAAGCGGTC
AGAGCCGTCA TTGAGAAATA TCGGGAATAC CAGATCGAGT ACACTTTTCT GAAGATAAGC
GAAAACCACG GGCTTCACAT GTTCGATAGT GCAACCGCAG GGGTGCAAAA GGGCCGACTT
GCCCCTCCGA GGGGCAAGAC GTTCAAGCTG AGCAAACATG AGATGCTGGT TTATCTGATA
GGGCAGAGGG AGCTGCGGCA AGACACCGAT GGTCATCCCA GGGGCGTCAT CCTTGATGTT
CACAAGGACA GTACATTCAA AGACATCACC TACCTTTCAG CCCAGCTCTA CTCATTTGCC
AGCCACAGCT GGCGCTCTTA CTTTCCCAAC CCTATGCCAG TAACCATTTC ATACAGCGAT
CTGATCGCTC GAAACCTTGG TTGGCTGAAC CAACTGCCCG GGTGGAACGA CTCCGTGATG
ATCGGAAAGA TCGGGCAAAG CCAGTGGTTC CTGTAGTGA

40 Gene Name: gene-20356

Gene ID: BBF.2017.48.40

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AGAAACGAAA GGTTGAGGAC CCCAAAAAGA AAAGGAAGGT GGGGAGCGGC
AGCATGAATA ACATACCCAT CAGGCTGAAC TTTTTCGCCC TGAAGAACCA GAACATTAGC
TTCAGGATCT ACAGGCAGGA CTTCAACGGC CAGAAAAAAC AGGACGGGTA CTACAGGACC
AAGCTGCCCA TCAACGACTC TTCTGACACC TACGCGGAGT ACTGGGTGAC AACCCAGCCC
AAGGATGGCT TCGAGAGGGT GTACTGCCTG GGTTCCTCAA ACCCTAAGCT CACCGTCCGA
ATCATGTGGG AGAGCTTCCT GGATAGGGTC CAGAAGTCCC TGAGCTCCGA CGAATATATC
CTTTACGGTA ACGGATTTAG CCGGAAGGTC GCCGTGATCA TCGGCAGGCA CAGGGAGGGC
AATGAGGTGA TCCAGATAGA GCCCTATTAC CTGAAGGCCG AGAAGAAGTT CGGCTTTCTG
GTGGACTTCG CATTTAAGAA GGCCAAGGAC GTGCCCTATA GCATCAGGGT TCAGCAGCTG
AGCCTGTCAC TGAACAAGTA TGGTAAGAGC AACGCCGACT ACTATAGCGA CAAGCTGGAT
AAGATAAAGT TCTTTATGCA GAAGTTTAAG CAGAGGCTTT TCCCATTTAG CTTGGATAAC
GAGGATTACG ACATCGAGAA CGAGCTGTAT CTGATGAGGA GCTACCCGCT CAAGATGAAA
ACCTACATAT TCTCTAATGG CAAGGAAAGC AACAGCCAGG TGCAGGGACT CAAAACCTAC
GGACCGCTGG CGAATCTCGA TAAGGAGCCA CTGTTCGTGT TCATGTTCGA GTCCCAGGAC
AGGAACGAGG CCCTGGAGCT CTATTCTAGC CTGCTGGGCA AGACGTACAC CAACATATTT
GCTGGCATGG AGAGCGTGTA CAAAATCAAA CTCGCAAAAG AGAATGTGAA GCACATCATC
ATCCCCAGCC TTACCAAGGA GGGTCTGCAA GTGGTGGAGC AAGAGCTGCA AACTATCGTG

GAGAGTCATC AGGACAAGAA GGTGATTGGG ATATTTGTAA TGAATGAAAA GGTGCCCTCA
TCCATCACCG GTTTCAGCCC CTACCACTAC GTCAAGTACA TCTTCACAGA GAAACGCATT
CCCCTCCAGA CAGTGAGGTG CGAGAGGATC GCTGCCAGGG ATGGCCTCAA ATGGAGCGTT
GGCAACATCG GCCTCCAAAT TTTCGCTAAA TTGGGCGGCA TCCCCTGGAA AGTCAAGCCG
AGTAACGATA AGTGCATCAT TTTTGGCCTG GGCTGCGCCC ACAAAAAAGA CGAACTGGGA
AACATTAACA AATACTTCGC CTACAGCGTG TGCATGGACA GCAGCGGCAT TTACCGAAAG
ATTAATGTGC TCGGTGATGC AAAGGAGCGC ACTGATTACA TCCTTCAACT GCGGGAGAAC
ATCAAAAGCG TGATAAGCGA GAATCTGGAC GGGAGCATTG AAAAGTGCGT GATTCATCTG
CCCTTCAAAA TTAAGAACGA CGAGATCAGG TACATAAAAT CCAGCGTGCA GGAGATCGCG
CACCTGTATT CCGACATAGA ATTTCAATTT ATCAAGATCA ACACGGACAA CAAGTTTTTC
GGATACGCTG AAAACAACAG CAAGGTACCC TACGAGAGCA GCTACATACA ACTGAGCAGC
AACGAGTTCC TGGTGTGGTT CGAAGGCCTG CAGTACGGGA AGGAGCTGGT GAAGAAAAAG
GTAGGTAACC CCGTGCACAT TGAGTTCATG CAGATCGATG AGTTGGATCC CGAAAAGAAG
CGGCGATATC TGCAGGATAT CATAAACCTG AGCGGTGCCA ACTGGCGAGG TTTTAACGCC
AAACTGTCTC CAATCAGCAT CTACTACCCC AACATCATAG CCAATTTCAT TTCAGAGTTC
AGGGAGTTCC AGCCCGAAGG CGACGTGGAC CTGACCAACT TTTACATTCC CTGGTTCCTG
TAGTGA

41 Gene Name: gene-20542

Gene ID: BBF.2017.48.41

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAGCGCAA AGTAGAGGAC CCTAAGAAAA AACGCAAGGT CGGCAGTGGC
AGCATGCATA ACATCGAAAT CAACACCTTC GTCAACAGCT TTGCCATTAA ACCCAACAAC
TCCATGTCCT TCCTGCTCGG CGCAGGCGCG TCTATATCCT CCGGGATCCT GTCTGGCGGA
CAGATGGTGT GGGACTTTAA ACGGAACCTC TATTGTGCGT CCAAAAACAT ACGCACCAGC
AATTTTCCCG ATATGAGCAA AAAGAATGCG CAGGACGAGA TCCAACGCTT TTTTGATGGG
CAGGCCGGAA ATCCTAGCCT GTGGTCCTCC GAGGAGTATA GTTTCTACTT CGAGAGGTGT
TATCCGGCGA GGAAAGACAG GGAGCTGTAC ATACAGAACA AGGTAAGGGA CGTCAAGCCG
TCATTGGGGT ATCTCTGCCT CGGGGAATTG ATCATACACG AGAAGATCGG TGTAGTATCA
ACCACAAACT TTGATGACCT GGTGTTGGCC GGCATCCATT CAATAAGACC GGACCTGAGT
GTTAAGACCA TCAGCAGTGC CCTCAAAAAT AGCACGGGAT TCTTCGTGAA CGACGGGTTC
CCGAACATCA TTAAGCTGCA CGGCGATTAC TTGTACGATA AGCTGAAGAA TACCGATAAG
GAGCTGCAAA AGCTCGAAAC GGAGATCAGC GGAATTTTTC GAGATGCCGT CAAGAGTGGC
GGGCTCATCG TACTTGGCTA CGCCGGCAAC GACAACAGCG TGATGAGCGT CCTGGAGGAG
CTCGTAAGCT CCGGGCAAAT CAGGTACGGC GTGTTCTGGT GCCAACCGAA GGGCTTCCCC
CTGTCCAAGC GAGCGCGGGA GTTTATTGAG AAGGCTTGCG CCTACAATGA GGAATCCGGG
GTTGTCGAGA TCAACAATTT TGACGACTTT ATGTACCGCC TGTTCCTTAC ACTCAACATC
CAAAACTCAT TTATCGACAG CATGTGGGAA CAGAGCGGCA TGAAGCAGCC GATCCTCTAT

GAGAATATCG GACGACACAA GTCCACCGCC GTGACGAACG CCCTGTGCGC CCTGCAGTAC
CCCCGAAAAT GCTACGTTTT CAACGCGAAT ATATCAAGCT GGAAGGAACT GCGCGAAACG
ATAAACGACA CGTGCGTGGC AGTGCTGTAT AAGGGCATGG TTTGGGCGCT GGGCAGCAAA
GCAGGCATCG TGCATGCGTT CGCCGGGAAG ATCAATGGAG ACATATACGA ACTCGACATC
CCGTTGTACA TGATGAAACT CGAGGATTCT GACATCCTGG GCATGTTTTA CGACATCATA
GGACGCGGCC TTCAGCGAAA GGGGCTGGTG AGCTACGGTA ATAGGAAACA TCACAAATAC
TTCAACCCCT CCAGCAAACG GTTCAAGAAC GGTCAAAACA TCTACGACGC GGTCAAGATA
TCACTGAGTT TCGTGGACGA TCAGCTCGTG CTCATCCTGC TGCCTACGGT GCATCTGCTG
AAACGCGACG GGACGGAGCT GGAGAAATTT GACTACCAAA AATTGGTGTC CCAGGAGATG
GCAACACACT ACAACAAAGT GGTGGACAGC GAGATAGAGA TCTGGCTGAA ATTCATCTCT
AATAACGGCA AGATAATCTT TGAGCTGGGG AACGCAATAC TGGAATTTAA CAACGTCCGC
ATCCAGTACT CTGGTAACGG TAACCTCAGC AAGTGCTACC AGGTGAGCGA GCCCGAGCTC
ACGTTCAGTT ACGAAAAGGA CAACTGTATC GCTACCAACC AACTGCGGGG TCTGATCAAC
TATGGACCCA TAGAGACTTA CGTGAACAAA GCCATCAGGT TGGCTGTACT CAGCCCTAAG
GAGTGTGCCG CGGACATTTG GAAACATCTG CAGAAGTTGA ATGAGCATCA CGTCACCTCC
CTTATTCAGG ATGCAAATTT TCTGCCGGAG TACACCGGCT TTCAGAACGT TTTTAGGTGC
AACCTTGACA TTCCCAATGG GAACGATGTG CATAGGTTCA AAGGCTACAG TATAGACAAG
GTCATGCAAC TCAACGCAAA GAGCTACTTT TACGGGATCT GCAAGTACAT TGATGCATTC
GAGACACAAA GGAGCCAATA CGACCTCCTC GTCATCTATA TACCTAAGCA GTTGACCCAC
ATCCGAGAGG CCAAGAATAA CTTCGAATAT TTCGATCTGC ACGACAGCCT GAAGATTTAT
TGCGCTGGTA AAGGTATAGT CACGCAGATC ATCGAGGAAC ACAGTGTTTA TACTAACAAT
GACACCGCCA AGATCATATG GGGACTCTCA ACGGCCATAT TCACCAAGAC CGCCGGAAGG
TTGTGGAAAC CCAGACGCTA TTCCATGAAC ACCGCTTACG TCGGCCTGTC ATATGTGCAG
AGCGTTAAGA ACAACGAGAA AGTCAGCATC GGTTGCAGTC AGCTGTTCGA CGCCGAAGGC
AATGGAATGA AGCTTTACCT CAGACCCTTG ATGAACCCCC AGATAATTCA AAATAACCCT
TTTATGCGGA GCGACGACGC TTGCAGGCTT ATGTCAAACC TTAAGCGGAT GTATGACGAC
AGTGTCCCGC TCTACAAACT GAATAGGATC GTGATCCACA AAACTACGTT CTTCACTAAA
GAAGAGATGG AAGGCATCAC CAAAGGGCTG GCTGGAGTGG ATGACATAGA GTTGCTCCAG
ATCCAGGAGT TCACAGCTTG GCGAGCAATA CGCTTCGACT ACGACAAGAT CGCACCGTTT
CCGATACAGA GGGGCACAGT GATTCTGGGG TGGGGCCACT TTAGTTACTT GGATACCTGG
AAGTGTACCA CCTAGTGA

42 Gene Name: gene-21311

Gene ID: BBF.2017.48.42

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAGCGAAA GGTCGAGGAT CCAAAGAAGA AACGGAAGGT GGGCAGCGGC TCCATGCAAG AACACCTTAA GACGAACATA CTGAACTTTA AATGGCCCAA CTCTGCTCCG ACCATCTACC TGACATTGGA GGACATTGAG GGGAGCCACC CTATCCACAA AAGCAAATTT TCTAGACAGA TAAAAGAAGT GTTCCCCGAC GCGGATTTGA GTAACAAGGA CCAGATCTTT ACGACATTCA CGACCGAAAT CCCAGACGCC CCAAGCATAA AACTGAACCT TGTGGACGGC CGAGAATTGC GGATCTATAA ACAGTTCCTC AAGCACAAGC TGCGGTCATA TTTCAAATCT AAGGACTACA TCGTGGTCAA GAATTTCGTG GGCGACGTTC AAGTGTGGAT GCCGAGCAAA AAGGGTAACA CCGCAGATTA CAACCTGTAC TATAAGTTTA GCTTTAAGAT CCAATTTGCC AAACTGACGG ACCTCCCCGA GCTGATCGTA AGCTACGATG GCACCTCCAA GGTGCTCACG ACGTCCGTTA AGGACATCGA AGATTCAGAG CTCATCAAAC GATGCGTCTA CGGCCAAAAG ACGTTTAACT ACCAAATGGA CTTGGACACC GAAGAGAAGC AAGAGTTTTA CAACGCGATA CAGTTTGACC AGGCCTACCC AATTTTCAAC CTTTCCCTGG CAAGGGCACT CGACATCCCC ATAGAGGAGC CAATAAGGCC GATCAACAAA TACCAAAAAT ACGTAGCCCT GATTAACAAT TTCGCAACTA ATTACCTTTT CAAGGAGGAC TTCAAGGTTA TCTTCCCGTT TAAAACAGAC ACGTTCATCG ACGTGCCTAT AAATCGGATA AATCACATCG ACCCCCAAGT CGGCCTGTTG GAATTCGGAA AAGATCAATA TGGCAACAAG AAAACCCACC TGGTACCTAA AAAGGCAATG AACATCTTGA ATCCATACCG GCGACCTAAT AATCAGAACA TCAAAATCTT TTTCATCTGT CACA-

CAAGCC ACAAAGACTC CGTGCTCAGC TTCTATCAGA ATCTGAAGGA AGGAGTAAAC ACGGA-
GAAGA ACTACTACAA AGGACTTGAA GCCTACGTGA ACATTAAGGC AAGTAGTAGC AAGGAG-
CATT TTATCGAGTT CACGAACGAG AATGACCCCA TCCCGGAGAT CGTGGAGAAG CTTGA-
GAGCC TCACATTTGA TCATGACAAT GTTCTCTACG CGGCGTTCTA TCTCTCCCCC TTCGA-
CAAAT TCACCCAGAA TCCGGAGGAC CGGGAAATTT ACATCCAAAT AAAGGAGTTG TTCCT-
GAACG AAGGTATCGT GACCCAAGTT GTCGATTACG AGAAAATGGT CGTCAATATC GAGAATCAGT
ATAACTTCCA GTTCAGCCTG CAAAACATGG CCCTCGCCAT TCATGCTAAG CTGGGCGGTG
CCCCGTGGAA GCTGGCCGTG ACCGACAAGA AGGAATTGGT CATCGGGGTT GGAGCGTTTA
CAAATCAAGG CGAGAACAGA CGCTATATTG CTTCCGCCTT CTCCTTTCAG AATAACGGCC
TCTTCCGCAA GTTCGAGTAC TTCGATCAAA GCGAAACCGA CCTCCTGGCT GGCAGTATCT
GCAAAGCCAT CCGCGACTTC ACCAGCGTAG CGGAGGCAGA TAAGGTCGTT ATCCATTTCT
ATAAGGAGAT GAGTTACGAG GAGCTTAAAC CCATCATTCG GGGCATGCAC ACGCTTGGGC
TGAAGATACC CCTTTACATA CTTAACATAA ACAAGACTGA AGCCGAGGAT ATTATCGCCT
ACGACCTGAA TTGGAACAAA AAGCTGATGC CCGTCAGCGG CACCTACATT CGCATCTCCG
AAAATCATTT CCTTCTCTTC AATAACGCAC GATATCCTAA TTCCCAACGG TACGCCGACA
CGGATGGTTA CCCGTTTCCC ATTAAGATTA AGGTCAGCTC TCCGGACGAG GATGCCTTTG
AAGATGCAGA TGTGGTCCTG GAGCTGCTTA CTCAGGTTTA TCAATTTAGT AGACTGTATT
GGAAATCACT TCGCCAACAA AATGTACCTA TCACCATCAA GTACCCAGAG ATGGTAGCCC
AGATTGCCCC CCATTTCAAC AACGGGGTGC CCGACGATGC CAAGGATGCT CTGTGGTTCC
TGTAGTGA

43 Gene Name: gene-6507

Gene ID: BBF.2017.48.43

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAACGGAA AGTGGAGGAT CCCAAAAAGA AGCGGAAGGT CGGCAGCGGC
TCAATGGCCT ATCCAATCGC TGACGACCGG CGAAAGTACT TCCACAGTCT TTTCGAGAAC
AAGGAGCCGT ACATCGGATA CAAGGCTCTG TGTCTGCTGG CCAAGAACGA CATCATCAAG
AGCGTGTGGA CGACCAACTT TGACGGGTTG ACTGTGCGGA CCGCATTCCA AAGTAACTTG
ACCCCCATAG AAATAACCCT CGACAACGCA GACAGACTGT TTAGGAACCA AAGCAAGAGA
GAGCTGCTGA GCATATCACT TCATGGCGAC TATAAGTATA GCACGCTGAA AAATACCGAG
AAGGAGTTGG ACTCACAGGA CGGCACCTTC AGCGAGCATC TGGGTAACTA TCACGTCGAC
AAGAACCTGA TTGTGATAGG TTATTCAGGG CGCGACAAAA GTCTGATGAA ATCCCTGAAC
GATGCATTCA CCAAGAGGGG CACCGGCAGG CTGTATTGGT GCGGCTACGG TGACAAGATC
AACACTGAGG TGGAAGAACT TATACGCAAC GTACGAACCG CTGGAAGGGA AGCCTTCTAC
ATATCCACCG ATGGTTTTGA TAAGACGCTG ATCGACCTTT CTAAAAGCGC TCTGGAGGAC
AACAGCATGA GCCTCGAAAG CCTTAATTCC ATCCTGAAAC TGGCAAACAA CGAGGAGCTC
TCAAAGATCG AATTTAGCCA GAGCATCACC AGGACCGACA AATACCTGAA GAGTAATCTG
CACGCAATTG TGTTCCCCAA GGAGATATTC CAGTTTGAAG TCGAGTTTGG CGACAACAAG
CCCTGGTCAT TCCTTAAAGA CAAAACTAAC AACACCGACA TATGCGCCAT CCCCTTCAAG
AGGAAGGTTT ACGCCCTGGG CACGCTCAGC GGTATATCTA GCGTGTTCAA AAACGTGCTC
AAAAGCGAGA TTAGGAGGGT ACCAATCTCC AAGTTCGACA TCGACAATGT GAGCAGCTTT

AGATCTCTCA TGATCCAAAC GGTGATCAAG CACTTTCTGT CATACGGAAT CTTCGACAGC
AACCTCAAGG ACAAACTGTG GCTTAGAAAT TCCGACAATT CCTTCGGGGA CAAGAAAATA
CACAAGGCGA TTTACCTCAG CTTCTACTTC GATAAGAGCA GCAAATTCGG CTACATTAGC
TTCAGCCCCA GCATACACAT AACCTCCGAT AACGAGATCA GCAAGGAGGT GAAACAAAGG
ATTAGCAAAG AGATCTTGGA AAAGCTCCGA AACGATAAGT TTGACGAAAT ACTGGAGTAC
TGGAACACCA TACTGTTCAA TTACAAAAAT CTTAAGTTCG AGTACCCCCT TAACAGCGGG
ACCGGATTCG AGTTCCAAAT AAGCCGAAAC ACTGCGTTTG CCGAAATCAT GGTGCTGGAC
CCGAACTATC GAGTCTATAA ACCAAGCGAT TACAACAACA AGCTGACCCA GTTCAGAGGT
GTGCAGTATC TGGAGCCGCA ACTGATCTTT CAGAACTCAC TGAGTAACTC CCACACCAAG
GACTACCACC CCATGAGGGC GTTGACCAAT AACAGGCCAT ACGACAACAA CTTGAATGGC
ATCATCTATT CAAACGAGGT CAATTTGGCC GTGATTTGCG GGGAAAACTA CTCCAAAAAC
CTCTACGACT TCCTGAACCA GCTTAACCTT AAACACCCCA CAGACAACAT CAACCCCGAT
TTCCTTATAG AATATCCTGG CTTCGCGAGC GCCTACAACC TCCCCATCAA CATCCCATAC
TATGAGGACG CGGACAAGTG GATTAACATA GATTTGGAAA AGAGCAACAA GTCCGACAGC
GAGAACGCCA TCATCGTTGC ACGCCTCATC ACAAGCAAAA TCGAGCAGAT CATAAACATA
CAGTCTCAGC ACACCATCGT CATCTTCATC CCCAAAGAGT GGCAGGCCTT CGAGAGCTTC
CAGGAAAATG GCGAGGACTT CGACCTCCAC GACTACATCA AGGCGTTTAG TGCATCCAAG
GGCGTGAGCA CCCAGCTCAT CAGGGAGGAG ACACTGTCAG ACAGGTTGAA ATGCCAGGTC
TACTGGTGGC TGTCTCTGAG TTTTTATGTA AAGTCTCTGC GCACGCCATG GGTCTTGAAT
AATCAGGAGA AAAACACCGC CTACGCCGGC ATAGGCTACA GCATTAAGAA GAACAGCAAT
GACACCGAGG TGGTGATCGG TTGCAGCCAC ATTTACGATT CTAATGGCCA GGGCCTGAAG
TACAAGTTGA GTAAAGTAGA TAATTACATC CTGGATAAGC AGAGCAATCC CTTCATGAGC
TATAATGACG CGTTTCAGTT CGGCGTGTCA ATTAGGGAAC TGTTCTACAA TAGCCTGGAC
AGGCTCCCCG AGAGGGTGGT TATCCATAAG CGGACCAAGT TTACGAACGA CGAGATAAAA
GGTATTACTG CCAGCCTCAA CATGGCGGGG ATTACCAAGA TAGATCTCAT TGAAATCAAC
TACGAAACGG AGGCTAGGTT TCTCTCCATG AACGTATTCA ACGGCCTTCT GGGCATAGAC
AAATTCCCTA TCAGTAGGGG TACGTGCATT ATTACGAATA AGTACGAAGC CCTCCTTTGG
ACCCACGGCA TCGTGCCCTC CGTGAAGAAT CCCATTCACA AGTATTACCT GGGCGGCAGG
AGCATCCCAG CCCCGATCAA AATTACTAGG CATTACGGCG AGAGCGATCT GAATACTATT
GCCATCGAGA TCCTCGGCCT CACCAAAATG AATTGGAATA GCTTTGACCT TTACAGCAAG
CTCCCTGCGA CGATTAACTC CTCAAATCAG ATAGCCCGGA TCGGTAAGTT GCTGGCGCGC
TTTGAGGGCA AGACCTATGA TTATAGGCTC TTTATTTAGT GA

44 Gene Name: gene-8284

Gene ID: BBF.2017.48.44

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAGCGAAA GGTAGAGGAC CCAAAGAAGA AAAGGAAGGT GGGCTCCGGA
TCTCTGGACA GTTTCCACCT CGTGCAGACA GAGAAAAAGG CTATCGCAAT GCCAAAGCAG
AAGCTTGCGG TTAATGCACT CCCCATTAGC CTGAAAGAGC AGGAGCAGCA CAAGCTGTTC
TTTTTTAGCA AGGAAAAGCA GGGCGAGCGA GCCCCGCTCA CCAGGAAAGA ATATCCTGAC
AGCTTCGCCA AGAGGTACCC CAAGAGCTCC AAAGAGTACG ACGTGCTGTA CACGGACTTC
ACCCCAGAGC CAGCTGAGGA TGGGTTTGAA ATTGATATCG ACCTGGAGGA GGCACCTGGC
CTTGCCAAGC ACTACTTGCA CAAAAGGATC TTTGAGGCCT TTAAGGGAGT AGCTGACTTC
AGAAAGCGGG ATTTCATCAA CGGTGTGGAG CTTTGGTTCA GGGACAAACC CGCCGACGAA
GTTAATTTCC GGGCCTACAA GAAGTTTAAG ATTACCACCC GCAGAACTTG GTTCTCCGCA
GGCTGGGCCC TGTTCATACA ATACACCGGC CATTCCTTTA TTCACCCGGT GGCGATCAAT
AGCGAAGAGG CCGCAGTGGA CACTACGGAA CTCACGCGGG TTGCTTATAA CCGACACATC
TTCCACTACG AGGAGATCCC CGAGGACAAA CTGAGTGAGA TAGATTTCAG TAAGATGTAC
CCCGTGGTGA ACTTCAACAT TAGGGATAAA ATGCAGCAGT TCCCCGTTAT CGATCCATTC
AAAAACAAGG TCAAGGAATA TGTCGACGAA ATAGACAGGT TCAAGAACAT GTATCTGATC
GCGCCAGCGG TTGAGGAGGT GCTTCCGTTT ACTTTCAACG ACGACAACTG GTGCGAGATC
AAGATCGGCA CCTACCATAC CGTGCCCAAT GCCGGTTCCA AATTGGTTTT CAGGGATGGG
CAAACCGAGA TACACCCGTT CTACGGTATC AGGAACCACG GCCCTTTCAT GCCCCCCAAA

CACAGCCACA TAAGGTTTTT GTTTATCATG AGCAAGAGGG ACATCAAGGG CGCTGGTAAG
CAATTCTATG AATACTTGAA GGGGGAGGTA AAAGGAGTGG ACGGGTTCAA CAGGTATGCT
AATATACCGT CATCCCTGAG GGGTGAGATG ATCGAGTTTG AGAACGAGCA AAACCCCCTG
CCGGAGATTA TCGACGGCTT GAACAACATG GAGCGAGAAG CGGGCGTGGC CTACTTCGCC
TTCTATATCA GCCCCATCGA CCGAGAAGTG AGGAACAGGA AGGAGAGGTT GGTGTACTAC
AGGGTTAAGG AGGAGCTGCT GAAGAGAAAG ATTGCCTCAC AAGTGGTAGA AAGGAGCACT
ATCGAGAAGG CCGACTTCCG CTACAGCATC CCCAATATCG CCGTTGCCAC AGTGGCCAAG
CTGGGAGGCA TCCCGTGGAA GCTTACTCAA CCCCCAGAAG CAGAGCTGAT CGTGGGCATA
GGCGCATTCC AGCCACGCGA GTTCGACAAG CGATATCTGG GCAGCGCCTT TTGCTTCCAA
GGCGACGGAA CCTTTAGCGG CCTGAGGTGT TTCACCAAGG ACGAACCCCA TATGCTTGCT
GGCAGCATCA GGGAAGCGGT TCAAAGGTAC GCCGATGAAA ACAGGCAAGT GGAACGGCTG
GTTATCCATT TCTACAAAAC CATGAGCTAT GACGAGAGGA AGCCGATCCT GGCCACCTTG
AAAGAACTCG GCCTGGACAT TCCCGTTGTG GTGGTCACTA TCAACAAGAC TGAATACGAG
CAGACAATCC TCTTTGACCT GAATTCTAGC ATGAGGCTGC CGCTGAGTGG TACCTATTTC
AGCCAGCGCA GGGACGACAT CCTGCTGAGC AACAACACCA GGTACCGCAA AGACAGCGAG
GTGAAGAGGG GTTTCCCTTT TCCCGTGAGA CTGCAGCTGT GGTGCTCCAA GGAGGGCCTG
CTGGACGACG AGGGTTTTAG GGAGCGACTG ATCACCCAAG TGTATAGGTT TTCTCGGCTT
TACTGGAAAA GCGTGTCTCA ACAGAATCTG CCCGTGACCA TTAAGTATCC CGAGATGCTG
GCCGAAAAGT TCCCATACTT TAACTCAAGG AGCCTTCCTA GCTTCGGCGA AAAAAGCCTG
TGGTTCTTGT AGTGA

45 Gene Name: gene-120

Gene ID: BBF.2017.48.45

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAGCGGAA GGTGGAGGAC CCGAAGAAAA AACGAAAGGT GGGCTCCGGA
AGCATGAACA ACACCATAAA CAAAATAGAC TTCGGCGCGT TTCTGAGATC ATTCAAGCAG
AACCTGGACG GTAGCTTTTC TTTCCTTCTG GGAGCAGGCG CGAGTGTGAG CAGCGGCGTA
CAGTCTGCAA GCGACTGCAT TTGGGACTGG AAAAAAGACA TTTTTCTGGC CCAAAACCTT
CAATTTGAGG AGTTTCTGGA CATCCATAGT GACTTCTGTA AAGATAAAAT CCAAAAGTGG
TTGGATGAGC AGGGCGTGTT TCCCAAGCGA GACTCAGAGG AAGAGTACGT GTTTTATGCC
GAGAAAGCGT ACCCAATGGA ACAGGACAGG ACCAAGTATT TCGAGAACCT TTGCGCGGAC
AAAACCCCCT ACATAGGGTA TAAACTGCTG ATGCTGCTGA ACAAATACGG AGTTCTGAAA
TCCGTGTGGA CAACGAATTT TGACGGTCTG ATAGAACGCG CAGCGCACCA AGCCGATCTG
ACGCCAATCG CCGTTACCCT CGACAACCCC GAAAGGATTA GCCGAAACGA GAGTAAATCT
GAGCTGCTCT ACGTGGCACT CCACGGTGAC TACAAGTATA GCAAGCTGAA GAACACAGCC
CAAGAGCTGG ACGCGCAAGA AATTCTCTTC ACCGAACGCC TGAAGTCTTA CTTCATCGAT
AAGAATTTGG TGGTGATCGG TTACAGCGGT CGAGACAAAA GTTTGATGCA CACCTTGTGC
GAGGCTTTTA TGACGAAGGG GTGCGGTCGG CTTTACTGGT GCGGCTACGG TAACAAGATT
ACCTCTGAAG TGCAGAACTT CCTCAACAGA ATAAACGATT CAGGTAGGGA AGCCGTGTAC
GTGGACACCG ATGGGTTCGA TGCCACCCTC GTGTCTATTA TGAAGTTTTG CTACGAGGAT
CAATTCGACA AGAAAATCGA AATCGGCAAG TATCTCAAGG GCCTGTCAAG GGTGAAGCAT

ATTATCCCTT TCAGCGTTGA GAATACCACG TTCACCGGCT GCGCCAAGAC CAACCTGTAC
CCCTTGATCA TCCCCCAAGA CATATTCCAG TTCGAGATAG AGAGCCCCGA AGGTAGCAGC
AAATGGACCT TCATTAAAGA GAAGATTAAG GGCAAGGACA TTATCGCTGC CCCTTACGAG
AAAATAGTCT ACGCATACGG GCTGCCAAAC TCAATCTACA ACGTATTCAG TAAGGAGCTG
ATCGGCGAGA TCAAGAGGGT TCCCATCAGC CTGAGTAACA TCAAAGACAA CAGCACCCTC
AAGAATATCA TCCTGAAGGT GCTGATATGT TCTCTGAGCA GTAACGCGGG ACTCAGGGCG
AGTATGAGCA AGAAGATCAT CTGGAATGAG AAAGAGAGGT TCCAGAGCAA CGTTTTTAAG
GCAATAAAGA TCGACATCGT TTTCATCAAT AGCGAAAAGT ACGCCCTCAT CTCAATCACC
CCTACCCTCT ATTTCAACAA GGAGGGCAAC TACACGACGC TGCAGAAGCA GGAAATTACG
CGGAGCTACA TTGACAAGCT GTACAATAAG ATTTATGAGG AAACCCTTTG TTACTGGGAG
GCCATCCTGT TTAAGCAGCA GACCAAGATC TGCTTCGACT ACCCGCTCAA TTCCGGGAAC
GGCTGTTTCT TCAAGGTTAG CTCTAACAGG GGCGAAGCCC TGTTCAATAA TCCGAATAAG
CCGTACGTGA TTACTAACGA CATCATACTT AAACGCAAAA TCTACGAAGG CATCATAATC
GACGAGCCCC TCCTGAACTT CTCAGGGTCA ACCAGCGCCC ACATCATTAT GGACTCCAAT
CCGATGCGCG GACTCAACAA CAATAACCCA TATGATCACT TCATTGCAAG CAAGTTTAGG
GACGTTTCTA TCCACATCGG AGTCGTGTGT CCCTGTACAT ATAGCGACAG GTTTTTTAGC
TTTCTGAACG AGCTGCAAAG TCCGATAAAG AATAACAATC CTAACTCAGA CTACATCCAG
AACTATAACG GATTCAGCCA GATATACGCA AGCATTCTTA ATATCCCAGC GATCAACAGC
CAATACTGGA TCTCATGCCG CGAGGAGCAG GATAACAGCA TCTCTTTGGC TAGGAACCTG
TGTAAATACG CGAACCAGAT GGCCACTAAC ATGCCAGGTA TAATAGTTAC CTTCTTCATT
CCTAACAGCT GGAGCAACCA CAAGAGTTTC AAAGAATGTG GCGAGGTATT CGACCTCCAC
AGTTACATCA AGGCTTTCGC CGCACAGCAC GGTTTTACAA CCCAAATCAT TGAGGAGCGA
ACTCTCACAA ATCTCTCCAT GAAAAAGGAG ATCTATTGGT GGCTGAGCCT GGCGTTCTTT
GTAAAGGCTA TGCGAGTACC ATGGACCCTG GCCAATCTGG ACCAGAACAC CGCCTTCGCC
GGCATCGGCT ACTCCCTGAG CAAAAAGCAA AGCGGCAAAT TCAATATCGT TATCGGCTGT
AGCCATATCT ATAATTCTGA GGGCCAAGGC CTGAGGTACA AGCTCTCAAA GATAGATAAT
CCAATCTTGG ACCGGAAAAA CAACCCGTAC CTGACCTATA ATGAGGCGTA TAAGTTGGGC
GTGAACATAC AGAATCTGTT CATTCAGAGC ATGGACAAAC TCCCGAAGCG AGTAGTGATC
CACAAAAGGA TCCCGTTCCT GGAGGACGAG ATAAAGGGCA TTACCGAGGC GTTGGCCCAG
GCCAACATCA CGAATGTTGA CCTCATCACT ATCACGATCG AAAAGAACAT CAGATGCCTG
GATCAGTTCT TCTACAATGG TCAAGCCAAG AACAGCAACT TCCCACTGCA TAGGGGCACA
TGCATGAAGC TCAGTGATAC CGAGTGTCTG TTGTGGACCC ACGGCGTGGT GGACTCAATT
AAGGCGGGCA GGAACTACTA CTCTGGTGGC AAGGGTATCC CCTCCCCCCT CCGCATATCA
AAGTTTTACG GCGCAGGCTC TATGAAAACT ATATGCAACG AAATCCTGGG GTTCACAAAG
ATGAATTGGA ATAGCTTTAA CTTCTATACC AAGCTTCCCG CGACCATCGA CACCAGCAAC

ACGCTGGCGC AAGTGGGGAA CATGCTCGAT AATTACAACG GTATTACATA CGATTACAGG TATTTCATCT AGTGA

46 Gene Name: gene-12508

Gene ID: BBF.2017.48.46

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAACGCAA GGTCGAGGAC CCTAAGAAGA AGAGGAAAGT AGGGTCTGGC
TCTATGCAAC TGAACTATTT CCCCATCCAG TTTGACTTTT CTGACTACCA GGTCATCACG
CAGCCCTACT CCGACGAGAG ATTGAAAGAA CTCAGGCAGG CCTACAACGC CAGCTATTCC
TTCTTTCGGG ACGGCAACCT TATCGTAATT TCCAATAAAG AGGACGAGGA AAACCAATTG
ACGGGCAACG TCGAAAACCG CAGCGTGTTC GACGATGCCA AAGTTACCGC CAGCATGGTC
AAGCATATAT TCTTTAGGAC GTTCAAGGAC AGGTTCCAAG GCTTCATCCC CGTGGACTTT
TACCCCTTCC GATTCTACAG CAGACAAGAG AAGGACGACC TTATTCTGAA CCATCTGCCC
GAAAAACTTA AGCATAAAAT CGCCTTTAAG AAACTGATCG AGGTGCAGCT CAGGGAAACG
AATCTTAATT CAACCCAGGG CTTTGCTTTC GTCGTCAACA TCAGGAGAAA TTGGGTGTTT
AACATTTCCT GTCTCGAGCT TTATCAGGAA GGCTTTGACC TCACAGATTT TGAAGTGCTC
CATGCGGAAA CGCTTCCCGG GTTGGACAAT ATCCTGGCCC CGAACGAGGA CTTCGTTGGC
CTTCTCAAGA GCATCAACGG CGAGACTGCC ATTGTGAGCA CTAGCGAGGG TGCCCGCTCC
TATTCACTGC AGGAACTCTT CATTCGCAAG ACTAAGCACA ACATACAGGC GTACCTCAAC
TTCGCCACCG GGGAGAAAAA GTGCGACCAG ATCCTTGCAG CCGTGTCCCA GGAACGAATC
CGGAAGCAGA ACCCCGTGAA TCAATTCAGC GAGATATCCA ATATCGCGAA GCATCTTTTT
TCAGACAAAG GCAATCCAGT GCTGTTCCAG AATATGGATG GCTTTTGTTT TAAAGTTGAC
ACCACGCCGA TGCAAGTACA AAACTCCATG AATCTGCAAA CTCCCACGTT CATCTACGAC

CACGCGGGTA CCAAGACGAA CACCCGCAAC GCGGACCAGG GGCTGAGCTA CTACGGCCCC
TACGATAGCC TCACCTTCGA CATTAAGAAG CCAAGAGTTC TCTCTATCTG CCATAAGACC
AACCGAGGCT CCTTTACGCG CTTCCTCCAC GACCTCAAAG ACGGGCTCCC CAATAGCAGC
TGGTTCAAGA AGGGCCTCCT GAAGAAGTAC GAGCTTCAAG AGGTGAATTA CCTCATCCAG
GAGATCAGCG ACTACAGGTT GGAGGACTAC CTGGAAGTGA TCTCAAACTA CGATGATGAG
AAGCCGCACC TGGCAATCAT CGAAATTCCA GATAGGTTCA AAAAACTGTC CGACCGGGAC
AACCCCTATT TCAAGATTAA GGCAAAGCTG CTGAGCCTTG AGATTCCCGT ACAATTTGTG
CGCAGCACGA CTTTGAGCAG CTACAGCGAA TACATACTTA ATCCGCTTGC ATTGCAAATC
TATGCGAAAC TCGGCGGCAC GCCTTGGGTT CTTCCGGCCC AACGCTCCGT TGACCGCGAA
ATCGTTATTG GCATAGGTCA CTCATGGCTT CGGAGTGGCA TGTATAAGGG TGCTGAAAAC
AGCAGGGTGG TCGGCATTAC TACGTTTATG TCCTCCGATG GCCAATACCT CCTGGGCGAC
AAGGTGAAAG ACGTGCCTTA CGAGTCTTAC TTCGAGGAGT TGCTGAAGAG TCTCAAAAGT
AGCATAAGCA GACTCTCCGA TGAGTATGCC TGGCAGGATG GCGACACAGT GCGCCTCATT
TTCCACATCT TCAAACCCAT CAAGAACGTT GAGTTCGATG TCATTAGCCA GCTTGTGAAG
GACATCAGCC AGTTCAACAT AAAGTTCGCG TTTGTGACCA TTAGCAAGTC ACACCCGTCT
ATTCTCTTTG ACACGAGTCA GCAAGGCGAG AAAAAGTACG GCTCTAACCA GGTGATAGGG
CAGTACATCC CTCAGAGGGG TAGCAATATC TTCATAGATG ACGAAACCAG CCTGGTGCAG
ATGCTGGGCG CCAGGGAACT TAAAACTGCC AAACACGGGA TGAGCACCCC AATCCAAATC
AAACTTAGGA CACCGCAGGG TAACCATAAC GACCAAGAAC TGAAGGATTT GATGTTTTAC
GATCTTAACT ACATTACCCA GCAGATCTAT AGTTTTACTT ACTTGAGCTG GAGGAGCTTT
TTGCCACGCG AGGAACCGGC CACAATGCTC TACTCCAACT TGATATCCCG ACTTCTTGGG
AAGATGAGGA GCATCCCTGA ATGGGATGCG GATAAGCTCA ATTATACCCT TAAAAGGAAG
AAATGGTTCC TGTAGTGA

47 Gene Name: gene-13533

Gene ID: BBF.2017.48.47

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGCGGAA AGTCGAGGAC CCCAAGAAGA AGAGAAAGGT GGGCTCCGGC
AGCGTGGGCG ACAAGACCTT CAGCTTCAAG GTGTATAGGA AACTGAAACA GCAGAACGAC
ACCAAGGAGG ACGAGATATA CCTTTACAAT TTGCCCCAAG GCGAAACCCT GAATGATTAC
AAGCCATATT GGATCAGTTT TACCCCGAAG GACGGATTCG AAGAATATAT CGCTAATTCT
TACTTGAGCA TCGGCCTGTC AAAAAAGTAC CTGTTCAATA GATTCGTGGA AACGCTCAGC
AACTCAAAAC TGCACTTCAC CTACAAGGTC AAAAGGAAAT TCACCGACTG GTACGTCGAT
TTCGTAATCG CGCAGTACAG CCAGGGAGAC AGGATCATCT ACATGAGCCC CTACTTCCTG
GAGGAGCAAA ACACCTACGG CTTCATCATC GACTTCAAGT TCAGCAAGAA GGATGGTATC
CCCTTCGATA AGGAGGTGCA AAAGCTGTCC CTTTCACTGG ATAGCAACGG CCGCAGCAAC
AAAAACTATT ACTCTGACAA ATTTAGGCTG GTGAACAATT TCATTAAGGA GATTTACACC
TCCATAAAGA ACATCGGGAC CAGTAATAAT CCTATCACCA TTTCCAGCAA CCTCATAGAA
ACCACCGTGT TCCACCTGAA CAAGAAAGAG TACATCTTTA GCAATAACAA CGTAAGCTCT
AGCCAGTTCC AGGGCGTGAG GAATTTCGGT GTCTATAAGA ATATCCCCCA GGACGTGATC
TTCGCGTTCA TATTCGAGGA TAGGTTCAGG AGCTTCGCCA ACGAGCTGTA TCTGAGCCTT
ACCGGAAAAT TGAACCCCGG GACCTTTCCC GGACTGGAGC AGATGTTCGG CATCAGCATC
AACACCAAAA ACGTGAGACA GATCAAGTTG GAGAACTACT CTCTGGATTC AATGCTTAGG
GTGGTGAATG ACGTTAAGAG CTTGCAGGAG AACAATCCCG ATAAGAAGAT CGTGGGAATC

TACGTGGAGG ACTGCACCAT CGACAGCGAG GACATCCCTG CGTCCAACAA CTACTACTTT
CTGAAGTATC ACTTTATCAA AAATGATCTG CCACTCCAGG TTGTGAATTA TCGGAAGCTG
GGCGAAAGGA ATTCTCTGAA ATGGAGTACC TCCAACCTGG CCCTGGCCAT GTTCGCAAAG
ATGGGCGGCA TCCCCTGGGT CGTAAAACCG TCTAATAAGA ACTGCTTGAT TCTTGGCATC
GGATCTAGTC ATAAGATAAA CCGGGAAACC GGCGATATAC TTAAATACTT TGCATACACC
ATATGTCTCG ACTCCAGTGG CCTGTACAAG GCCCTTGAGG TGCTGGCCGA CGAGGAGAGC
GAGGTGAGCT ACCTTGAGAA GCTTACTGCC AATCTGGTCG CCATACTGAA GGAACAAAAG
ACCAATTACG GCACCTGTGT GCTGCATCTG CCCTTCAAGA TTAAGAAAAA AGAGGTAGCC
GCCATTAGTG ATGCCATAAA ACAAATCAAC GACATCGAGC TGGTGGTGGT AAAGATCAAT
GTGGATAACA AGTATTTCGG ATACTCCTTC CACAACACAT TGGTGCCCTA CGAGAGCAGC
TTCGTGAAGC TTTCTAAGGA TGAGTATCTG GTGTGGTTCG AGGGCCTGCT GTACGGCAAA
GAGATCGTAG ATAAGAGGTT GAGCAACCCC GTGCACATCC AATTCTTGAA CATCACCAAC
AGGAAGAACT TCGATGAGCA GGCGTTTCTG CAGGACATTC TGAATTTGAG CGGAGCCAAC
TGGAGGGGCT TCAACGCCAA AAGCATCCCT ATCTCAATTT ACTATTCTCA AATTATCGCG
AGGTACACCG AGGCCTTCGA AAACATCGAC GGTTACAAGG AGGGTACTAT CTCTAACGAC
AAACCCTGGT TCCTGTAGTG A

48 Gene Name: gene-17059

Gene ID: BBF.2017.48.48

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AAAAGCGAAA AGTGGAGGAC CCCAAAAAGA AGCGGAAGGT GGGCAGCGGC
AGCATGGACA ATTTGGCTCT CTCTGCGCTT CAGCTGGACA GTAGATTGGA TCACTGTATG
GTATATCAAT ACAGGATCGT GTACCATAAG TTCGACGAAA CAGAGGCGGG TGAAAAACTG
GCAAGAAAGG CCGCCTACGA ACTGTGGAAG GTAAACAACT TCGGACTGCT CACCAACCTG
GGTGCCAGTA GCATCCTGTC CCTTAAGAGC CTGAGTCAGC TGTCTATCGA TTCACCGCTG
TTGCAGGCAA GTTTGAAAGC TGACGGCCAG TTGGAGCTGG ATTGCGGTAA CGAACAGCAT
CAGGAGGCGC TGCAGAGACT CGTGAACCAG GACATAAACA AAGCGGCTTG GAACCTCAAA
CAAGCGAGCG AGGGGAAGCT TGATTGCCGA AAATCACCAG GCGGGCACGC CGAAATCTTC
GAGCCAAGTC ACAGTAGTCG GATCAAGGCC CACAGTACCT ATTTGGATGC CTTCTGCACC
GTAAGGCTGA TTCCCGAAGT GCTGTCAGAC GGGACAGTGC TGATAGGGTT GCATCTTAAG
CACAGCCTGA CCGCGAAGGC GGACATCTCT CTTCAGTGGG TCATTGATCA TAGGCCCGAT
TGGCTGATAT CCATAGAGAA GGTGCGCCAC AGGTATTACG AGCCCGGCAA AGCACCCCTC
GTTGCGGAGT TCGTGAAAGT CGATGATTCC ATCAACGGAT CATCCCTTCT CCCACACTTG
GGCAAATCCC TTGTCGCTTA CCACCAGGAG AAAGGGCTGC TTTCAGCCGG ACAGCTCGCA
GAGGCAGCCA CCAGCTCACT CATCAAAGTG CGCTACGGAC AGAAGGAGGC AGACCACGTT
GCTAGCTTGG TGGAACCCAT GTTTGATTTC GATACTCTGT CAAAGATTGA CAGCCCCTTC
CTGAATAGGC TCGCCAAAGA CCTGAAGTGG AGCTTGGACG ATAGAATAAA GACAAGCGCG

GAGATGGTCA AGAGGCTCTA TCTGCCCGGG TTTAATCGAA AGTTGGTACA AGTTGACTAC
CAGAATCTGA GCAGGAAGAG GTTCAACCAC AACCTTATGC TCCAGTTCGC GGATGGGGCA
AGGAGCGGCC ATGAACAAGA CGTCCTGAAA TACAAGGCTT TCGCCGACAT GACCAGGGCT
AGGGTAATCC CACTCGTGGT AGGAGAGAGG AACAACACCG AAAGCAATAG ACAATTGCTC
CGGAACGCCT ATAACGCACT GAGGCAACTT ACCAAGGCCG AATTGCCCCC CTTCACGTCA
TTTCCCCCCA GCATCGGAAA CGCCGACGAG TTGGACGCAC GGCTGCACAA GAAATGTCCC
GACAACGCCA TCCTGCTTAT CGGGCTCACA GAGAAGAGTG ACAAAGCCGC GATCAGGGAC
ACGGCGTTCA ACTACGGCCT GGCCACCCAG TTCATGAGGC TCGATCACAA GCCCAAGGTT
TACGACAGCT TCTACTTCAA TAACGTCGCA GCGGGCCTGT TCTCCAAGGG AGGAGGGCAA
CTGTGCGCCG TGAACGACAT GCCCGGTGAG ACTGAACTGT TTATCGGTCT GGACATGGGC
GGCGTGAATG TAAGGGCGCC AGGTTTCGCA TTCCTGTTTC TCAACTCTGG CGCGCAACTG
GGCTGGCAGC TGGCTGACAA GCAGCAGGGC GAGAAAATGC AGGACGACGC TCTCAGCAAT
CTGCTGGAGA AGTCTCTCAA AACCTACCTG AGGAGCACCG ACGGGCTTTT GCCAAGGAGG
ATAACTCTGC ACAGGGACGG AAGGTTTTAC GAGAGCATCA ATGTGATAGA ACAGTTTGAG
CAGAAGCACG GGGTCAAGCT CGATGTTCTG GAAGTCTTGA AAAGCGGAGC CCCGGTGCTG
TACCGGAGAG AACGCAGTGC GGACGGTAAG AAAGTTTTCA GCAACCCAGG GGTTGGAGAT
GCCGTGTTCC TTAGCGACAG GGAGGTCATT CTTAGCACTT ACAGCGGCGA GGAACTTGGG
AAGTCATGGG GTAACAAGGT GAGTGTGAGG CCACTTCGAC TCCGAAAGAG ATACGGCGAA
ACCGCATTGA GCGTGTTGGC CCATCAGGTG TTGGTCCTGT CTAGGATCCA TGGGGCCAGC
CTCTACCGAC ACCCCCGACT TCCGGTGACC ACCCACCACG CGGACAGGTT CGCAACCTTG
CGGCAAGATG CGTGCATAGA CGCACTTAGT AAGATGGATA GACTGTGTCC GGTGTATCTG
TAGTGA

49 Gene Name: gene-17650

Gene ID: BBF.2017.48.49

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAGAGGAA AGTGGAGGAT CCCAAAAAGA AACGAAAGGT CGGCAGCGGT
TCTATGAGCG AGCTGGAAAC CAACATCTTC CCAATCACCA ACTTGCATGA GCTTGAATCT
AGGTTCAGGT TGTATAGGGT GAGGGGCCTG AGCATCAACC AAGAGGAGTA CGACCCCAAC
ACCCAGACAT TGGTGAGGAA GCTGAGCTAC AGCATGAGAT CTCCCGTAGC TGTGATACTT
AGGAACAGCG ACCCGTTCCT GGCACTTCCA ATCGACGCAC CCGAGCCCAT CTCTCCGTAC
CCGCTCGTGA GAGCCACTGC TGTGTTCGAA AAGACGGACG AGGTATTTAC TCTCGATTAC
GAAAGCCCAA CTCCCGAGAC AGATGCGCTG CGAATAAGGT TCCTGCAATT TATCATCCAA
GGCGCGCTGT TTAGGAATCC CAGCCTGTGG CAGCCCTCAG CTGGCACCCC CTTCTTCGAG
AGGAGCCCCG TGTTGGAGAA GGCCGGCATT TGCGCGTACC GAGGCTTCTC AGTGCGAGTC
GTGCCCATAG AAGGTGGTAA ACTGGGAATC TGTGTGGACG TTAAGCACAG GTACGTCAGC
AAAAACCCCA TCGAAGCAAA CATCAAGCGC GAGGAATTCA GGAAATACAA GAACGGTAGG
TGCATATACC ACTACGGCCA CAACTGGTAC GAGATCAAGT TGCAAGACCA CACTGGGCTG
TCCGTGTCAG AGCAGATGAT CAGCAACGGG ACGGCCAAAC CCATAAGCTT GTATCAGTTC
ATTATGAATA ACGCGCCCAA GCCCCTGCCC AGGGAGGTCA TAGACATGCC TCCCGACTCA
CCCGCAGTCA AATACATGAC CAGCAGGGAT GAGGTGCGCT ACGTGCCCTC CATCCTTTGT
TATCCGGTCT TTGACACCTC TGACCCCAGG GTGAAGCCGA CGCATAGGGG CACAATCCTC
CTCCCTAACG TGAGGCGACA GTATATCCAC AATTTCGTGA ACTCACACCT GACCGATGTG

CGATCCAAAG ACATGGCAAT CCGAATCAGC AGCAAGCCAG TTATCGCCCC TACCAAGATT
TTCCTGCCGC CTGACCTGGC ATTCGGCAAC AACACCGTGT TCAGCGTAAG AGGCACACCC
GGGACCACGT ATGTTAGCCT GGAGCAGCTG GGCCAGACGC GGATAAGCGC CCTCTTCAAT
CAGAAAATAG GCCCTTATGA CAGCAGGCCG CTGGATAGGC AGTACATGAT TCTGCCGAAA
AGCGTGTGGG ACTCCCACGG GCCAGTATTT CTGAATGACT TTAAGAAAAT CATGAACGAG
CTGTATCTGC ACGAACTGCC CTACAATCCC ATCGTCGTGA CCTACAACGA CTTGAGCGCC
AAGACCTACG CGCTTCAGGG AAGGGCTATT CTGGACGCCG TGGACAGCGA ACTGAGAGAG
CCGGGATACG GCGTGGTTAT GATACACGAA ACGGTGGACC GCCGGAATAG ACAGCACGAC
CAGCTTGCCG CAATGGTGAT GAGGGAGCTG CGGAACAGGA GGCTGTATGT GAGCGTGATC
CATACCACGG TGACGAAGGA CTGTTACCAA TTGCCCCAGA ACGCCCCCAT TGGCAAGGCC
TACTGCCCGG TAGCAGGCAA GCAGGGCAAA CTCAATGGCT ACTTGAGGAA CGTGGCCATT
ACCAAGGTGC TTCTGACCAA CGAGAGGTGG CCCTTCGTTA TATCTACCCC GCTGCATGCG
GACTTTACCG TTGCCTTCGA CGTGCAGCTT AACACCGCTT GCTTCACATT CATCGGCAAG
AGCGGCTCCG ACATCCGGAC CGTTCTCAAG ACCAGTAACC AAAAGGAGAG GTTGAGCAAG
GCACAAGTAA GGCAGACGCT CCTGGAAGTG CTCCGCCAGG AGGTTGGCTT CGGTCGACGG
ACCATGCAGA CCATAGTGGT TCAGAGGGAT GGCAAATTGT TTGCCAGTGA GATCGCGGGA
GCAAAAGACG CTATAGAGAT AGTGAAGAAA GAAGGCATCT TGCCCTCCGA TGTGTCACTG
AATTTCATCG AAATCCCCAA GAGCAGCGTC GCCCCATTTA GGCTGTTCGA TAGCAGCCCC
AGGCCAGGGC AGCCTGAAAT GGCGAACAAC CCAAGAATCG GCTCCTACTT TATCGCGACG
AATTACGACG GTTACATTTG CACCACCGGC AAGGAGTTTT ACCATCCCGG TACGGCAAAT
CCTCTCCACG TGAAGTACAT CGAGGGAAAT ATGCCATTTG AGAAGATCCT GGAGGACGTG
TACGCCTTGA CTTGCTTGGC GTTGACCAGG CCCGAGGACT GCACAAGGGA ACCCTTCACC
ATGAAACTGG CCGATATCCG ACTGAGGGAA CATGCCGGAG GCTACGACGA AGATGCATTG
GCGTATGATG ATGAAAATGA GAACGACGAG GATAACGAGA ATGAATAGTG A

50 Gene Name: gene-20377

Gene ID: BBF.2017.48.50

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AAAAGCGCAA GGTGGAGGAT CCAAAAAAGA AACGGAAAGT GGGATCTGGC
TCCATGAACT ACACAGAGGC CAAGACCGCC AATAGCCCCT TGTTCCTTAG CGAGATTAGT
AGTTTGACAC TTAAGAATAG CTGCCTGAAT TGTTTTAAGC TGAACCATCA GGTCACCCGG
AAAATAGGCA ACAGGTTCTC TTGGCAGTTC AGCCACAAGT TCCCTGACGT CGTGGTAGTG
TTCGAGGACA ATTGCTTTTG GGTGCTGGCT AAAGATGAAA AGAGTTTGCC TAGTCCACAG
CAGTGGAAGG AAGCACTGTC AGACATACAG GAAGTGCTGA GGGAGGACAT TGGGGACCAC
TACTACAGCA TTCACTGGTT GAAAGACTTC CAGATAACCG CCCTGGTCAC CGCGCAGCTG
GCTGTGCGGA TTTTGAAGAT ATTTGGGAAG TTTAGCTACC CGATCGTGTT CCCCAAGGAC
AGTCAGATCT CTGAAAACCA GGTCCAGGTG CGAAGGGAAG TGGATTTCTG GGCTGAGATA
ATCAACGACA CGGACCCAGC AATATGCCTG ACGGTGGAAA GCAGCATCGT TTACTCTGGC
GACTTGGAAC AGTTTTACGA AAATCATCCG TACCGACAGG ACGCCGTGAA ACTTCTCGTA
GGGCTGAAAG TGAAAACTAT CGAAACCAAC GGGATCGCGA AGATTATCAA AATTGCCGGG
ACCATCGGAG AAAAGCGGGA GGAACTGCTG ACCAAGGCAA CCGGGTCCAT AAGCAGGCGC
AAATTGGAGG AGGCACACCT GGGCCAACCT GTGGTGGCCG TGCAGTTCGG CAAGAATCCG
AGAGAATACA TCTATCCCCT TGCCGCGCTC AAACCGTGTA TGACCGACAA AGACGAGAGC
CTGTTTCAAG TGAACTATGG CGAGCTTCTG AAAAAGACTA AGATTTTCTA CGCCGAACGG
CAGGAGTTGC TGAAATTGTA TAAACAGGAG GCGCAAAAGA CGCTGAACAA CTTCGGCTTC

CAGCTCCGGG AGCGGTCAAT CAATAGCAGG GAGAACCCCG ACTTTTTCTG GACCCCCTCA
ATTTCCCTTG AACAAACGCC CATCTTGTTT GGCAAAGGTG AGCGAGGTGA GAAACGAGAA
ACCTTGAAAG GCTTGAGCAA AGGCGGCGTG TACAAGAGAC ATAGGGAGTA CGTCGACCCC
GCGAGAAAGA TTAGGCTGGC CATCCTGAAG CCGGCCAATC TCAAGGTTGG GGATTTTAGG
GAGCAGCTCG AGAAGCGACT GAAGCTCTAT AAGTTCGAAA CCATCCTTCC CCCCGAGAAT
CAAATCAATT TTAGCGTAGA GGGCGTGGGC TATGAAAAAC GAGCCCGCTT GGAAGAGGCC
GTGGACCAAC TGATTAGGGG GGAGATACCC GTGGATATCG CTCTTGTCTT TCTTCCGCAG
GAGGACCGAA ACGCCGACAA CACCGAGGAG GGGAGCCTTT ACTCATGGAT CAAGAAGAAG
TTCCTTGACA GGGTTGTGAT AACGCAAATG ATCTATGAGA AAACGCTTAA CTATAAGAAC
AATTACAAGA ACATCCTCGA TCAGGTGGTG CCTGGAATCC TTGCGAAACT TGGTAATCTG
CCTTACGTGC TCGCAGAGCC ACTGGAAATC GCCGACTACT TCATTGGCCT GGATGTGGGT
CGCATGCCTA AGAAAAACCT CCCCGGGTCA CTTAACGTGT GCGCGTCCGT AAGGTTGTAC
GGGAAGCAGG GCGAGTTTGT GCGGTGCCGA GTCGAAGATA GTCTCACCGA AGGTGAAGAG
ATCCCCCAGA GAATCCTGGA GAATTGTCTG CCCCAAGCCG AGTTGAAGAA CCAGACCGTG
CTGATATACA GGGACGGTAA GTTCCAGGGC AAGGAGGTGG ATAACTTGCT GGCCCGAGCC
AGGGCCATTA AGAGCAAATT CATACTTGTC GAATGCTATA AAACGGGCAT CCCCAGACTG
TATAACTTCA AGCAAAAACA GATCGACGCG CCCAGTAAGG GCCTGGCGTT CGCTCTGAGT
AACAGGGAGG TGATCCTGAT CACGTCCCAG GTTAGCGAAA AGATCGGCGT GCCGCGACCT
CTGAGGCTTA AGGTACATGA GCTGGGAGAA CAGGTAAATC TGAAGCAACT GGTGGACACC
ACACTCAAGC TGACCCTGCT CCACTATGGA TCTCTTAAGG ACCCGAGGCT GCCCATCCCC
CTTTACGGCG CTGACATTAT CGCGTATAGG AGGTTGCAGG GAATATATCC CTCTTTGCTG
GAGGACGATT GTCAGTTCTG GCTGTAGTGA

51 Gene Name: gene-20403

Gene ID: BBF.2017.48.51

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AGAAAAGGAA AGTCGAGGAC CCCAAAAAGA AGCGAAAAGT GGGCAGCGGC
TCCTTGGACA ATTACATACT GACCGAGTAC AAGGCCGGCA TCCACGCCAG CGAGATCAAG
ATACACATCT ACCGGATGCC CGTCAAGGAT CTTGAGAAAA TCGACTATGA GTACGGGAAG
TACACACGCG ACCTCAGACA GAAAAACAGA AAGACGATAT CCTTTTACCG CTCTCTGATC
GGCAGCTTTG AGAAGCTCAC CATCGTGCCC AAGGGATACG AGAAGTACGA GTATAGATCA
ATTAAACTCG ACCAGAGTGA GGAGTCACTC CAGGAGAGGA AACTGCTGGA GAGGCTGATC
TTCGACGGCC TTAGGGACAG CAATAGGAAC CACTTTATGA GCACCGAGCA GAGCATCATC
GAGAAAGAGC CCATCAAGTC CCTGAGCAAG TGCAAAATCC ACCGGGGTAT CTACATAGAC
ATCACCGTGA AAGAGAAAGG CGACATCTTC ATCGGTTTCG AGCTGAAGCA CTCCATCCAG
AGCACCCACA CGATTATCAA GGCTCTGAAG GAGAAGAAAC TGAACAAGGG CGATAAGGTG
TTTGACTTTC TGAACAGCGC CCACTACGAG TTCGAGGGGA TTAGCGACAA AACCATCAGC
GACCCCCTTC CCGAACTGGG CAACAAGAGC ATTATCCAGC ACTACAAAAC GAAACCCAGC
ATCTACTGCC ACCTCGTGAA AAAACCGAAC ATGCCCGCCA TCCTGGTACG CAGCAAGAGC
GGCAAGGTGT ATCCTTACCC CCCACAGCTG CTTAAGAAGG AGTGCCTGAT GAAGGATGTG
CCGGCTAAGG AGCACAGCTC TATCAAGCTG AACCCCAACG ATAAGATCAA CTACAGCATT
GAGATCATGA AGAGAATCAT AGATGCGTTC GAGAACAGGT ATTTCCCCAT CGGCTTTGAA
AAGAACAACC TGAATATCGC CAAGCTCGGA TACAGGAGGA GGCTGGTCCC GGATCCCCTG

CTGAGGATTG GCAACGGAGC CACATGCAAC CACAGGGACC TCAAGGGTGC CTTCCTTAGG
CACAAGATTT ATGACAGCGT GAGCTCCCCT ATCTACTACC AGCTTCTGCT TGACCAACCC
TTCGAAAGGG AGTGGCAGAA AAAGATGAGC GAAGCGTTCA TTACGAAGAT GGAAAACCGG
AGCAGGCAGT GGGGCATAAA GCTTCAGTGT ACCGGGAACC AGATCCTCCC TACCTCTAAC
CCGTACGCGC TGAGACTGCA TCTTAAGGAC ATCAACCTGG ATACCGACAT CATTAGCGTG
GTCCTGTTGG ACGAAACCAA ACAAGAAGGC GAGGAGGTTT ACTCTACCAT CAAAAAAGAG
CTGGGTGGCA CCAGGGGCGC ACATACCCAG GTAATCCTGA TCGATAGCCT GAAGAACGAA
TACACTATCC CCCAGATACT GTTGGGAATC TACACCAAGG CTGGATTGCA GCCCTGGGTC
TTGCACCAGC CGTTGCACGC CGACTGCTAC GTTGGCTACG ACGTGAGCCA TGAAAATGGC
AGGCACACCA CTGGCATAGT GCAAGTGTTC GGCAAAGACG GGTCACAGAT CTTCAGTCAG
CCCATTAGCA GCGCGGAGGC CGGAGAGAAG GTGTCAAAGG AAACCATTCA GACTATGGTG
ATACACGTTC TTTACTATTA CCAGAAGAAA GTTGGCAAGA TGCCACAGCA CATTGTGTTC
CACAGGGACG GCCGAGGATA CGTAGAGGAG ATAGACTGGA TTAAAGACAT ATTGAGTAAT
AGGGACCTCA CCAACGGCCA ATCTATCGCT TTCGATTACA TCTCAGTGAT CAAAGAGTGT
GGTCGGCGCA TGGCTTACTT TGACGACATA AAGAAGAAGT ATGTGAACGT GCCCGGGATT
GCCTACCTGG ACGACAACGC CCAAAAGGCC TATCTTTGCA GCACCAATCC ATACGAAAAA
GTAGGGATGA GCAAACCTAT TAAGATTGTG AAGAAGATTG GCGAGATGAC CCTGGAGCAG
ATCGTAGAGG ACATCTATCA CCTGAGTTTT ATGAATATCG ACACCGATAG GAAGGTGAGG
CTGCCCGTGA CTACCAATTA CGCCGATAAA TCTTCAACGT TTTTCTCTCG CGGCTATCTG
TCATCACAAA AGAAAGGAAT TGGCTTCGTA TAGTGA

52 Gene Name: gene-9026

Gene ID: BBF.2017.48.52

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAAAGGAA AGTGGAGGAC CCCAAGAAAA AGCGGAAGGT CGGGAGTGGC
TCCGTGGCCG CTTTGAAGCG CTACTTTAAT GACAAGAACC TGATCGTGAT AGGCTACTCT
GGCAGGGACA AGAGCCTGAT GAGTGCGCTT ACCGAGGCTT TCTCTGAGAA GGGCTCTGGC
CGCATCTACT GGTGCGGCTA CGGCAGCCAC ATTTCCCCCG AGGTGGAAAG CTTGTTGAGG
ACCGCGCGAG AGGCAAACCG CGACGCCTAC TATATCGACA CCGATGGGTT CGACAAAACC
ATGTTCAGCC TGGTAATAAA CTGCTTCCAG GCGGATATCG AAAAGAAGAA AGAGATAATG
AGCATCCTGG AGTCTGCTCC CGAGGACAAC GATACCAGCC CGTTCTCAAT TCACATCACC
AGGACGGATA AATACCTTAA GTCCAACCTC TACCCGATCA TCTTTCCTAA GGAGCTGTTT
CAGTTTGAGA TAGAATATCA TGAGGGCGAA CGACCATGGA CCCTGCTGAG AGAGATCACC
AAAGACCAGA ACATTATCGC CGTGCCCTAC AAGCAAAAAG TCTACGCCTT GTCAACGGGA
TCAGCTATCA ACAACGTGTT TGGTAGCCGG TTGAAATCAG ATATAGAGAG GATTCCCGTG
TCTATGGATG ACATTGAGCG CAAGTCTAGT TACAGGGAAC TCTTCCTGAG GGCCACCCTT
CAGTCTATAG CCATTATAAG GGGCCTGAAC GTGGACATAC GACACAATAC CCTTTGGCGG
AGCGACATCT TTAGGAACGA CAATGGCACC CTCATCCACG AAGCGATCGA GTGTTCCCTG
GTGTTTGTGC CCCAACAGAA GTATGCCCTG TTGAGCTTGA GGCCCACCAT CTACATAGAG
AACTCTCATA CGGTTAGCAA GGAGAAAAAG CAGGAGTACG CCAGGATCTA CCTGGATAAG
ATGTGGAATA AAGCGTACAG CACGAAGTTG GCCCAGTGGG AATCTATAAT CTTTGGAGAC

ACGAGGCTCG CCTTCGAGGT GCCGCAAAAT TCAGGATCCG GGTTTAAGTT TCTGATAAGC
CACAACTGCG GCTTCAGCGA AATCCAGTAT CAAGACAACA CCGAAAGGGG ATACAGTAGC
AAGAGCTACG ACAACAAGAG GACGATCTAT AGGGGCTTGC AGCTGAAGGA ACCCGAGCTG
GAATTTGTCA ATACGTTTGC AGACCGGCCC TTCCTGGACA GCAACCCCAT GCGAGGCCTG
AGCAATCACA GGCCGTACGA CAGCTGGCAG AAAGACGTTC TCTTGCAGAA CGTGCGGTTG
GGCGTGATTT GCCCGAACAC GCACACCGAC CGATTCCACT CTTTTCTGCA GCAGCTTAAC
ACCACAATTC AAGCCAATGA CGATAGCGAC TACATTCAGT CCTACACCGG TTTCCATAGC
ATTTACAAGA CTCTGCTGGA AATCCCCGAT AACGGGACCG ACAAATGGAT AAACATCGAG
GATACCCCCA AGGACACCAT CAGTCTGGTT CAGAGTATAT GTCACCAAGC GAACCGACTG
GCCGACAAGT ACCCGGGCAT CGTGGTGGTG ATTTTCATCC CCGCATTTTG GTCTATCCAT
CGACAGTTCA AACACAACGG GGAGAGCTTC GATTTGCACA ACTACATCAA GGCCTACGCC
GCACAACATA GCTTCACTAC CCAAATCATT GAGGAAAAGA CGCTGCGCGA CCACATGGTC
TGCGAAATTT GTTGGTGGCT GTCACTCGCA CTGTTCGTTA AGGCTATGCG AATCCCGTGG
GCACTGGCCA ATTTGGACTC TGACACCGCT TACGCGGGTA TAGGGTACTC AGTCAAGACC
AACAGCAAAG GCAACGTCGA CATAGTGCTT GGATGTTCAC ATATATACAA CGCAAAGGGC
CAGGGACTCA GATACAAACT CTCTAAGGTC GAGCAGCCCC AATTCGATGG CAAGAAAAAT
CCTTACCTTA CGTATGAAGA GGCCTTCAAG TTTGGAATTA CCATACGCGA GTTGTTCGTC
AAAAGTATGG ACCGGCTTCC CAGGAGGGTT GTGATTCACA AGCGGACGCC GTTCAAAAAG
GAGGAAATAG AGGGAATCAC TCACGCGTTG ACTCAGGCTG GCATTAAGGA CATCGATCTC
ATTACGATCA ATTACGAGTA CGACGCCAAG TTCATAGCGC AGAAGGTATA CTATGACAAC
ATCAGCGACG ATTCATATCC CGTAAGTAGG GGCACATGCA TCAAATTGTC CAGCCGAAAT
GCGCTGCTGT GGACACACGG CGTGGTTCCC TCAATCCGGG AAAGACGACG CTACTACCCC
GGTGGGCGCT GTATTCCCGC ACCCCTGAAG ATAACAAAAT ACTACGGTAA AGGCGATCTT
CCGACAATCG CCAGCGAGAT TATTGGATTT ACTAAGATGA ATTGGAACAG TTTTAATCTG
TACACGAAAC TGCCCGCCAC CATAGATACG AGCAATACAT TGGCACAGGT CGGCAATCTG
TTGCATCAGT ATAACGGCGC AACTTACGAC TACCGATATT TCATCTAGTG A

53 Gene Name: gene-9487

Gene ID: BBF.2017.48.53

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAAAGAAA GGTGGAGGAC CCTAAGAAGA AGCGCAAAGT GGGATCCGGC
TCTATGTTGG AAACGAATAT CAGGGTGGTG CGGCCTGGTC CGCAGCTGTG CGTTCCTGTA
CGCAGGGTGA TCGTGTCCGG TCAAACCTTG GCTCCCGACC TCCTGGAGAG GCTGTGTAAT
CTGCTGCGAA GGAGGTACGG CATTAGCGCC GCAAGAATAC CGGGCTCCGT GAGCGAGCTG
TTCGTTGCGA CCGACCGCCA GGTGGAGAAG GTGACACTGG AAGAAGATAA CTGGCAACTG
ACCGCCGTGG ACTCCAACGA CCCTACTCGA ATCATGTCCA TCTCTAACAC GGACGATGAG
AGCTTTATAA GCATCCTGAT CGAACGCGCG CTCCTTGCCC AGATCGCCAG TCGAAGCCTC
TTTTGGACCC TCGACTCTCC TCGAATTTGG TATGAGAAGA ACCCGTTCCA AAGGAATGAA
GGCGTAGCCG TCTACCACAG GTACGAGGTG GATGCGCTCC CCCTCGGCGA CGCAGGCATT
GGCATCTCAG TGGATGTTTC AACGGCCTTT TTTAGCGAGC ACACCCTGGA GTACTACTTC
GCCCCCAACC TGATTAGCGG CGAGAGCAAG ACGCGACAGG ACGAATTCCA CAAGTTCACC
GGCCGACAAG CTGGTCAAAA GGGGACGCTG CTTTACAATA ACGGCAGGAG TAAGGTGAAG
TGCTATTTCG AGAACAATAG GGTGGGCCTG ACATGTGGCG CAACCGGCCA AATGAAACTC
GAGGGAATCA CGTATCCCAG CCTGTACCAC TACTATGCGA GCAAGTATAG CGCATTGCAG
ATCAACGAGA ACGATGCCGC AGTGCAAGTG TCTTTCCCTG GCTTGGACCG CCCAGTTCCG
GTAGCCGCCA GGCTCCTGTC CCTCCGAGTG ATGAACGACG ACGTGCCCGA TGGTCTGAGC
TCCGTCGACA AGATCCCTCC AAGGAACCGC AAGTACCTTA TCGAGCAGTT TTGGAAGTGC

CTGGAGCCCA GACCCTTCGG GAATGTGGCC CCTGGTGTAT TCGACGGCTT CTGGCGCCCC
AACAACGAAA GGGTGCATTA CATCCAGCTG CCCGAGATTA ACTTTGGACA AGGCCAAAAA
GCAGAACCGC CTGACGTACG CTCCGTTGCA TCCATCAAAA ACTATTTTAG GCGACGACTG
GAATTGCTGG GTCACGCGGG GTGTTACCAC TTTCCGCCCT CAGCCCCCAG GACAATCTTC
TGCGCCTACC CGCAGTCATT GGGTGAGGAG ATCCCGGAAA AGTTGGTGAA CGGGATCGTC
AATGTGCTGA ACAAGTGGAC CGGCCTCAGC TTCTGTAGCA ACCTGGTAAG CTACAGCACG
GCCAGCGAGG CGTACGGTAA ATTGAGGAGG GCCGAGAGTG CCGGCGTGGT CCTGTTCATC
TTGGACGAGG AGCCGGCAGT CTACTACGAC GCGAGCTTCA ATCTTGAGGG CTGGAGGGTA
AAGCGCGTAA CCGAGCCTGT GCTGCGCCAG CAGCATAAGT ATCTGACCAA CGGCGTGTGG
GACCGGAAGA GGCAAGAGTA TAGTTTGGGG AGGGGGCAGA GTCGCTGGGA AAGCTTCATC
AATTTGATCG GATTGGACGT TATCCAGCAA CTCGATGCCA TTCCGTATAG GATCCCCAAC
ATCGGCCCCT ACGAAGGCCA GCTGATAATC GACGTGGGGC ATGACAGGCA ATTCTTCGCC
GTGTCACTGC TTATTGTGAG ATCAGAGGAC AAAGTGCCCG CATTTAACAT CAGCAGCCAG
GTCCAGCACA AGGCGGATCA TAAGCACGAA AGCATTAACC CGGTGCTGTT GAAGGACACC
ATCATTAACG TGTTCAAGAC CGCCAAACGG AGGACTTTTG ATCCTCTGAC TAGCCTGTTG
ATCATGCGGG ATGGCAACGT GCAGGGCAGC GAGATCGGCG GGATAGACAA CGCCCTGGTC
GAACTTAGGC AACTTGGCAT AATCTCCCCC GATGCGAGGC TGGACATCGT GGGCGTACAC
AAGGAATCTG TAAGCTCCAT CAGGCTCTGG GACGTTGACG TAAGGGGGGA GGTAAGCAAC
CCGATCGAGG GCACCGGTCT GTCAGTCAAC TCATCTCTGT ACCTGGTGGC GTGCACAGGT
GAGGCCACGC TGACCCAAGG CACCGCAGAG CCCGTGGCCA TCGTCGCAAA CAACAGGTGC
CTGAGTATTG CCGATGCAGC CCTGAGCGCC TTTCTGGCAG CCCAACTGAA CTGGAGCAGC
CCGGGAGTCG CCCAGCGCCT GCCCCTGCCT CTGAAAAGAA CAGATGAGGA ACTTACCGCT
AGGAGCGATC AAGAAATTAG GAGGATAAGG TAGTGA

54 Gene Name: gene-9711

Gene ID: BBF.2017.48.54

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AGAAACGAAA AGTGGAAGAT CCCAAGAAGA AAAGGAAAGT TGGTAGCGGC
AGTATGATAA TGAGCCTGGA GAGCAATATC TTCACTTTTA GCAACCTCGG GACACTTACC
ACGCAGTACC GACTGTATGA GATCAGAGGC CTGCAGAAAA GGCACCAAGA GTACTACCAG
AACAGGCAAA TCCTGATCCA CCGACTCTCC TACCTTCTGA AAAATGCCGT AACTATCATA
GAGCGCGACG AGAAACTGTA CCTTGTTGTA GCTGCCGATG CCCCGGAACC ACCCAATAGT
TATCCCATCG TTAGGGGCGT CATCTACTTC AAGCCCACCG GCCAGATTCT GACCCTGGAC
TACAGCCTCC GAACACCCCA GAACGAAGAG ATCTGCCAGA GGTTCCTCCA TTTCATGGTA
CAAAGTGCCC TGTTTCAAAA CGCGAATTTG TGGCAACCCA GCGCCGGAAA GGCTTTCTTC
GAGAAAAAGC CCTCATTCGA GTTCGGATCA ATTCTGTTGT TTCAGGGATT TAGCGTTAGG
CCCATATTCA CCAAGGACAA GATCGGCCTG TGTGTAGACA TCCACCATAA ATTCGTCAGC
AAAGAACCCC TCCCTAGCTA CCTGAACTTC AACGAGTTCC AAAAATACAG AGGCGTGTCA
TGCATCTACC ATTTCGGCCA CCAGTGGTAC GAGATCCAAC TCTCTGAACT CTCCGAGCTT
AACGCGACGG AGGCAATGGT ACCCATCGAG AATAAGTTCG TGACCCTTAT TAACTACATC
ACCCAGCAAG CCAGGAAGCC CATCCCGGAG GAGCTGGCAA ACGTGTCACA GGACGCAGCC
GTCGTGCACT ACTTTAACAA TCAGAACCAG GACAGGATGG CGGTGACGAG TCTGTGCTAT
CAGGTTTACG ACAACTCTTA TCCAGAAATC CGAAAGTACC ACCAGCACAC CATTCTGAAG
CCACACATCC GCCGCAGCGC GATCCACGGA ATAGTGCAGA AGTATCTCGC GGAGCTCAGG

TTCGGCGACA TAACCCTGAA GGTATCAACT ATCCCCGAGC TGGTGCCCCA GGAGATGTTC
AATCTGCCCG ACTATTGCTT CGGCAACGAT TACAAACTGA GCGTGAAAGG AAGCGAGGGC
ACAGCCCAGA TTAGCCTCGA CCAGGTCGGG AAGCAGCGCC TTGAGCTGCT GAGTAAGGCT
GAAGCTGGTA TCTACGTGCA GGAAAAGTTC GACCGCCAAT ACATTCTCCT GCCCCAAACC
GTGGGGGACA GCTTCGGGAG CCGGTTCATC GACGACCTCA AAAAGACCGT GGACAAGCTG
TACCCCGCTG GAGGAGGGTA CGACCCGAAG ATCATTTACT ACCCCGACCG AGGCCTCCGG
ACCTACATCG AGCAGGGTAG GGCTATACTG AAAACAGTTG AAGAGAACGA GCTGCAGCCC
GGCTACGGTA TCGTAATGCT TCATGACAGT CCGGATCGAC TGCTCAGACA ACACGACAAA
CTCGCAGCTC TGGTCATTAG GGAGCTGAAG GACTACGATC TGTACGTGGC CGTCATCCAC
AGCAAGACCG GGAGGGAGTG CTATGAGTTG AGATATAACA ACCAGGGCGA GCCCTTCTAT
GCAGTAATAC ATGAAAAACG GGGGAAGCTC TACGGCTACA TGAGAGGGGT GGCGCTCAAT
AAGGTGCTTC TCACCAACGA GAGGTGGCCC TTTGTGCTTT CTACCCCCCT GAATGCGGAC
GTGGTGATCG GAATCGACGT CAAGCACCAC ACCGCCGGTT ACATAGTCGT CAACAAGAAC
GGGAGCAGGA TCTGGACTCT GCCCACGATC ACGAGCAAGC AGAAGGAGAG GCTGCCCAGT
ATCCAAATAA AGGCGAGCTT GATCGAGATC ATCACTAAGG AGGCCGAGCA AACAGTAGAT
CAGCTGCACA ACATAGTGAT ACATAGGGAC GGACGAATAC ACGAAAGCGA GATCGAGGGC
GCCAAGCAGG CCATGGCCGA GTTGATTTCA AGGTGTACGC TGCCTGTGAA CGCCACACTC
ACGATCCTGG AAGTGGCTAA GAGCAGCCCC GTTAGCTTTA GGCTGTTTGA TGTCTCCAAT
ACCAATTCTA AGGACCCGTT TGTGCAAAAC CCACAAGTCG GGTGCTACTA CATTGCCAAC
AGCACTGACG CCTACCTGTG TAGCACGGGG AGGGCGTTTC TCAAGTTTGG CACCGTGAAC
CCCCTGCACA TAAGGTATGT GGAAGGTACG CTCCCCCTTA AACTGTGTTT GGAGGACGTG
TACTATCTGA CAGCCCTGCC TTGGACGAAA CCCGACGGGT GCATCAGGTA CCCCATTACC
GTAAAGATCA ACGACAGGAG GCTTGGGGAG GACGCCAGTG AGTACGACGA GGACGCCCTG
CGCTTCGAGC TGTTCGAGTC TCTCGAGTCC GAGGATGACT TTGACGAGAT GACCGACAGC
GACTTTAATC AGGAGGAGAC AATGGTGTAG TGA

55 Gene Name: gene-12723

Gene ID: BBF.2017.48.55

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAAAGAAA AGTCGAGGAT CCCAAGAAGA AGCGGAAGGT GGGGTCCGGG TCTATGCTCA CACAAGAACA ATTTATACGC AACTTTAGCG TTATGGCCAA TGGTGAAGTA GACTTCTTTC TTGGTGCCGG TGCATCTATT GCGAGTGGAA TCCCAACTGG GGGTGGCTTG ATTTGGGAAT TTAAGAGGAC ACTGTACTGT AGCGAGTGCG GCATCAGCGC CGAAAAGTAC AAGGACCTGT CACTCCCAAG CACGCGCAAA ACGCTCCAGG ACTACTTCGA CATTAAAGGG TATTGCCCCA AACAATATGC GCCTGAGGAA TACAGCTTCT ATTTCGAGCA ATGTTACACC GATCCCATGG CCCGAAAGAG GTTCATCGAG AATATGGTTA GTGGGAGGGA GCCAAGTATA GGTTACCTTT GTCTCGCGGA GGCCGTTATG CAAGGCAAAG TTAAAAACAT TTGGACTACC AACTTCGATA GCCTTCTGGA GAATGCCCTC CATAGGCTTT ACCCCATGAA CAACGTTTTG GTGTGCTCCG AGGCTAATAG AGGCAGTGTG TGCCTGCTCA ACCCGACGTA CCCAGTCATA GGCAAGCTCC ACGGCGACTA TCGCTATGAT TGGCTCAGGA ACACCGAGGA CGAATTGCAG CGACTCGAAA CCAGCCTTAA AGGTTACGCG TCCAGCCAAC TTACAGGGAA ACAACTCGTC GTTATAGGAT ATAGCGGGAA CGATGAGAGC ATTATCAGTT TCCTCAAGGA TTGCATAGAT AACCCGGCAC TGCTTACCAA GGGTCTGCTG TGGGCTGTAC GACGCGGTTC CTGGGTAAAC CCGAGGGTTA ATGAGCTGAT AGAACGGGCG CACAAAATTG GGAAACCAGC CGACGTGATC GAGATCGATG GCTTCGACCA ATTGATGTTC TCAATATACC AGATCCAGAA CTACCATAAT GAGATTATCG ACGGCCAAGG CAGGCTCCTC CAGGTCGGAT CTGACATCCG CCTCACGGGG

AAGCCCGTGG ACAGCTTTGT CAAGCTGAAC GCTTACAAGG CTGAGTACTG CCCCCTTTGT
AACGTGTTCG AGACAGACAT CACATCCTGG AAGGAACTTC GGACCATAAC CGGCAGCAGT
GACATTATCG CCGGTCTGTT CTCCAAACAT ATCTATTCTC TCTCTTCCGC AGACAAACTC
AAGACCGTGT TCAGCAAGCA CTTTCTCTCT AGCATTAACA AGGAGGAGGC TCCCGAACGG
GACATTCGAC GGAACGAGAG TGTGTACATT GGATTGATTT ACCAGCTTAT TAAGCGGACC
CTGCTTTCAA AAGGGATGGT GTCCTTCGCT AAGAATAAGG TCTATAACCC CGACAGCTGC
CGCAGCGAGC AAGGCTACCA AGTTTTTGAC GCCCTGGAGA TCGCGGTCAG CTTCGTTGAT
GGAAACCTGT ACCTGAATCT TATGCCCACG GTACATGTGA GAGGCTCAAA TGGCGAGAGT
CTCGACAAAG AGTCCTACCA AATACAAGTC AACCATGTGG TCAGCACAAT CTACAATAAG
CAATACAATG AGAAACTGCG GTTCTGGGAG AGCTTGTGTC TGGACAGTGG TAGAATAATC
TTCGAGAACG ACGGCTTCAG CATATCATTT GTCGCTCCCG CTGTCTCCCT GGGCGGCAAC
AATCGGAGAG CTAAGTGGCT TTCCATGCCG TCCTGCAAGT ATGACGAACC ACTCATGTGC
TTCTCAGACA CTGACAAAAG CAAACGAGTT ATTAACCAAC TGAAGGGACT CTGCCAGTAC
GGGCCAATCG ACTGCTCTTA TATGCGGGAT AGCACCACAA GGCCCAGCGT TAGGCTGGCC
GTTCTGAGCC CGAACCAGGA CATGGACCGA ATTCTTGCAC ACCTCAATAA ACTCAACACC
CACGTCCAAA ACAGGGGCAG CGATAATTTC CTGCCCCACT ATGAGGGCTT TGAGCAAGTT
TACAGAAGGG CTCTGAGCGT CCCTACGAAG GAGCAGAGCA ACATCTGCAT CGGATACAAC
GTGAACGCCA TCCTCAAAAT GTCTCCTGCA GAGTTTCTGG CTTTTATGAA GCGGGGTATA
GAGAAATACT CCCTTCGGTC AAGCGATTTC GATATACTCG TTATTTACAT CCCAGAGTCA
TTCGCGCATT TCCGGACAGC AACCGAAATT AGTAGCGACT ACAATCTGCA CGATGCGCTC
AAACTGTATG CCACGGATAA GGGGATTATC CTTCAACTCA TAGAGGAGAA ATCTGTGAAG
TCATACGACC CCTGCAAAGT AATGTGGGGC TTGTCCACCT CACTCTACGC GAAGGCGACA
GGGGTACTTT GGCATCCAGA GGCAATTAGA AATGACACGG CCTACATAGG GATAAGCTAC
GCTTTCAGCG AAGAGAAAAG GATTTGTATA GGCTGCAGTC AGCTGTTCGA CTCAACCGGG
ACAGGTATTC GGATGGTCCT TAGAAAGATA AACAATCCGA TATTTCTGGG GCGATCCAAC
CCCTACATGA GGGAGGACGA CGCTCGAATT ATGATGACCG AGCTCAGGGA GCAGTATTAC
CACAGCGCAC CTGTGAATAC TCTCAAGAGG GTCGTGATCC ATAAGACCAC GCCCTTCATA
CGGGATGAGA TAGCCGGTAT AATGCAGGCA TTTAACGGCA TCGAGGTCGA GCTGGTTCAG
ATTCAAGACT ATTGCTCTTG GAGAGGCATA CGCTTCGGCG GTGAGCCTGG GAAAACGGCG
TTTGGGTTCC CGGTGAAGCG AGGTATGGCC GTAAAACTCG ACCGAGAAAG CTTCCTGCTC
TGGACCCACG GCTGCGTGAT TCACCCGGAA CTGTCAGGCA CGCATAACTA TTTCAAAGGT
TCACGCGGTA TCCCAGCACC CCTCCTGGTC AGGAGGTTTG CGGGTAACGC AAGTGGCGAC
ACATTGGCAA AAGAGATTCT GATGCTTACG AAGATGAACT GGAACTCCGG TGACAGTCTG
TACAAAACCC TTCCCGTGAC CCTGGATTTT GCGAAAGTTC TCGCCCGCAT GTCTAAGCAA
GATGAGGCGA TCTTTGATAA GGCGTACGAC TTCAGGTTTT TCATGTAGTG A

56 Gene Name: gene-14016

Gene ID: BBF.2017.48.56

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAGCGGAA GGTTGAAGAT CCAAAGAAGA AGAGGAAGGT GGGGTCTGGG
TCAATGCTCC TTAATCATCT CCCAATCGAG TTCTCCAGCG CACAGTTCGC TGGACACGAA
ATTGCTTATG TCGACGGCGA GCAGTTGAGG TCCATACGAC AGAGACTCAC GCGCACGCAC
TTCGTGTTGA GGGATGGGGA CAATGTTCTC CTCTTCCCGT ACGAACATGG AACCGCGACC
GAGGGAACCA GGCGAACATT CGACACGGGC GTTAATTTCA GCGTAGCCAA CGCCCTGGCG
CGCAACGGCA TGCTTCTGCG ATTCTTCCAG CACTCTAGAA GTATTTCCGG CGTCCGACCG
GTGAAATTTG TGAAAGACAA CCAGAATCTG CTCACGGGTG ACGTAGGCCG GTTGTTTGCT
ATATGTCCGG AGTACAGTTT CGACATCCGA CCCCTGGCAC CTCAAGACGG CAGCCTTGTG
AACGGGGTAC TGGTAAACTT CTCAGCCCGA TTTTTGGTGA AGCCCTCCCT CGACGAATTG
ATTGCGCAGG GGCTCGACCC ACGGGGCCTG TATGTTGTTA AAGAGGCAGA AAGAGAATCA
CCCTACATCC TGCCGATGTT TAATCGGAGA TTGGTAGGGC GGATCCAGGA CGTGGTCGGA
GGTATCGCCA AGCTGGTGGA CGAGCGCGAA CAGGACCTCC CTGTACATGA ACTTCATGTC
GAGGCCAACC TGGTCAACTT CGAGAAAGTA GGCAGAGCAC TGCTTGGCCG GGATTACGAG
CGAGTGAGTC GACAAGTGCT TCCCACCCTC CATAAGGTGA GCGGCGCAGA GAAACAGCTC
GATCGCTTGG TCCAGCTGCT GACGAGCTTC AAAGACCTCC AGGGTGACAT CCCGTGTTGC
GACGGCCTGA CCGTTAGACT GGCAGGCATA CTTACAGATG TGCCCTTCGG CAGTGAGGTG
GGCCAATTCC GCAAATTGTC CGCGCCACAG TGCAGCCTCC GCCCAGGGGG AACTATTACG

GTGCCGTGGC CCGTGGACGG CAAACTCAAT GCCAACGGCC CCTTTGATGC AGACGCCTTC
AGCAGGAAGG AACCAACAAT CGGCGTTCTG TTTCCGGAGC AGCACAAGGG TAGTGTAGAG
GAGCTGGCCG CTAAACTCAG GGACGGCGCA CCGTCTGATG GAAAGTACCC AAGTCCATTT
CCCCAAGGAA TGCCCCGGAA GTATAGACTT AGGAAGATGA CATATGAGCT GACGCCCACG
AAAGTTTCAG GGGACAGGGC CGCAGCCTAC AAGAATGCCG CGCTTGCAGC CGCCCAACAA
GAGCTTGATC TCGCTCTGGT GGTCATATCT GAATCAGATA AGGCGTTGCT TGGAGCCGCC
AGCCCCTACT ACACTGCGAA AGCCACATTG ATGAGCCAAG GCGTGCCGGT GCAGGCTATT
ACCATTGAGA CTATCAACAG GCTCAACCCC TACACCTTGA ATAATCTGGC ACTTTCCCTT
TACGCAAAAC TCGGCGGGAT ACCTTGGACC CTGTCAGTTC AACAGCGACT GGTCCACGAG
ATAATTGTAG GGATAGGGTC TGCGAGAGTG GGCTTCGACC GCCTCTCAGA GCGGGAGAGG
CTTGTCGGCA TCACGACCGT GTTCTCCGGG GACGGATCAT ACCTTCTTGG CAATGCAACG
ACGGAAGCCA GCAGTACCGA ATATCGATCT CGCCTTCTGG AGAGCCTTAG GGCGACTTTG
GCAGAGTTGC GCAGACGATT TGGCTGGCAG CGGGGAGATA AATTGAGGAT TATCTTCCAC
CAAAGCTATA AGCGGTACAA GGAAACCGAA GCAACCGCCG TTAGCGACCT GATCGCCGAA
CTTGATGAAT TCGATGTGGA ATTCGCGTTT GTGCAGATCA GTAGCGATCA TGACTGGAAG
TTGTTCGATG AGAGTGCCAC AGGCGTTACG TATCAGTCCC GGCAAAAGGG AGCGAAGGTG
CCGGAACGCG GAGTCATAGT CCCTCTCGGA CCTCGCGCTG CGCTGATCAC GTTGGTGGGT
CCGCATCAAC TGAAAACCGA TCTGCAAGGG TGCCCCTCCC CCATACTGGT GTCTATCCAC
CCGAGCTCAA CTTTCAAGGA TTTGAGTTAC GTGTCAAAAC AGGTGTTCGA CTTGACCTTT
ATGAGTTGGC GAAGCTTTAA CCCAAGCACG CAGCCCGTTT CCGTGAGTTA TCCCAACATG
GTGGTGGATC TGCTCGGTAA TCTGCGGCAA ATCCCCAACT TCAATCCCGA CATTCTGACG
ACAAAACTGA GGGAGTCTAG GTGGTTTCTG TAGTGA

57 Gene Name: gene-1475

Gene ID: BBF.2017.48.57

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAGCGAAA AGTTGAGGAC CCCAAGAAAA AGCGCAAGGT CGGGAGCGGA TCTATGATGG GAGCCTCTGA TGAGTATTCC TTTTACGCTG AAAAGGCCTA TCCCATAGAA GCGGACAGGC AAAAGTACTT CGAACAGCTG GCGTACAACA AAGCCCCCTA CATTGGCTAT AAACTCTTGT GTCTGCTGAA TAACGCGGGG CTGATAAAGT CTGTTTGGAC CACAAATTTT GATGGCCTGA CGGAAAGGGC CGCTCACCAA ATGAACATCA CCCCCATCTG CATTACCCTG GACGACCCCG AGAGGATTTT TAGGAATGAG AACTCTCACG AACTGCTGTA TATCGCCCTT CACGGCGATT ACAAATATAG CAAGCTCAAA AATACCACCC ACGAGCTGGA CACCCAAAAC AATATCTTCC GCGACGCACT GAAGCGATAC TTCGTGGATA AGAATCTTAT TGTCATAGGA TACAGCGGCC GAGATAAAAG CCTGATGAAC GCACTTAAAG AGGCATTTTC CCAATCCGGC TCCGGGCGAC TGTACTGGTG TGGCTTCGGG GACGATATAT GCAGCGACGT TAAGGAATTG ATAGATATCG CCAGGAGCAA TAATCGGATT GCCTACTTCA TCCCGACGGA CGGCTTCGAT AAGACCATGC TCCAACTTAG TCGCGCCTGT TTCGAGGACG ACATTGTGAA GCAGGAGGAA ATCAAAAAGC TGATCAAGTC CACGATCAAG AAGGACGAGA CTAAGACCAG CTTCCGAATC GAGAGCAGCA GGAACGATAA ACTTATTAAG TCTAATCTGC ATCCCGTGGC GTTCCCCAAG GACGTGTACC AGTTCGAGAT TAAGACTAAC GGCGAGCATC TGTGGAACAA CATAGACCAG ATCATTGGCG GCAATAAGGA CATAGTTGCC GTACCGTTCA AAGGTAAGGT GTTCGCTGTC TCAAGCATTG CGAAAATCAA GGAGAGGTTC GGGGGCTATA TCAAGGGGGA AATATTGAAA

GACCCGATTG GCGTCGATGA CATCCGCAAA GTATCTGTGT TCCAGCGGCT TATGATGAAA AGCATCCTGA TTGGAATCTC TGAGTTGGCA AATCTGGAAA CTGATGGAAA GTGGCGCCTT TGGAAAAAGA ACACCCTGAG GCGAATCGTA AACGGCACGG AGTATTTTAT CGCCGACGCT GTAGAGCTGT CCTTTTTCTT CGGAAAAGAT ACCAAGTTTG CCTATCTCAG CATCAAACCG ACCATTTACA TTTATACACA TAGCGACGAA TTCATACCGA AGGATATAAA GCTGCAATTC ACAAAGGAGA AGTTCGACCG ACTCTATAAT GCACAATACG ACCAATCCCT GGAGGAGTGG AATAATCTCA TCTTCCACAA CAACAGCCTG AGGTTCACCT TTCCCGTACT GACCACCTCC GACATGAGCT TTAGCATCAG CAACAATGTG GCCTTCTCAG GAATTAAGGT TTTGAGTGAC AAGTATAAGA GCTACCCCGT TTCTATCGAG CAGAAGCGCA TAGTTTTCAA GGGCGTGGAG TTCCTGGAGC CCCAGCTGCT GTTTCAAAAT AAGAACAGCA ACTTCAAGTC ACGCGACTTC CATCCCATGA GGGGATTGAT TAACCACTAC CCCTTCGACT ACCAGAACAA TGGGATCACC AACACGTTTA ATGTCAAACT CGGCGTGTTG TGCTCCTCTA AGTACTCTAC TAGGCTGTAC GAGTTTCTCA TGAAATTGAA TGCCCAACAT AAAGCGCCCG AGAAAAACGA GTACATAATT GACTATGCTG GATTCAACCA AATCTACAAC ATCCCTATTG AGATACCGCT GGTAAACGAC GAGAAGTGGA TGGACGTAAA GTTTAATAGC AGCGTGAGTA TCAAAGACGA CGCTCTCAAC CTGGCAAGAA TCATATGCAC CCAGATCGAG GCGCTTCACG AGTCTTACAA AACTGACATG ACCATCGTGA TCTTCATTCC CAACGAGTGG CAACCCTACA GACATATCGA GGAGGACACA TGGGTTTTTG ACCTCCACGA CTACATCAAA GCATATAGCG CTCAGAAAAG AATTTCCACG CAGTTCATAG AGGAAGATAC TCTGAACGAT TCATTGACGT GCCAGATATA TTGGTGGCTC AGCCTTAGTT TTTACGTGAA ATCCTTGCGG ACGCCGTGGG TTCTGAATGC TAACAATAAT GAAACCGCTT ACGCGGGCAT CGGCTACAGT ATAAAGAATA ACAACGGTGA GGCGTCAATT GTCCTCGGGT GTAGCCATAT TTACGACAGC CACGGCCAGG GCCTCAAGTA CAAATTGAGC AGAGTGCAGG ACTGCTACAT CGACAACAAG CGGAACCCCT ACCTGAGCTA CAATGAGGCC TACAACTTTG GCATAAGTAT CAGGGAGCTC TTTCTGCACA GCATGGAGTA TCTGCCAAAA AGGGTAGTAG TGCATAAACG CACCGAGTTC AAACCCGACG AAGTGAATGG CATTGTCGAC TCACTGCAGA TAGCGGGTAT CGAGAATATA GACCTTATCT CCATCAACTT CGAGCGGGAA GTTAAATTCA TGTCCACTAA ATCCAACTAC GGGCAGTTGC AAATCGATAA CTTTCCCATA CGCAGGGGCA CCTGTATCGT GGTGAACGAC TATGAAGCCC TTCTCTGGAC CCATGGAATT GTGCCGAGCG TTAAGTCCGA TAACAGGACC TTCTATCTGG GCGGACGATC TATTCCTAGC CCTCTTATCA TTAAGAAGCA TTACGGTAAG AGCGATATCA ACGTTATCGC TACAGAGATA CTGGGTCTTA CCAAGATGAA TTGGAACTCT TTTGATCTCT ACACGAAGCT GCCGGCCACC ATCGATAGCT CTAATCAAAT CGCGCGGATC GGGAATCTGC TGACTAGGTT CGAGGGCAAG ACCTATGATT ACCGGTTTTT CATTTAGTGA

58 Gene Name: gene-15288

Gene ID: BBF.2017.48.58

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGAGGAA GGTCGAAGAT CCTAAAAAGA AAAGGAAAGT CGGGTCCGGT
AGCATGCCCA CCCAGTTCCA GGAGGTGGAA GTGATACTCA ACCGCTTCTT TGTAAAGAAA
CTGTCTCGGC CCGACCTTAC GTTCCATGAG TACCAATGCC AGTTCACCCA GGTTCCAGAG
CAAGGCAGCG AACAAAAGGC CATCAGCAGC GTGTGCTACA AGCTCGGTGT GACCGCCGTG
AGGCTGGGCT CATGCATCAT CACCAGGGAG CCCATAGACC CTGAAAGGAT GCGCACCAAA
GATTGGCAGT TGCAGCTGAT CGGATGCCGA GAGCTGAGCT GCCAAAACTA CCGAGAGAGG
CAAGCTTTGG AGACTTTCGA GCGAAAAATC CTGGAGGAAA AGCTCAAGGA AACATTTAAA
AAGACCATCA TCGAGAAGGA CTACGAGTTG GGCCTGATCT GGTGGATATC AGGCGAAGAG
GGACTGGAAA AAACCGGTCA CGGGTGGGAA GTGCACAGGG GCAGGCAAAT AGACCTCAAG
ATCGAAACGG ACGAAAAGTT GTACCTGGAG ATCGACATAC ATCACAGGTT CTACACCCCC
TTCAAGCTGG AGTGGTGGCT GAGCGAATAC CCCAACATCC AAATCAAGTA CGTGCGCAAC
ACGTACAAGG ACAAGAAGAA ATGGATACTG GAGAATTTCG CCGACAAGAG CCCCAACGAG
ATTCAGATAG AGGCCCTTGG CATCAGCCTT GCGGAATACC ACCGGCAAGA AGGTGCTACC
CAGCAGGAAA TCGACGAGAG TAGGGTTGTG ATCGTCAAAA AGATCTCTGA CTACAAGGCG
AAACCCGTGT ATCACCTGTC TCAGAGGCTG TCCCCGATAC TGACCATGGA AACCCTTGCC
CAGATCGCCG AGCAGGGTCG GGAAAAGAAG GAGATACAGG GCGTGTTCGA TTACATTAGG
AAGAACATCG GCACGAGGCT GCAGGAGAGC CAGAAGATCG CCCAGGTCAT TTTCAAGAAT

GTTTATAACC TTAGCAGCCA GCCCGAGATC ATGAAGGTGA ACGGTTTTGT AATGCCACGC
GCGAAGTTGT TGGCAAGGAA CAATAAGGAG GTCAACCAGA CCGCTAGGAT CAAGAGTTTC
GGCTGCGCTA AGATCGGAGA AACGAAGTTC GGATGTCTCA ATCTGTTCGA CAACAAACCG
GAGTACCCGG AGGAGGTACA CAAGTGCTTG CTGGCGATTG CGCGGAGCAG TGGGGTCCAG
ATAAAGATAG ATAGCTACTT CACGGGGAGC GACTACCCGA AAGATGACTT GGCCCAGCAA
AGGTTCTGGC AACAGTGGGC GGCACAAGGA ATAAAGACGG TGCTGGTCGT GATGCCCTGG
TCCCCTCACG AGGAAAAGAC AAGACTGCGG ATCCAAGCTC TTAAAGCCGG GATCGCAACT
CAATTTATGA TCCCCACGCC CCAGGATAAC CCATACAAAG CATTGAACGT TGCTTTGGGT
CTGCTCTGCA AAGCCAAATG GCAACCCGTT TACCTGAAGC CCCTGGATGA CCCCCAGGCC
GCAGACCTGA TCATCGGCTT CGACACTTCT ACCAACAGGC GGCTCTACTA CGGTACAAGC
GCCTTCGCGA TTCTGGCGAA CGGCCAGTCA CTGGGCTGGG AGTTGCCTGA CATCCAGAGG
GGCGAGACAT TTAGCGGCCA AAGTATATGG CAAGTAGTGA GCAAACTTGT GCTGAAATTC
CAAGACAACT ACGACAGCTA CCCTAAGAAA ATTCTGCTTA TGAGGGATGG ACTGGTTCAA
GACGGCGAGT TTGAACAGAC CATAAGAGAG TTGACCCACC AAGGGATCGA CGTGGACATC
CTGAGCGTGC GAAAGAGCGG TAGTGGCAGG ATGGGAAGAG AACTGACAAG CGGCAATACT
GCCATCACCT ATGACGACGC CGAAGTGGGA ACCGTGATAT TCTATTCTGC CACCGACTCA
TTCATACTGC AGACAACCGA GGTAATTAAG ACAAAAACGG GCCCACTCGG TTCCGCGCGA
CCGCTCAGAG TGGTTAGGCA CTACGGGAAC ACCCCGCTTG AACTGCTCGC GCTGCAAACG
TACCACCTGA CCCAATTGCA TCCCGCCAGC GGCTTTCGGA GCTGTAGGCT CCCCTGGGTT
CTGCACTTGG CAGACAGGAG CAGCAAGGAG TTCCAACGGA TCGGTCAAAT TTCATTGCTC
CAGAACGTGG ATAGGGAGAA GCTGATTGCA GTGTAGTGA

59 Gene Name: gene-18034

Gene ID: BBF.2017.48.59

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AGAAGAGAAA GGTTGAGGAT CCCAAGAAAA AGCGGAAGGT CGGCAGTGGC
AGCCTGGGAG CCGGTGCCAG CATCAGTTCC GGCATCCAAA GCGCTAATGA CTGCATTTGG
GACTGGAAGT ACTCTATCTA CCAAACTAAC TCCGGCAGTC AACGAGTGGC CCTCGTGGAC
CCTAAGAAAT CCGACGCCTC CAAGTCTATC ATCCAGAAGT GGCTGGATAA TCAACCGAAA
TTCTCACAGA TCGAAGCCCA TCAGGAGTAC AGCTTCTACG CCCAGGCGGC TTACCCCATT
GAGGCGGACC GAATCAAATA CTTTCAGAAT CTCTTCCAGG GGAAGTCCCC CTATATCGGC
TACAAATTGC TCTGCCTGCT GAACAAGTAC GGTGTAGTGA AATCTGTGTG GAGTACCAAC
TTCGACGGCC TGGTCGAACG GGCAGCACAG CAAGCCAACA TCACCCTGAT CGCCATCAAT
CTTGACTGTG TTGACCGCAT ATATCGAGCA GAAAGCGTGA ATGAACTTCT GTATATCGCG
CTCCACGGGG ACTACAAGTT TAGTACCATA AAGAATACCG CGAATGAGCT CGACAGCCAG
CACACCGAGT TCGTATCTGC CATGTGCCGG TACTTCGTCG ATAAAAACTT GATCGTCATG
GGATACAGCG GACGCGACAA GTCACTTATG GACGCCCTGG TCCAAGCGTT TAGCAAGAAG
GGTGGTGGGA GACTTTATTG GTGCGGCATG GGCGAAACCA TCACGATCGA GGTGCAAAAC
CTGATACAGA GAGTGAGGAC CGCAGGCCGG TCAGCTTATT ATGTAGATAC CTCTGGGTTT
GACAACACCA TGCTGTCACT GGTAAAGTAC TGTTTTTCAG AGGACGTCGC CAAACAGCGA
GAAATAAACG AAATTTTGAA AATTGTGGAA CCGGAGCAGA TTACTCCGTT TGAGATTCAA
AAGAGCCAGA ACAAACGGTA TCTCAAGAGC AATCTGCTGC CAATCGTGCT TCCCAAGGAA

CTCTTTCAGT TTCAGATCTC TTATAACGAC ACGGCGGACA GGTGGGGATT CTTGCGCGAG
AGGATTAAGG AGCGGGAAAT CATAGCAGTC CCGTACCAGG ACAAAGTATA CGCAATCAGC
ACAGTCTCCA TCATTAACGA CGTTTTCAAG GACTGTCTCG TAAGCGAGAT TGAGCGCACG
TCCATCTCTC TGAATGAGAT CGAGCGCAAT GGCTGCTTCA AAGAGCTGTT CCTCAAGGCT
ATTCTCTACG GGTTTAGCCA AATCCGGAAT CTGGGCATCA ACTACCGCCA CGGCATCATT
TGGAAGAAGG AGGCGCTCTA CACTGAGCCC GGCAAGACCG TACACGAGGC CATAGAATGC
GGCTTGTCTT TTATACCGCA AGCGAACTAC GCTTTGATTA GCATCACACC AAGTTTGCAC
ATCGAATCCA GCAGCCCGAT CGAAAAAGAG AAGAAACAAG AGTATAACAG GCGGTACCTT
GACAAGATGA GGAATAAAGA GTACGAGGAA AAGATCCAGG AGTGGTGCAA CATACTGTTC
TCCGGTAACA AGCTCGTTTT TGACATCCCG CTGCAAAGCA ACAACGACTT GAAGTTCTTC
ATTTCCAGTA ATAGGGGTTT CGCCGAGGTA TACAATTACG GTAAGGACAT CGAAAAGAGC
TACACGCCCA ATGCTTACAA TACGAAACAG ACCATTTACT ACGGCATGCA AATCGAGGAG
CCTCAGTTGG AGTTTATCAA CTCCATAATC AGTAGGCCGT TCTATGACGT TAACCCAATG
AGGGGCCTCT CAAATCACAA ACCATTCGAC GCGGACTACT ATGACAAGTT CCCCCAGGAT
GTGTGTTTGG GCATTGTGTG TCCGACCAGC TACAGCCTGA TGTTCTCAGA ATTCCTGAAG
CGCCTGAACA CTAAGATCCC AGCACCGAAG TCATCCGACT ACATCCACAA CTATATTGGC
TTTAACAGCA TCTACAACTG CAGGCTGGAC ATACCGGACA TCAATGCCGA TCGCTGGGTG
AGCATCGGCG ACAACCCCCA GAACGCGGAG GAATTGGCCC GCAACATCTG TATGGAAGCA
AAAAAGCTGA GTGAACAATA TCCGGGCATC GTGGTTAACA TATTCATCCC TACTATCTGG
AGCAACTACA GAAACTTTAA ACACAACGGT GAATTCTTCG ATCTGCATAA CTACATTAAA
GCATTTGCGG CACAAAATCG CTTCACCACG CAACTCATCG AGGAGAAAAC TGTTTGTAAC
ACGATGATGT GCGAGATATC CTGGTGGCTT TCCCTTGCCC TTTTCGTTAA GACCCTGAGG
ACTCCGTGGA CACTGGCTGA CCTTAACCCC AACACCGCCT ACGCGGGGAT AGGGTATTCA
GTTAAAAAGC AGGCCAAGGG CAGGACAGAG ATCGTACTGG GGTGTAGCCA CATTTACAAT
GCGCAGGGAC AGGGACTCAA GTACAAACTG AGCAAGGTCG AGCACCCACA GTTCGACAAA
AAACGGAACC CATTCTTGAG CTTCGAGGAA GCCTTCAAAT TCGGGATGGA TATTCTTAAT
TTGTTCCAGA GTGCAATGGA AAAACTGCCG CAGAGGGTGG TTATTCATAA ACGGACGCCT
TTTAGGGAAG AGGAAATAGA AGGGATTACC AGCGCCCTCA AGCGGGCAGG GATCACGGAG
GTGGACCTGA TCACTATAAC GCAGGAGCGA AACATTAAGT TTATAGCACA GGTTGTCTCC
TTCGGCCAAC TCAATACCGA CGGCTATCCC GTCAACAGAG GCACTTGCAT CAAGCTTAGC
TCTCGCAATG CACTCCTTTG GACCCACGGC GTCGTCCAGA GCATTCGAGA CAAAAGACGG
TACTACCAAG GGGGGAGGTG CATTCCGAGC CCGCTGAAAA TCACTAAGTA TTACGGCAAC
GGCGATCTCC AGACTATAGC TAAGGAGATC ATCGGTTTCA CGAAGATGAA TTGGAATAGC
TTCAACTTCT ATACGAAGCT GCCAGCGACC ATTGACACTA GCAACACCCT GGCCCAAGTG
GGCAACCTTC TCAGGAACTA TAATGGCACC ACCTACGATT ATCGCTACTT TATCTAGTGA

60 Gene Name: gene-18626

Gene ID: BBF.2017.48.60

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAGAGGAA GGTGGAGGAC CCAAAAAAGA AACGAAAGGT GGGGTCTGGC
TCTATGCCAC ACACCTCCCT GCTGTTGAAC TTTCTGCCAG TCTCTCTTAG CGGCGACACA
CGCATCCATG TCGGCTACCG GCCATATAAC GAGGATGTGC TGCGGGAACT GAGGGAGGAG
TTCGGCGAAA GCCACGTGTT TAAAAGGGAC TACCAGGAGG ACACGATAAG CGAGATACCG
GTCATCCCCG GAGCCGAGCC CCTTAGCGAC AAATCTACTG GCGTGGATCT TGCCGAAGCA
CGATGGCTGT GGAAACCACT TCTGAACGCT GCATTGCTTC GCCTCTTCAG CGGAAGCAGA
GAGATCACCT CTGATTATCC AGTCAGCGTG CTTGGTAACC CCAAGAACAA CTTCATCAGC
CATGCCAATC TCCCCGACTG GGTGAGAATC CTGCCCCTTC TGGAATTCGA GAGCCGAACC
CTGTTCGGTG GTAAATCCGG TCCGCAGTTT GGGCTTGTTT GCAACGCCCG AACTAGGCAC
CAGGTCCTGG CAGGCTGCGA CCATCTCATT GAAAGAGGTA TAAGTCCCAT TGGCCGCTAT
GTTCAGATCG ACCAGCCACA AAGAGACTCC AGACTTGCGC CACGCGGTCT GACTGTTGGT
AAGGTGAGCT CTATCGATGG GGACACGTTG ATCCTGGAGG ATCACCGAAA GGGCTACGAG
CGCGTGAAGG CAAGCGACGC TCGCCTTACC GGCAATCGGG CGGACTTCGA CTGGTGCGTG
AACGCGCTGT TGCCTGGACA AGGTCAAGCA ACGCTGAGCA GGGCGTGGGA CGCCATGAGC
GCCCTGAATC AGGGACCCGG CCGCTTGCAA ATGATCAATC AGACAGCTGA ATATCTGAGG
ACCGTGAACC TTGAGGCGGT TCCTGGGGTA GCATTTGAGA TCGGCGAGTG GCTGAGTTCT
ACCGATGCTC AGTTTCCTGT GACCGAAACC ATCGACCGCC CTACCCTCGT GTTTCATCCC

TCCGGCCGAC CCAACGACAC TTGGAACGAG AGGGGGATAA AGGACAATGG CCCGCACGAC
CAGAGGACAT TCACCCCCAA ACAGTTGAAT ATCGCCGTGA TTTGCCAGGG CAGATTTGAG
GGACAGGTAG ACAGATTCGT GGGCAAGCTG CTCGATGGCA TCCCGGACTT TCAGTTGAGG
AACGGCAGGA AGCCCTACGA CGACGGTTTC CTTAGCCGGT TTAGGCTGGA GAGGGCCAAC
GTGCAAACCT TTCAGGCTAA CAGTGCGTCC CGCGAGGCTT ACGAAGCAGC GTGTGAGGAC
GCTCTGAAAC ATGCCGCTGA TAACGGCTTT GGCTGGGATC TGGCTATCGT TCAAATCGAG
GAGGATTTCA AGGCGCTGCC TGGGCCCCAA AATCCCTACT ACGCCACCAA GGCAATGCTC
CTCCGGAACA ACGTAGCCGT GCAGAACATC AGGATCGAAA CAATGAGTGA GCCTGACAAA
AGCTTGGTCT ACACTATGAA CCAGGTTTCT CTTGCTTGCT ACGCAAAGCT GGGTGGTAGA
CCTTGGCTCC TCGGTGCCCA ACAGAGTGTC GCGCATGAGT TGGTGATTGG ACTGGGCAGT
CACACCGAGC AACAAAGTAG GTTTGATCAG TCCGTGCGAT ACGTAGGCAT CACCACCGTA
TTTTCCTCCG ATGGAGGCTA CCATCTGAGC GAGCGAACCG GAGTAGTGCC CTTTGAAGAT
TACGCCAAGG AGCTGACAGA CACCCTCACT AGGACCATAG AGAGGGTGCG AAGGGAGGAC
AATTGGAAGA ACACTGATAG AGTTCGCCTG GTGTTCCATG CTTTTAAGCA GATTAAGGAC
ATCGAGGCCG AGGCCATCAA ACAGGCAGTG GAATCTCTTG ATCTGGAGAA CGTTGTGTTC
GCATTCGTCC ATGTGGCCGA GCACCACCCT TATTTGATCT TCGACCAAAA CCAAGAGGGA
TTGCCCCACT GGGAAAAGAA CAGGAGCAAG CGCAAAGGCG TCTTGGGACC CAGCAGAGGC
GTGCATATAA AGTTGGCGGA CAGCGAATCC CTTGTGGTAT TTGCTGGTGC TAGCGAGTTG
AAGCAGGCGG CACACGGTAT GCCTCGGGCC TGTCTGCTGA AGCTGCACAG AAACAGCACC
TTCAGGGATA TGACCTATCT GGCGAGACAA GCCTTCGATT TCACCGCCCA CAGCTGGAGG
GTGATGACCC CTGAACCATT TCCGATCACA ATAAAGTACA GCGACTTGAT AGCAGAGCGA
TTGGCGGGCC TCAAACAAAT AGAAACCTGG GACGACGATG CCGTGAGGTT TAGAAATATT
GGCAAAGCCC CCTGGTTTCT GTAGTGA

61 Gene Name: gene-7150

Gene ID: BBF.2017.48.61

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AGAAGCGAAA GGTCGAGGAC CCGAAAAAGA AAAGGAAAGT GGGGAGCGGC
AGCATGCAGC AGGAGATCCA GCTTAACATC ATCCCCTTCA CCGCCCCTGT GGAAGAGGCA
GAGTTCGCTT TTTACACCGC CAAGCAAGAC GGCTACTGCC CCATCCATAA GGATGACCTG
AACGGGGCCA TCGAAGGCCT CGTGGATGAA TCAGATCTGC ACTACGGCAA CTGGCTGTAC
ACTGACTTCG CTCCCGCCAA AGAGAACGCC ATCATAATTA GCGTCAATCT CAATGACTGT
AAGTACTTCG CCCAGCACTA CTACAGGCAC CTTATCAGGA CCCACTTCAA GGGAGTGGCC
GACATCATGA GGAAGAATTT CACCAACGAA ATCGAGGTCT GGTTCCACAA TACCAAAGCC
AGCTCTACCA AGTTTAAGGT CTATAACCAG TTTACCCTCA AGGTACAGCA CAACAGGGTG
ACGGACGGAC CGGAACTTGT CGTGTCCTTC GACGGGACGA CGAAGGTGCT GAACAAGTCT
ATCGCCGAGA TACACAACTT CAAAACGGAG CTTTACAACT GGATAAACTG CAACGGCGAG
CTTAATCGCT GGAAATACCT GACCGACGAT CAGAAGCTGA ATCACGAAAA GAACTACCCG
GTAGTGTCAA ACACACTTAA ACCGCATTTC GACATTGCCT TTGACGTTCC CGATTTTAAG
AACCGGTATC CCAAATACTT CACTCTTCTG AATGACTTCT ACAACAACTA TCTGAATACA GACGC-
CTTTA CTGCGATCTT GCCGCTTTCC GCTGACGGAT TCTTCAAGCC AAATGGCCTG TCAGT-
GCAGA GGATCAACGG CACTAGCAAT GAGCTGCAAT TCGGCAATGG CGTCGGCGTG GAGCC-
CAAAA GGGATCTCAA GCGCCTGAAG CCGTATAAAC CCGTGCCCAA ACCCAGCAAC GTAAAGTTTT
TCTTCATCTA TCACAAGCCA GATAGGGAGC ATGCGGTCAA AAACATCTGG CAGTATTTCA

AAGACGGATA CAACGGCCAA TACCCCTTCC CCAAGATGGA GGAATACATA TCTCAGCCCT
TCGAGCTTGA GGAGAATGGA TCTATCTCAT TCGACAATAT CGACGACGCG GTAAGCGTTG
TCCAAAAAGC CATCAAGAAC AAGGATCGGC TGCCCGACAC TAAATACTTT GCGGTATACA
TCTCCCCCGT ACCAAAATGG GAGAAGGACC CTAAACGGAA TAGTATCTAC CATCGGATGA
AAGAGATACT CCTGTACGAG GGGATCACCA GCCAGGTGAT CTGGAAGGAG AACATTAGCA
AACCGGCTTT CAACCTCTTC TTGCCTAACA TCGAAACCGC CATACTGGCC AAGCTGGGAG
GCGTCCCCTG GAGGCTCAAG AGGGACACCA CGAACGAGTT GATCGTTGGC GTGGGTGCTT
TCTACTCAAT CACGCGGAAG TCCAAGTACG TGGGCTCTGC ATTTTGCTTC AATAACGAGG
GCATCTTTAA GGGGTTCGAC TGTTTCGGTG CCAATGACAC CGACTCAATC GCGGGCTCTA
TCAGGGAGGC CGTGGGAAAG TTTATCGCGT CTAATTACAA GGCCACAAGG CTGATCATTC
ACTTCTATAA GGACCTGTCA AAGAAGGAGC TCAAACCAAT CATCGATACA CTTCACGCCC
TGGGCTTGCC CATCCCAGTG ATAGTCGTGA CCATCAATAA AACCGAGAGC AAGGAACTCC
TGGCATTTGA TACCAGCTCA CAAAAGCTCA TGCCCTACTC TGGCACCATC GTGAAGGTGG
GAGCCAAGGA GTATCTGCTG TTCAACAACA CGCGATACGA GGAAGCATCC GCCCCAACGG
ATCGCGAGCA CCACTTCCCG GTGAAAATCA GCTTTTTCTC AGACAAGGCG GAGCTGTTGG
ACGATCCCGC ACTGATCAAC CAACTGATCG ACCAGGTGTA CCAGTTCAGC CGCATGTATT
GGAAAAGCGT GAGCCAACAG AACTTGCCCG TAACCATTAA GTATCCCGAG ATGGTGGCGG
AGATTTTCCC ATACTTTACC CACGATAAAT TGCCCGATCA TGGAAAGGAG AGCCTGTGGT
TCCTGTAGTG A

62 Gene Name: gene-7179

Gene ID: BBF.2017.48.62

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGCGGAA GGTTGAGGAC CCAAAAAAGA AGAGGAAAGT TGGCAGCGGG
AGCATGGAAA ATCTGACCCT GAATATCATC CCTTTCAGCC ACCCCGTGCA GGAGCTTGAG
ATCGGCTTCT ATAAGCAAGA GAAACAGGGA TGCTACAGCC TGTGGAAGGG CGAGTACCCG
CAGTCATTCT GGGACGACTT CAACGAGGAA ATGCAAAATT GCGACAAACT CTACACCAAC
TTCATTGACA CGGAAAACTG TGATTACAAA GCCAGTGTGG ACTTTAGCAA AAACAGACGC
CTGGCGGTCC ATTACTACAG CAGGCTGATC TACAACTACT TTGAAACAGT GGCAGATGCC
GTGAAAATCA ACTTCGTGAA AGATATCCAG ATATGGTTCA AGGACGAAAC CAAGAGCACC
GCCGTCTATA CCAGTTACAA GCGGTTCACG ATCAAGGTCC AGTTCCATAA GGTGACCGAG
TCCCCAGAGC TGTTGATCAG CTTCGATGGC AATACCACGG CCTATAACAA AAGTCTGGCC
GAGTTGGACG ATTTCCCTCC CGAGCTGATT AACTACGTTA AGTACAATAC CCAAGTGGTG
AAGTACGAGT TCGCCGAGGA CGCTATTAAG CAGCATATCG AGGAGCTGTA CCCGATCCTG
AGCAACCCCA TCAGGGACTA CCTTAAGATT GCCAGGCCCG ATTTTAAGAG GGGCAACAAG
TATAAGCCCT ACTACAAGAA CATTACAGAC TTCTATCACA ACCACCTGAA CTCCAAAGAG
TTTAAAGCTA TCCTGCCTAT CTCCGAGGAC GGTTTCTACA AAATGCCTAA GCACAAGGTT
CACAAAACCA GCTTCAATAG CAATAAACTG AGATTTTTCA ATAACACGGA CATCGTGCCC
CACAACGGGA TGAAAAACAT CGGCCCCTAT AAGGCGTCCC CCCACCCCAA CGTGAGGTTC
TTCTTCATCT ACCATAAGCC AGACCGAAAC TTCGCCGTCA AGACGCTGTA CGAATACTTT

ACGGAAGGGT ACAAGAGCCC AGAGGGCTAC CTTTACTTCA AGCCTCTCAA AACCTACATT
AAACAGCCCT TTCTCATCGA CAAGGATACC TCTATCGCGT TCGAAAGCCC GGAAAGCGCT
CTGCGCGAAG TCAAGCAGGG TTTGCTTAAC CTGGAAAAGC AGCCCAATAC GAAATACGTC
GCTATCTATG TGACCCCCAT ACATAAGACC GAAACCGACG AGCAGAGGAA GATGCTTTAT
TACCAGGTCA AGGAAGAATT GCTCAAGCAC GACATATCAA GCCAGGTGAT ATACAAGGAC
AACATTGGAC ATAAGGATTT TAGTTTCTAT CTGCCCAATA TCGCAATCGC CCTGCTGGCC
AAGATCGATG GAATCCCCTG GAGGCTGGAC AGAGACACTA AGGAGGAACT TATCGTGGGC
GTAGGCGCAT TCACAAGCCT GAACCACAAT ATCAAATATG TAGCTAGCGC CTTCTGCTTT
AACAACAATG GGGAATTCAA GGGATTCGAC TGCTTCAAAG CGAATGAAAC CGAACTTTTG
GCTGGCACCA TCGGCAAGCA AATCCTGAAG TATGTGGTGG ACAACGGCGA GAGCGCCAAG
CGCCTGATAA TCCACTTTTA CAAAAAGATC AGTAACAAGG AACTCGAGCC CATAAAGAAA
ATGCTGAACA AGCTGAACCT GACCATCCCC GTAGTGATAG TGACTATCAA CAAGACGACC
TCAGAAGATA ACGTGGCGTT TGACACCAGC AGCCATAACC TGATGCCCGT GAGCGGCACC
TACCTCAAAA TAGGATGGGA CCAGTACCTC CTTTTCAACA ACACGAGATA CAACGCCAGC
GACACCGAGA AGGATAACCC CTTCCCTGTA AAGCTGAGCT TCTCTAGCAC CGTAGACAAT
TACTTCGACG ACAGGAAGGT GGTCGAGGAA TTGATCGACC AGGTGTATCA GTTCTCCCGC
ATGTATTGGA AAAGCGTGAA GCAACAGAAT CTGCCCGTTA CCATCAAGTA CCCCGAGATG
GCGGCAGAGA TCTTCCCATT TTTTGAAGGC GATAAGCTGC CCGACTTCGG AAAGAATAAC
CTTTGGTTTC TGTAGTGA

63 Gene Name: gene-7338

Gene ID: BBF.2017.48.63

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAGAGAAA GGTGGAGGAC CCGAAGAAAA AACGAAAGGT TGGCAGCGGC
AGCGTGCAGC AGACAGTGGA GCTCACCCTC TACACAGAAA AACATCCCGA CACCCACCCA
GAGCTCGTTT ATGCCGACGA GTGTCCCGAC CTGTGGCAAC AGCACAGCGA GCTTACGGGG
GACAAATCTC TGTTCTACTC TCTTACGAAC CCGGCAGAAT GCAAGGGAAC CCAGTACACA
GTGCAAATCA ACCTGAATAA CCAGAAGCAG CGAAGGATCG CCAAGCACAT AATTAGCCAG
CAACTGTATA ATCACTTCCG CCAGACCCAA ATCGCTACCT TCGACAAGAT CGACAATGTG
GAGGTGTGGA CCAAGAACAC CCAACAGCCT ACCCAGAATT GCACGGAGTA CCTGAGGTTC
AGCCTTATAC CCCAATACGC CGTGTTCTCT GACTCATGGG AGCTGGTCGT GTCCTCAAAT
GGCATATCCA CCGTGTATAA CAAGCCTTTG AGCGCACTGG ACCTTCAGAC CGACCGATTC
AAGGTCGTCG TTGGAGGGGA AGTGGTCAAG TACAAGAACC TGAGCCCCAA TCAAAAGCAA
CAAATAGACG AGGCCTTCCC CAAAATCAAT AGGGAACTGG CCGCTGAACT GCATATTAAC
GAGAAACGCT TTCTCAATAA AGACAAGTAT ACGACCACCT ACAACCACAT TAACAACTTC
GTGCGACAGC ACCTTCTCAC ATCCGAGTTC CAGGCACTGT TTTGTCTGAG CGGCGAGATG
TTCAACGTAC CCGAGGAGCG GATCGGCCAA GTGGCGAAGG GGGCGAACCT GTTGCAGTTT
AAGGACGGCA AGACCGGCAT TGACCCATTC AGCTGTGTGT TCGGCAGCAA GAGCATGGAC
GCACTCGGCA TCTACCAACC CAGCCTGAAG CCCCAGGTGA AATTCTTTTT TATCGCCCAG
CAAAGCGATA TCAACGTGTG CAAAAGCCTG TACGATATTT TCACGAAGGG ATACAAGCCC

TACGTGGACA CAGCCACTGG CGAGCAGAGG TACGTGTTCC CACCCCTGGC GACGTGCATC
AAGCAGCCCT TTTCAACCGA CCCCAAGGGG AGCATTTACT TCAGCGACCC TCAAAATGCC
CTGAGCGAGA TCAAGAGCCA GCTTAACAAT AAGCCTCTTG ACCCCCAAAC GCAGTATGTG
AGCATATACG TGTCACCCAT CCCTCGCGAC GCCGTCAACA ATCCCTACTA CGGTCTGTAC
TTTCAGATTA AGGAGCTGCT GCTCGAAAAG AGGATAACCT CTCAGGTGAT CTATAAGGAC
CGCCCCAACA ACCAGTACTT CAACTTCCAT CTGCCCAATA TCGCGACTGC CATCCTGGCA
AAAATAGGCG GCATCCCGTG GCAGTTGAAC TCCCACACGA CGAACAAAGA TCTGGTGATA
GGCGTGGGCG CCTTCCTTAG CGAAAAAGTT GGCGAGAGGT ATGTGGGCAG CGCGTTCAGC
TTTAACCCCA ACGGCCTGTT TAAGAACTTC GACTGCTGTA AAGCGAACGA TCTCGAATCT
ATCGTAGCCG GGATCAGAAA GGCCATCGGA CACTTCGTTG TGGACAGCGA AACAAACCCC
CAGAGGCTGA TCATCCACTA CTACAAGACC ATGTCAAAGA GGGAGGCCAG GCCCATCACG
CAGATGCTGA ACACGCTTGG CCTCAACATT CCTGTATTGA TCGTCACAAT AAACAAGACG
GAAACCAGCG ACATTGTTAT GTTTGATGAG AAACAGCAGG GCTACATGCC CCTTTCAGGC
ACCGTACTGA AGATAAGGAA CGATGATTTC CTGCTCTACA ACAATTCCAG GTACAAAGAG
AACGAAAAGT CAGATATGCT TTTTCCAGTG AGGATCCGCC TGAGTAAGAT CGTAAACCAA
TCCGACAAAG ACATCCCAAT GACAGACGCC TTCAATTTGC TCAACCAAGT GTACCAGTTC
TCACGCATGT ATTGGAAAAG CGTTAAGCAG CAAAATCTGC CGATCACGAT AAAGTATCCA
GAGATGGTGG CCGAGATAGT GCCACACTTT TCAGAAGCCG AATTGCCGCA GTTCGGAAAG
AATAATCTGT GGTTTCTGTA GTGA

64 Gene Name: gene-11534

Gene ID: BBF.2017.48.64

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AAAAACGGAA GGTCGAGGAT CCCAAGAAAA AGAGAAAAGT CGGTAGCGGC
AGCATGAACT ACACAGCCGC CAACACGGCC AACAGCCCAT TGTTTCTCAG CGAGATTAGC
AGCCTTACCT TGAAAAACAG CTGCCTCAAC TGCTTCAAAC TGAATTACCA GCTGACTCGC
GAAATAGGCA ATAGGTTCGG CTGGCAGTTC AGTAGGAAGT TCCCTAACGT TGTGGTGGTG
TTCGAGGACA ACTGTTTCTG GGTTCTCGCT AAAGATGAAA AGAGCTTGCC CTCTCCTCAA
CAGTGGAAGG AGGCTCTGAG CGACATCCAG GAAGTGCTGC GAGAGGATAT CGGTGACCAC
TACTACAGCA TCCACTGGCT TAAAGACTTC CAGATCACCG CCTTGGTGAC CGCCCAGCTC
GCCGTGCGAA TTCTGAAAAT CTTCGGTAAA TTCAGCTACC CCATCGTGTT CCCCAAGGAC
AGTGAAATTA GTGAGAATCA AGTGCAAGTA AGGCGAGAAG TCAACTTCTG GGCCGAGATC
ATTAACGATA CCGACCCCGC CATTTGCCTC ACCATCGAAA GCAGCATCGT CTATTCCGGC
GATCTCGAGC AGTTCTACGA AAATCACCCG TACAGGCAAG ACGCCGTGAA GCTGCTGGTG
GGCCTGAAAG TTAAGACCAT TGAAACCAAC GGCACCGCTA AGATCATCAA AATCGCTGGC
ACTATAGGGG AAAAGCGCGA ATACCTGTTG ACTAAGGCCA CGGGAAGCAT ATCCCGGCGA
AAGTTGGAGG AAGCCCACCT CGCACAACCC GTGGTTGCGG TGCAGTTTGG TAAAAACCCT
CAGGAGTACA TATACCCCCT GGCTGCCCTC AAACCTTGCA TGACCGACAA GGATGAGAGC
CTGTTCCAGG TCAATTACGG CGACCTCCTG AAGAAAACCA AGATCTTCTA CGCTGAACGA
CAGAAATTGC TTAAACTGTA CAAGCAGGAG GCGCAAAAGA CTTTGAATAA CTTCGGTTTT

CAGCTTCGGG AAAGGTCCAT CAATAGCAGG GAAAATCCAG ACTTCTTCTG GACGCCCCCA
ATTTCATTGG AGCAGACCCC CATCCTGTTT GGGAAGGGTG AGCGCGGTGA AAAGAGGGAA
ACCCTCAAGG GCCTTTCAAA GGGCGGAGTC TACAAAAGGC ACAGGGAGTA CGTTGATCCT
GCCAGGAAAA TTAGGCTGGC CATCCTTAAA CCGGACTCTT TTAAAGTGGG CGACTTCAGG
GAGCAGCTGG AGAAGCGACT CAAGCTGTAT AAGTTCGAAA CGATTCTCCC CCCTGAGAAC
CAAATCAATT TTTCTGTGGA GGGTGTTGGG AGCGAAAAAA GGGCCCGACT GGAAGAAGCC
GTAGACCAGT TGATAGGTGG CGAGATCCCC GTGGATATCG CCCTCGTCTT TCTGCCCCAG
GAGGACCGGA ACGCGGACAA CACCGAGGAA GGCTCCTTGT ATAGCTGGAT CAAAAAGAAA
TTCTTGGATC GGGGGGTGAT AACACAGATG ATATATGAGA AAACTCTCAA CAATAAGAGC
AACTACAATA ACATCCTGCA CCAGGTGGTT CCCGGCATAT TGGCAAAGCT CGGAAATCTG
CCGTATGTGC TGGCCGAGCC TCTTGAAATC GCCGACTACT TCATCGGCCT GGACGTCGGA
AGGATGCCTA AGAAGAATCT CCCTGGTTCA CTGAACGTGT GCGCGTCCGT TAGGCTCTAC
GGAAAGCAAG GTGAATTCGT CCGATGTAGA GTCGAAGATA GCTTGACCGA GGGGGAGGAA
ATCCCCCAAA GGATTCTTGA GAATTGTCTG CCGCAGGCAG AACTTAAGAA CCAGACCGTC
CTGATCTACA GGGACGGGAA ATTCCAGGGT AAGGAGGTGG AAAACCTTTT GGCTCGGGCA
CGAGCCATCA ACGCCAAGTT CATCCTGGTA GAGTGCTACA AGACCGGCAG CCCGAGACTT
TACAATTTCG AACAAAAGCA GATTAATAGC CCCAGCAAGG GGCTGGCGCT TGCATTGAGC
AACCGGGAGG TCATCCTCAT CACCAGCCAC GTTAGCGAAC AGATCGGCGT GCCTCGGCCT
CTCCGCCTGA AGGTGCACGA ACTGGGAGAA CAGGTGAACC TCAAGCAACT TGTGGACACG
ACCCTGAAAC TGACTCTGCT GCATTATGGC TCTCTGAAGG AACCTCGGCT TCCAATCCCC
TTGTACGGAG CCGACGCGAT CGCGTATAGG AGGTTGCAAG GAATCTATCC AAGCCTGCTG
GAGGACGACT GTCAGTTCTG GTTGTAGTGA

65 Gene Name: gene-13519

Gene ID: BBF.2017.48.65

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGAGGAA AGTTGAGGAT CCCAAGAAAA AACGAAAAGT GGGTAGCGGT
AGCGTTCCAG GCGGTAGGGG ACCGCTGCTC GTGCTTAACT TCCTTCCCGC TCGCTTCGAC
GGCCGAGTTG ATGCGGGCAC CCTCCCCTTC GAAACCCCTG ATAAATTGAG GGCCATTAGG
GAGGAACTGA GAACTTCCCA TGTAGTTGTA ACGCGAGGAA AAGAGGTCGT ATGCGTGCCC
TTCGTTAGTG GCGCGAAATT GATCGGCAAA CGAACCACTA TCACCGCAGC GGGACCCGAC
CTCGTCGTAC AAACGTCCCT TCTCGAATCC AGCCTGAGGC GGACCTTGAC CGAAAAATGG
AAGTACGAAT TGCGCAGGGA AAACCCGCTC ACCTTTGTGT CAAGGACGCC AGGAAGGGAT
CTGCTGGAGA AGGCCCTTGG TCGGGAGTTG CCGGGACTCC ATGTGTTCCC CGCTTACAGC
CTGGACGTGC GCAGATACGG TCCTGGGGGG TTCAGCGGGG TTGTTGTAGG ACTTAAGACC
CGCTATGAGA TCGATCTGCC TGTCGGAGTG CTGCTCAGGA GGGGCGTTCA AGTAAACGGC
CTTTATGTCC TGGCTGAAAG CCCCCTCGCG CCTACGTGGC CCTTCCAAGA TCCCCACACC
AGAAGGCGGC TCGTGGGACA AGTTGTCGCG GTGGATGGCG ACAAATTGCG AGTGAGGTGT
AGGGACGGGG AGCTGGAACT TGATGCCGCC GAAGCATGGA TTGAGCCCAA CACTGCCAAC
TTCTACGCCG TCCTGCGGAA GGCGTGCGGA CGCTCTTACG AACGAGACTT TCACGCCCTG
GAAGCCCAAG TCGTGTCCCT GACTAACGCC CAGCAGCGAA TCGCCGATAC CAACAGGATC
GCCGCCAACC TGATAGGCCT TGGTAAATTC GACATCAGTA ACGGCTTGAC TGCCGAGCTG
GGGAAACCAC TCAGACTGAC TTCCACTCAA CATCCACACG TTCGGACTCT GGCCGAGCCC

ACATTTGTGT TTGACCAGAG CGGAGACAAA ACCGCGCCTT TTCCCGAAAC CGGGCTGACC
AAGTGGGGCC CATTGGACGC TGAGAGCTTT ACACCCAAGG CACCACATAT CGCCGTGGTG
GTTCCGCGGC AGTTTCAGGG TCGCGTCGAA ACGCTGGTTG AGCGGTTCAG GAACGGCGTG
AGGGGCAGCA ACGCCTATGC CGAGGGCTTT GTCCGAAAGT TTAGGCTCAC CGACTGTACC
TTCAGCTTCA CCGTTTTTGA CGGTGACGCT ACTGACGCAG CCGCATATAG GCAAGCGTGC
CTTACCGCCC TGAGTAATGA CGAGCAAATT AACCTCGCCT TCGTGTTCAC ATCAGCCGTG
CAGGAGCATC AAACGGGGGA CGACAGTCCC TATCTTGTCA GCAAATCCAC CTTCATGAGC
CAGGGTATCC CCGTGCAAGA GTATCAAGTG GAGAACATCA TCGGGGATTC AAACTTGGCT
TATCCCCTGT CCACGATGGC GCTGGCGTGC TACGCCAAAC TGGGTGGCAC CCCTTACGCC
ATAAGCGATC GAGGACGACC TATGGCACGA GAACTGATCT TCGGCATCGG GTCTGCCCAG
GTAAGCGACG GAAGGATGGG CGAAACAGAG CGATTTGTGG GCATTACCAC CGTGTTCAAT
TACGACGGTA GGTACTTGGT TAGCAACGTT AGCCGCGAGA CACCCTACGA AAGGTACCCG
CAAGCCCTGC TTGACGCATT GCGGACTTGC ATTGCCGACG TGAAGGTTAG GCAGGGATGG
AGGTCCGACG ACTTTGTGCG GCTTGTGTTC CATATCTTCA AACCTCTGAA GGACAAGGAA
GCACGCGCCG TAAAAGAGCT GGTGACGGAG CTGACGTCTG AATATGCCAG CGTGGAGTTC
GCTTTTGTGA CAGTGGTGGA CGATCACCCG TGGCTGGTGC TCGATGAAAA CTCTGATGGG
GTTAAGGTTG GGCGAGGGAC TAAGGGCAAG CACGTAGCTC GGAGGGGTTT TGCCCTGCCG
ATTTCCAAAA GGGAGCTTCT TGTGACGGTT AAAGGTCCCC GGGAAATGAA ATCCGATAAG
CAAGGGGCTC CCAAGCCCCT CTTGCTCAAG CTCCACCGCG AAAGCACCTT TACAGACATC
GACTACCTGG CTTCCCAGGT CTTTCAATTC ACCGCCATGA GCTGGCGCAG GCCATACCCT
ACCAGCAAAC CCGTGACTAT AAGCTACAGT GACCTGATTG CGGGACTTCT CGGAAAGCTG
CGACACGTGA CGAACTGGAA TAGCGACATG ATCTACATGA AGTTGCGCTT CAGCAGATGG
TTCCTGTAGT GA

66 Gene Name: gene-14983

Gene ID: BBF.2017.48.66

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAGCGCAA AGTCGAGGAC CCCAAGAAAA AGCGAAAGGT GGGCTCTGGC
AGTATGATTA ACAAACTGCA ATTCGACGAG TTTCAGAGGG CCATAGGTAT TTCTAAGAAC
GACACCTTCA GTCTTTTGCT CGGAGCGGGT TGCAGCATCA ATAGTGACAT CCCTAGCGCG
GAGGACTGTA TATGGGAGTG GAAGCGAGAT ATTTACAAAA CAAATAACAG TTCTAGCTTC
GGCTGGATTG ACAATTACAA GAATCCCAAG ACTCAGGAGA TCATTCAGAA CTGGCTCAAC
AACCAAGGCA TCTATCCCGA ACGCGGCTGC AAAGAGGAGT ACAGCTTTTA CGCCTACAAA
TGCTATCCCA TCGACGAACA TAGGCGACAG TATTTTCAGA AAATCTGTAG TGGTAAAAAG
CCATCCATCG GGTACAAACT TATTCCCCTG CTTGCCCGAA AGGGCATGCT TGATAGCGTG
TGGACCACGA ATTTGGACGA CCTCGTGGTG ACCGCCTGTA TAGGCAACGG GATCCAGGCG
ATCGAAATCA CGCTCGACTC CGTGCAAAGG TTGAACAACC GGCCTCAGAA CCGACATGAG
CTTCCTGTGA TCAAACTCCA CGGAGATTTT AAGTATGGCG ATCTTAAAAA CACCGAGGAG
GAACTCCTCA ATCAGGATAA AACGTTCAGG GAGAGACTTA TTGAATACGT ACAAGACAAG
CACCTGATCG TGCTCGGCTA CAGTGGCCGA GACACCAGCC TGATGGACAC ACTTAAAGAG
GCCTACTCAA AACAGGGGGG TGGAATTCTG TACTGGTGTG GATATGGTGA CAACATAAAC
TCCGATATCG CCGAACTGAT TCAAATAGCC ACTAAAAATG GCCGACGAGC CTTTTACATC
CCCACTGATG GTTTCGATTC TACGCTCCGG AAAATCACAC AGATAGTGGT CGAGGATGAT
AACAACCTGA AAAAAGAGCT TCTCGAGCTT CACCAGACCA GCAATATCAA TGACACTATC

ACACCTTTTG ATCTGAAGTG CGAGAGGGTG AATAAGCTGT TGAAGTCAAA CATATTCCGG
ATTAGCTTTC CAGACGAAGT GTTCGTTTTC GATGTGAGCA TCAGCGATAA ACCCTGGAAG
TTCGTGGACG AAAGGACTCT TGAGCGCAAC GATATTAGCG CCGTTCCCTA TAACAAGCAA
ATCTGGGCAT TCGGTAGGCT TGACATCATA AAAGACATCT TCAAAGACGT GATGAACTCA
GACATTCAGC GAAAACCCCT GGCAAACATC AAGATATACA ACACGGCGGT TAGTCGGCTG
TTGCTTACTA CGATTTGCAA GATACTGGCG CTGCAGAGCA ACCTTAAGAC CGACTATAAG
GGTAAGATAT GGACCGAGAA CAACAGTAAG TCCATTTCCG GCCACATAGT ATACAATGCC
GTGCTGCTGT CCTTTGATCG GATAAGCGGT GAGTATTACC TTAGCCTCAA CCCCGACTTC
GTGCTGGCTA ACCCCAACAT TGAGAAGAGT AGCATACAGA CCATAGGACT GTTCTTCTTC
CAGAAGCTGT GGAATCAGCA GTTTAACGAG TACATTAACT ATTGGAGGGA AATTTTGTTG
AAAAAGAATA ATGAGTACGA GTTCCCCATA AATAGCGGAA CCGGCTTCAA GTTCAAGATC
AAGAACATCC CAGTGTTCAC TAACATCTGC GACCTGAATA ACCCTCGCAT CAACAATCAC
AACGTGTCCA GCCACCATCT GCTGCTTCAG GGGGTGCAAT TTAAGGAAAT CCCGCTGCTT
TTCAGCACCA ACAATGGCAA CCGCACGGCC ACCGACACCC ACCCTATGAG AGGACTTCTC
ATAAACAAAC CGTATGAAAC GGGCGTCAAC GACTTCCTCG AAAAGTCTAT CACCCTGGGA
ATCATAAGCC CCAGTCAGGA CGCCCTCAGG TTCTACCAAT TCCTGGAAAA CCAGAACTCT
AAAATCAAAA AGCACAACGA CAAGGACAAC TACATAATAG ACTACGAAGG GTTTTTCGCC
ATCTACGGCG TTAGTCTCAG CTTCCCAACA CCTAACGACA ACGAGTGGGA AAGGATCAAC
GAACCGCTGA TTATGGGCAT CAAGGAAACC GCCCAACAGA TAAAGCAACT GATATGCGAC
AGCATCGTGA AGATCTCAAG CACGACCAGG AGAAAAATCA TCGTCATCTA TATCCCCCAA
CGCTGGGAGC CCTACACCTC TTACCAGCTC GATGGTGAGT CATTTGACCT CCATGACTAC
GTGAAAGCGT TCTGCGCGGA GAAAGGGATT ATGAGCCAAC TCATTCGAGA AAAGACCATT
AACGATACTA TCCAAAAATG CCAGATACAT TGGTGGTTGT CTCTGTCATT TTTCGTAAAA
TCCTTCCGGA CCCCATGGAT TCTCGCAAAT ACTAACAACA CCACCGCCTT CGCGGGTTTG
GGGTACAGTG TAGAAAACAA GAAGGATATT AACGGACATA TTGTGCTGGG GTGTAGCCAC
ATTTACAGCT CAAACGGAGA AGGGCTCAAA TACAAGCTGG CCAAAATAAG TAATGATAAG
ATTCAGTGGA GGCATAAGAA GCCGCACCTC TGCTACGACG ACGCGTATGA GTTTGGCAAG
TCAATTGTGA ACCTGTTCTA CGAATCTATG AACGAACTGC CAAAAAGGGT GGTCATCCAC
AAGAGGACCT TCTATACCGA TGAAGAGAAA CAAGGGATCA TAGACTCCAT TAGCGACAAT
AAGAAAATAG AGAGCATCGA CCTCATCGAG ATCAACTTTG AAAACAATAT AAAGTACGCC
TCTAGCAAAA TCCACGACGG AAAGGTAGAC ATTGACGGAT TTAGCGTATC TAGGGGAACT
TGCATACAAC TCAGCTCTAA GGAGGCGCTC CTGTGGGCGC ATGGAGTGAT TCCTAGCGTC
ATTAACCCTA ACTGGAACTT CTACCCTGGC GGGAGGTACA TACCTAAACC ACTTAGGATC
ATTAAACATT ACGGTACAGG TAGCTTGGAA CAGATCGCGA ACGAGATTCT GGGCCTGACT
AAAATGAATT GGAATAGCCT GAACATGTAC AGCCAATTGC CTGCCACAAT TTCAAGCTCC AAT-

GATATAG CTAGGATAGG TAAATTGATA GGGGCGAACA GTATGCACGA ATACGACTAC CGAT-ACTTCA TCTAGTGA

67 Gene Name: gene-16946

Gene ID: BBF.2017.48.67

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAAAGAAA GGTGGAGGAT CCAAAGAAAA AACGCAAGGT GGGTAGCGGC
AGCATGCCAT CAGCCGAGAG GTGCATCTGG GAGTGGAAGA GGGAAATCTT CATCACTAAA
AACCCCTTGC TCAGGGAAAC CGTCGGCGAG CTGTCCCTCC AGGGCACGAA GGACCGAATC
CAAAAATGGC TCGATCAACG CGGCGAATAC CCCGCACTGA ACTCCCCAGA GGAATACTCA
TTTTATGCCG AGGAGTGCTA CATCACCGAA CAAGACAGGC GGAGCTTTTT TCAGCAGTAC
GTAGAGGTCG CCAAGCCGCA CATAGGTTAT AGATTGTTGC CCCTGCTGGC ACAGACCAAG
ATCATAAAAA CTGTATGGAC GACTAACTTT GACGGGCTTG TCGCCAGGGC CTGTCATTCC
AACGACGTGG TGTGCATCGA AGTCGGACTC GACAATACCC AACGCATTAC GCGCCAGCAT
TCTGAGGGGG AGCTGCGGGT TGTAAGTCTC CACGGCGACT ACCGATACGA TGAGCTTAAG
AATACAGATG AGCAGCTCAG GTACCAGGAG GAGGCGCTTA AAAACAATAT AGAGCACGAG
CTGCAGGACT ACGACCTGGT AGTGATCGGT TACTCCGGCA GGGACCGGAG CCTCATGAAC
GTACTCGAAA ACATATTCAG CAAGGCCGTG AAGTCAAGGT TGTTTTGGTG TGGCTACGGC
GAAACGATAA GCCAGCCCGT TATGGAGTTG TTGGAGCTGG CCCGCAAGAA TAATAGGGAC
GCATTCTATG TCAGCACCGA AGGCTTCGAC GACACCGTTG AAAGAATCAG TAGGAAGCTG
CTTGACGGCA ACATGCTGTC CAAAGCCTTG GCTGAGATAC AGGAAACCAC TTGCATCACC
AACCAATCTG CCAAATTCAC CGCACCTGAA AACGACATCA GCAGCCTTAT TAAGTCAAAC
GCATACCCCC TCCTGAAGCT CCCCTCTCAG TTCCTTAAAG TGACCCTCAA ATACCCGGAG

GGGTCCTTTA GTTACATTGA TTGGCTTAAC TCCAAGGTTG ACTTCAAGGA GGTTGTGTTG
TCTAAGATAG ACAAGGAGAT AATCGCGTTC GCGGATGTTG ATAAGCTGAG GAAGTATCTG
GGCGAGTTCT ACCTGTCTAC GCCCACGGTG GTGAACTTTA GCAAAACGGA CGTGCTTAAC
GATACTCGCA TTCAGAGTCT GGTGAGGCGC GGACTTATAC AGTCCATCGT AAAAAACCTG
AACCTGTCCA GCGACCAGAA CAAGCGAATA TGGAATCCAG ACGTGAGCTC CATCGAATTC
TACAACGGCA AGAAGTACAA AATCATCGAC GCGCTCATCC TCAATCTTAG TTTTATCAAA
GATGACATCT ACCTCACGTT CAAACCCGAT CTGCTGGTCC TTAACCTCGA CGAGAGCCTG
CCAGACAACG ATATAGTTAA GACTATCAAG AACAAAAAGT TCGGCTACCA GCACAACAAA
GAGTACAGTC AGATCCTGGA GAAGTGGGCC AACCTTATAA CGAAGAAGGA TTTGGTCGTG
AGTGGCGGGA GCGTGTTCTT CCTTGGGAAG AAACCGCTGT ATGCCGGACT TGTGTCTTAC
GCCGCGAGGA AACTCCCAAC AGATTATAAC AAGCACGCCA CCCAGAAAGG ACTGATCATT
CAAGACGCGA AACTGATTTT TTGCAGCAAT TCCATCTCCA ATGAGATTTC TCACATCAAC
CCCCTGAAGG GGCTCGTGGA AAATCGCCCG TGGGACTACA AAAACACCAG CTCTGGGCTG
TGCCCCGAGA TCTGCATTAA CGTGATCTCA ACCAGGCAGG ACGCGGGTGT GGTGAGCAAC
CTTCTCCGAG GTATTCACGA GAAGTCCTTC CCGGAAAAAT CCGAGCAAGA TTACTTGCAC
CCCTTCCATG GGTTCACAAA CGCTTTCGGG GTGCCCATCA CGATCCCTAA GATCGGTGAG
AATACGTGGC GCTTTGTGGA CGAAGCACTG AGTGCACAGA AGGCCATCGA TAACGCGAAG
AACCTCGCGA ACCGCATTTG CTATGAACTT GACAGCCTGA AGAAGCTTGA ACTGCGGACG
GGCACCGTCG TGATCATATA CATCCCCAAG AGATGGGAAG CATTGACATC CATCAAGTCT
GAGCATGAGT ACTTCGATCT GCATGATTAC ATCAAGGCCT ATGCTGCGCA ACAGGGCATT
AGTACGCAAT TCGTGCGCGA GAAAACGGTT AATTCAAGCC AAAGCTGCCG GGTAAAATGG
TGGCTCAGCC TGGCGTTCTA CGTGAAGGCT ATGCGCACTC CGTGGCGGTT GGAGAGTATT
GATAACCAAA CGGCTTTCGT GGGGATAGGG TACAGCATCA ATCGCAATAT GCATCCCGAG
AATTCCAAGC GGATAATTCT TGGATGCTCC CACATATACT CCGCCCGAGG CGAAGGCATG
CAGTTTCAAC TTGGGCGAAT TGAAAATCCC ATTATCCACC ATCACAATCC CTACATGAGC
GAGGAGGACG CTAGACGCAC CGGCGAGAAG ATACGACAAA TGTTTTTTGA TGCCAAGATG
CAACTGCCAC GCAGGGTCGT CATCCACAAG AGGACCGCTT TCACTGAAGA GGAACAGCGG
GGGTTCATAC AAGGATTGGA AGGCGTTGAG GACATCGAGC TGATCGAAAT TAACTTCGAG
GACTCCCTCC GCTATTTGTC TAGTAAGTTT GTAAACAGCA AGCTGGAAAT CGACGGGTTC
CCAATCGCTC GGGGGACCGT AATCGTGCAA AGCAGCAACA CCGCGCTCCT GTGGGTGCAT
GGTGCAACCC CTAGCGCGCA AAATCCAACG TTTAAGTATT TCCAAGGCAA ACGACGGATC
CCCGTGCCCC TTGTCATAAA GCGCTACGTG GGGCAGAGCG ACATTAGCCA GTTGGCGAAC
GAAATATTGG GCCTCAGCAA AATGAACTGG AACACCTTTG ACTATTACTC CAGGCTTCCT
GTAACCCTTG AGAGCGCCAA TGATATTGCC CGGATCGGCG TGTATTTCAA CAATTTCTCC
CCCATGAGCT ACGACTATCG GCTCCTCATA TAGTGA

68 Gene Name: gene-18861

Gene ID: BBF.2017.48.68

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGCGAAA AGTGGAAGAT CCGAAAAAGA AGAGGAAAGT GGGCAGCGGG
TCTATGAATA ACGTGATGCA GGAGTTTCCC GTCGCAAGCT TCCCCACATT CTTGTCCGAG
ATCAGTCTGC TTGACATCAC ACCGAAGAAC TTTATCTGCT TTAGGCTCAC CCCCGAAATC
GAGCGCAAGA CCGGTAACAG TTTTAGCTGG CGCTTCAGCC AAAAATTCCC TGACGCCGTC
GTGATTTGGC ATAACAAGTT TTTCTGGGTA CTCGCTAAGC CCAATAGACC AATGCCCAGC
CAGGAGCAGT GGAGAGAAAA GTTGCTGGAA ATCTGCGAGG AACTTAAGAA GGACATAGGC
GACAGAACCT ACGCCATTCA GTGGGTTAGC CAGCCCCAAA TAACCCCTGA GATCCTGTCT
CAACTCGCCG TCAGAGTGTT GAAGATCAAC TGTAGGTTTA GCTCTCCCAG CGTAATTTCT
GTCAATCAAG TTGAAGTGAA GAGGGAGATC GACTTTTGGG CCGAAACAAT TGAGATTCAG
ACCCAGATCC AACCCGCTTT GACCATCACC GTGCACAGTT CATTCTTCTA TCAACGACAC
CTGGAAGAGT TCTACAATAA TCACCCTTAC AGGCAGAACC CCGAGCAACT GCTCATCGGC
CTCAAGGTGA GGGACATTGA AAGGAATAGC TTCGCGACGA TTACTGACAT TGTGGGCACC
ATAGCGGACC ACCGCCAGAA GCTGCTCGAG GATGCCACTG GAGCTATTAG TAAGCAAGCC
CTTATAGAGG CCCCAGAGGA GCAGCCCGTG GTCGCCGTAC AGTTCGGTAA GAACCAACAA
CCCTTCTACT ACGCAATGGC CGCGTTGCGG CCTTGTATCA CCGCCGAAAC CGCTAGGAAG
TTTGACGTGG ACTACGGCAA ACTGCTGTCC GCCACCAAGA TACCCTACTT GGAGCGGAAG
GAGCTGTTGG CTCTCTACAA AAAGGAGGCG GGTCAATCTC TGGCGACTTA TGGTTTCCAA

TTGAAAATCA GCATCAACAG CAGGAGGCAT CCGGAGTTGT TTTTCAGCCC AAGCGTGAAA
CTGAGCGAAA CCAAACTCGT ATTCGGGAAA AACCAAATAG GGGTGCAGGG GCAAATTCTT
AGCGGATTGA GCAAGGGTGG GGTGTACAGA AGGCATGAGG ACTTCAGCGA CCTCTCAAGA
CCTATACGGA TCGCTGCGCT TAAATTGTGC GACTACCCTG CGAATTCATT TCTGCAAGAA
ACCCGGCAAC GCCTCAAACG GTACGGTTTT GAGACTCTGC TGCCCGTCGA GAATAAGAAA
ACCCTGCTGG TAGACGATCT GAGCGGGGTC GAAGCACGCG CGAAAGCCGA GGAAGCCGTT
GACGAACTGA TGGTGAACCA CCCCGACATC GTGCTCACTT TCTTGCCGAC CAGTGATAGG
CACAGCGACA ACACGGAAGG CGGCTCATTG TATAGTTGGA TTTATTCCCG ACTGCTGCGG
CGAGGGATTG CTTCACAGGT TATCTACGAG GACACGCTTA AGAGTGTGGA GGCGAAATAT
CTCCTTAACC AGGTGATCCC CGGAATATTG GCAAAACTCG GCAATCTGCC GTTCGTACTT
GCGGAGCCCC TGGGAATCGC TGACTACTTC ATAGGCCTGG ACATCTCCAG GTCAGCAAAG
AAACGGGGGT CTGGAACCAT GAATGCCTGT GCCAGCGTTA GGCTGTATGG TAGGAAGGGC
GAATTTATCA GGTACAGGCT TGAGGACGCA CTGATCGAAG GGGAGGAAAT ACCTCAGCGC
ATTCTGGAGA GTTTTCTGCC AGCCGCTCAA CTGAAGGGCA AGGTAGTGCT CATTTACAGG
GACGGCCGAT TCTGTGGTGA CGAGGTCCAG CACTTGAAAG AGAGAGCAAA GGCTATAGGA
AGCGAGTTCA TCCTGGTTGA ATGCTACAAG AGTGGGATTC CACGACTGTA TAACTGGGAA
GAAGAAGTCA TAAAGGCACC AACTCTGGGA CTGGCCCTTA GGTTGAGTGC GAGAGAAGTG
ATTCTGGTGA CAACCGAGCT GAACAGCGCA AAAATCGGGC TTCCTTTGCC TCTGCGACTC
AGAATTCACG AAGCCGGTCA CCAAGTATCT CTCGAGTCTT TGGTAGAAGC CACACTGAAG
TTGACCCTCC TCCACCACGG CAGCCTGAAC GAACCGCGGC TGCCTATACC ACTGTTTGGT
TCCGATCGAA TGGCCTACCG GAGACTCCAG GGCATATATC CCGGATTGTT GGAGGGGGAT
CGGCAGTTCT GGCTTTAGTG A

69 Gene Name: gene-21126

Gene ID: BBF.2017.48.69

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAGAGAAA GGTAGAGGAC CCAAAGAAGA AGCGGAAGGT GGGCTCCGGT
TCAATGAACC TGACTCTGTT CAACGAGATC CTCCCCATCA ACATCAGCCA ACTGCCCAAC
CAGTACTTCT ACAAGCTGTG CACTGCCGGC GACGTGGACC TGGATTCTCT GGGCAGGAGC
ATCAAGTACC GGATCCAGAA ATACTTCAGA GGAATCTGGG TGTGGAGTAC CAACGACCAA
CTCCTCATTT CAGACAAGCT CATCGAGTAC CCCGAACTGC AAAAGTTCAC CCAGTATCTG
TGGACCGACC AGTCTAACCT CACATTCAAC CAGCTCGAGG GGATAGAAAT CGAGAACATT
AGGTGTTGCA CCCCCCAAGG TATCGCTGAT TTCTGTAGCC AAGGACTCAT CAAAAAGTAC
GACCAGCAGA TCAAGAAGAT ACTCGAACAG TCCAAGACAG CACGGAGAGA CTATCATATC
AAACTGATCC ACAAGTTCGG CTCCTGGGTG GTGAACAATC AGCCCTGCAT AAGCCTGAGC
CTGAAACAGG AGATCGATTT TAACGGAACT CTCCAGGACT ACCTGACCAA GTTCCCCAAC
TCTAACATCA TCGGCCTGCA TGTGCTCGAC ATCACTAAGC CTTTCAACAC CGCACAGGAG
GTCATCAAGA TTCTCGGTAT CTTGGGTGAG GGAAATCGGC GGCAGCGCCT CCTGACTTGG
GTCAAGGAGC CAACCATGAA AAAACTCGTG GAAGAGGCCC CAGATAGTGA GCTCGTAGTT
GAGATCGGGA ACAAGAAAAA ATCCTATCAT TACATCATTT CTGCCCTGCG CATCAGAGTC
CTCAACCAAG ATTACCTGAG GCTGGGGATT AGCGAGAAGC TGCAAATAGT CAGTGAAGAG
AGGTTGAAGT ACATCGAGCC ACTTTTCCGC ATACTGCAAT CAGAGGGCTT CCTGGACAAG
GTGTATACTA GCCAGCGCAA CCCCGAGCTG TTTAGGTCAT GCAGCGAGGA ATGGGGTTAC

AATCCCCTGC TGAAGTTCAA GAATAACGCC ACTGTTGCGG CGGAATCCGT GCAGTCCACG
GTCCAGGTGG TGCAGAAACA CGGCGAATTC AGGAAAGCCG ACAAAAGCGA AATTAGGATC
GCCATACTCA ACACACTGAA GAGTGAAAAC AGCACCAAAT TGATTGAGAT TTTCCGAAAC
AACTTTAAGC GAAGCTTTAA CCAGAATTTG GAGGGAATCG GTAATCAGCT TAAGTATAAA
CTCAAGTTGG TGGGCCAGCC CATTGCACTG GATCTCAGTA AGAACTCCCT CAGCCTGCTG
GACAGCAAAA TAGGAGAATT GTCTAAAAAG AAGCCGGACA TTGTGATCTG TGTGATCCCT
AACTTCCTTA GCAAGGGCGA GGACGGGCGG ACACTTTACG ACGATTTGAA GCAGACGTTC
CTCAAATACA ATCTCCAATC ACAAATGTTG CAGGAAAAGA CTCTCACGAC GTCATTTGCC
ACAAAGAACA TCGTGTTGGG CGTGCTGGCG AAAATTGGAA GCGTTCCCTA TATTCTGCAA
GAACCGCTGA CGTACACGGA CTTTGTCGTA GGTTTGGACG TGAGCAGGCG ACGCAAGAAA
AATCTGCAAG GAACCAACAG CGTAGCCGCC ATGACCCGAA TCTACAGCAA TCAAGGCGAA
CTGGTCCACT ATAGCATCCG CGACGCAACC ATCGACGGCG AGATCATTCC CAAGAGGATG
CTCTACGACC TCTTTCCACT TCACGAATAT CAGGGCAAAC GCGTGGTGAT TCACCGGGAC
GGAAACTTCC CCGAGGAGGA GCGCCAGGCA CTCGAGGAAA TTGCCGAAAA GATTGACGCG
AAGTTCTACT TCGTAAGCAT TATCAAATCT GGCAATCCCA GGATCTACGG TAGGACCAAA
AACGAAGAGG GCATCGGCAG TTATCGCAAG GCACCTAAGG GTAGCATTTT CCTCCTCAGC
GAAACGGAGG CCTTGCTTAT CAGCAGCGAC TTTCCGGACC GCTTCAGGGC CACGCCACAG
CCTCTCAGAA TTAAGACGTT TGGCAACTTT CCCCTTCAAA GCGCCGTCCA TAGCGTTCTG
TCACTCACCT ATCTGCACTA CGGTTCCGAG CGCCCACCGA GGCTGCCGGT GTCTACCTAC
TACGCAGATA GCATTAGCAC TATGGTATCC AAGGGCATTA AGCCCAAGGA CGTTGACGGC
AATATACCCT TTTGGCTGTA GTGA

70 Gene Name: gene-4617

Gene ID: BBF.2017.48.70

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGAGGAA GGTGGAGGAC CCGAAGAAGA AGCGCAAAGT GGGTAGCGGG
TCCATGAAAG AGTTTAACGT CATTACCGAG TTCAAGAACG GCATAAACAG CAAATCTATT
GAGATCTACA TCTACAAAAT GATGGTCCGA GATTTCGAGA AGCGACACAA TGAAAATTAC
GACGTGGTGA AGGAGCTGAT TAACCTTAAC AACAACTCCA CCATAGTGTT CTACGAGCAG
TATATCGCCT CCTTTAAGGA GATTGAGAAA TGGGGGAACG AGCAATACAT AAATGTGGAG
AAGAGGGCTA TCAACCTGGA GTCCAACGAG AAGAAAATTC TGGAGAGGCT CCTGCTGAAG
GAAATCAAAA ATAACATAGA CAATAACAAG TACAAGGTCG TCAAGGACAG CATATACATC AATAAGCCAG
TGTACAACGA GAAGGGCATC AAAATTGACA GGTATTTCAA TCTGGACATA AACGTTGAGT
CAAACGGAGA CATTATCATC GGGTTTGACA TCTCCCATAA CTTCGAGTAT ATCAACACTC
TGGAGTATGA AATAAAGAAC AATAATATCA AGATTGGGGA CCGGGTAAAG GACTACTTCT
ACAACCTGAC CTATGAGTAC GTGGGGATCG CCCCCTTTAC TATCTCCGAG GAAAACGAGT
ACATGGGCTG CTCAATCGTC GACTATTATG AGAACAAGAA CCAGAGCTAT ATTGTGAATA
AACTGCCTAA AGACATGAAG GCCATCCTGG TAAAGAATAA TAAGAACTCT ATATTTCCCT
ACATCCCGAG CAGGCTTAAA AAGGTGTGCA GATTCGAAAA CCTTCCCCAG AACGTGCTGA
GGGACTTTAA CACGAGGGTG AAGCAAAAGA CAAACGAGAA AATGCAGTTC ATGGTTGACG
AAGTGATCAA CATCGTGAAG AATTCCGAGC ATATCGACGT CAAAAAGAAA AACATGATGT
GCGATAACAT TGGGTACAAG ATCGAGGATC TGCAACAGCC CGATCTTCTC TTCGGTAACG

CCAGGGCCCA GAGGTACCCC CTCTATGGCC TCAAAAACTT CGGGGTGTAC GAAAACAAGC
GGATAGAGAT CAAATACTTC ATAGACCCCA TCCTCGCCAA GTCAAAGATG AACTTGGAGA
AAATCTCCAA ATTTTGTGAC GAGCTGGAAC AGTTTAGCAG CAAGCTGGGC GTGGGGCTCA
ACCGGGTTAA GCTGAACAAC ATAGTTAATT TCAAAGAAAT CCGCATGGAC AATGAGGACA
TTTTCAGCTA CGAGATAAGA AAGATAGTGA GCAACTATAA TGAAACTACC ATCGTAATCC
TGAGCGAGGA GAACCTGAAT AAGTACTACA ACATCATTAA GAAAACATTC AGCGGCGGAA
ACGAGGTGCC CACCCAGTGC ATCGGTTTCA ATACGCTGAG CTACACGGAG AAAAACAAAG
ATTCTATCTT CCTGAACATT CTGCTGGGGG TTTACGCCAA GAGTGGCATC CAGCCCTGGA
TCCTGAATGA GAAGTTGAAC AGTGACTGCT TTATCGGCCT GGACGTGTCT AGGGAGAATA
AGGTCAATAA AGCGGGAGTC ATCCAGGTGG TCGGGAAAGA CGGCAGGGTG CTCAAAACTA
AGGTGATCAG CAGCAGCCAA AGCGGAGAGA AGATCAAGTT GGAAACCCTC AGGGAGATCG
TGTTTGAGGC AATCAACAGT TACGAGAATA CGTACCGGTG CAAACCCAAA CACATTACTT
TCCACCGAGA TGGAATCAAC CGCGAGGAAC TGGAGAACTT GAAGAACACC ATGACCAACC
TCGGTGTTGA GTTCGACTAC ATCGAAATTA CCAAAGGCAT TAACAGGAGG ATCGCCACTA
TCAGCGAAGG TGAGGAATGG AAAACGATTA TGGGGAGGTG CTACTATAAG GACAACAGCG
CGTACGTGTG TACCACCAAG CCTTACGAGG GAATCGGCAT GGCCAAGCCC ATCCGAATCA
GGAGGGTGTT CGGCACGCTC GACATAGAAA AGATTGTCGA GGACGCCTAC AAACTGACCT
TTATGCACGT TGGCGCAATT AACAAAATCA GGCTTCCCAT TACTACGTAC TACGCAGACC
TGAGCTCCAC TTACGGCAAT CGGGATCTTA TCCCCACAAA CATCGACACT AACTGTCTGT
ACTTTATATA GTGA

71 Gene Name: gene-5297

Gene ID: BBF.2017.48.71

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAAAGGAA AGTGGAGGAC CCAAAGAAGA AGCGGAAGGT GGGCAGCGGT
AGCATGCAAG GCACTATATC CATAAACGAG GTGAGGATCC AGCTTAATAC TATTAAGAAT CTTTCAGTGT
TCAAGTGCAG CCTCAGCGGA ATTAGCACCC GCCATAAGAA CCAGATCGAG TTCATCCTTC
GCAGCGAGCA AAACCGAGTT AGCATCTTTG AGGGTGAAGT GATCTTTGCG CTTCCCGTCG
AACAGCAGAA CCTCGAAAGA GATAAGCAGG CTCTGTTCAG CTTCCTGGTC AAACAACAAA
GGGATCTCAA TCTGAAACAG CTGAGCCTGG TGCCCCTGAG GGAGGTGCCC GAGCGCGTTA
TCGAGCGACT GACTTTCGCA ATGGTTAGCT ATCAGGCCAT GAAGCAGGGC ATCTTCTCTA
TCTATGGTCA TACATTTTTT CGCCCCACCC TTATGACGGA TAGGCTTGCG CACAAGGCGG
TGGAAGTCAC GACGTGCATC GAGGATGGCT TCCTCAAGTT TTATCTGGAC CCGACGTACA
TTGCACTGAC ATGCATAACG GACACAGCAC GCGAAAATAG GGAGAACCTG GAACTGGTCG
GGCTCTGCTC TTTCCGCAAC AAAAACCTTT GTAGCCTTGT CAGGCCGGAC GGCTCATGCA
ACTGCCTCAT ACCTGGTAAG TTGGGGTATT ACGTCCAGGA GATGGGGATT AAGGACGTTG
AGGATGATAG CAAGGACTTT CTGGCCAAAC GGTTCAATAG CTGTCCCCGG TTTAGTGAGC
ACACGCGCTT TATACAAGTG AAGGCGAGTA AAAGAGGCAC GAAGTACTCC CTGTTCCCTT
CTTACGTAGT TTTTTCCAGG TTGTCCCGAA TGGACCTGTC CGCTAAGCCA GATGTGCGGT
CCAGTTATCG GAAGGCCACA TTGATGGACT CTCACGAAAG GCTTAACTTG ACCAACGACT
GGATAAGACA AATTTTCATG ATCGGGCAGA AGGGCCTTCA AAATTGGGGT GTTATAAAGG

TCAACCAGAC CGAGATTCCC GTTGAAATTG TACTCACAAT TGCCCACGCG ATCGCGCCCA
AGACTTCTCA AGGCATCTAT AAGGCTATAT TCCTCCCGGA CCAGCAAATT ACGAATGACA
GCAATAACCC AACGCCTCAA ACGCTGAGCG GGGGTTGGCT GTTCACGAAT AGGGGTGCGT
TCGACAGGAG GGATCCTAAT AGGCCTTTTA AAGTAATCAG CCCCTACATC ATCGTGCCCA
ACAATGAGCA AAGCATCAGC TCTTGCCGCC AGCTGATCAA CTACTTCAGC AACGGAAGGT
ACAAGGCCCG GTGCAAGGGT GACAGAGACT TTATTGGTAT TTCATTGCCC GAAAACAAGG
GCAAGTACAA CACATCATTT GTCAATGCTT TCGAAGAGGA GGACGGCCTG TATTTCGTTG
AAGAAACGAT ACAGGGCTAC CAGAAGGCGC TGCAAGACAT TGTTAGAGAC TGGAATATCA
CGTCCAAGCG GGACATCAAT AAACACGCTA TAGTGATCAT ACCGGGCGAG AACGATATTG
ACGACAATCC TTTCTATTAT CAACTGAAAA AGGCGTTCGT AGAGGAAGGG ATTCCCAGCA
CCTTCATCAC GTACGAGACT ATGAACAAAA TCAACGACCC CGATATCGCG TTCGGGCCAA
TCATGGACAG CCTGTGGTTG AACATTTACA GCAAAATGGG GGGCAAACCG TGGCGCCTCG
CTAATAGCCT CGGCAACGTG CACTGCTTTA TCGGTATTGG GTTTGGAATT AACCCCGAAA
CCACCGGAAA CCACATATTC GCAGGGATCG CCCACATCTT CGACAACTAC GGGAGTTGGA
TAGACGTAGC GAGTGATTCC GCCAACCTCT CCCAAAACGA TCTGAACTCA TTCGAGGGCA
CGGAAAAGTA CACACAGGGG AGTGCTAGCT TTAAGATCAG TCAGAGCGTG TCCCAGTCCA
TTGTGTATAA CGCATTGAAG CTGTACCAAC AGAAGCAAAC TAAGACCCAC GAAAACGCCA
CAAACATCGT CCTGCACAAA CTGGGCCAGA TCTACGAGTG TGAGGTCATC GGGTTCCTCG
AAGGAATTCG CCAAGTGCTC GGGAGTCTGG GCGACTGCAA GCTGGGATTG CTGCAAATTG
AGCAGGAGCA CCATCTGCGC CTCTATGGCG CAGCAGCCCA AACCGGCAAG GAGAACAACA
CGATCTTTCG CGGTTCAGCA CTTCAACTCA ACCCGGAGAA GCTGGTTATC GCGTCCACTG
GCCGCTCTTA CCGGCAGACG AGCTCCGGGC TGTTTATGAA TTATCCGGGC ATCGGCACCC
CCCAGCCGCT CCTGTTGACT TCTATCGTAC CGAATCAGCA GATCCTGCAG AAGTACGGCT
GTAACGCAAA CCAATTCTAC TCAAGCGAGG ACCTGGCGAA ACATGCAATG GCCCTGACGC
AACTTCACTG GGGGTCACTG AAGGATAATG TAAGATTGCC GATTACCACG CTTTACGCGC
AAAAGGTCGC CGACTTGATT AGCAAGACCA ACATGCGGAT CAATCCAGGC TTGGGCTACT
TCCGACCCTG GTTTCTTTAG TGA

72 Gene Name: gene-6521

Gene ID: BBF.2017.48.72

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AAAAGCGAAA GGTGGAGGAC CCAAAGAAGA AACGCAAGGT GGGCTCCGGC
AGCATGAATA ACCTGACACT GGAGGCCTTT CGGGGCATTG GCACCATCAA GCCACTGTTG
TTCTATCGGT ACAAGCTGAT CGGCAAAGGG AAAATAGAGA ATACCTATAA GACGATACGC
AACGCACAGA ATCGGATGTC TTTCAACAAT AAGTTTAAGG CCACCTTCAG TAAGGATGAA
ATCATATACA CCCTGGAGAA GTTCGAGATT ATCCCGACGC TGGATGATGT GACGATCATC
TTCGACGGGG AAGAAGTGCT TCCTATAAAG GACAACAACA AGATTTACAG CGAGGTAATA
GAATTTTACA TTAACAACAA TCTCCGGAAC GTTAAGTTCA ACTATAAGTA CCCGAAGTAC
AGGGCTGCCA ATACAAGGGA GATCACGGGC AACGTGATCC TCGACAAAGA TATGAACGAA
AAGTACAAAA AGAGCAACAA AGGCTTCGAA CTCAAACGGA AGTTCATAAT CAGCCCCAAG
GTCGACGATG AGGGTAAGGT CACATTGTTC CTGGACCTGA ACGCGTCATT TGACTACGAC
AAGAACATCT ACCAGATGAT AAAGGCCGGA ATAGATGTGG TAGGAGAGGA GGTCATCAAC
ATCTGGAGCA ATAAGAAGCA GCGCGGTAAG ATCAAGGAAA TCAGCGACAT TAAGATAAAC
GAACCCTGCA ACTTCGGCCA GAGCCTGATA GATTACTATA TAAGCAGCAA TCAGGCGTCA
CGGGTGAATG GATTTACGGA GGAAGAGAAG AACACAAACG TCATCATCGT GGAAAGCGGC
AAAAGCCGCC TGTCATACAT ACCGCACGCG CTCAAGCCTA TCATAACGCG AGAGTATATC
GCCAAGAACG ACGAAGTCTT TAGCAAGGAG ATAGAAGGGC TCATCAAAAT CAATATGAAT
TACAGGTACG AGATTCTCAA GAGGTTTGTC TCCGACATCG GCACTATTAA AGAACTGAAC

AATCTGCGCT TCGAGAAAAT CTATATGGAC AATATAGAAA GCCTGGGTTA CGAGCAGGGT
CAACTCAAGG ACCCCGTGCT CATCGGCGGC AAGGGTATAC TTAAAGACAA AATACATGTG
TTCAAGAGCG GCTTCTACAA ATCCCCCAAT GACGAAATTA AGTTTGGCGT GATATACCCG
AGAGGCTACA TAAAAGATAC CCAGAGCGTT ATCCGAGCCA TCTACGACTT TTGCACCGAG
GGCAAGTACC AGGGAAAGGA TAACATATTC ATCAATAACA AGCTCATGAA CATCAAGTTC
TCCAATAAGG AGTGCGTCTT TGAAGAGTAC GAGCTCAATG ACATAACCGA GTATAAGCGG
GCTGCAAATA AGCTCAAAAA GAATGAGAAC ATAAAGTTCG TGATCGCAAT CATCCCCACT
ATCAATGAAA GTGACATTGA GAACCCCTAC AACCCCTTCA AAAGGGTCTG TGCCGAGATC
AACCTCCCCA GCCAAATGAT CAGTCTCAAA ACTGCAAAGC GGTTCAGCAC CAGCAGGGGC
CAATCTGAGT TGTATTTCCT GCATAACATC AGCCTCGGCA TTTTGGGCAA AATAGGCGGC
GTACCCTGGG TAATTAAGGA CATGCCAGGC GAGGTCGATT GTTTTGTGGG CCTGGACGTG
GGCACAAAAG AGAAAGGAAT CCACTACCCC GCATGCAGCG TGCTGTTCGA CAAGTATGGC
AAACTCATTA ACTACTACAA GCCGACGATC CCGCAGAGTG GAGAGATCAT TAAAACAGAC
GTGCTGCAGG AGATCTTTGA CAAGGTTCTG CTGAGCTACG AGGAGGAGAA CGGCCAGTAT
CCCCGCAACA TCGTGATACA CAGGGACGGC TTCAGCCGGG AGGACCTGGA GTGGTATAAG
AACTACTTCC TGAAGAAAAA CATCGAATTC AGCATAGTAG AGGTCCGCAA GAACTTTGCC
ACGCGACTTG TAAACAACTT CAACGATGAA GTGTCCAACC CAAGCAAAGG TTCATTCATT
TTGAGGGACA ACGAAGCGAT TGTCGTCACG ACGGATATTA ACGACAACAT GGGAGCGCCC
AAACCGATCA AAGTTGAGAA AACGTATGGC GATATTGACA TGCTCACAAT TATCAACCAA
ATTTACGCAC TGACACAGAT TCACGTGGGG TCCGCGAAAT CCCTTAGACT GCCTATAACC
ACGGGCTACG CCGATAAGAT CTGCAAGGCT ATCGATTACA TCCCGAGCGG CCAAGTCGAT
AACAGGCTGT TCTTTCTGTA GTGA

73 Gene Name: gene-8259

Gene ID: BBF.2017.48.73

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGAGAAA GGTCGAGGAC CCGAAGAAGA AAAGGAAAGT GGGCAGCGGC
AGCCTGAAAA TCAAAATTCT CAAGGAGCCG ATGCTGGAGT TTGGCAACGG CGCTCACATA
TGCCCCAGGA CCGGTATCGA AACCCTGGGA GTGTACGATA AGAGAGATGA ACTGAGGAGG
AGCGAGCTGC GAATAGGCAT TGTGGGTCGG GGCGAGGGCG TGGACCTTCT GGATGAGTGG
CTCGACAAGT GCAAGCGCGG CATCGTGGGT AAAGAGGAAA CCAAGTTCCC CAACTTGTTC
AGGGGCTTTG GGGGCGTCGA TGAGTACCAC GGTTTCTACA CCAAGATTCT GAGCAGCCCC
CAGTATACCC GGACTTTGCA GAAAAGCGAG ATTAACAACA TCAGCAAGAT CACCGCCCGA
GAGGACAGGG TAGTGAAGTG CGTGGAGCTG TACTACGAGC AGATCCGATT CCTGTCAGAG
AACAGGAGCA TTGACGTGAT CGTGTGCGTC GTTCCCAATG ATATTTTCGA CAGCCTTACT
AAGGCCACCG GAGACAAAGA CACCGAGTCC CTGGAGGCCT ACCTCGAGCA CAACTTTAGA
CGGTTGCTCA AGGCCCGCTG TATGCACCTT GGGATACCCT TGCAGCTTGT GAGGGAAAAG
ACCATCCTGA GCGTGAAGCC TAGCATAGAC CAGCAGGACC TTGCCACAAA GGCTTGGAAC
TTCTGTACGG CCCTCTATTA CAAGGGGAAT AGGACTGTAC CATGGCGCCT GGTGGAGGAT
AAATTCAAGC CTAAGACTTG CTACATCGGC ATTGGGTTCT ATAAGAGTAG GGACGGCGAA
ACGGTGAGCA CATCACTTGC ACAGGTATTC GACGAGTTCG GCCACGGGGT CATCCTTCGG
GGAGCACCAG TTAGCCTGGA CAAACGAGAC AAGAGGCCCT ACATGGACGA GTCTCAGGCT
TACGAACTGC TGGACAGTGC CCTGGCGGAG TACGAGAAGG CCCTGATGCA AAAGCCCGCT

CGAGTGGTGA TCCACAAGAG CTCCAGGTTC CGGCCCACCG AGGTGAGCGG CTTCAGCAGA
GTGCTGAACG CGAAAGGAAT CAGAACGAAG GACCTCGTGA GCATCACATC AACCGACATC
CGCCTGTTCA GCGACAAAAA CTATCCCCCC ACCCGCGGTA CCTTGTTGTC CCTGTCTGAA
ACACAAGGAG TACTGTATAC CAAGGGAATC GTAGATTTTT ACAAGACCTA TCCGGGCATG
TATATCCCTT CACCCCTGAG GGTTGAGGCG TTCGAGTCCG ACAGCTCTCT TGAGGACTTG
TGTAAGGAAA TCCTGGGCCT GACCAAAATG AATTGGAACA ACACACAACT GGACGGCCGA
CTGCCCATTA CCCTGGAATG CGCCAATAAG GTGGGCGATA TCATGAAGTA TGTGGACGCA
TCCGAAAAGC CACAGGTTGG TGTGGCGCTG TTTATCTTCA TGTTGGAGCA ACTCGTACCC
GGCTGGAAGC TGCCTAAGGT GAGTACATGG GTAGCACGGG TAATTTTCCT GAATATTGTA
CAGGTGTCTA TCGCTCTGCT TGCCGGGATT ACTTGGAATA AATGGATGAT GGGCCACAGT
TTGTTGCATA CCTCTGATGC CCTGCCCCCC TTGCTCGCAG GATTCGCCGC CTACTTCGTT
AACACCTTCG TGACCTACTG GTGGCACAGG GCCAGGCACG CCAACGACAC CCTTTGGCGA
CTTTTTCACC AACTGCACCA TGCGCCCCAG AGGATCGAGG TGTTTACTAG CTTCTACAAA
CACCCAACGG AAATGGTATT CAACTCTCTT CTTGGCAGTT TCGTGGCCTA CGTCGTTATG
GGGATCTCCA TCGAAGCTGG CGCGTATTAC ATCATGTTTG CGGCTCTTGG CGAGATGTTC
TACCACAGCA ACTTGCGAAC ACCGCATGTT CTCGGTTATC TCTTTCAACG CCCTGAGATG
CACCGGATCC ACCACCAGAG GGACCGACAC GAGTGCAACT ACAGCGATTT CCCCATCTGG
GACATGCTGT TCGGCACCTA CGAAAATCCC AGGAGAATAG ACGAACCACA GGGGTTTGCC
GGCGACAAGG AACAGCAATT CGTTGATATG CTTTTGTTTA GGGACGTGCA TTCCCTCCCC
GGAAAGACAC AACCAGCTCC CGTACTCGTC AAACCCGACG TGAGGTAGTG A

74 Gene Name: gene-11357

Gene ID: BBF.2017.48.74

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAGCGGAA AGTAGAGGAC CCGAAGAAAA AACGCAAGGT GGGCTCCGGG
TCTATGGCCA ACCATACCTT TAACATCCTG ACTTTCAACC ACCCCCAGGA GGAACAGACC
TTCTACTTCA CGGACCAGGA GCAAGACAAC CTGACCCGCA TCTACAAGAG CCTGGTGCCC
GACGAGGTCA TCGAGAAATA TGGCGAGCAG GATCACTACT ACACCTCTTT CACCGTAGAG
AAGGATGGTT TCCTGGCCGT CAGCAAGCCC ACAACGCCCC TGTTCGAAAC CAAGACTACG
GAGGCGGGCG AGGAGAGGAG CTATACCATC AGGAATTCAA CGTTCAGCAG CAGCGTGTTG
AAACGGTACT ACAACAGCCT TATCCACAGC CACTTCAAGG AGAAGGGCTT CCTGGTGAAG
CCCAACTTCG TGAGCGACAC GGAGGTGTGG CTGCCTAGCG CCAAGCAGGA CACGACCGGC
AAATACAAAA TATTCGACCG CTTTAGCCTG AAGGTGCAGT TCAAGACAGT CTCTGATTCC
CTGGAGTTGC TCGTCACGTT CGAGGGGAAG TCAAAGATAT TCAAAGTACC TGTTAGCACC
CTGCTGGAGG ATGTGAGCCC CACGGACATC AACTGGGTTG TGTACGAAAA GGGATTGTAC
AGGTTCGACG AACTCCCGGA CAGCGGCAAG AGGGAGTATG ACAAGGTTTA CCCCGTGTGG
ACCTTCGAGA TCAGGGACGC GCTTATGCAG GGCACCGAAG CCCCAGACAA GACCAACAAG
TACAAAAAGT TCAGGGAGGG CATCGACAAG TTCTATAACC AGTATCTGAA CACAGAGGAG
TTCAAAGCCA TCATTCCAAT CACGTCTAAT GGCTTCATCC CGGTCAATAA GATCAATGTC
GGTAGTGTGA ATAATAGTAG CAACAGGCTG CTGTTCGGGG AACAAAAGAG CGGTATCGTG
CCAATGGACG GCATGAAGGA ACATGGCCCA TTCGACTTTT CCAGCACCAG CAAGATCCAT

TTCTTCTTTA TCTTTCATAA AGACGACCAG CATATCGCCC AAAAGATGGA TGGCTATTTC
AAAGGCAGCG AGTTCGGGTT CAAGGGACTC ACCAAATTCA TACACACCCC CTATCACACC
GAGAAAGGAT TCTCAATCAG GTTTGAGGAC CGCGACAATC CGTGGCCCGA GATCTACGAA
GCCGTCACTA ACAAGCACTT CGAGTCCGAC ATACAATACA TTGCGATCTA CATCAGCCCC
TTCAGCAAAA ACAGCCCCGA CAAGAGTCGG CGCAAAATCT ATTACAAGCT CAAAGAACTG
CTCTTGAAAG AAGGCGTGAG CAGCCAGGTG ATTGACGGCG AGAAGGTGAT GACCAACGAG
AAGTATTACT ACAGCCTCCC CAACATAGCA ATCGCCATTC TGGCCAAGTT GAATGGCACC
CCTTGGAAAC TGGACACCAA GCTGAAGAAC GAACTGATCG TGGGAATCGG CGCCTTCCGC
AACAGCGAGG TTGACATTCA ATATATCGGC AGCGCGTTCT CTTTCGCAAA CAACGGCAAG
TTTAATCGCT TTGAGTGCTT CCAGAAGGAC CAGACGAAAG AATTGGCGGG AAGCATCATA
CGGGCGGTGA AGGAGTACGC CAACGTAAAC ACCGGCATTA AGAGGCTTGT GATCCACTTT
TACAAAAGCA TGCGACAGGA TGAGCTCCAG CCGATCGAGG ACGGCCTTAA AGACCTCGGC
CTGGACATTC CGGTATTCAT CGTATCTATC AATAAAACAG AAAGCAGTGA TATCGTGGCG
TTCGATAACA GCTGGAAGGA TCTGATGCCG ATGAGCGGCA CATTCATTAA AGTGGGGTAC
AACAAATTTC TCCTGTTCAA CAACACCAGG TATAATCCAA AGTTTTACAG CTTCCACGAC
GGGTTCCCCT TCCCCATCAA ACTTAAGATT TTTTGCACTG AAAAGGAACT CGTGGAGGAG
TATAAAACGG TTAAAGAGCT GATCGACCAG GTGTACCAAT TTAGCCGCAT GTACTGGAAG
TCTGTCCGCC AGCAGAATCT GCCCGTGACC ATTAAGTATC CGGAAATGGT GGCCGAAATG
TTGCCTCACT TTGACGGGAA TGAGATACCT GAATTCGGTA AGGACAACTT GTGGTTCCTG
TAGTGA

75 Gene Name: gene-12090

Gene ID: BBF.2017.48.75

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGCGAAA AGTAGAGGAT CCAAAGAAGA AACGGAAGGT CGGCAGCGGA
AGTGTGAACC ATTACTATTT TTCCGAATGC AAGGCGGACG AGAAAGCCAG CGACATAGCC
ATCCACCTTT ACACCGTGCC CCTGTCCAAC CCCCATGAGA AATACAGCTA TGCGCACTCA
ATCGCCTATG AATTGAGAAA ACTCAACTCA TACATAACCG TGGCCGCGCA CGGTCAGTAT
ATCGCGTCTT TCGAGGAGAT ATGCCACTGG GGCGACCACA GGTACATACA GCACGAACAT
AGACCAATCC AGTGCAGCCT CCCGATGGAG AGGACCATAC TGGAAAGACT CCTCAAGAAA
GAGCTCGAGA ATAGGTGCAA AAGCAGCTAT AAGATGGACA ACGACCTTTT CCGGTTGGCT
AACGAGCAAA GCATGCACGT GGGCGAGATC AGCATACACC CAGCGATCTA CATCTCATTC
AGCGTGGAGG AAAATGGTGA CATATTTGTT GGCTTCGACT ACCAGCACCG GTTCGAGTAC
CGCAAAACAC TCCAAGACGT CATCAACAAC GATCCCTCCC TGCTTAAGGA AGGCATGGAA
GTGGTGGACC CCTTCAATAG AAGGGCCTAC TATTACACTT TTGTGGGCAT GGCCGATTAT
ACCGCCGGAC AGAAAAGCCC CTTCCTGCAG CAGTCTGTGA TCGACTATTA TCTCGAAAAG
AATGAGCTGT GGAAGCTCAA GGGTGTGCAC GAAAAAACCC CCGTGGTGCA CGTCAAGAGC
CGGGACGGTC ACTTGCTCCC GTATCTGCCG CATCTGCTCA AATTGACATG TTCATACGAA
CAGCTCTTGC CCAGCATGAC CAAGGAAGTC AATCGCCTGA TTAAGCTGAG CCCCAACGAG
AAGATGAGTA AGTTGTATAC GGAGATGTTT CGATTGCTCC GGCAGCAACA GGTGCTGACC
TTCAAGAAGG AAAACGTGCG AGCCGTCAAC CTCGGCTACG ATGTGAATGA ACTTGACAGC

CCGATCATGG AGTTCGGACA AGGCTACAAG ACAAACGAGA TCTATCGAGG CCTGAAGCAG
AGCGGAGTAT ACGAGCCCAG CTCAGTGGCC GTGAGCTTTT TTGTTGACCC CGAGCTTAAC
TACGACCCCC AGAAGCGGAA AGAAGTAGGT TGCTTCGTCA AAAAACTGGA GAGCATGAGC
GAGGCCCTGG GAGTAAAACT GAACATAAGC GACCAGCCCC GACAACTTTA TGGCCAGCTC
CCCAAGGACT TTTTCAAGCA GGACAACCTC TCATATCATT TGAAATCTAT CACCGACCAG
TTCAGGGGAA CGGTGGTGGT TGTTATCGGC ACTGAAGAGA ACATCGACCG GGCATACGTT
ACAATCAAAA AGGAATTCGG CGGCAAGGAG GATCTGATGA CCCAGTTTGT CGGCTTCACC
TCCTCCCTCG TCACGGAGAA CAACATTTTT CACTACTACA ACATCCTGCT CGGCATCTAT
GCGAAAGCTG GTGTTCAGCC CTGGATACTC GCCAGCCCAA TGCACTCAGA CTGTTTCATT
GGACTCGACG TAAGCCACGA GCACGGTAAG CACGCATCAG GGATAATACA AGTGATTGGA
CGGGACGGCA AGATTATCAA ACAAAAGAGC GTTGCGACAG CAGAGGCCGG AGAGACTATT
GCCAATAGCA CGATGGAAGA AATCGTCAAC GAAAGCATTT ATTCCTACGA GCAGATCTAC
GGGGCCAAAC CGCGCCACAT AACATTCCAC CGCGACGGGA TCTGTCGCGA GGACCTCGAT
TTTCTGCAAG CGTATTTGCG GAGTTTCCAA ATCCCATTCG ACTTCGTAGA AATCATAAAG
AAGCCGCGAC GCAGAATGGC GATATACTCT AATAAGAAGT GGGTCACGAA ACAGGGAATA
TACTACAGTA AGGGCAACAC CGCTTATCTG TGTGCCACGG ACCCCAGAGA ATCCGTGGGT
ATGGCGCAAC TTGTCAAGAT CGTACAAAAG ACTAACGGAT TGAGCGTTCA CGAGATAGTG
AGCGACGTGT ATAAGCTGTC CTTCATGCAC ATACACAGTA TGCTCAAGAC CAGGTTGCCT
ATCACGATAC ACTATAGCGA CCTCAGCTCA ACGTTCCACA ACCGGGGCTT GATCCATCCC
CGGTCCCAAC ATGAGAGAGC ACTCCCGTTC GTGTAGTGA

76 Gene Name: gene-15946

Gene ID: BBF.2017.48.76

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAAA AGAAGAGGAA AGTAGAAGAT CCAAAGAAAA AGCGAAAGGT GGGAAGCGGC
AGCATGACCG GCGAGACTAA AGTGTTGGTC GGGAGGCAAC CCTTCGACGT GGATCGGCTG
AATGAACTCA GGGACGAATT CCGGGAAACG CACGTGTTCA GAAGGGATGG CATCGACGAT
GTCATTGTTG ATGTTCCGGT CGTGGCCGGA CAGAAGCCCA TCGGCAACGT CCAGGAGGAA
ATAGACCTGG CTAGGTACCA AAAGGTGTGG CCCTCCCTCC TCAGTGCTGC TCTTGTCCGG
GCGTTTAGCG GCGTAAGGGA CATCCTGAGC GATAGGCCCG TGAGCGTGGT GGGGAGCACA
CTGCGGGGTC TGGTTCAACA TCCGGAACTC CCCGAATGGA TGCAGAAACG CACACTCCTT
AGGTTCGACA CCCGGACCAT CTATGCTGGT GATAAAAGAA CCTTTGGCTT GGTGTGCGAG
GCCAGATTGA AAAACCTTAT CCAAGGTAGT TGCGCGGAGC TGCTGGCACT TGGAGTTTCC
CCACTGGGTC GATATGTCCA AGTCGAGGAG CCACATTACG ATCCCAGGCT TATGAAAAAA
CGGCGCCTTG TGGGCAGGGT ATCAGCGATC TCCGGCGATA ATCTGGTGCT GGAGGACCAT
GCCGAGGGCT TTCCGACCGT GAGTGCAAAG CTGGCATTTC TGGAGGCGCG AAGGGAGATT
TTTGACGACT GTGTGCGGAG GATTTTGAAC TCTGATGCGG CCTCCGTGCT GAACAAGGCC
GAAGCTACTG CTGCCTCATT TCACTCAGGG CCAGGTAGGA AAGAGCAAAT AGAGGAGGCT
CTCAAGTATC TCAGGGAGAA GGTGAGCCTC GAAGCTGTAC CCGGAGCGAA ATTCGTGATC
GGGCCGATGC TGAGTAGCGG CAACAAGGGC TTCCCCATCA CGGAGATGAT CCCGAAACCC
ATTCTCGTGT TCGATCCGAG CGGTACACGG AAGGATGAGT GGAACGAAAG GGGCATTAAG

AAGAACGGGC CCTACGACCA GAGGACGTTT TCACCTAAGC AGTTGAAGGT GGCGGTCATT
TGCCAGGCGA AGCACGAGGG CCAGGTGGAT GGATTTATCG CGAAGTTCTT GGAAGGTATG
CCAGACGTTA TGACGGGCAA GAACCGAGTT GCTAGATATG GTGACGGTTT TCTGCGGCGA
TTCGCCCTTG AGAAACCTTC TGTGACCTTC TTCACAGCGC CCTCAGCCAA GGCGAGCGAT
TACCTGGTGG CCAGCCGGGC TGCGCTGACC AAGGCAACGG ACGAGGGTTT CAAATGGGAC
CTCGCGCTTG TGCAAGTGGA GGAGGAGTTT AAGGGATTCG ACGACGAGAG CAACCCCTAC
TATGCCACTA AATCCGTGTT CCTGAAGAGG GACGTGCCGG TCCAAAGTGT ACGACTCGAA
ACCATGGCTC AGGCCGACAG CCAGCTGATT TTCTCTATGA ACCACATGAG CCTGGCGACA
TACGCCAAGC TCGGTGGTAC CCCCTGGCTT TTGGCGTCAC AGCAGACGGT AGCGCATGAA
CTGGTTATCG GTCTTGGCAG CCACAGCGTG GCCAACAGCA GGATCGGTAG CCAGCAACGA
TTCGTCGGGA TTACGACGGT GTTCTCCTCC GACGGGAGCT ATCTGCTCTC AGACCGCACG
GCGGTTGTCC CCTATGAGGA GTATGCGACT GCGCTTTACG ATACGCTCAA ACGGAGCATC
ACTACGGTGA GGAAACAAGA CAACTGGAGG TCTACGGATA AAGTCCGCCT GGTGTTCCAC
ATGTTCAAGC CCCCCAAGGA CACCGAGGCC GAGGCTATAA AACGGACAGT GGACGATCTG
GAGCTGGAGA ACGTGACTTT CGCCTTCGTG CATATCGCCC CATCTCATCC CTACCTCATC
TTCGACAATA CACAAAAGGG AATTGGTTTC CGGGACCCCA AGAAGGGGAT ACTCGGACCC
GAGAGAGGTC TGCACTTGAA GCTGGGGGAC TACGAGTCCT TGATCGTATT CAGCGGCGCA
AGCGAGCTGA AACAGGCAAG TGACGGGATG CCCAGGCCAT GCCTGCTCAA GTTGCACCGG
CTTAGCACGT TCACTGACAT GACGTATCTG GCGCGACAGG CATTCGAGTT TTCAGGTCAT
TCATGGCGAA TGCTCTCCCC AGAACCGTTC CCTATAACTA TTAGGTACTC CGACCTGATC
GCCGAAAGGC TCGCAGGCCT CAACGCCGTC CCGGGTTGGG ACGCGGAGGC TGTCAGATTC
GGCCAAATCG GCCGCACGCT CTGGTTTCTG TAGTGA

77 Gene Name: gene-16426

Gene ID: BBF.2017.48.77

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAACGGAA GGTGGAGGAC CCGAAGAAAA AGCGCAAAGT AGGTAGCGGC
AGTATGCGAT TGGGGCACAT AGGCAACGGC TGTTACAGGG AAGGCGTTAA AGCACAATTC
CAGACACGAG AGAGGGAGGA TGCCGGTTCA AGGGCTGCGG CTGCCCAACC CCCGATTAAG
CAATTCGGAT ACACCGATAG ACTCGGCCTG AACCTCGCCC CCATAAGGTT TTCTAGCGAA
GAGTTTGAAG CCGGACGGAC GGTGTACCGC GACGAGGAAC AGTACCGAGC TCTTAGGGAA
GCCCATCAAG CCACCCATGC CTTTAGGTAT GACGCAAGGG ACGCGGCTAT ATACGACATC
CCTATGGCAG AAGGGGTGGC GCCTCTGGGT ACTCCCGTGA GGATCAAAAC TAAGGACCAC
CTCGCTCTGC TCGGCAAAGC GGCTAACCAC GCGCTGCTCG ATTGGCTCGC ACCACGCAGA
ACCATTCTGC GGAGGGCCAG ACCTCTTCAG TGCTGGGGCA ACAGGAAGGC CTCACTGTTG
TCAGCCGCCG TGCGGGATCA AGGACTTGCC GAAACAAAGG GTCTGGATGT TCTGGTAAGG
CATTCTTTTG ATTTGAGGGC TTTGGGCGCA CCTCACCAGG GTGCTGAACC GTACCTTGCC
CTGATGTTGG ACGTGAGTAC GAGCAATGAG CTGGAGATAC CTGTGGGCGA GCTTCTGCGC
GAGAGATTCG ACCCCATCGG TCGATACGTT TGTGCCAGAG CCGACTCTGG CCAAGATAAC
GTACTTGCTA GGTTGGAAAC ACTGGGTAGG GTCGTGGGTG TGGATGGTGG TAAGCTTCAA
CTGAACGACT TTACCGGAGA AGAATTCGTG GACGCTGATT CAGTCACGTT GGAGCCTAGA
TTGGAGAATC TCGATGCGCT CATTCGCCAC TTCTATCCCA GGGATGCGCC AAAAATCCTG
GAGGGCCTTC GCAAAAGGAG AGTGCCTTTC TCCACCGCGA ACGACAAGCT GGCGAAGATA

CGAGAAGTGC ACGGAGGAGT AGCCGGCCAC CTTGAAACGA TTAGGATCGC TGGCATGGCT
ATAGAGGTGG GTGCCCTGCT GCAGAGAGGC TCTAACCTGT TTCCCCCACT CATAAGCACG
GACCGGCCTG GATTTCTGTT CGGCGCTCAA GGTAGGGAAA CTGGCGCGTT CCCCGACGTG
GGGGTGAAGC AGCATGGGCC CTACAAGTAC ATGCAACACG AGCGCAATGA ACCTGTGATC
GCCATCATCT GCGAGAGTAG GTTTCGGGGT CGGATAGACC AACTCGCCCG AACACTTAGG
GATGGTGTCG CGGAAGATGC CTGGCAAGAC GCAATGAGGG GCAGAAATAA GGTGCCGGAA
AACCCCTTTA GAGGCGGGCT GATCGGTAAA TTGAGATTGT CTCGGGTGCA GTTTGAGTTC
GAAGAAGTAA CCGAGCCCAC TCCCGAAGCC TATCGCGAGG CCATCCTTCG GCTGCTTGCG
AGACTCCCAG AGACACCCGA CCTCGCGTTG GTTCAAATAC GAGCGGATTT TAAGCAGCTC
CGCAACGACA GGAACCCATA CTTCGCTGCA AAGGCCGCAT TCATGACGGT GGGAGTGCCC
GTGCAGTCCG TACAAGCCGA GACTGCGGAC ATGCAGCCCA GTAATTTGGC CTACATGGCC
AACAACCTGG CCCTCGCCGC CTACGCAAAA TTGGGCGGTA GTCCGTTCGT GATCTCCACA
CGCATGCCGG CGACGCATGA GCTCGTGGTT GGCTTGGGCT ACACAGAGGT GTCAGAAGGA
CGCTTTGGAC CGAAGTCCCG ATTTGTAGGC ATCACCACCG TGTTCCAAGG GGATGGTAGG
TACTTGGTGT GGGGGCAAAC TAGAGAAGTA GAATTTGAAA ACTACGCCGA CGCTCTCTTG
GCGAGTCTCA AGACTACCAT CGACACAGTG CGCAAGGACA ATAACTGGCA GCCACGCGAT
CGAGTGAGGT TGGTATTCCA CGTGTATAAG CCCCTTAAAC ATGTCGAGAT CGACGCTATC
AAACAGTTGG TGCAGGAGTT GCTGAAGGGC GAACATGAAG TGGAGTTCGC ATTTCTGGAC
ATCTCCCGCT TCCACGATTT TGCCCTTTTC GATCCTTCCC AAGAGGGCGT GAATTACTAC
GCTGACCGCA GACGACTGCT GAAAGGCGTG GGCGTCCCCC TTAGGGGTAT CTGCCTCCAA
CTGGACGAAA GGAGCGTGCT CTTGCAGCTG ACAGGCGCTA AGGAGGTTAA GACCAGTGAA
CAAGGTCTGC CCAGGCCCCT GCGACTGACG TTGCATTCCG AGAGTGATTT TAGGGACCTC
ACATACTTGG CGCGACAGGT GTACAGCTTT AGCTACCTCT CCTGGCGCAG CTACTTCCCG
GCCATAGAGC CGGTGAGCAT TACCTACAGC AGACTTATTG CCAATGCACT TGGCAACCTT
AAGAGCATCC CGAACTGGAA CAGCACATTC TTGACAGCTG GCCCACTGAG GTCAAGGATG
TGGTTTCTGT AGTGA

78 Gene Name: gene-18203

Gene ID: BBF.2017.48.78

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAGAGGAA AGTGGAGGAT CCGAAGAAGA AACGAAAGGT CGGCAGCGGC AGCATGTATC TTAACCTCTA CGAAATCAAG ATCCCCTACA GGGTTAAACG ATTGTACTAC TTCAATAAGG AGAACGACCC CAAAGAGTTC GCCCGGAATC TGAGCCGAGT GAACAACATA CGGTTCAACG ACAGTAAGGA CTTGGTGTGG CTCGAAATCC CCGACATCGA CTTCAAGATT ACACCCCAGC AGGCGGAAAA GTACAAAATA GAAAAGAATG AGATAATTGG GGAGAAGGAG GACAGCGATC TGTTCGTCAA AACCATTTAC AGGTACATCA AGAAAAAGTT CATCGACAAT AACTTCTACT ATAAACGGGG AAATAACTAC ATTTCAATCA ATGATAAGTT CCCGCTCGAT TCTAATACAA ACGTTAATGC GCACTTGACA TATAAGATTA AACTGTACAA GATAAACGAA CGGTATTACA TTAGCGTGCT TCCAAAATTC ACCTTCCTCA GTGACAAGCC AGCCCTTGAG AGCCCCATCA AGAGCACCTA CCTGTTCAAC ATTAAAAGCG GCAAGACGTT TCCCTATATT AGCGGGCTCA ACGGAGTCCT GAAAATTGAC CTGGGCGAGA ACGGCATAAA GGAGGTCCTT TTTCCGGAGA ACTACTATTT CAACTTTACC TCCAAGGAGG CCGAGAAGTT TGGGTTTTCT AAGGAAATCC ATAACATCTA CAAGGAAAAA ATCTTCAGCG GCTACAAGAA AATCAAACAG AGCTTGTATT TCCTCGAGGA CATCATCAAT ATAAACAATT ACAACCTTAC CATGGACAAA AAGATCTATG TGAACATAGA ATACGAGTTC AAAAAGGGCA TCAGCAGAAA CATAAAAGAC GTGTTCAAAT ACAGCTTTTA CAAAAATGAC CAGAAGATCA AAATTGCGTT CTTTTTTAGC AGCAAGAAGC AAATCTATGA GATTCAACGC AGCTTGAAGA TGCTGTTCCA GAACAAGAAT AGCATATTCT ACCAGACCAT

CTACGAGATG GGGTTCAGCA AGGTGATTTT TCTCCGCGAG CCTAAGACTA ACAGCAGCGC ATTTATGTAT AACCCCGAAA CCTTCGAGAT TAGCAACAAA GATTTCTTTG AAAACCTGGA GGGGAACATT ATGGCAATCA TTATACTCGA CAAGTTTCTG GGCAATATCG ACAGTCTTAT CCAAAAATTC CCTGAGAACC TCATCCTTCA ACCCATACTC AAAGAGAAAC TGGAAAAGAT TCAGCCGTAT ATCATTAAGT CCTACGTCTA TAAAATGGGA AACTTTATTC CAGAGTGCCA ACCATACGTC ATAAGGAACC TGAAGGACAA GAACAAAACC CTCTACATCG GCATCGACCT GTCCCACGAC AACTATCTCA AGAAGTCTAA CCTCGCCATC AGCGCCGTAA ACAACTTCGG TGACATTATC TACCTGAACA AGTATAAGAA CCTTGAGTTG AACGAGAAGA TGAACCTCGA TATAGTCGAG AAAGAGTACA TACAGATCCT CAACGAGTAC TACGAGCGCA ATAAGAATTA CCCCGAAAAC ATCATTGTTT TGCGGGACGG ACGCTATCTC GAGGACATAG AGATCATAAA GAACATACTG AACATTGAGA ACATCAAGTA CAGCCTCATC GAAGTTAACA AGTCCGTGAA TATCAACTCC TGCGAGGACC TTAAAGAGTG GATTATCAAG CTTAGCGACA ACAATTTCAT ATACTATCCC AAAACGTACT TTAACCAGAA AGGTGTAGAG ATAAAGATAA TAGAGAACAA TACCGACTAC AATAATGAGA AAATACTGGA ACAGGTGTAC TCACTGACGA GAGTGGTGCA TCCCACCCCC TACGTAAACT ACCGCTTGCC CTACCCCCTG CAAGTCGTCA ACAAGGTCGC CCTTACCGAG TTGGAATGGA AGCTTTATAT CCCTTACATG AAATAGTGA

79 Gene Name: gene-18719

Gene ID: BBF.2017.48.79

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAGCGGAA GGTGGAAGAT CCGAAGAAAA AGAGGAAGGT TGGCAGCGGG
AGCATGACTG AGGACTTGTA CCTCGACTAC GACGCGTTCC TGCGGAGCTT TAAAAGAAAC
ATAGATGTGC CGCACTCCTT TCTCCTGGGA GCTGGTACAT CCATTAGCAG TGGCATCCAG
ACCGCCTACG ATTGTATCTG GGAGTGGAAA AAGGACATTT ACCTCTCCAA GAACATCAAC
GCCGCTGAGT TCTATAAGAA CCATAAGGAC GAGGCGGTAA GAAAGAGCAT CCAAAAGTGG
CTGGATAACC AAGGTGAATA CCCAGTTCTC GACAGCACGG AGGAGTATTG CTTTTATGCC
GAAAAGGCCT ATCCCATCCC CGAGGACCGC CGCAAGTATT TTCTGTCTCT TATCGAAAAT
AAGGAGCCCT ACATAGGGTA TAAGCTCCTC TGTCTGCTGG CCGAGCGCAG CATTGTAAAG
GCTGTCTGGA CTACTAATTT CGATGGCTTG ACCGTCAGGG CTGCTCATCA GAACAAGTTG
ACGCCCATTG AGATAACCCT CGATAACTCT GATAGAATAT TTCGCAACCA GTCTACCAAG
GAATTGCTCA CAATTGCGCT GCATGGTGAC TACAAATTCT CTACGCTGAA AAATACGGAG
AAGGAGCTCG ACAACCAGAA CGACACATTC AAACAGCAGC TGGGGACGTA TCACGTGGAC
AAGAATATGA TCGTAATAGG CTACTCAGGG CGCGACAAGA GCCTCATGGA CGCCATCAGC
GAGGCCTTCA GTACGCGGGG TGCAGGGAGG CTTTATTGGT GCGGCTATGG CGAAACGATC
CCCAACGAGG TTAGCGAGCT CATACTGAAA ATCAGGTCCC AGGGTAGAGA TGCATACTAC
ATATCAACGG ATGGATTTGA CAAAACGCTG ATACACCTGT CTAAAAGTGC GTTCGAGGAC
AACCCCGAGA TTACGAAAAA CATCCAACTC GCGCTCGAAA ACAGCGCGGA CGAAGAGTAC

TTTAAGACTG ACTTTTCACT GAACTTTAGC AAGCCGGATA AGTTCATCAA GTCAAACCTC
CACCCCATCG TGTTCCCGAA AGAAATCTTT CAATTCGAGC TTGACTTCAA GGAGGACAAG
CCTTGGCAAC TCCTCAAAAC TATTTCACGC GAGACAAACA TTTGCGCCGT GCCGTTCAAG
GGTAAGGTGT TCGCACTGGG CACGCTTACT GACATTGGGA ACGTGTTCAA GAACCGCCTG
AAGAGTGATA TAAAGCGCGA AGCAATTAGC ACCTCCGACG TGGATAATGT GAGTGCCTTT
AAATCTCTGA TGCTGCAGGC TGTGCTGAAG TTTTTCATTG GTATCGAAGG CGTGGAGTCC
AACCTCAAAG ACAGATTGTG GCTTACCAAC GCGGAGCAGC TCGTGGGTGA TATTAGTGTG
CATAAGGCTA TCCACCTCAG CCTGTACTTC GACAAAAACA AAGGATTCGC TTACCTGTCC
TTCACCCCCA CCGTACAACT CATCTCTCCT GAGGAAATCA GCAAAATCCA GAAGCAGAGA
ATCTCTAAGA GTAAACTCGA GAAGCTGTTC AATGACAAGT ATGACGAGAT ATTGGAGTTC
TGGAACCAAA AGCTCTTTAA CAATAGCCAA ATCAAGTTCG AGTACCCGAT CAGCTCAGGT
AGTGGGTTTG AGTTCAAAAT CTCCGCCAAC ACCGCATTTG GGGAGATAAA CGTATTGGAC
CCCAACTTTC GCTCCTTTTC CCCTAGAAAT TATGACCCGA AGCGCACACA GTTTAAGGGC
GTGCAGTTCC TCGAACCGCA GCTGATATTC CGCAACATCA GTACTAATGT GGAATTTAAG
GACTACCACC CGATGAGGGG GCTGGTGAAC AACCGACCGT TCGACGTGAA CCTGAACGGT
ATAATTCATT CTAACGAAAT AAACCTCACG GTCATCTGCG GCAAGTCATA CGCCAACGAC
CTGTATGAAT TCCTGAGCAA GCTCCAAGTG AAGCACGCCA CTGAGAATGT CAACCCGGAC
TATCTTATTG AGTATCCGGG CTTCCAAAGT GTGTTCAATC TGCCACTCAA CATACCCCAC
TTTGACTCTT CCGAGAAGTG GTACGACATC GACTTCGTAG CTGACAATAA CGGGGAGAAC
CACGAGAATG CCATTAAGCT TGCCAGACTC ATCACCACCA AGATCGACCA GATTGCCTCT
ACACAGAACC AGAGCACGGT CGTGGTGTTT ATTCCAAATG AATGGCAGTT GTTTGAGGGG
TACCTGAATC AGGGGGAGAG TTTCGATTTG CACGATTACA TCAAGGCATT CAGCGCTAGT
AGGGGCATTT CAACGCAGCT CATCCGCGAG GATACACTGG CGGATACGTT GAAGTGCCAG
ATCTACTGGT GGCTGAGCCT CTCATTTTAC GTTAAAAGCC TGCGAACTCC TTGGATTCTG
AATAATCAAG AAAAGAACAC GGCCTACGCC GGGATCGGTT ATAGCGTGAC TAAAATACAG
GACCGGACGG AAACGGTGAT CGGCTGTTCC CATATTTACG ATTCCAACGG CCAGGGGCTC
AAGTATCGGT TGAGTAAAAT TGACGACTAC TTCCTTGACA ATCGCAATAA TCCATTTCTT
AGCTATAAGG ATGCGTTCCA ATTCGGTGTG TCCATACGGG AATTGTTTTA CCAGTCCCTG
GACAAATTGC CTGAGCGGGT AGTTATACAC AAGCGGACCC GATTTACCGA TGATGAGATC
AATGGTATTA AGGCATCTCT GAACAAGGCG GGGATTAAGA AGATTGACCT GGTGGAGATT
AACTACGAAA CGGACGCCCG CTTCGTGGCC ATGTCCGTAT ACCAGAATGC ACTTCAGGTA
GACCGATTCC CTATCAGTCG GGGTACTTGT ATAGTCACAA ATAAGTACAC TGCCCTTTTG
TGGACGCACG GGATTGTCCC AAGTGTACGG CAGCCAAACT ACAAGTTCTA CCTTGGCGGT
AGAAGCATAC CGGCTCCGAT CAAGATCACA AAGCATTATG GTGATAGTAA TATAGACGTT
ATCGCCACCG AAATCCTTGG GCTGACCAAA ATGAACTGGA ACTCCCTTGA CCTTTATAGC

AAACTTCCCT CTACGATCGA CTCCAGCAAT CAGATCGCTC GGATTGGCAA ACTGCTCTCC CGGTACGAAG GCAAGACGTA CGACTATCGA TTGTTTATCT AGTGA

80 Gene Name: gene-351

Gene ID: BBF.2017.48.80

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AGAAAAGGAA AGTGGAGGAC CCCAAGAAAA AGCGCAAGGT TGGCAGCGGG
TCCCTGGAGA ACCTCACCAT AAACATAATC CCCTTCAAGC ACCCCAGCAT CCAAAAAGAA
TTTGGCTTCT ATACCGAGAA GAAGGAGGGC TATTTCCCCA TTCATAGGAC CGAGTTGCCC
AACGAGCTGT GGGACAACCA GAAAGAGGAA GTGGTGAAGC ACAAGTTCTA CTACACGAAC
TTTGAGGACA CGGAGGATTG CGTTCTCAAG ACCAAGGTGG ACCTGTATAG TAGCACTAAG
TTTGCCAAGC ATCTGTACAC GCGATTGGTG TACCAGTATT TCATTGGGAT AGCGGATGCA
ATCCAGTTCA ACTACGTGGG TGACATAGAG GTTTGGCTGC TGGATGCGAA AGCCAGCACC
ACCAAATACA ATAGCTACAA CAAGTATACC CTGAAAATAG AGTTTAGCGG TCTGACCAAG AGC-
CCCGCTC TCCTCCTCAG CTATGACAAC ACTAGTAAGG TAGCGACTAC GAGCATAGAC GAAAT-
CAACA TTCCCACCGA GTACTTCAAG ACCGTCGTGT ATAACAAAGA AATCCAGAGG TTCAAG-
TACC TGACCGAGGA CGCGAAACAA CACCTCGATC AAGTGTATCC CCTGCTCAAC ATAC-
CGTTGA AAAACCATCT TGAGATTCCT CACACCGTTC CCCGCAAGGG CAACAGGTAT AAGC-
CCTACT TTAACCACAT TACGACTTTT TACAATAACT ATTTGAACAC CGACGAATTC AGGGC-
CATCC TGCCCCTTGA TGAGAATGGA TTCTTCAATA TCCCAGAGGA CAGCATTTTG AAAAC-
TAGCA AAAATTCTAA CAACCTCCGG TTCTATAAGA AAGTCGGAGT AGATCCCAAG GCTG-
GAATGA AGAAGCCCGG TCCCTACAAG GCCAGCCCCC ACGACAACGT GAACCTGTTC TT-
TATCTATC ACAAACCCGA CGCACATGAA TACGCCAAAA CGTTGCATGA CTACTTCATG GAGGGGTACA

AAAAGTTCTT TCCTCCCCTC AAGAACGTTA TCCGGCAGCC GCTGTTCCTG GACAAAGGCA
CCTCACTTGC ATTTGAGAGC TTCGACAGCT GTATCGCCGA GCTGAAAACC CATCTGTTCG
ACCTCAAGAA AAAGCCCAAT ACCCGGTACG TGGCCATCTA CGTGAGCCCC ATCCATAAGG
AGGACGAGGA CAATAAACAC CTGTACTACC AGGTCAAAGA GGAGCTGCTT AAACATGACA
TCACCAGCCA GGTGATTTAC AAAGAGTCCA TCAAAGATAA ATACTTCGGC GCTTTCCTCG
AGAATATCGC ACCAGCTTTG CTTGCAAAGA TCGACGGCAT TCCCTGGCGA CTGGACAGGG
AGTTGAAACA GGAACTGATC GTAGGCGTCG GCGCCTATAA AAGCAGCGTC ACCAACACAA
GGTTCGTTGG AAGCGCCTTT TGCTTTAACA ACAAAGGAGA GTTCAAGAGC TTTGACTGCT
TCAGGGAGAA GGAATTCGAT CTGATTGCCG GGAAAATCGG CAAACAGGTG CTCACCTTCA
TTGAGGAGAA CGAGAACAAG TTGGAGAGGC TGATCATCCA TTATTTCAAG CCTTTCAACA
AGGATGAGAT AGATCTCGTG CAGGAAACCC TCGGCCTGCT GAAGCTGGAA ATCCCCATCA
TCATCGTGAC TATCAATAAG ACCGAGAGCT CCGATTACGT CGCTTTTGAC ACCAACGACG
ACGCCCTGAT GCCCCTGAGC GGCACCATTA TCGAGATAGC ACATCTGAAG TATCTGCTGT
TCAATAACGC GAAGTACAGC AGCATCGGCT TCGCCAAAGA CCACCCCTTC CCCGTTAAGC
TCAGTCTGTA CTGCACCGAC CAGGATTACT TCGAGGATAT CGCCATCGTC AAGGAGCTCA
TAGATCAGGT TTATCAGTTT TCTAGGATGT ACTGGAAAAG CGTCAAGCAG CAAAATCTGC
CCGTGACAAT CAAATACCCC GAGATGGTGG CCCAAATCTT CCCACACTTT GAGGGCGATA
AACTGCCTGA TTTTGGAAAA AACAATCTCT GGTTTCTGTA GTGA

81 Gene Name: gene-4660

Gene ID: BBF.2017.48.81

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAGAGGAA GGTTGAAGAT CCCAAGAAGA AACGAAAGGT GGGGAGCGGC
AGCGTGAGGC TGGTAAACCA GAAAGAGAAA CCGGAAGGCG ACTACGTGTA TGGCTACACT
CTCCCAATAG ACCCCAGTAA CAGGAACATG AGGCAGCCCT TCTGGATAAG CATGGATAAA
AAGGAGGGCT ATGAAGCTCA TTTCGTTGGC CCCTATGAGA ACATTGAGTT GACCAAGAGC
GTGATCTTCT GGGACCTTCT GAGGAGGACC AGGGAGCAAC TCAGCAGCGA TAAGTTCACG
GAATCAAGAA AAAAGTTCTT TAAGGAGATC TACTTCCCCC TTAACCTCTA CAATGAGGGC
AGCCAAGGGC TCGCCGTGCA ACCCTACTAC CTGAAGATTG ATCAGCAATT TGGACTGCTG
GTGGATTTTC AATTCAAACT TGACAAAGAT TTCACCTTCA GCCGGAAGAT TCAACAGCTC
AGTCTGACAT TGGATGGGAA GAACCGGAGG AACCTCAACT ACTACGTCGA CAGGATAACC
AAAACCAACC AATTCATCAA GGCCCTCTGG AACATCATTG GCACCTTCTC CCATAATGAA
AACAAGGAAA ACTACACGCT GAGGAACGAC TTCTACCCCT GCGCCGCAAG CAGGCTGCGA
TCTCGAATGT ATCTCTTTTC CAATGGCAGT GAATCCAGGA GCCAGTTCAA TGGCTTGAAG
GAATACGGCC CACTCCGACC CCTGACAGCC AATCCGACAC TGCTGTTTGT GTTCCGGGAA
CAAGACCGCG ACGCCGCGAG AAAACTGGCT ATGGCACTTA AAGGCAGCAA AAAGCAAGAT
CAATACAGCT TCCCCGGGTT CAACTCCCTG TTTAAAGCGG ACCTGTTGAT CGACGGAAAT
CCCATGGTCT TGAAAGACTT TTCTATCGAG AGCAGCAGGG AGGTGTTGGC CAGGGTGACA
ACATCAACAT CCAGCTTGTT GCCCATTTTC ATCCTGCCCA ACCGCGAGGG CGACGGCTAC

CTGGAGCACA AAGCCATCTT CGCCGAGAAC GGCATACCTA CTCAAGCGTG CACACTCCAA
GTCATTCAGG ACGACGTGAC CCTTAGGTGG AGCGTCCCCA ATATCGCCCT GCAAATATTC
TGCAAAGCGG GTGGCTGGCC CTGGAAAGTG CAGAGCCCCG TAACCGACAA CGCCCTGATT
ATAGGCATAA GTCAGAGCCA CAAGTTGAAT TATAGTGACG GTAAGACAAC TGTGGACAAG
CACTTCGCTT TTAGCGTGCT GACTGATTCA AGCGGCCTCT TTCAGAAAAT TCAGGTGCTG
AGCGAGCAAA AGACGGAGGA AACCTACTTC GAACAACTGA AGCTGAATCT CAAAAGCATC
CTGAACGCCA ATAGCAAGAA CTACCAACGC ATCGTGATCC ACACCTCATT TAAGCTCAAA
TACAAAGAAA TAAGTGCAAT CGAGGAAGTT GTTAGCGAAT TTGCAAGGAA CAGCAACAGC
GCCGACTGCA AGTTCGCCGT TGTGAAGGTT AATCACAAGC ATAGGTACTT CGGGTTTAAT
CGGGAAGTGA ATAGCTTGGT GCCCTACGAG GGAACCGTGT GTAAGCTGGG CGATAGAGAG
TACCTGGTCT GGTTCGAGGG TATCTATCAG GAGAAGCCGA CCGTTACCAA AGCATTTCCG
GGTCCCACCC ACATCGAATT TCTTAAAATC GGGTCTAATA ACGTGATTAG CGACGACCTT
TTGTTGCAAG ACCTGATGAA CTTGAGCGGA GCGAACTGGA GAGGCTTTAA TGCGAAGAGT
GCTCCGGTAT CCATCTTTTA CTGCCACCTG GTGGCCGACA TCGTGCATGA TTTCCAAATC
AAAGGCCTCC CTATGCCCGC CATAGATCTT ATACGACCCT GGTTCATCTA GTGA

82 Gene Name: gene-5460

Gene ID: BBF.2017.48.82

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AAAAACGAAA AGTAGAGGAC CCTAAAAAGA AGCGGAAAGT AGGGTCAGGC
TCTATGCTTC AACTGAACGG CTTTAGCATC GAAATCGCCG GAGGTTCCCT GACTGTCTTG
AAATCTAAAA TCGCGCCTAC CGACGTTAAA GAAACCCGCA GGAGCCTGGA GGACGACTGG
TTCACCATGT ATCACGAGGG CCACTTGTAC TCACTTGCAA AAAACAGCAA CGCATCCGGC
GGATTGGGTG AAACCGAGCT CCTGGTCCTG TCTGATCATC TGGGTCTTAG GTTCGTTAAG
GCTATGTTGG ACCAAGCCAT GAGGGGCGTA TTCGAGGCCT ACGACCCCGT TAGAGATAGG
CCCTTCACAT TTCTGGCGCG AAACGTAGAT CTCGTAGCCC TCGCGGCAGA AAACCTCGAG
TCCAAGCCCA GCCTTCTCTC CAAATTCGAG ATCAGGCCCA AGTACGAACT GGAGGCCAAG
GTAGTGGAAT TCAGACCGGG CGAGCTGGAA CTTATGCTGG CGCTCAATCT GACTACACGG
TGGATCTGCA ACGCCTCCGT AGACGAGCTC ATTGAGAAGA ACATACCGGT CCGAGGAATG
CACCTGATCC GACGGAACCG GGAGCCGGGA CAGAGAAGCT TGGTTGGCAC CTTCGACCGC
ATGGAAGGCG ACAACGCCCT GCTGCAGGAT GCTTACGACG GACAAGACAA GATAGCAGCC
TCACAGGTGA GGATCGAGGG GAGCAAGGAA GTTTTCGCGA CCTCTCTGAG GAGGCTCTTG
GGCAATCGCT ATACCAGTTT CATGCACTCC GTGGATAACG AGTACGGCAA GTTGTGCGGG
GGTTTGGGGT TCGACGGCGA ACTTAGGAAG ATGCAGGGAT TTCTCGCGAA AAAGAGTCCT
ATACAACTGC ACGGAGGTGT AGAAGTGTCC GTGGGGCAGA GGGTACAACT TACCAATCAG
CCTGGGTATA AGACAACAGT TGAGCTTTTG CAGTCAAAGT ACTGCTTTGA CAGAAGTAGG

ACGAAGCTCC ACCCCTACGC CTGGGACGGG CTTGCTCGAT TCGGCCCATT CGACAGGGGC
AGCTTCCCGA CGCGATCCCC CAGGATTCTG CTCGTGACAC CCGACTCCGC GAGCGGTAAA
GTCTCTCAAG CTCTGAAGAA ATTCCGCGAC GGGTTCGGCA GCAGCCAGAG CAGCATGTAT
GACGGCTTCC TCGACACCTT TCACCTCAGT AATGCTCCTT TCTTCCCCCT TCCCGTGAAG
CTGGACGGCG TGCAGCGCAG CGACGTGGGC AAAGCTTATC GAAAGGCGAT CGAAGATAAA
CTCGCACGCG ACGACGACTT CGACGCCGCC TTTAACATTC TCCTGGACGA GCACGCCAAT
CTGCCGGACA GCCATAACCC CTATCTGGTC GCCAAGTCCA TCCTCCTCTC CCACGGCATC
CCAGTGCAAG AAGCACGAGT GAGCACTCTG ACGGCCAACG AATACAGCCT GCAACACACC
TTCAGGAATG TCGCCACAGC CCTGTACGCC AAAATGGGTG GTGTCCCATG GACCGTTGAC
CACGGGGAAA CCGTGGACGA TGAGCTGGTA GTAGGAATCG GAAACGCGGA GCTTAGCGGG
TCTAGGTTCG AGAAAAGACA GAGGCACATC GGAATCACGA CAGTGTTTAG GGGGGACGGC
AACTATCTGC TTAGCAACCT CAGCAAAGAG TGCCGATACG AGGATTACCC GGACGTACTC
CGGGAGAGTA CGATCGCCGT GTTGAGGGAG GTTAAGCAAA GGAACAATTG GTTGCCGGGT
CAAACCGTGC GAATCGTTTT CCACGCCTTC AAGCCTCTGA AAAACGTGGA GATTGCCGAC
ATTATCGCGA GCTCTGTAAA GGAGGTAGGC TCCGAACAGA CCATAGAATT TGCATTCTTG
AATGTTTCCC TCGACCACTC CTTCACCCTT CTGGACATGG CTCAAAGGGG AATAACGAAG
AAGAATCAGA CCAAGGGGAT ATACGTTCCC AGGAGGGGCA TGACAGTCCA GGTTGGGCGC
TACACCAGGC TTGTAACCAG CATCGGTCCG CACATGGTAA AAAGGGCAAA CCTTGCCCTC
CCGCGACCCC TGTTGATTCA TCTGCACAAG CAGAGCACCT ATCGGGACCT GAGCTATCTG
AGCGAACAGG TTCTGAACTT TACCACCCTG TCCTGGAGGA GCACCCTCCC CAGCGAGAAG
CCTGTTACCA TTCTCTACTC ATCACTGATA GCCGACTTGT TGGGAAGGCT CAAGTCAGTG
GATGATTGGA GCCCCGCAGT GTTGAATACC AAACTGAGGA ATAGCAAATG GTTCCTGTAG
TGA

83 Gene Name: gene-841

Gene ID: BBF.2017.48.83

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AAAAGAGGAA GGTTGAGGAC CCCAAAAAGA AACGCAAAGT GGGCAGCGGA
AGCATGTCCG GCCTTTTCCT GAACTTTTAC CAGGTAGACA TCCCCACCAA ATCCGTACCG
ATCCACAGCG TAGAGTATAG CCATTACAGT TCAAAGGAGG CCTTTATCGC GTTGAAAGAA
AACTTCCCCT ACTTTAGCTT CTACCGGGAT GACGACCGAA TACTGATCTG GAAGAAAGAC
AAGGATGCCG AGCTCCCCGA GAAGAACTCA TTGATTGAAA TTGATTTCAC CGAGAAAGCG
AAGGTCCTCA GCAAAATACT CGAGAGGGCC ATCATTGACT TCATCGAGCC AAAGGGCTAC
AAGATATTCA AGAACAAGTA CAGCAACAGC TGGGAAATAG TGAGCATGAA GGACATCCTG
AATGGTGGGA TCGAGGGACT CAGCATCAAT CGAATCGTGC ATTTTTCCCC CTGCTTCTTC
TTCAAGGAGA ACAAACTCAT GCTGGGTTTC AGCCTTAGCA CAAGCCTCAA AAACGTGTTT
ACCTGGAATA AGGCGGACTT CGAAAGGTAC GGCTTTGACA TCAAGGGCCT TAAAGGCGAC
GAGGAGCGGA TTTTTGCCAA CAAGCAATCC CTTAAGAGGT TCCTGGAAAC CAAGGGCGCA
GTTGCAATGT ATGACCAAAT TATCGCAAAG GAAAACAAGA ACGCGAAAAT GTTTAGCATC
ATCGACGGCT TCTATCGGTG GCTGGAGAGG AACAAGACTG AAATCCAGCT TCCATTCGGA
CTGAAGATAA ATTCAGTGTC TAAAAAGTAT CTGCCGTTCG AGGATGAGCT GATCAAGAGC
GAGATCATCC CTAAGCCCCA AAGGTATTTC TATAGCAATA GGAAGAACAC CCAGAGCCTG
CGGTACTATG ACGAGATGGT TAAGACTTAT CAGCCCTACT CTCTGGAGCT CTACCAAAAC
AAACAGATCA ACATCGGAAT CATCTGCCCC AGCGAGTACC AGGGAGAAAC GGAGGGGTTC

ATAAAGAAGA TCGAACTGAA GCTCAAGGAA GTATTCCATT TCAACAGCCT GATCTTTCAC
TTCAAGACCA TTACGAACAA GGACCTCGCG TCCTATAAGG AGGTTTTGTA CGACGATGAA
CTGCTGAAGT GCGACCTGAT TTACGTCATC GTGAATGAGG CCCAGGAGAA ACTCTCACCT
AATAACTCCC CTTACTACGT GTGCAAGGCC AAGTTTATAG GCAATGGCAT ACCTACGCAA
GACATTCAGA TTGAAACCAT CCGGCAGAAC TTGAATGCGT TCACAATGAC GAACATCTCA
CTTAACAGCT ACGCCAAACT GGGAGGCACC GCGTGGACCA TCGAGAAGGA GGACAAACTT
AAGGACGAGC TGGTCATTGG CATCGGCTCC ACCCTGTCAG AAAACGGCCA GTTCGTGCTC
GGTATCGCAC AAATCTTCCA TAATGACGGG CGCTACATGG CGGGTGACTG CAGCCCCCTT
TCTACCTTCT CCAACTACGC GGAGAACCTG GAGGATCACC TGTACAAGAC CCTGAAGCCC
CTGGTGGAGG AGATGAGCAA AAGCGGCACC TTCCGGCTGA TTTTCCACTT GTTTAAAAGT
GCCTCTGAGG AGTACGAGAT ACGCGCGATC AACGGCCTGC AGAAGAGGCT GGCGAACTAC
AATTTCGAAT TTGCACTCGT TCACCTGGCC TATGGACACA ACTTCCGACT CTACTACAAC
GACGGCAACG GCGACATTAA TCAGGGCACA TATATACAAC TGTCAAAACA CAGCGCCCTG
CTCCACTTCG TTAGCAAGTC AGACTTGCCC CTGAAAATCG ACCTGGACAA GCGGTCTACT
TTCACCAGCC TGTTTTATAT CGCCAAACAG GTGTACTGGT TCAGCCATCT GAGTCACCGC
AGCTATATGC CCAGTAAGAG GACCGTGACC ATCATGTATC CGTCAATCAT GGCGAAGATG
ACCGAGGAGC TTAAGAAGGT GGAAGGATGG GACTACGAGC GCCTGAAAGC AGTAAGCGAT
AAGCTGTGGT TCATCTAGTG A

84 Gene Name: gene-136

Gene ID: BBF.2017.48.84

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AGAAAAGGAA GGTGGAGGAC CCAAAGAAGA AACGGAAAGT TGGCAGCGGC
TCCATGAGCG TGGCGATCGT GAGCCCCCAA ATGTACAAGA GTCTGAGCGA GGTGTTTCCT
CTGACCGCCT CCCAACTGAA CTTTATGTGC TTTAGGCTGA CTCCCGAAAT CGAAAAGAAG
GATGGTAATA GGCTCAGCTA CCATTTCAGT CTGAAGCTGC CGGAAACTGT TGTGATCTGG
CACCAGCCCT ACTTCTGGGT GTTGGCGAGT AGTAACAGGC AAATCCCCAA TAAGGACGAG
TTGCAAGAAA CTCTGATAAG GATCCAAAAC GAGGTGGATG ACTTCAAAGA ACGACTCTTC
GGTTTCCAGA GCGTTCGCCA CCCCCAACTC ACCCCCTTTA TCATCAGCCT CTTCGCCGTT
CAGGTCCTCA AGAAAACAAA GTTCGACTAC CCCATTGCAT TCAGCAACAA CGGTGTAATC
GTCAGGAGGG AGCCCGACTT TTGGACGGAG AGCATAGAGC TTCAAGACAG CCTGCATCCT
GCCCTCACGC TGACCGTAAG TTCATCAATA GTGTTCCGCG ACAACCTCGC GGAGTTCTAT
GAAAAACATC ATCAAAGGGA GAAGCCCGAG CAGTTTCTGA TCGGCCTGAA GGTGCAGGAA
ATAGAGAGGG GCAACAATGC GATCATCGTG GGACTCGTCG GCACCATCGG CGAGCACCGG
GACCAGCTGC TTGAAAAAGC AACCGGGAGC ACTAGCAAGC AGGCGCTGCG AGAGGCACCG
GACAACCAGC CGGTGGTTGC GATACAGTTC GGCAAGGATA CGAAGCAGTT CTACTACGCA
ATGGCCGCGT TGCGGCCGTG CGTAACCTCA GAAACGGCAA ACCAGTTCGA GGTAGAGTAC
GGTAAGCTCC TGAAAGCTAC AAAGATAAGC CACCAGGAGC GAACCAATCT GCTGGCCTCA
TACAAAAAGA CGGCCCAGGA GTCATTGGCC GCTTATGGCA TCCGCCTGGA GCTGAGTGTG

AATAGCAGGG ATTACCCCAG CTTCTTCTGG CAACCCCCCG TGAAGATCGA AGATACCAAA
CTTCTGTTTG GCAACGGCAT AACCGGCAAG CGGACTGAGG TGCTCAAGGG GCTTTCTATA
GGGGGCGTGT ACCGACGCCA CGGGAAATTC CAGGACAAGT CAAAAGTGAT CCAGATCGCG
GCTCTTAAGC TTTGCGACGT GACCGTTAGC TTGTTCCTGA AGCAACTTAC TCAAAGGCTG
GCAAAATACG GCTTCCGAAG CGAGATAATC ACCAAGAAGC CTCTGTCAAT CAAGAACCTT
GCCACCGCCG AAGCCAGGGC TGCTGTTGAG AAAGCGGTCA ATGAGCTCGT GGAAATACCC
CACGACATCG TGCTTGCCTT CCTGCCTGAG TCCGACAGGC ACACCGACGA CACGGATGAG
GGTTCCTTCT ATCACCAGAT CTACTCCCTT CTCCTCCGGA GACAAATAGC CTCACAAATT
ATCTACGAGG ACACCCTGTC CAACTCTGGG AACTACCAGT ACATCCTGAA CCAGGTCATT
CCGGGGATCT TGGCGAAACT CGGGAATCTG CCCTTCATTT TGGCGGAAAG CCTCGATATA
GCGGACCACT TCATCGGACT TGACATCAGC AGAATCTCTA AGAAAACACA GGTCGGGACA
CGAAACGCGT GCGCCAGCGT GCGACTTTAC GGACGCCAGG GTGAATTTAT CCGCTACCGG
CTTGAGGACG ACCTGATCGA CGGCGAGGCG ATTCCACCCA AGCTGCTGGA AAGGTTGCTG
CCTGCGACCG AGCTTGCGAA TAAAACCATA CTGATCTACA GGGACGGGAG CTTCGTGGGC
AAAGAGGCCG ACTATCTTGT GGAGCGAGCC AAGGCGATAG ACGCGAAGTT TATCCTCGTC
GAGTGTAAGA AATCCGGCGT GCCGCGCTTG TATAACTTGG AGCAAAAGAC CGTGATCGCG
CCGAGTCAGG GACTGGCACT TCGACTGAGC AGTAGGGAAG CAATACTCGT GACCACCAAG
GTGCCCGATA AAGTGGGCCT GGCTAGACCC ATCCGGCTCA CAATCCACGA AAAGGGCCAT
CAAGTAAGCA TCGAATCCGT GCTGGACACT ACACTCAAGC TTACTCTTCT TCACCATGGC
GCGCTGAAAG AACCGCGACT GCCCATGCCC CTGTATGGGA GCGACAGGAT GGCATACCTC
CGGCTGCAGG GGATACGGCC TAGCGTTATG GAGGGCGACC GCCAATTCTG GCTGTAGTGA

85 Gene Name: gene-14123

Gene ID: BBF.2017.48.85

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGAGAAA GGTGGAGGAC CCAAAGAAGA AACGGAAAGT TGGCTCTGGG
TCAATGAACC TGACCGTAAA CCTCGCCCCC ATCAGCGTGC AGGGCGACTG CTCAGTCCTG
ATTGGCAGAC AGCGCTACGA CGAGCAGAGG CTGGCTGAAC TTAGGTCAGA CTTTCGGGGC
ACCCACGTGT TTCGGAGGGA CGGTCCAGAT AGCATGATTG ACATCCCCGT GGTCCCCGAC
GCGGCACCTC TGGGCAACCT GAGGGAAACG ATCGACCTTA GGCGGTACCA GCGGCTGTGG
CCCATGCTTC TGCAGGAGTC CCTCATCCAG CTGCTTGGTA AGCGCCCCAT CCAGTCCAGC
AAGCCCTTGA AGTTCCTGGG AGCTCGCTCT CCTCTGATCG AGCACCCGGA TCTCCCTGAG
TGGTTGAGGC GGGTGAGCGT TACCGAGATC CACACCCGAC ACATCACCGT GGACGGCAAG
CAAATCTACG GTATCGTGTG TGATGTGAGG GCCAAGTCTT TTATCCTCGC CACATGCAGC
GAACTTCTGA AATTCGGCGT GACCATCCTT GGTAGATACG TCCAAATAGA ACAGCCCGCG
ATAGACGAGA GAACCATGCC TAAAAGGAAG CTCATCGGCA GGGTAAGGTC CATCCAAGGG
GATGATCTGC TTCTTGACGA CTGTGAGGCC GGCTTCGAAA AAGTCGCTGC GAATGAGGCA
TTTCTCGAGC CGCGGAAGGA AAATTTCGAG GACTGCGTGC GCCAGGTGCT GAAGCGGGAC
GCCGAGAGGG TGTTGGAGAG GTCAGCTCGC GCCAGCCAAA ACCTGGCCGC AGGCCCTGGG
AAACTGGAAC ACATCGACGG AATCATCAGG TATCTTAGGG AGAAGAAGCC CGCAGCGGTG
CCCGGCTGCC ATTTCGTGAT CGATGCCATG CTCAACACAA ACGGCCACAT TTTTCCACCC
GGGGAAACAA TGGACAAACC CTTCCTCTTG TTCGACCCTA GCGGTTCACG GAGAGAGGAC

TGGCCCGAGA AGGGCCTTAA AGATCACGGC CCCTATGATG AACAGGTGTT TTCCCCCAAG
TCCCTGAAGA TCGCTGTTGT GTGCCAAAGC CGGTTGGAGG GCAGAGTGGA CGAGTTTCTG
GCGAAGTTTC TCAATGGGAT GCCGAAGGTC TTTCAACCCG GCAAGAGCTT CGCCCGCTAC
GGCGACGGAT TCGTGAAACG ATTCAGACTG AACAAGCCCG AGGTGCACTT CTTTCTTGCA
GATGGCAACT CCGACGAGGC ATACGCCGTG GCCAGCCGCG AGGCACTCGA TAAAGCGAGG
GATAGCGGGT TCGAGTGGGA CCTGGCGATT GTGCAAATTG AGGAGGAGTT CAAGTCACTG
GCCGACGGCT CCAATCCCTA CTACACCACT AAGAGCATCT TCTTGCGGAG GGACGTTCCG
GTGCAGAGCG TCAGGCTGGA AACCATGAGC CTGTCAGATA ATGACCTGGT GTTCCCCATG
AACCACCTGA GCCTCGCTAC CTACGCCAAG CTGGGGGGCA CGCCCTGGCT CCTGGCTAGC
TCACAAACCG TGGCGCACGA ACTGGTGATC GGACTGGGTA GCAGCACCAG CTCCGAATCA
AGGCTGGGCA GCCAGATGAG ACATGTGGGA ATCACCACCG TGTTCAGCAG TGACGGCAGC
TATCTGCTTT CTGATAGAAC CGCCGCAGTG CCCTTCGAGC AGTACCCACA AGAGTTGAGG
AAAACGTTGC GAAAAACAAT CGAGGCCGTC AGGGCCGAGG ACAATTGGCG GAGTAGCGAC
AAGGTGAGGT TGGTATTCCA TTCATTCAAG CCGTTCAAGG ACAGCGAGGT AGAAGCCATA
GAGGCGCTGA CCACCGACCT GGGCCTGGGC GACGTGAAGG CCGCCTTTCT GCACATTGCG
CCCGACCACC CGTTCCTTAT CTTCGACCAC GACCAAATGG GGATCGCCGC ACGAGGGGGC
AAAAAAGGCG TGTTGGGCCC TGCTAGGCAG TTGCACATCC GGCTTAGCGA CGCTGAGAGC
CTTGTGGTGT TCGCAGGGGC CAGCGAGCTT AAACAGGTGA CGGATGGTAT GCCGCGACCC
GCGCTGCTCA AGCTGCACCC CAAAAGCACC TTCAAAGATA TGACCTACCT GGCAAGGCAG
GCCTTTGCCT TTAGTGCCCA TAGCTGGCGG ATGCTGTCCC CCGAACCTTT CCCAATTACT
ATCCGCTACA GCGACCTGAT CGCCGACCGC CTGGCGGGAC TCGCGTCTGT TAAGGGCTGG
GACCCCGATG CCGTGACGTT CGGCGCTATC GGTCACAAGC CTTGGTTCTT GTAGTGA

86 Gene Name: gene-16842

Gene ID: BBF.2017.48.86

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGCGAAA AGTAGAGGAT CCAAAGAAAA AGCGGAAGGT CGGGAGCGGC
TCCATGGCGT TTAGGCCCGG TGAACGAGTC AGACCGCAGC TCGCGCTGAA TGCGATCAGG
GTCCTTACAC CCCCTGGCAC CATCCCCGCC AGTGTAGTCC AATTCGACAG AGCGCTGCTG
CACGCATATC TTGACAGACC CGAGAACGAC GTATTCGCTA CCCGACACGG GGAGACTGAT
ATGGCGGTCG TACCCCTGAC CAGCGGTGCG AATCTGCCAA CGGACAGAAT GGGGCTTCCA
GCTGCAGAGC ACCTCAGGCT GGTATCTGCG CTGACAAGAG AAGCTGTGTT TCGCCTCCTC
GCGGCCAGCC CGGAAGCGGA TCTGCTGATC CGGCGACGCC CACCGACCGT CGCGGGGAAG
AGAGAAAACG TACTTGCAGA GGACATTGGG CTCCCGGACT GGTTGAAGAA AAGACTTGTG
CTGGAGTTCG ACACGCGCAT ATTGCAACCA CCGAGAGGGG ACGCCTACGT GGTGCTGACG
TGTAGTAAAA GGCTGCGCAC GACAATAGAC GCGAGTTGTC GCACCCTTCT GGAACTCGGT
GTACCACTGA CGGGTGCCGC AGTCAGCTCC TGGAGGGAAG ATCCTGACCC CAAGGTGAGC
CGGCGATTGG CCTACGCTGG GCGCGTTGTA GAAGTAGGGC AGGACACGCT CACTCTGGAC
GACCACGGAG CTGGTCCGAG TGTTGTCTCC AGCGAGGACG TGTTCCTCGA GCCGACTCGA
GCAAACTTCA ACAAGGTGGT GGAAGTGATA ACCCAGGGTA ACTCCGAACG AGCCTTCAAG
GCCGTACAAA AAGCAGAAGC CGAATGGCAC GGCGGGAGGC GGACAATCGA AATAGTGCAT
GGTGTCCTCA ACCAACTCGG CAACCGGTCA ATGGTTCTTG CCGATGGCGT GCCTCTGCGG
CTCGGGGGCT TGATAGACCA AGCGGTCGAT AGCGACGCAT TCCCCCCAGC CGAGGCGGTG

TGGCGCCCTA AGCTCTCATT CGACCCCGTG CACAGCCCCG AGACATCAAA TTCCTGGAAA
CAGCAGTCAC TGGACAGGAC GGGCCCTTTC GATAGGCAAA CCTTTGAAAC AAAAAGACCG
CGAATCGCGG TTGTCCATCA GGCCGGAAGA AGGGAGGAAG TGGCTGCGGC TATGCGCGAT
TTCCTCCACG GAAGGCCTGA TATCGCCAGC GATACGGGCC TGGTTCCCCA CGGTTCAGGA
CTCCTCGGAC GCTTTAGGCT CCACGAACCC GAAGTGAGAT ACTTTGAGGC CGCAGGCCGC
GGGGGACCCG CTTATGCCGA CGCAGCACGG AGTGCGCTCA GGGACGCGGC GTCAAGGGAC
GAACCATGGG ACCTCGCAAT GGTACAGGTA GAGCGGGCGT GGCAAGATCG CCCACATGCC
GATAGCCCGT ACTGGATGAG CAAGGCAACG TTTCTCAAGA GGGATGTGCC GGTGCAAGCC
CTTAGCACAG AAATGTTGGG TCTTGATGCA TTTGGGTACG CGAACGCACT TGCGAACATG
TCACTTGCAA CGTATGCGAA ACTGGGCGGT GCCCCGTGGC TTTTGTTTGC CAGGTCACCA
ACCGACCATG AACTGGTGGT CGGGCTCGGA AGCCACACTG TAAAAGAGGG CCGAAGGGGT
GCGGGTGAGA GGTTTGTCGG TATCGCGACC GTATTCAGCA GCCAGGGCCA TTATTTCTTG
GATGCCAGGA CAGCCGCGGT CCCGTTTGAA GCCTATCCTG CTGCCTTGAG CGACAGCATC
GTTGACGCGA TCAAAAGGAT TGGACGAGAG GAAGCCTGGC GACCAGGCGA GGCCGTCAGG
TTGGTCTTTC ACGCCTTCAC CCAGTTGAGC CGAGAAACCG TTCAGGCAGT GGAGAGAGCA
GTAGCAGGCA TCGGGGCCAC CAACGTAAGC TTCGCGTTTC TGCACGTTGT CGAAGATCAC
CCGTTTACCA TGTTTGACCG AGCGTGGCCA GACGGAAAGG CGACATTCGC CCCTGAAAGA
GGTCAGGCGC TTCGACTCTC CGAGCGCGAA TGGTTGTTGA CACTTACCGG CAGGCGCGAA
GTTAAGAGCG CCAGTCACGG GCTGCCTGGG CCGGTTCTGT TGCGACTTCA TGACAGCAGC
ACCTATAGAG ACATGCCCGT GCTCGTCCGA CAAGCATCCG ACTTCGCCTT CCACTCTTGG
CGCAGTTTTG GACCCAGCGG ACTCCCCATC CCGTTGGTTT ACGCGGACGA AATTGCAAAA
CAGCTCAGCG GCTTGGAAAG AACCCCCGGA TGGGACACGG ATGCGGCTGA GGGTGGCCGG
GTTATGAGAA AGCCTTGGTT TCTGTAGTGA

87 Gene Name: gene-18839

Gene ID: BBF.2017.48.87

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAGCGCAA GGTTGAGGAC CCGAAAAAGA AGAGGAAGGT CGGCAGCGGG
AGCATGCAGC TGAACTACTT CCCCATAAAG TTTGAGTTTG AAGAGTACCA GATAAAAACT
GAGCCCTACA GCGAAGAACG ACTTAAAGAG TTGAGGGCCA GTTACAACGC CACCCACTCC
TTTTTTAGAA ATGGAGACAA TATATGCATT AGCAACAAGG AAGGCGAGGA CATTAGTCTG
ACCGGCGAGG TGATACCGAA AAGAATTTTC GACGACAGTC AAGTGACCGC CTCATTGATA
AAGCACTTGT TTTTCAGGAC GTTCAAGGAG AGGTTCCCCA ACTATATTCC TGTGGACTTT
TACCCCTTCC GCTTCTTCTC CGCCCAGGCT AAAGACGACA TCATCTATAA CGCCCTGCCC
GGCAACCTCC GGAAACGAAT CGCTTACAAA AAGCTGATCG AGGTTCAGTT GCGGCTGACG
GAAATAAACG GCATCAAGCA GTTTGGCTTC CTGATCAACA TTAAACGAAA TTGGGTGTTC
AACAAGTCAT GCTTCGAGCT CCACTCCGAG GGCTACAACC TGATCGGGGT GGACGTGCTG
TACGCCGAGG AACTGCCGGG GTTGACCGAG GTGCTGGCCC CAAACGAGGA GCTTTTGGGC
GTAATCGCGG AAATCGTGGA CGACAATGCC AGGATAGAAA CCAACGAGGG CATTAAGGAG
TTCCCTCTGA ACCAGTTGTT CATCAAGAAA AGCAAGTACA ACATTGGCAA TTACCTTAGC
TTCGCGATCT CTCAGCAAAA GAGCGACGAA ATAATGAATC TTATCGAGAG CAAACGCTCC
GACATCTACA ATACCAAGGG TCTTTACGAC GAGATCTTGA AAATTGCGAA CCATCTTTTT
TGCGAGAACA GCGCACCCAT ACTGTTTCAT AATAAGGACG GATTCTGCTT TACTGTCGAT
TCCCAGCCGC TCAGTGTGAC GAACAGCATG GAACTTAAGA CTCCAACATT CATATACGAT

CCAGCGGCCA CCAAGACGAA TTCTAGCAAT CCCGACTTGG GCCTGTCCAA TTACGGGCCC
TACGACTCCA GCATTTTTGA CATAAAGATA CCCAACGTGT TGTGCATCTG CAATAGGAAT
AATCGAGGCA ACTTTACAAA GTTTCTGTCT AACCTGAAAG ACGGGATACC TCAAAGCCGC
TATTTCCAGA AAGGCCTCCA GAAGAAATAC GACCTCCAGG ATGTGATCCT CAATATCCGA
GAAATCCAGG CCTATAGTAT CGCCGACTAC CTTAACGCCA TCAGGGACTA CGATGAGAAC
AAGCCTCATC TGGCGATCAT CGAGATCCCT GCCAGCTTCA AGAGGCAGGC CGACGTGGCG
AACCCCTACT ACCAAATTAA GGCCAAGTTG TTGAGCCTGG AGATTCCCGT GCAATTCGTT
ACCAGCGAAA CCATCGGTAA CCACAACGAG TATATCCTGA ACTCTATCGC GCTGCAGATC
TACGCAAAGC TCGGCGGGAC CCCGTGGGTC CTGCCCTCTC AACGCAGCGT TGACAAAGAG
ATAATCATCG GAATAGGCCA TTCCTGGCTT AGGCGCAACC AGTACGCTGG CGCAGAACAG
AATAGGGTAG TGGGGATCAC GACCTTTATG AGCTCCGATG GCCAGTACCT TCTGGGTGAC
AAGGTCAAAG ATGTTGCCTT CGAGAACTAT TTTGAGGAGC TTCTGAAAAG CCTGAAGCAA
AGCATCCAGA GGCTCAGCAC AGAGCAGGGC TGGTCAGATG GCGACACCGT GAGGCTGATA
TTCCACATAT TCAAACCGAT AAAGAACACT GAATTCGACG TGATCAGTCA GCTTGTCAGA
GACATCACGC AGTACAAGAT TAAGTTCGCA TTCGTAACCA TCAGCACTGT GCACCCTTCC
ATGTTGTTCG ACATTAATCA GTCCGGTATC GCCAAATACG GTTCCAATAT CATGAAGGGA
CAATACATAC CAAACAGGGG CAGCAACGTT TTCCTGGACG AAAAGACATG CATCGTACAG
ATGTTCGGCG CGAACGAACT GAAAACGGCC AAGCAAGGCA TGAGCAAGCC CATCCTTATA
AACATTCGCA CCCCCCAGGG GAACTACAAT TCAAGCGACC TGAACGATCT CCTGTTTTAT
GACCTGGGGT ACATCACACA ACAGATATTT AGCTTTACCT ACCTCAGCTG GCGGTCCTTC
TTGCCCGGTG AGGAGCCGGC GACTATGAAG TACAGTAACC TCATTTCCAA ACTTCTCGGG
AAGATGCGGA ACATCCCTAA CTGGGACGCC GACAATCTTA ACTACGGCCT GAAACGGAAA
AAGTGGTTCC TGTAGTGA

88 Gene Name: gene-18998

Gene ID: BBF.2017.48.88

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGAGGAA AGTGGAGGAT CCAAAGAAAA AGAGAAAGGT GGGTAGCGGA
AGCATGACCG AGGCCTTCCT CACAACCAGG AGGGGCTTCG TGCAAAAGCT GACGCTGACC
AGGTACGATT ACCTGAACTG GATCATCGAG TCCGAGGCGC AGAAAGCCAA GCTGAAGAAC
TGGCTTAAGA ACAAGAGCGG GTTTCTGACC CACGAGATCG AGGATACCTG TTTCTTCACC
TTCGAGAGGC TTCTGGAGGA GAGTACTAAG CAGTATAGAG CCTCCGGCGA GAAAACTCTG
TCTGCCCCGT TCAAGAACAC GCAACTGATC TCAAATCTGA TCGGTACCAT ATTGAAAAAG
GAGTTGAGCA AGAAATACAA GCAATTCTTT AGTCAAAACA TCTTCATCGT GAGCACCATC
GATCTGTATC CATTCAATCT CTTGAAGGCG TTCGAGTTCA ACATCGAAGT GTTTGACAGC
GGCCACTTCC TTATCCACGT CAACCCAGTG TCTAAAATTG TAAGCAGCAA GGTTGTGGAC
AAGGAGTATC TGGACTACCT CAAGAAAAGC AACCTCAACA ACAGCAAAAC CACCGAGATG
GAGTTCGCGG TGATCAACCA TGAAAGGAAT TTCAGACTTA AATTCGATCT GCTTGACGAA
TGCATCTTTG AGAAGATAGA GAAGCTGCAC AGCGAGAAGA ATATGTTTAC AGCCACTTTT
GATTACCATT TCCTGGCCAA CTTCAGCCCC GAGATCTTCG GCAAAATCGT GGAACATACT
AGCAAGGATC TGAAGCAGGC CATCATGTTC CTGAATGACA TACTGAGCAA TATCAAGCTG
CCGAGCTTTC TCAATCTGCA CGAGGAACGA TACTTTAAGG TCAATATCTC CGAATTGGAC
CGAAAGAATA ATCTTCTGAT TGGAAGCAGT TTCGAGGTAA TAACCATATA CTCAAAAAGC
CAGACCCAGT ATGGACTGAG GATTGAGTTC ACTCGCGACA GCATAAGCCG GGACGAGCTT

ATAACAATCT TTCTGAAAAA CGAGGAGCTG ATCGAGAAAC TCAACGACAT TAAAGTGGTC
CCCGCCACCA TCAACGCAAA AATCGAACAA AAGACCGGCT GGAAAAACCC CTACATCACC
AATGTTTTCA TCGATAACGT GGGTGCCTTC AGCACCAGCA GCCTGCAAAG CGCCTCATAC
TTCCACGGCA TCTACAAGGC CGTTAACAAC TGGAATATCC TGCCCATCGT GTACGAGGAC
CTCGACATCA AAGTATTCGA GAACCTGATG CTGCACGCCT TTAACAAGAA CGCCACCGAA
TTCAAGATCC TGGAACCCAT CATAATCAAG TCCACGAACG AAATCGACAA ACAGGAGGTG
CAGAGGAGCA TCAAAAACCA GGCCGGCAAG ACCATGATCG CAGTGTTCTG CAAGTACAAG
ATACCCCATG ACAGCTTCGC CCCCCTCAAG GGCTTCAAGT ATCAGATCTA TCAAGGCGAC
ACCACGGACA ATAAGCAGAA TAGGGCCAAA CTGAGTAACT TCACGTGCAA GTGCCTGGAG
AAAATGGGAG GGGTGATTGC GGCAATCGCG GACACAAGCA TAGCCGAGGA TGGATATTTC
ATTGGCATCG ACCTTGGCCA CACCACAAAT GGCAAGGAAA AGTTCTCCAA CCTCGGAGTG
AGCTTGTTTG ATAGCCTGGG CATCCTGTTG GGCGATTACG TGGAGAAGGA GATTCCAAGA
AGGGAAAACC TCATCGACAC GAACTGCCTC AATGCTTTTA AGAAACTTGA CAAAATGCTG
GAAGCTAAAA AACTGAACAA GCCCAAACAC CTGATCATCC ATCGGGACGG CAAACTGCAC
TTCAAGGATA TCAACATTCT CGTAAGCTGC GTGGAAACCG TGTGGGGTAA GATAAACGTC
GATATAGTCG AGATCATTAA GAGTGGCTTC CCCGTGATGG CTATAAAGGA CGAAACCAAC
AAACCAATCA ATCCCATAAG CGGGACCAGC TACCAGGACG ACATCCATAA GTACGCCATA
CTCGCCACAA ACGTACAAGC CGACGAACAG TCAGCCGTAA TAAACCCGAT AATCATAAAA
CACAAATACG GAGAGCTGGA GTTTAGCAAA ATAGTTGAAC AGGTGTACTG GTTCACGAAA
GTGTATACCA ATAACCTGTA CAATAGTACC AGGCTCCCAG CGACTACACT CAAGGCCAAC
AACGTGGTTG GCACGTCTAA GAAGCTCCAC AGAAGTACAT ACTTGGGCTA GTGA

89 Gene Name: gene-19803

Gene ID: BBF.2017.48.89

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAACGGAA GGTGGAGGAC CCTAAGAAAA AACGAAAGGT CGGAAGTGGC AGCGTTCCAG TGTACCTTAA TCGGTTCCTG CTGGACCACC TCACATCACC CTTGTCCTTG CCGGCGTTTC GGGTCGAACT GGACCCTCCC CCTTCCAAAG ATGAAGTGCA CCCGCTCCTG GCTCTCGTCG GTCGGGAAGC GGGAGGGCTC GTGAGGTTCC AGAACAGGCT GATCGGCTGG GAGGCTCCAC GGGCCCTCGA AGGTCAGGTT AGGCGAGGCA AGCAGTCATA TAGACTGGTG CCCCTTGGCC GGCAGGCACT CAATCTTAGA AAACCCGAAG AAAGGCAGGC GCTCGAGAAT TTGTATAGGA TCCGACTGGA AAACATCTTG AAAGCCCTCG CCAAACGACA TAGGGCTAGA GTCGAACGCA GGGGCAACGG CCTTTTTCTG TGGAGGCCAG AGAATCCCCG AGAGGAGAAG GAGGGGTGGC ACCTTTACCG GGGAAGCCTG TACCGCATAC ATCTCTATCC TGACGGCGAA GTGATACTTG AAGTCGACGT GCAGCATCGA TTTCAACCCA CTCTCCATCT CGAGGAGTGG CTGCAACGAG GCTATCCACT CCCTAGGCGC GTGACTAACG CCTACGAGGA CGAGAAAGAA TGGGCACTCC TGGGCATCGA AGAGGGGAAG GATCCCCGCT CTTTTCTCTT GGATGGGGGC GAGTCATTGC TTGACTACCA CCGCAAGAAG GGACGATTGG CAGAGGGGCA GGACCCCGGT CGAGTGGTCT GGGTTGCTAG AGGTAAAGAA CGCGAGCGGA TCCCACATCT GAGCGTCTTG TTGAAGCCAG TCATCACCAT GGAGCTGCTG GCGGAAGTCG CTGAGGTCAC GCAGGAG-GCC TTGCCTGCGC TTCAGCTCGA ACCCGAGGAA CGGCTGAAGG ACATTAGGCG CTTCGCT-GAA CCTGTACTGC AAGCGTTCGG CAAACGCGAA ACTGCAAAAC CCCTTGAAGG CAGAGC-

CCAG CGATTGCCGC GACCCAGTTT GTTGGCACGG GGAAAAAAGC GAGTGGGCAA AG-
TAGCGGAC GTACTCGAAA AGGGAGCATT GTCACCGGGC GAGACACGGT TGGCCCTGCT
CGCATGGGAG GGCGACGGGA AGGCCAAAGG CGGGCTCGCG TACTTGGAGG AGAGGCTTCA
GGGCGTCGGG TCTGCATCCG GCATCAAACT TGAACTTAAA CGGCGATTTC TGCCCCGAGG
CGATAACCTC GAAATGGCAC AGGTGTTTGA GGAGCTCTCC CAGGAAGGAG TAGGTGCCGG
TCTGCTTCTG ACTCCGCGCC TCACAGAAGG GGAAAGACGC GAACTGAAAA ATACTGCGGC
GAGCCATGGG CTCGCTCTCC AACTCCTTAA CCCGTTTGAC CCTGGCGACA TCTACAGGGT
GAATAACGCT CTGCTTGGAT TTCTCGCGAA GGCCGGGTGG CTGTTCCTGA GACTGGAGGG
AACTTATCCG GCCGACCTGG TGGTGGCCTA TGACGCAGGC GGGGAGAGTC TCCGATTCGG
CGGAGCCTGC TTCGCCCACC TGACTGATGG CACGCATCTG GGGTTCAGTC TGCCAGCCGC
TCAGGGTGGT GAACGGATGG CCGAGGAGGT CGCGTGGGAG TTGCTGCGAC CCCTGCTGTT
GAGATACCGG AAAGCGAAGG GCCAGACACC AGGGAGGATC TTTCTGCTCC GCGACGGTAA
GATTCAAAAG GAAGAGTTCC GAAAAGTGGA AGAGGAACTG AGAAAGCGCA ATATTCCCTA
CGCGCTGTTT AGCGTCAGAA AGACGGGGGC TCCCCGACTG TTCAGCAAAA ATGGGCCGCT
CGGTGACGGT TTGTTTTTGC GACTGCCAGA GGAGGAGGGC GGGTTTCTGT TGCTTAGCGC
CGAGGGTGGG AAGGGCACCC CACGGCCGGT TAAGTATGTG TTGGAGGCGG GAGAAGTGGA
CCTCAACCTG GAGGAAGCTG CCAGGCAATT GTATCACCTG AGTCGCATCT ACCCGGGCTC
CGGTTACCGA TTCCCCAGGC TGCCCGCACC GTTGCATATG GTTGATAGGA TGGTGAGGGA
GGTTGCACGG CTCGGCGGCA GCCATAACTT GAGACTCAAA GAAGAACAAC TGTTTTTCCT
GTAGTGA

90 Gene Name: gene-20968

Gene ID: BBF.2017.48.90

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AGAAACGAAA GGTGGAGGAC CCAAAAAAGA AGCGGAAAGT GGGGAGTGGC AGCATGTTCG TGGAACTGAA CGCCTTCCCC ATCGACATCC GCAATATCGG TATCGTGGAG GCCTGCGAGG TGCCGTACGA CAAGGAGGTG CTTTATAGCC TGCATGATAA CCCACAAAAA GATTACCATG CTATCAGAAA CGGCAACCAG ATATTGATAT TTTCTAATAG CAAAAACTAC CCCATCCAGG GTACAATCAA GGAGATAAAT CTTGCACAGG ACTACCGCAT CCTGTTTTTC CTTATTAAGG AGTCCATTAT CAAGATCCTG ACGCAGATCA AACGGGAGCC TTTCAAGTTC AACCCGATTG AGTTCATCTC ACCAAAGGAG AACATCACCG AGAATATCCT GGGAATCAAT TACCCATTTC AAATAAACGC CAAATATTCA ATCGATACCA GAATCATTCA GGGGGTGCCC TGCCTCACCA TTGATTGCAG CACGAAGAAA TACAACAAGG AATCCCTGAT CTACTTCATT AACGACGGCT TCAACCTGAT TAACAGGTAC GTGATCTCAA AGCAAAACGA GAAGTATAAG CGCGTAGGTA AGATACTGAG CATTGACAAC AACATCGTGA CTGTTCAGAG CTGCGACAAG ATAAAGAAGT ACTCCGCCGA GGAAATCACC TTGGAGGCGA ACTCTAAGAA CACCAAGGAC TATCTGGCAT ACAAGTTCCC CTATAAGTTC GAGCAGATCC AAGAAAGCAT TAAGAAGGCG ATCAGTACCT TCACCCAGGG GACCTCTAAG CAGATAAACA TTGGCAAGAT CTGGGACTTT TTCAGCCAGA AAGGCATCTT CCTGTTCAAC GGCCACCGAA TTAACATAGG GCTGCCTCCC GACATCTCCC AGCAATGCAA GAACCTTGTG TACCCGCGCT TTTTCTTTAG CAACTCCCGA GAAAACAATT CCAAAGAGAA CGGCCTGAAG GATTATGGCC CTTACACCAG GAATTACTTT GACAGGAATA

ACCCCAGCAT TTGCGTGATT TGCAACGCTA AGGAACAAGG CAAAGTGGAA CAGTTCCTGC
ACAAATTTCT GAAGGGCATA CCCAATAGCC ATAACTTTAA GACGGGCTTC GAGGGCAAGT
TTCATATTGG CCTCTCTCAG ATAGAATTTT TCACGACCAG CGACGACAGC CTGGGCAGCT
ACCAGTTGGC TATCCAGAAG GCAATCCAAA CGAGGACTAA CCAAAACTCT AGCCAGTGGG
ACCTGGCCCT GGTGCAAACC AGGCAGTCCT TCAAGAAATT GTTGGTGGAG CAGAATCCGT
ACTTTATTAG CAAGAAAATG TTCTTTCAGC ATCAGATCCC CGTTCAAGAC TTCACCATCG
AGCTGACCAA TCAGAACGAC AAAAACCTGG AGTATTCTCT GAATAACATG GCTCTGGCGT
GCTATGCGAA GATGAATGGA AAGCCCTGGC TGCTTAAATC AAGCCCTACT ATCAGTCATG
AGCTGGTTAT TGGCATCGGG AGCAGCAACA TCATCATCGA GGAGGACAGT CTGAACCAGA
GGATCATGGG CATCACCACC GTGTTCAGCG GCGACGGGTC TTACATGGTA TCAAACACTA
GCAAGGCGGT GGCGCCCAAT GAGTACTGTT GCGCCCTCAT AGACACACTT GAGCAAACGA
TCAAGAAGCT GGAGAAACTT ATGAACTGGC AGAGCAATGA CACCATTAGG CTCATCTTTC
ATGCCGCCGT AAAGACCTTC AACAAAAATG AAATCCTCGC CGTAAAGGAA GTGATCAAAA
AGTATAGTGA GTACAAGATC GAGTACGCTT TTCTCAAAAT CAGCAGCGAC CACGGTCTGC
ACCTGTTCGA CCACTCAACT AAGAATGAGA ATAAGGGTAA ATTGGCTCCC AAGAGGGGTA
AGTATTTTGA ACTGAGTAGC CATGAAATTT TGCTGTACCT CGTGGGGCAG AAAGAGCTGA
AACAGGTGAG TGATGGCCAC CCCCAGGGCG TGATCGTGTC CTGCATAAG GACAGCAGCT
TTCAGGACCT TAAGTACCTC TCTAATCAGA TTTTCAGTTT TAGCTCCCAC AGTTGGAGGA
GCTACTTTCC CTCTCCCCTG CCCGTGACAA TTCATTATAG CGATTTGATC GCGGAGAACC
TGGGCTGGCT TAACAAGCTG AGCGGCTGGG ACGATACAAT CCTGCTGGGC AAACTTGGAC
AGACCCAGTG GTTTCTGTAG TGA

91 Gene Name: gene-2579

Gene ID: BBF.2017.48.91

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGAGAAA GGTCGAGGAC CCGAAGAAGA AGCGAAAGGT AGGAAGCGGT
AGCATGAAAA GCAACTTCTT CCCCATCCAG TTCAACTTCG ACGACTTCCA TATCCAGAGG
CTTCCCTACC AGAAGGAGGT GCTGGACAAG CTTCGGCAAC AACACAATGC GACCCATAGC
TTTTTCCGCA GGGACGATTT TATCTATATT AGCCCAGGGG TAGAGGCCGC AGCGAACCTG
GGCGACGTAG TACGCCTCTC TATTACCAAG CACCCCGAGG TCGTTGCTTC TCTTGTTAGG
CACATATTCT TTAGGACAAT CAAGGATAAG GTCCCCGGTC TGCTGCCAAG CTTTCACCCA
TTCACCTTTC CCGCCAAACA GGACAAATAC GATCTGGCCC TGAACATGCT CCCCGAGCGC
CTGCAGAATG TTATCACCTA CAAGAGGATA ACCGAGGTAC AGCTTCGATT CAACGAAACC
GAGGAGCAAC CCCAGTTCGT CGCCGTAGTT AACCACAGGT ACCAGTGGAC TATCGACCGA
ACTTGCGAGC AATTGGTAAA CGAGGGTCTG GACATCCTTG GCCTGGAGGT GAACTCTAGT
ACGAGCCCTG ATTATTCAGA CGGAGTTGTG GCACCAGAGC TGACACTGTT GGGCAGGGTG
ATGGCCGTGA ACGGGGATCA CGCCACAGTA GGGACCAACC AGGGTCCGAC AGAGTATGCC
CTGTTCGAAT TGACCTTGTT CAAGTCCAAG GAGAACATAG TGAACTACCT TGGATCTTTG
GTGGGCGAGG GTAAAGCCGA ACAAATAGTC AACCATATCA AACAAGATGA AGCAGAAGG
CTGCAACCGG ACGTTGTGAT GAGGGAGATC GAGGAAATGG GAGTGTGGCT GTCTAGGCTG
GCCTACAGAA ACTTTGACTC CTTTTGCTTC ACCATCGGAA CGAACAACGC TGTCAGCGGC
CAAGCTGGTA TCAGACTGGA GGAGCCAAAG CTGATATTTG ATGTCTCAGG TACGAACATA

CACGCTACCC CCACAACCGG GCTCAACACC TTCGGCCCCT ATAGTAGAAG CACGAGTTTC
GACGTTAACT CTCCGAAGAT TCTGGTTGTG TTTCACCAGC GGAACGCAGG CCACTTCGCA
GAGTTTCTCG CACAGCTGAA GGGCGGGATC GCTCAGCACG CATACTTTGC TAACGGGATG
GTCAGGAAGT ATGGCCTCAC GGCAATGGAG TACCGGATTG CCGAGATCAC TGACTACACC
GTGCCCCAAT ATCTTACCGC CATCAATAAG CTGCTTAGGG CGGAGAACGG AAGCTTTGAT
ATCGCCATCG TGGAAACCTG TGAGGATTTC CGGAGGCTGC CTCCCATGGA TAATCCGTAT
TTTCAGGTTA AGAGTTTGTT GTACAGCCAT GGAATCAGCA CCCAATTCAT CAGAGCGGAA
ACCGCTCAGA AACCGATTTA TTCAATAGAT AGTATCGCGC TCCAAATGTA CGCCAAATTG
GGCGGAACAC CATGGACGGT GCCAATAGGG CCGAGCGTAG ATCACGAATT GGTGATAGGC
ATCGGTAGCT CCATATTGCG CAGCAACCAG TATGCTGGTG CAACCCAAGC TCGAATAGTG
GGGATTTCTA CCTTCTTCAG CGCCGACGGG AAGTACATAA GCAATAGAAA GACCCAGGAC
GTGCCTTACG ATCAGTACTT CGATGAGCTC TTGCATAACC TTAAAGTCTC CATCGACGAG
ATTTCCAATA ACTACAGCTG GAGCTCAGGC GACCGCATCA GGATCATATT CCACATCTTC
AAGCCCATAA AACACATCGA GGCAGACGTC GTCGCAAGCC TGATGGAACA GTACCAGGAG
TTCGATATAA AGTTCGCTTT TGTGACCTTT AGCGAGTTCC ACCCGTATGT GCTGTTTAAT
GAAAATGAAA GGGGGGAATT TGATGCGTAT AGGAAGGTTT ACAAGGGCAC CCATGTACCG
TGGCGCGGTT ACAATGTTCT GCTGGATCCT CGGTCATGCC TGGTCCAGAT GCTGGGACCC
CATGAGATGA AAACCAGCCG GCACGGCGCT TCTAGGCCCG TCCTTGTGAG AATCCACCGC
AGTTCTACGT TTGTAGACCT CGCGTACGTC GTGCAACAGG CCTTTAAGTT TACTAGGCTC
TCATTCCGCA CGTTCTACCC TGTGCATAGC CCTGTGACGC TGCTCTACAG TAATATGTTG
GCCCGACAGC TCAAGGACCT GAGGGGCATT CCGGGTTGGA ACTACGATGT AGCTAGCAGG
CAGTTGAGGC ACAAGAAATG GTTCCTGTAG TGA

92 Gene Name: gene-4481

Gene ID: BBF.2017.48.92

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AGAAGCGAAA GGTCGAGGAT CCCAAAAAGA AACGGAAGGT TGGCTCCGGG
TCTATGGGCA GGCAACTCCA ACTGAACTTT ACCCCGCTCA GGGTTAGGGG CGACGCCATC
AGACTTCAGG CGCTGCCTTT CGAGGACGCT CAACAATTTA GGAATCTGAG AGATGAGCAT
CGAGCACACT ACGCTGTGAC GAGAAGGAGC GACCACATCG TGGCCCTCCC ACTTACACTG
AATGCCTCCC CAATCGGCGA GGAGAAGATC GTGAGCGTTG TGGAGCATGC GAGTTTGATT
CGGCCCCTGC TTGAACAGAG GTTGGTGACC CTTCTGTCCA GTAACCGGAG GCCGGTGGCC
CGGTATAATC CGATCACCAC CATTGGAAGA ACCTTGCCAA CGGGCTTCAT AGAAGCCGAC
CGACACCTCC ATTTGCAGTC CCGCGTGCTT ATTGCTATCC GCTCCCTCAA GCTGCCGGAC
GCCGAGCCCT TGGGATTGCT CTGGGACATC GAAATCCAGA AAACATGCGC GACTAGCCTT
GCCGTCCTGC ACGCACAAGG GGTACGGCTG GACGGGCTCA CAGTGGAACG GCTTGTCCCG
GTGGAGGACG TGCGAATGTT GCCTTATAGG CGACTGGTGG GCAGAGTAGG CGCGCTGACC
GATGGCCACG CCCGATTGAG CGAGCGGTTC CAGAACGTCG AAGAATTGCT GCCCCTGGAC
GAGCTTTACC TGGAGGCCAG TCCGGAGAAC CTGAGGCACC TTCTGCAGCA TTTCATGCGC
AACACAAGCG GGCGAGTGCA AGGGAAGATA GACGAGATCG TGTTCGAGAA CTCACGGGGA
CGCGCTCGGA TGGAGCACAT TGCCCGGATC TCCGACTGGC TTAGAGGCCT GGGCGAGATT
GAACTGCAGG AGGGTTTGTC TGTAGGCATC GGAAATCTGC TCTCTGAAAA GGACGCCCAG
AACTTTCCCA GGTTCACTGA GGGAACGACC CCAACCTACG TGTTTGACGC TGGGACGCTC

AAGAGCGAGT CAAGGGCCGC AGTGGGCCTC AGTAAATTCG GGCCCTACAG CCGGCATGTA
TTTACACCGA CTCGACCCAA CGTTTGCGTC ATCTGCGACC GCGCAAGAAG AGGACAGTTT
GAGCTGTTCC TGCGGAAATT CCGGGATGGC CTGACTGTTG ATGGGAAGTC CCTGCCGTTT
GGTCGCGGGT TTCTGGGAAT ATATGGCCTT CAGGATATCA ACCTGACCTT CGTCGAGGCG
GATGCATTCA CCGCGGACGC GTACCATGCT GCCGCAAGCA AGGCAGTACG GATGGGAGCC
GAGGGCGCAC CGTGGCACCT GGCACTCGTG CAAACAGAAC GCGACAGTCG GCAACTGGCT
CCCCCCAAGA ATCCGTATTT GGTAGCGAAG GCGGCGTTTC TGTCTAATCA AATTCCTACC
CAGTTTGTGG CGTTCGAGAC ATTTTCTATG GCGCCTCTGA ACCTCGCGTA CACACTGAGC
AACCTGGCGT TGGCGGTTTA TGCCAAGTTG GGCGGCATCC CATGGCTGAT CAAGAGTGAT
AAAGGTATAG CCCACGAGGT CGTCATCGGG TTGGGTAGTG CCGCGATCGG GGAGTCCCGA
TTCAGCCGGA AGGAGAGGAT TGTCGGCATC ACAAGTGTTT TTCGGGGTGA CGGCGGGTAC
CTCTTGTCTA ACCTGTCCAA TGCCGTGCCC ATGAGCAAGT ACGGCGAAGC ATTGACCGAA
TCTCTCCAGG CGACCCTGCA GAGGGTTCGC AATGAGATGA ACTGGATCAG GGGGGACAGC
GTTCGGGTCA TAGTTCACGC TTTCAAGCCA ATGAGGAACA CGGAGGTGGA GAGCGTTAAG
GCTGCGCTGA AAGAATTCAG CGAGTTCGAT CTGCAATTTG CTTTCCTTCA CGTTAAGCAA
GACCACCCGT ACCTCCTTTT TGACGACGAC AGCATCGGTA CAAAAGGGCG AGGCGAGAAA
ACCCCCGTGC GAGGCTTGTT CGCGGAGGTC GGACACAACG AGACACTGCT GACCCTGACC
GGACCACAGC AGCTGAAAAG ACCCACCGAC GGGCTGCCGA AACCGCTTCT GCTCAGCCTC
CATAGGGACT CTACTTTCAC AGATATAATC TACCTCACGA AACAGGTGTA CTGGTTTAGC
AATCACTCAT GGCGGTCTTT CCTGCCAGCA GCCATGCCGG TGACGATATA CTACAGCGAC
CTGGTGGCTG GTTTGCTCGG CAGACTGGAT AGGCTGGGCT CTCGCTGGTC ACCGAGTGTA
ATGCTGGGCA AGATCGGAAC CACAAGATGG TTCCTGTAGT GA

93 Gene Name: gene-5004

Gene ID: BBF.2017.48.93

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGAGAAA GGTGGAAGAT CCCAAGAAAA AGAGGAAGGT GGGTAGCGGG
AGCATGAGGG AAACCAACAT CTACGAGCTC AGCGGCCTCG AAACCGTGAG TACCAGCTAC
AGACTTTTCG AGTTGCAGGG CGCGCCAGAG TTCTCTCCTG AGTATTATGC TGGTGTGAGC
CGCCTCGTGA GGACGCTTAG CAGGAGACAC CAGGCACCCT TCACCAGTAT CCAACGGGGC
GAAACCATGT TGCTCGCTGC ACCCGAGGCC CTGAGCGGTG ATCTCGCAGA ACACCATAAT
CTGGCACGCT GGGTGGCGAC CCTGAAGTCA CTTGGAGATA GCATAGAGAT AGACTGCAGC
GTGAGCGGAG ATGAGCTGGA CCCCATAAGG CTGCGATTCC TGAACTTCAT GATCCAATCT
CCATTGTTCA ACCACGGCGA GCTCTGGCAG CCCAGGGCCG GTGATGCCTT CTACTACCGG
AAGCCTGCCG ACACGTTCGA CGGAATCGAA CTGTTTGAGG GTATTGCCGT GAGGGCCGTG
CCCTACCCAG GAGGCGGGTT CGGCGTTATG CTCGACGCGA GGACTAAGCT GATCTCACAG
CGGGCTGTGG GCGCCTACGC GGACCCGAAT TTCATAAGGA GGCTGAAAAA CACTAGCTGC
CTGTACCGAA TGGGAGACAT CTGGTACGAG ATAAAGATCA GTGGCGCGAA TCAGACCGTT
TCTCACCCCA TCCTGTTTAA GGACAACCAG CCCGTGTCAC TCAAAGCCTA TCTGCACGAA
CAAGCACGGC AGCCAATCCC CAAGTCTCTG ATTGATCTTA AAGGTGACGG CGTGGTGTTG
ACCTATCGCG GCAGCGATAG CGCCGAGGTC AAAGCGGCAC CCGCGGAACT TTGTTTCCCC
ATAGTAGACA CCCATAGCAA GAGGGGTGCC CGGCACCAGA GAAGGAGCAT CCAAGCCCCA
CACATCCGAC GCAGCAAGGC TTACCGATTC AAGCAAAGGT TCTTGCGGGA CATCAAAATA

GGAAATGCCG TGTTGAGCGT GGCCGACCAA CCCGCAGCCC TCAAGACCAG GCCCATCGAC
TTGCCCGAGC TGCAATTCGG CTCCAATAGG ATTCTGTACG GCACGGACAG GGGCGGGGAC
CGAATCGACC TTCGCCAGTA TGCCAAGAAT CGGCGAACGC TGCTGGAGCG CGCAGACGTG
GGCTTCTTTG AGACTTCTCC CCTGGAGCCC CAATGTTTGG TACTTCCTAA GAGCGTGATG
AACGCATGGG GCAACGAGTT CGTTAGGGAC CTGACTGCCG AAGTGAAGCG ACTCCACCCC
ACCGGTAACT ACAAGCCAAC CGTAATCGCG TTTGATGATG TCAGCGCAAC CGTGGACGCC
AGGAGCCAAG CAGAAGCCAT CTTCAAGCTC GCGGAGGACG GGGATCTCCC TCCAGGCGAC
TGCGCCATTA TGATACACCG AACCAAAGGA AAGGCAAGAG CGCAGGAGGA GCTGCCCGCA
CTTCTTATAA ACAAGCTGAG AAAGAGCTAC GGAGTGAATG CCGCCATATT CCACGCGACT
GTCCCCGGCA ACGCCTACCG AAGGGAAAGC GCCTCTGATG GCGCTCGCTA TGTGCGCAAG
CGGGATGAGA AGGGAAGGTT TAGTGGATAC CTGACCGGAG CGGCGCTTAA CAAGATTCTT
CTGCCCAACG CCAAGTGGCC CTTCGTGCTC AAGGACGAGT TGGTGGCAGA TATAGTGGTG
GGCATAGATG TGAAACATCA CACCGCAGCT CTCGTTTTGA TCGCCGAAGG CGGGAGGATT
ATCAGGCACA CTCTTCGCCT CAGCACCAAG AACGAGAAAC TCCCTGCTGG TATCGTGGAA
ACGAAGCTGG TGGAACTGAT TTCAAATGAA GCACCACACC TGAGCAGGCT CACCAAAACA
ATCGCCATCC ATAGGGACGG CAGGATTTGG CCCTCCGAGC TTAAGGGATT GCGAGCAGCC
TGTAGGAAGC TTGCCGACGA CGGCCACATC GATCCTGCGT TCGATCTGAA CGTGTTCGAG
GTGAGCAAAA GTGCCCCTGC TAGGCTTAGG CTGTTTAGCG TCGACCGCAG TGCTGGCAGA
AAGCCGAGGA TTGAAAACCC GGAACTGGGG GACTGGATGA TGCTGACAGA AACCGACGGC
TACGTTTGCA CGACCGGTGC TCCGCTGTTG AGAGGTGGTG CGGCTAGACC CCTGCATGTA
AAACAGGTCG CTGGTGATAT GAGCTTGCAG GACGCCCTTT CCGACGTGTT CCGACTGAGC
TGTCTGACCT GGACTAGGCC CGAGTCATGT TCTAGGTTGC CTATCAGTTT GAAGCTCTGC
GATATGCTGC TGATGGACGA GGGAACTGCC CACGACGAGG ACGAAATCCT TCATGCTAAC
GACGACACCC CAGCCGTTAG CGCCTAGTGA

94 Gene Name: gene-13471

Gene ID: BBF.2017.48.94

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGAGAAA GGTAGAGGAT CCCAAGAAGA AACGAAAAGT AGGCAGCGGC
AGTATGGTCG CGCTGAGGCT GAACGGCGTA CCCATCTTGT GCGCCGCTGA CGTAACCGTG
GCCGTGGCGA AGTTGCCGTA CACGAAGGAG AGCCTGGACG AGTTGAGGAA GGAGCATGCG
GGGAGGTATT TGATTAGGAG AGGCGGAGAT GACGGGCAGG AAATCATGTC TGTTCCCTTG
CTTGCTGATG CTCCGCAGCT GAGTGATGCC GTTGTGGAAG TTAAGCTGTC AGAAGCCCAC
TGGTTGCTCG CCTCACTCGC GGTGGAGGCC CTCACCAGGT TGTTCACAGA ACTTGGTAGA
CCTATCCTGC GGTCCCGGCC ATTGCGGCTG CTCTCCCAAA AGCCGGCCAA TCTTTTTCCG
GAGAACGTCG GACTGCCAGA CTGGCTGCAA AGGAGGGTTG TGCTGGATTT GGAGACTAGG
AAGATCTGGC GGCAGGATGG TGACCCGACA TTGGTGCTGC TGTGTGATGT GCGGACTCAA
AACTTTATCG ACGTGCCAAC GGATAAACTG ATGGCCACCG GCGTAAGCGT TATGGGTCGC
TACGTTAGCC GAATGGTGAG CTCTGATGAT CCCCGGATCA CCTCACATCT GAAGCTCGCC
GGCAGGGTCA TTAGCATAGA GGGCGACCGA CTGCTCCTCG CCGACTTTGG CGAGGGACCG
GATAGTATAA GCATTGCTCA TGCCTATCTG GAAAGACGAC GGGAAAATGT CGACTGGTGT
GTTCAACAGC TGAACCCCGC GAAAGCAGGG CAAATCCTGA TGAGCGTGCA GGCCGAGGCT
GCGAAATTCT TGAACGGACC TGGCCGATTC GAGCTGATCA AGAGGACATT CGATTATCTG
CGCACGCAGA GTATAGAGCT TGTGCCCGAC GTGAAGCTGG AGTTGGGGGA CTTGATTGGC
ATGGGAGCCG CACGCTGGCC CTTCCGCCAG GAAACAATTA AGAAGCCTAC CCTGGTGTTT

GATCCGTCTG GTGTCAAGAC CGATACCTGG AACGAGCGAG GGCTTGACAA ACACGGACCC
TACGACCAGA GGACCTTCAG CCCCAAGGAA ATGAGGATCG CCGTTATCTG CAGGGAAGCA
GACGAAGGTC GGGTTGAAGG ATTTCTGGCC AAGTTTCTGG ACGGGATGCC ACACGTTATC
GTCGGGGAGA ACCGAAAACC CTATGAAAAG GGATTCATAA GGAGGTTCGC CCTGAGTGCC
CCGAAGGTGC ACACTTTCAC CGCTAAATCT TCTAGTGTGC CGGACTACCT GAATGCGTGC
CGAGCGGCCC TGAAGTTTGC CCACGACCAA GGCTTTGAAT GGAGCTTGGC AATCGCGCAA
ATCGACAAGG ACTTTCGGGA ACTCCTCGGT CCTGACAATC CCTACTTCGC GATCAAGGCC
GCGTTTCTCA AGCAGAGGGT GCCCATCCAG GAGTTGACGC TCGAGACAAT GAGCACCCCC
GACAGGCAGC TGGTGTACAT TTTGAATAAC ATAAGCCTCG CAAGCTACGC CAAGATCGGC
GGCATTCCGT GGCTGCTTAA GAGCGGTCCT ACCGTGGGCC ACGAGCTGGT CATTGGTATT
GGTAGCCAGA CCGTTAGCAG TAGTCGATTG GGCGAGAAGC AACGGGTGGT GGGCATTACC
ACCGTATTCA CCCACGATGG CAGATACCTT TTGGACGACA GGACGCGAGC CGTGCCATAC
GGCGAGTACG AAGCAGCTTT GTCCGAAACG CTGACCAGGG CCATAGAGAG GGTAAGGACG
GAAGATAACT GGAGGTCAAC CGACGCGGTG CGACTTGTAT TCCACGTGTT CCAGCAAATC
AAAGACTACG AGGCCGACGC AGTGGGGAAA CTGGTCGAGA ATCTCGGCTT CTCAGATGTC
AAGTACGCCT TTGTGCATGT CGTTGACAGC CACCCCTACA CCCTGTTTGA CGAACACATG
CCAGGCGTTA AGTTTGGCTA CGAGATGAAG GGCGCCTACG CACCTGAGAG AGGCCTGTGC
ATCAGTCTTG GCAGGGACGA ACGCCTCCTC AGCTTTACCG GGTCTAGGGA GGTTAAACAA
ACCCATCATG GCCTCCCAAG GCCAACCCTT CTTCGACTGC ATAGGAACAG TACCTTCCGG
GACATGACCT ATATCGCCAG GCAGGCTTTC GACTTCGCAA ACCACTCATG GAGGATGCTC
ACCCCAGCGC CCCTCCCCAT CACCATCCAC TACGCCGAAC TGATCGCCCG GTTGTTGGCT
GGTCTGAAAG ACACACCCGG CTGGGACGAG GACACAATGC TCGGCCCAGT AGGTAGAACC
CGATGGTTTC TGTAGTGA

95 Gene Name: gene-21289

Gene ID: BBF.2017.48.95

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAACGCAA AGTAGAAGAT CCTAAAAAGA AGAGAAAGGT CGGCTCCGGG AGCATGGATT ACATACTTGA ATTCGACGAG TTTATTCGAA GCATCAAGCA GAATATTGAT ACAAAGTATT CATTCCTGTT GGGGGCTGGC GCTTCAGTCG AATCAGGTAT TCCGTGTGCC AGCGAATGCA TCTGGGAGTG GAAGAGGGAT ATCTTCATCA GCCAAAATCC GACCCTGGCT GAGATGCACA ACAACATCAA GAGCCAGAAC ATTAAGCGCA GCATCCAGAA CTGGCTCGAT AACCAGGGCA CCTACCCAAA GGAGGGCGAG GACATCGAGT ATTCCTACTA TATTGAGAAG GCTTTCCGGA TTCCCGACGA CCGGAGGAAG TATTTCGAAC GAAACATCAC CGGCAAGACT CCGTCACTGG GCTACCATAT CCTGTGTCTG CTGGCGGAAC GCGAGATAAT CAAGTCCGTT TGGACAACAA ACTTCGACGG CTTGATCATT AAAGCCGCCC ATAAGTACCA GTTGGTGCCC ATCGAGGTCA CCCTCGAGAG CCAAGATAGA ATCTATCGGA CGGATGCCAA CAAGGAGTTG CTTTGCATAG CCTTGCATGG GGACTACAAG TACGGTCCGC TGAAGAATAG TAAAGAGGAG CTGGACAGCC AGTCTGACAT CTTCGTGAAT GCCCTTTCCT TCGAGGCGTC TAAGCGCTAT TTTGTGGTGA TGGGATACAG TGGGCGCGAC AAAAGCCTCA TGCAGGCTAT TGAGCGAAGC TTTTGCAGAA GCGGCGCTGG CCGCCTTTAC TGGTGTGGAT ACGGCCGGAA TATCGCGCCT GAGGTACGCG TGCTGATCGA GAAGTTGAAC TTGTATGGAC GCGAAGCGTT CTATATTCCC ACGGACGGGT TTGACAAGAC GATGTTGAAC ATAGCCCATA TGTGTTTCGA GGATAAGGAA TTGCAGGAAG AAGTGGAGAA ACTCAAAGCG GATCTCGGTG CGGGGTATGA GTGTCGCACC

ACCACGTTCA GCCCCTACAA GGAAGGGGTG AATAAGATCG TGGACACAAA TGTTTACCCG
ATCAAATTCC CCGACAAGTG CTATCAGTTC GAGGTGAAGA ACAGCAGCGT AATGAACCTC
TGGGATTACT GCAAGCAGCT GATAGACTAT AACATTGTGG CCGTCCCCTA TAACGGAATG
ATCTACGCCT GGGGAAACCG CAACAGCATC AGCAACATGT GCGGACCAAA TGTGAACGGG
ACGATCGAAC TCGTTCCTCT CACTAGGAAA ATCTTTTTCG ACAACGGCAC TCTCAAGTCA
ATGCTCCTTA AAACTTTGCT CATCGTGATT GGAAAGCACT CCAATTGCAA GTATAACCGA
AACAAAATCT GGCGAGAGTC CAAGAAAATC AACTACACTA TTAACGGCAA AAACATTGAA
GCGTACCAAG GCATTAGGTT TAGCTTGTTC ATGGACTGGA AATACAGCTA CCTCACCCTG
ACCCCCGCTT TCTACTACAA AGACAGGAAC AACGTTAGCA AGGAGGAGAA CAAAGAGTTC
AGCGACCGGT TTATGGAGCA AATATGTAAG ATGCAAGCCA ATAAGAATTA CGCCGCGTAC
ATAAAACACT GGATTAACAT TATCTTTCCT GATGGCAAGT CCATCATTTC CATGTACCCG
TGTAACAGCG AGAGCGGATT CGAGTTCACC ATTGTTAATA AGTCACTGCT GGTCGGACTG
CGGAGTAGGC AAGCACTGCA TAATCCTGAC GATGACATGA AGAAACGGAT TTGCATCGGT
GGAGCTGAGT TGGCGGACAC CGAGCTCAAG TTCTACAATC CGGCTCAGAA TGCAATGCAC
ACCGACTTCC ACCCCATGAG GGGCCTTATC AACAATAAGC CCTACGACTT CTACATGAAT
AACAGGCTGT TTAAATCTAA CATCTCCCTG GGCGTGATCT CTCCTGTGGG TTCAGAGAAA
AAGCTGGAGG ACTTCCTGGA CCGACTCAAC AAAAAGCACA AAGTGAACTA CAACGTCGAC
TATGTCATAG ATTATCCTGG GTTTCAGTCC GTCTACGGGG TTGGCCTTTC TGTCCCTCTG
ATCGCAGAAT GGGCGTTGTT GGATGATAAA ATGCTGAATA AAGCCAACCT GTATCAGAGC
TGCCTTAACT TCGGGGATCA GATCAAGAAG AAGATTGAGT ACCTCAAGAG CCGCGACAGC
GTGGACGTGA TCATCATATA CATTCCGAAA GAGTACGAGC TGTTCACCTT CTTCAACGAC
GGAAATATCC ATTATGATCT GCACGACTAC GTGAAAGCAT TCAGCGTGCA GAGGCACATT
AGCACCCAGT TCATACGGGA GAAAACAATT GACTCTGAGC TTGACTGCCA GATCGCGTGG
GCCCTCAGCC TCGCTATCTA CGTTAAAGCA GGCCGCACTC CGTGGATTCT CAGTGGCTTG
AGGACTGATA CCGCCTTCGC CGGCATCGGC TATAGTGTGG ACCATATAAA GACCGACAAC
CAGACCCTTA TCGGCTGTAG CCATATTTAC GGGGCAGATG GCCAAGGCCT CCGGTACAAG
CTCTCCAAGA TTAAGGATGT GACCTTCGAC AGCAAGAACA ATCCCTACCT GTCCGAAAAC
GAGGCCTACC AACTCGGCCT GAATATCAAG GAACTTTTCT TTGATAGCTT CAAGACGTTG
CCCCAACGAG TGGTCATACA CAAAAGGTTT CCGTTCCAGA AGCAGGAGAT CGATGGCCTG
ACTAAGTGTC TTGGGTCCGC GGGAGTGAAA GACATAGACC TCATCGAAAT CACCTTGGAG
GATCGATTTA GGTGCTTTGA ATACGACAGG CGACTCCAGA TTGACGGCTA CCCCGTGAGG
AGGGGCGTGT GCTTCGCCAT CAACGAGAAC ACCGCCTATC TGTACACCCA CGGTATTGCA
CCAAGCGTCA AGAATGCCAA TCTCCGCTAC ATACAGGGCG GTAAGAGCAT CCCTGCCCCC
CTGAAAATCG TTAAGCACTA CGGGAACGGC GACCTGGCCC AAATTGCGAC AGAGATCTTG
GGCCTGTCAA AGATGAATTG GAACAGTTTT GGTCTGTATA GCAAGCTTCC GTGCACTATC

CAATCTAGCA ACGCTATCGC TCGCGTAGGG TGGCTGCTCT CCCAGTATGA GGGCGTAGTT TACGACTATA GGAATTTCAT GTAGTGA

96 Gene Name: gene-17982

Gene ID: BBF.2017.48.96

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGAGGAA GGTAGAAGAT CCAAAGAAAA AGCGGAAGGT CGGGAGCGGG
TCCATCACCA GCTACCCTTA CGCTAGGAAC AAGGCCGACA TGATTCGCAA GGTTAATTGG
AATCTGATCG TGTTCGACGA AGCCCACAGG ATGAGGAATG TCTATAAGAA GTCCAATAAG
ATCGCCCGAA CCCTGCGCGA GGCCACTGCC GGCTATCCCA AGATCCTGCT CACTGCAACC
CCCCTCCAAA ACTCCCTCAT GGAGCTCTAC GGATTGATAT CTTTTATTGA CCCCCACATC
TTCGGGGATG AGACAACTTT CCGCAGACAG TTTAGTCGCG GCACCAAGGA AATGAGCGAG
ATGGACTTTA TCGACCTGAA ACAACGAATT AAACCCGTGT GTCACCGCAC CCTGAGGCGC
CAAGTCACAG AGTACGTTAA CTACACTCAG CGCATTCCGA TCACCCAGGA GTTCATGCCC
ACCAACGAAG AATGGGAGCT GTACGAGAAG GTCAGCGCCT ATTTGCAACG AGAACATCTC
TTCGCGCTCC CCGCGTCACA ACGAGCACTT ATGACCTTGG TAGTGCGCAA ACTGCTCGCC
AGTTCTTCAT TTGCTATTAG CGATACCCTG CTGAGCCTCA TCAAGAGGTT GGAACAACTG
CTGGAACAGC TGGACTCCGG CAAGACGGAG ATTACCGTAG AACACTCCGA TGTCTACGCG
GACGTGGACG AGTTTGATGA TACAGTGGAG GAGTGGGAGG AGGACGACCA GCCTTCTTAC
ATAGATAAAC TGAGCCCAGA CGAGATGAAA CGGTTGATTC AGGAGGAAAA GGAAGAACTG
GAGCAGTACT ACAGCCTTGC AAAAAGCATT AAAGAGAACT CAAAGGCTGA GGCCCTCCTC
ATAGCGCTTG AAAAAGGGTT TGAAAAGCTC AGGATGCTGG GGGCTAATGA GAAGGCCGTG
ATCTTCACAG AATCCCGACG CACACAGATG TATCTGAGAG AATTCCTGGA GAGAAACGGC

TACGCCGGGA AGATAGTGCT GTTCAACGGT GAAAACCAAG ACGAACAAGC GAAGCAGATC
TATGAGCAGT GGTTGGAGAA GCACCGACAC GACGACAAGA TTACGGGCTC TAAGACGGCG
GACATGCGAG CCGCGCTCGT GGAGTACTTT AAGGAGCAGG CTAGTATAAT GATAGCGACC
GAGAGCGCCA GCGAAGGCAT CAATCTGCAA TTTTGCAGCT TGGTTGTGAA CTATGACTTG
CCATGGAATC CGCAAAGGAT AGAGCAACGG ATCGGGAGGT GTCACCGCTA TGGTCAAAAG
CACGACGTGG TGGTAATAAA CTTTCTCAAT TGTAAAAACG AAGCGGACAA GAAAGTAGAT
GAGATATTGT CCGAGAAGTT TCGGCTGTTT GAGGGCGTAT TTGGCAGCAG TGATGAAGTC
CTGGGGTCCC TCGAAAGCGG CGTGGATTTC GAGAAGAGAA TCCAACAAAT CTACCAGACT
TGCCGAACCG CGGAAGAAAT TGAGCAAGCG TTCAAGAATC TGCAAGCTGA GCTCGACGAG
CAAATTCAAC TGAAGATGAA GGAAACCCGA ATGCATCTTT TGGAAAACTT CGATGACGAG
GTGAGGGAAA AGTTGCGCGA CCATTATCAC CAAACCTCCC TGCATCTGAA TAGGATGGAA
AGGTATTTGT GGAACCTCAG CAAGTACGAG GGGGCACGCG AAGCCATCTT TGACGACGAA
ACGCTGTCCT TCGTGAAGGA CTACGAAACC TATCAGATGA TCAGCCAGGC GAAGAAACAA
AACAGTCCAA ACGTGCATCA CTTTCGATTC TCCCACCCGC TTGCGCAGAA GTGGATCGAA
CAGGCCAAGA GCAGGGAATT GTTGCCAAAG GAGATAACGT TCAGGTACAG CGACTACAAG
GGCAAAGTCT CCATCTTGGA AAGACTCATC GGCAAGGAGG GTTGGTTGAG TCTGGATCTG
CTTCACGTCC AGAGCCTTGA GAGCGAACAA CACCTCATCT TTAGCGCCAT CGACACCGAG
GGCGGTCAAC TGGACCAGGA GATGTGCGAG AAAATGTTCG AGCTGCCCGC TGTGGAGGGC
GAGGAAGTAG AGATATCCGA CTCCATCCGA AACACACTCA GACGAATCTC AGAGGGCCAG
CAAGAGGCAA TACTGAATGA GATTATGGAA CGGGCGTCCG CCTACCTCGA CTCAGAACTC
GAGAAACTGG AAAAATGGTC ACAGGACCTC AAGAATAAGC TGGAGAAAGA CATTGATGAA
ATGACGGTGG AGATCGAGCA TCTTAAACGG GAAGCTAAAT TGACACGCAA CCTGGCAGAA
AAACTCGAGA AAAACAAACA GATCAAGGAG CTTGAGAAGA AGCGCAACGA AATGCGCCGG
AATCTCTATG ACCAACAGGA CGAAATCGAT GAACAAAAGG ACCGCCTCTT CGAGGAGGTA
GAGAAAAAAC TTGAACAACG GACTGCGACG GAGCACCTCT TCACTATCAA ATGGCGGATC
GTGTAGTGA

97 Gene Name: gene-18958

Gene ID: BBF.2017.48.97

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAGAGGAA GGTCGAAGAT CCCAAAAAGA AACGAAAGGT TGGATCAGGA
TCTCTTCACC TTAACTACCT CCCATTGCGC TTTACCGCCG ATATATTCAA GGGTGGTGCT
TTGACATTTC CCGAAGGCAG CGAGAAAAAC TGGACCAGCG ACGATCCAAT CAGCAAGGAG
CTGAGCAAGT TGCGAGAGAA ACACGGAGAT AGTCATGTAT TCCACCGGAT GGGAAACAAA
ATTGCATGTA TCCCCGTTGT GGAGAACGCC ATTGCTATAG GCACCGAAAC GGATTTCAAC
ATCATTAGTG ACTTTCAGCT GGCTAATGCT CTTGCTCGCA GCGCCCTCCA CAGGTACTTC
AAAGCTGCGG GAAGGGAGAC TGTAATTGGG TTCCGACCCG TAACCCTTCT CTTGGAAAAA
CACAACTTGG CCAGCAACAG GAAGGACGTG TTCGGCATTT TCCCCGAGTA CACTCTGGAC
GTCAGGCCTC TTGCACCACA TGAGGGCGAC ATAGCGAGCG GAGTGCTTAT CGGCTTTGGA
ATAAAGTATG TTTTCCTTCA GAACGTAGCC GAGCTGCAGG CACAAGGGGT GAGTGCCGCA
GGGATGTACG CCGTGAGGCT GGTAGACGAG AGCGAACATC AATTTGACCG GGCCTACCTG
GGAAGGATTG ATCGGTTCAC AAAAGATAAC GTGACGCTCG TTGACAGCGA TTACGCGGAA
TATCCCGCCG ACCAGTGTTA CTTCGAGGGA AGCAGGACCA ACATCGAAGC CGTGGGCCGA
AGTCTCCTGG GGAAAGACTA TGATGCCTTC AGCTCAAGCC TTTTGCAGGA GAGCTACAAA
GTGACCGGAG CCCCCAACCA AACCCAACGA CTGCACCAGT TGGGCGCGTG GCTCGAGGCC
AAGAGTCCGA TCCCCTGCGC CGTTGGTCTG GGAGTACGGA TTGCAAAAAA GCCGCATGAG
TGCTCACGAG GCAACGACGC CGGGTACAGC CGCTTTTTCG ACAGCCCCAA GTGCGTGCTG

CGGCCTGGCG GCTCTCTGAC CGTGCCCTGG CCGGTCGACA AGCAGATAGA TCTCAATGGC
CCTTACGACG CTGAGAGCTT TCCCAACAAG AGGGTACGAA TTGCCGTCAT CTGCCCTCAG
GAATTCACCG GGGATGCGGA AGAGTTCCTC CGGAAGTTGA AGGAGGGCCT TCCTAACGCA
CCGGACGGCA GTCCGTTTCG CAAGGGCTTT GTTCGAAAGT ACCATTTGTC TAGCTGTGAC
TTCACGTTCC ATGAGGTTAA GCGGAGCTCA AACAGTGACG ACATCTACAA GGATGCGTCC
CTTGAGGCAC TGAAGCAGAA GCCAGATATG GCAATCGCCA TAATCCGGTC CCAATATCGC
GGGCTGCCCG ATGCTTCTAA TCCCTATTAC ACGACAAAAG CTAGGCTGAT GGCCCAGGGC
GTACCAGTTC AACTGCTGAA CATAGAAACC ATCAGGAGGA AGTCTTTGGA CTACATTCTG
AATAACATCG GTCTTGCCAT GTATGCCAAA CTTGGAGGAA TCCCTTGGAC CCTCACCCAG
AATAGCGACA TGGCGCACGA GATCATCGTC GGGATAGGGT CAGCCCGGCT CAATGAGAGC
AGGAGGGGTG CTGGCGAGAG GGTCATCGGG ATCACGACCG TGTTCAGTGG TGACGGACAG
TACCTCCTCG CCAACAACAC CCAGGAAGTT CCCAGCGAAG AGTACGTAGA CGCATTGACT
CAGTCTCTTA GCGAGACAGT ATCAGAGCTT AGGAGCCGGT TCGGTTGGCG CCCTAAAGAT
CGAGTGAGGT TCATATTCCA CCAGAAGTTT AAGAAGTACA AAGACGCAGA GGCGGAGGCG
GTTGATAGGT TTGCACGCTC ACTGAAAGAT TTTGACGTGC AATACGCCTT CGTGCATGTG
TCTGATTCTC ATAACTGGAT GCTGCTGGAC CCAGCTAGTC GGGGGGTGAA ATTCGGCGAT
ACGATGAAGG GCGTCGCCGT CCCTCAGCGG GGACAATGTG TGCCCCTGGG GCCAAACGCT
GCGCTGCTTA CTTTGAGCGG TCCGTTCCAG GTAAAGACCC CACTGCAAGG CTGTCCGCAC
CCCGTGCTGG TGTCAATTCA TGAAAAGAGC ACTTTTAAGT CTGTTGATTA CATAGCCCGC
CAAATCTTCA ATCTCAGCTT CATCAGTTGG AGGGGCTTTA ACCCTAGCAC CCTCCCAGTG
TCCATTTCCT ACTCCGACAT GATCGTAGAC CTCTTGGGAC ATCTTAGACG CGTTAAGAAT
TGGAATCCGG AAACCCTGTC TACCGCTCTT AAGGAACGAA GGTGGTTTCT GTAGTGA

98 Gene Name: gene-18963

Gene ID: BBF.2017.48.98

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAGAGAAA AGTGGAAGAT CCCAAAAAGA AGCGAAAGGT GGGTAGTGGG
AGCATGAATT TCCAGCTGTG CGACCAACGC AAAGCCATTA TCGCCGAACC AGGCCATCTG
TTGGTCCTCG GTGGGCCAGG AAGCGGGAAA ACTACCGTCG CCCTCTTCAA GGCCAAGCAG
AGATTTAGCA CTCTGAAACC TAGCCAAGAA ATCCTGTTCC TGTCATTCAG TAGAGCTGCC
ATCCGACAGG TCCTGCTGCG GTGCAAGGAG ATTCTGAAGC CCGCAGAAAG ACGCGCTGTC
GCCGTTCAAA CCTATCATAG CTTCTGCATG GACATGCTGA GGGCGCACGG TAGACTGCTC
CTGGGCCACC CCGTGCGATT CATGTATCCC GGCGACGAGA GGCTTCAAAA GGCCGCATTC
GAGGGGGACT GGGAGGCGGA AAGACAAAGG CAAGCCAAAG AGATGGGCAT CTTTTGCTTC
GACCTTTTCG CGCAAGGCGC AGCTGAGTTG CTCGAGAGGT GTGCCGCACT TAGGAAGCTT
ATAGGGGACA GCTTCCCCAT GATAATAGTG GACGAGTTCC AAGACACCGA CGACAACCAA
TGGCGGATCG TGGCGCAACT TGCCAAGGTA GCGGACATCT TCTGCCTTGC CGACCCCGAC
CAGAGGATCT TTGACTACCG GGACGACATC GACCCCCTTC GGATCGAGGG TTTGCGGACC
ACTCTTGCCC CCAGGGAGTT CGATCTTGGC GGTGAGAATC ACCGCTCCCC GAACGCAGGG
ATATTGAACT TCGCCAACGC TGTGCTGCAT AACCAGAGCC CCCTGCCCGA TACCAGCGAC
ATCATGCAAC TGCGGTACTG GCCTAGAGCG TTCGCGAGCA CCGTGCATGC CTGCGTAGTG
TTTACCTTCA GCGAACTCAG GAAACTGGGC GTGGAGAACC CCAGCGTGGC AGTGCTGAGC
CGATCCAACG GGCTTATCTC CGATGTGAGC GCCATACTGG CTGAGAAGCA CGCGTACAAC

GGGAGGGAAC TGCCAATCGT GGAACACGAC GTGGTTTGGG ACGCGGAGCT GTCTGCGGCA
GCAGCCGTCG TCGTTGCGTC CACCCTGGAG TGGCCAACAG CCGCTGCAGA GGTTGCTGTT
GCCAGGACAC TTGCGCTCAT AGCAGCCTAT TACAAGCTGA AGAACGCCGA GGAACCCACC
AAGAGCGCGG CTGAGGCTGC CCAAAAGTAC GAGGCGGCTG CAAGCAAGGT GGCCAGTGAG
GAAACCCCAA GGATCAAAGC CGCGAAAGAA TTGCTGGCCG CTCACCAAAG TGGCATCCAG
ATGGTGGGCG ACCCGGTGGC CGATTGGAAG TCTGCGAGGA GGGTATTGCA AGAGATAAGC
GCCCTGGGTG AGTTGTACAG GGAGGTCCGG CTCGTGAGGT TGTTCCGGGC AACCGACGCC
TTGGCTTCCG GCCTGAGCAA TAGGTGGTTG GCTACTGGAA GCTACGAGGG CGTGTCCGAC
CTGGTGAAGG GCATCCTTGA GCAGGAGAAA CTGATTGCCG TGGAAAGGGA CCCAAGAGGC
TGTATACTGA TGAACATCCA TAAAAGCAAA GGTAAGGAAT TCGACGGCGT GGTACTCATT
GAGGGGGCAT TTAAGTCCCA TTTCTTCGAT GAGCGGAAGG AAGTCAGCCC CTATGAGAGG
TCCAGACGGC TCCTGAGAGT CGGTCTGACC CGCGCTAGGC ATAGGGTGAC AATCCTTAGA
CCTCAGGGAG CGAGGCCCCT TGTGGATCCC ATCTAGTGA

99 Gene Name: gene-8487

Gene ID: BBF.2017.48.99

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAACGAAA GGTAGAAGAT CCCAAGAAGA AAAGGAAAGT GGGAAGCGGA
AGCATGGAGA ACCTGGCTCT TAGTGCGCTG CAACTGGACT CTAAGCTCGA CCGCTACATC
GTGTGTAGGT ACAGAATCGT GTACCAGAAG CGGGACGAAA CCATTCCCGG CGAACAGTTG
GCCCGGAAGG CGGCCTACGA GATCCAGAAA GCGAATGACT TCGCCCTTTT GACCAACCTC
GGCAATCAAC ACATCGTTTC CCTCAAGCCC ATCTCACAGA GGGGCATTGA AAGCACCCAC
CTTCAGGCGA ATCTCATCGA GGACGGGGAC CTGGAGCTCG ATTGCTCCAT CGAACAACAT
CAGCAGGCAC TCCAGCGGCT CGTGAACCAG GACATCAATA AAGCTGCGTG GAAGCTTAAA
AAGAGCTCAC AGGGCAAACT CGATTACAAA AAGGCAGCTA GCGGGAACAC CGAGATCTTT
GAGCCAATTC ATAGCACTCG AATCAACGCC CGAGCCACGT ATCTTGACGC TTTTTGCTCA
CTGCAGCTTA GCCCCGAGGT GCTTGCTAAT GGAACCGTAC TGATAGGGCT GCATCTCAAG
CACAATCTGG TAGCAAAGTC TGACATCTCT TTGCAGTGGA TCATTGATAA AAGGCCCGAT
TGGCTGCAGA GCATCAAGAA GGTGCGGCAC AGGTACTTCG ATCCCGGCAA AGCGCCCCTG
GTCGCCGAAT TCCTGAGGGT GGAGGACTCC CTGAATGGCA ACAGCGTCTT GCCCCACATG
GGCCAGAGTC TTGTTTCATA CCACCAAGCG AAGGGACTCT TGTCAGAAAG ACAGCTCGCA
GAGGCCACTA AGAGCGTGCT GATAAAGGTA AAATACGGCA AAAACGAGGC GGACCATATC
GCATCTCTGG TTGAACCAAT GTTTGATTTC GACACGCTCA GCAAGATCGA TAGTATCTTC
CTTAACAAGT TGGCAAAGGA CCTGAAGTGG AGCCTGAACG ACAGGATACG CACTTCCGCG

AAAATGGTGA AAGGCTTGTA TCTCCCAAAC TTCAACTGCA AGCTGGAACA GGTTGACTAT
CAGATCCTTC ACAGGCAGCG ACTTAATCAC CAACAGATGC TTCAATTCGC CAACGGGGCG
AAATCTTCAA GAGAGCAGGA CGTGCTGCGA CATAAGGCGT TCGGCAACAT GACGCGCACA
CAAGTTATCC CGCTTATTGC GGGCGAGAAG AACAATACAG AACAAAATAA GCAGCTCCTG
TGCAACGCAT ACCAAGCATT GCAACAACTG ACCACCACGG AATTGCCTCC GTTCACCAAG
TTCCCCAACC CCGTAGAGAA CGCAGCCGAG CTGGACGCAA GACTGAATGA ACGGTGTCCC
CCAAATGCGA TACTGCTCAT CGGCCTTATC GACAAAAGCG ACAAAGTGGC GATCCGCGAC
ACCGCGTTTA GCTACGGTCT TGCAACCCAG TTCATGCGCC TGGATCACAG ACCGAACGTC
TACAGCCCCT CATATTTCAA CAACGTGGCG GCTGGTTTGT TTTCCAAAGG TGGCGGGCAG
CTCTGCGCCA TTGATGACAT GCCGGGTGAA ACCGACTTGT TTATCGGGCT CGACATGGGA
GGGATCTCTG TAAGGGCACC AGGCTTCGCG TTTCTGTTTC TGCGATCTGG TGCGCAGTTG
GGGTGGCAAC TCGCGGACAA ACAACAGGGA GAAAGGATGC AGGATGAGGC CCTGATGTCA
CTGTTGGACA AGTCTCTCAC CACCTACCTG AGAAGCTGCT CTGGTGAGCT TCCTAAGCGC
ATAACCCTCC ATAGGGATGG CAAGTTCTAC GAAAGCATAG AAGTGATCGA GCAGTTTGAG
CAGAAGCACG GCGTGAAAGT AGATGTGCTG GAGGTTCTGA AAAGCGGTGC TCCGGTTTTG
TATAGACGAA GCCGCATGGC CGACGGAACC AAGGAGTTTA GCAACCCCAA TGTGGGCGAC
GCGATCTATC TCAGTGATCA TGAGATGATC CTGAGCACGT ATAGCGGCGA AGAACTCGGA
AAGATATGGG GTGACAAGGT CAGCGTCAGG CCTCTTAGGC TGCGCAAGAG ATACGGTGAT
GTGAGCCTGG AAACCCTGGC ACATCAAGTG CTCGTGCTGT CTAGGATACA CGGCGCTAGC
CTGTATCGCC ATCCTCGACT GCCCGTGACC ACGCACCACG CCGACCGATT CGCAACACTG
AGGCAGGAAA CATGCATAGA CGCCCTCTCT AAGATGGACC GGCTCTGTCC GGTCTACCTG
TAGTGA

100 Gene Name: gene-18016

Gene ID: BBF.2017.48.100

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AGAAGAGAAA GGTCGAGGAC CCGAAAAAGA AGCGAAAGGT AGGTAGTGGT
TCCATGGTCG GCGGCTATAA AGTCAGCAAT TTGACAGTGG AAGCGTTCGA AGGTATCGGG
AGTGTCAACC CGATGCTGTT TTACCAATAC AAAGTCACCG GAAAGGGAAA GTACGATAAT
GTGTATAAGA TTATCAAAAG CGCACGGTAC AAGATGCATT CTAAGAACCG ATTCAAGCCC
GTGTTCATCA AGGACGACAA ACTGTACACC CTCGAGAAGC TCCCGGATAT AGAGGACCTG
GATTTCGCAA ACATTAACTT CGTGAAAAGC GAGGTTCTCA GCATAGAGGA TAATATGTCA
ATTTATGGCG AGGTGGTGGA ATACTATATC AATCTCAAGC TGAAAAAAGT GAAGGTGTTG
GGAAAATACC CCAAGTACAG GATCAATTAC AGCAAAGAGA TTCTCAGTAA TACGCTGCTG
ACACGAGAGC TCAAAGACGA GTTTAAGAAA TCAAATAAGG GTTTTAACCT GAAACGGAAG
TTTAGAATTT CCCCCGTGGT GAATAAGATG GGCAAAGTGA TACTCTATTT GTCCTGCAGT
GCTGATTTCA GCACCAACAA GAACATTTAC GAAATGTTGA AAGAGGGCTT GGAGGTTGAG
GGGCTGGCCG TTAAGAGCGA GTGGAGCAAT ATCAGTGGCA ACCTGGTGAT CGAGAGCGTA
CTGGAAACCA AGATATCCGA GCCCACTAGC CTGGGCCAAT CCCTGATAGA CTACTATAAG
AATAACAACC AGGGCTATAG GGTGAAGGAT TTCACCGATG AGGATCTGAA TGCCAACATT
GTCAACGTGA GAGGAAATAA GAAGATCTAT ATGTATATTC CGCACGCGTT GAAGCCGATA
ATCACCCGGG AGTACCTGGC CAAGAACGAT CCAGAGTTTT CTAAGGAGAT CGAGCAGCTT
ATCAAGATGA ATATGAACTA CCGATATGAA ACCCTCAAGT CATTTGTGAA TGACATCGGG

GTCATTGAGG AGCTGAACAA CCTGAGCTTC AAAAACAAAT ACTACGAAGA TGTGAAACTG
CTGGGTTACT CCAGCGGCAA AATAGACGAA CCCGTCCTGA TGGGGGCAAA AGGGATCATA
AAGAACAAAA TGCAGATTTT TTCCAATGGA TTCTACAAAC TCCCCGAAGG CAAGGTACGA
TTTGGCGTTC TGTACCCAAA AGAATTTGAT GGCGTGTCAA GGAAAGCTAT CCGCGCCATT
TATGACTTCA GTAAGGAGGG CAAATACCAC GGCGAAAGCA ACAAGTATAT CGCGGAACAC
CTGATAAACG TGGAGTTCAA TCCAAAGGAG TGCATATTTG AGGGATACGA ACTGGGCGAT
ATCACCGAAT ACAAGAAGGC GGCTCTGAAA CTTAATAACT ACAACAATGT CGACTTCGTA
ATCGCAATAG TCCCGAACAT GTCCGACGAA GAGATAGAGA ACAGCTACAA TCCGTTCAAG
AAAATATGGG CCGAACTGAA TCTGCCCAGC CAGATGATTA GCGTCAAGAC GGCCGAAATC
TTTGCCAATA GCAGGGATAA CACGGCGCTT TACTATCTGC ATAACATCGT CCTCGGTATC
CTGGGTAAGA TAGGAGGGAT TCCCTGGGTG GTTAAAGACA TGAAGGGCGA CGTGGATTGC
TTCGTTGGAC TCGATGTCGG CACCAGGGAG AAGGGCATAC ATTACCCCGC CTGCAGCGTT
GTGTTTGACA AGTACGGCAA GCTTATTAAC TATTACAAGC CTAACATCCC GCAGAACGGA
GAGAAGATTA ACACAGAAAT ACTTCAGGAA ATTTTCGACA AGGTGCTCAT AAGCTATGAG
GAGGAGAATG GAGCCTACCC GAAGAATATC GTGATCCACA GGGACGGCTT TAGCCGAGAG
GACCTTGACT GGTATGAGAA CTACTTCGGT AAGAAAAACA TAAAGTTTAA CATCATCGAA
GTCAAAAAGT CAACTCCGTT GAAAATCGCC AGTATAAACG AGGGAAATAT CACGAATCCT
GAAAAGGGTT CCTACATCCT GCGCGGCAAC AAAGCCTACA TGGTGACCAC AGATATTAAG
GAAAACCTGG GAAGCCCAAA GCCCCTGAAG ATAGAAAAGA GCTACGGCGA CATAGACATG
CTCACAGCTC TCAGCCAAAT ATACGCACTC ACGCAAATCC ATGTGGGGGC GACCAAAAGC
CTGCGCCTCC CAATCACCAC CGGCTACGCC GACAAGATTT GCAAGGCGAT CGAGTTCATC
CCCCAAGGGC GCGTGGACAA CCGCCTTTTC TTTCTGTAGT GA

101 Gene Name: gene-4720

Gene ID: BBF.2017.48.101

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AGAAACGAAA AGTGGAGGAC CCCAAGAAAA AGCGGAAGGT GGGCAGCGGC AGCATGAACA ATCTGATGCT GGAGGCGTTT AAGGGCATTG GCACCATCAA GCCCCTGGTG TTCTATAGGT ACAAGCTCAT CGGCAAGGGG AAGATTGAGA ATACCTACAA GACGATCAGC AACGCCAAGA ATAAGATGAG TTTCAATAAC AAGTTCAAAG CGACGTTCAG TAAGGGAGAA ACCATCTACA CCCTTGAGAA ATTCGAGGTC ATGCCCAATC TTAACGATGT GACCATTGAG TTCGACGGAG AAGAGGTTCT CCCGATAAAA GACAATAATG AAATTTACTC CGAAGTCGTG CAATTTTACA TCAACAATAA CCTTCGAAAG ATCAAACTGG ATAACAAATA TCAGAAGTAT CGAG-CAACGA ATACCAGAGA GATAACTGGC AACGTCATAC TCGACAAAGA CTTCAAGGAG AAG-TACAAGA AGTCTAAGTC AGGGTTCCAG CTCAAGCGCA AATTCATAAT TTCCCCCAAG GT-GAACGACG AGGGTAAGGT AACCCTGTTC CTTGACCTGA ACAGCAGCTT CGACTATGAC AAAAACATTT ACCAGATGAT CAAGGCCGGG ATGGACGTGG TGGGGCAGGA AGTGATTAAT ACGTGGAATA ATAAGAAGCA GAAGGGCAAG ATTAAGAAGA TTTCTGAGCT GACGATCTCA GAGCCTTGTA ACTTCGGCCA GTCCCTTATC GATTACTACG TTTCCCTCAA CCAAGCTGTG AGGGTGAAGA ACTTTACGGA AGAGGAAAAG AACACAAACG TTATCGTCGT CCAGGTGGGA AAGGGCGAGG TTGAGTATAT TCCGCACGCG CTCAAACCCA TCATTACTAG GGAGTACATA AAGAAATACG ATGAGGCCTT CAGCAAAGAG GTAGAAAACC TGATCAAAAT CAACATGTCA TACAGGTACG AAATACTGAA AAAGTTCATC GACGACATCG GCTCTATAAC CGAACTGAAC

AACCTTAAGT TTGAGAACAC GTACATAGAT AACATCGAGT CACTGGGCTA CCAACAGGGA
AAGCTGAACG ATCCCGTGCT GATAGGCGGC AAAGGCATCC TGAAGGATAA GATACATGTG
TTCAAATCCG GCTTTTACAA AAGCCCCATT GACGAAGTCA AGTTCGGCGT GATTTACCCG
AAAGGCCACA CCAATGATAG CAAGTCCACC ATCCGGGCGA TTTATGATTT TTGTACCGAC
GGGAAATACC AAGGCAAGGA CAACATCTTC ATTAACAACA AACTGATGAA TATCAAATTT
AGCAACCAGG ACTGCGTGTT TGAGGAGTAC GAGCTCAATG ACATAACGGA GTATAAGCGA
GCCGCGAATA AGTTGAAAAA CAACGAGAAC ATCAAGTTTG TAATCGCCAT CATCCCCGCG
ATTGATGAGA GTGATATAGA AAATCCCTAC AACCCTTTTA AGCGGGTCTG CGCCGAGTTG
AATCTGCCCA GCCAGATGGT AAGCCTCAAG ACCGCGAAAA GATTCGGCAC CAGCAAGGGT
AATAACGAGT TGTATTTTCT GCATAACATT AGCCTGGGTA TCTTGGGTAA GATAGGAGGG
GTCCCTTGGG TCATTAAGGA CATGCCTGGG GAAGTTGACT GCTTCGTGGG CCTGGATGTG
GGCACCAAAG AGAAAGGGAT CCACTACCCC GCATGCAGCG TCCTTTTCGA CAAGTACGGC
AAGCTGATTA ACTATTACAA GCCCACAATC CCGCAGAGCG GCGAGATCAT CAAGACAGAC
GTGCTGCAGG AGATCTTCGA TAAAGTGCTG CTGAGCTACG AGGAGGAGAA CGGGCAGTAT
CCTCGAAACA TCGTGATTCA CAGGGACGGG TTCAGCAGGG AGGACCTGGA GTGGTATAAG
AACTACTTCA TCAAAAAGAA TATAAACTTC ACGATTGTAG AAATCAAGAA AAACTTCGCC
ACCCGCGTCG CGAACAACAT AAACAATGAA GTGTCCAACC CATTTAAAGG GAGCTTCATA
CTGCGCGAGA ACGAGGCCAT CGTTGTAACC ACCGACATCA AAGATAATAT CGGCGCTCCG
AAACCAATCA AAGTCGAAAA GACATACGGC GATATTGACA TGATGACCAT AATCAACCAG
ATCTACGCCC TCACGCAAAT CCACGTCGGA AGCGCGAAAT CTATGAGGCT GCCGATCACG
ACCGGCTATG CCGACAAAAT ATGTAAATCC ATCGAATACA TCCCGAGCGG TAGGGTGGAC
AACAGACTCT TCTTCCTGTA GTGA

102 Gene Name: gene-5971

Gene ID: BBF.2017.48.102

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAACGGAA GGTGGAGGAT CCAAAGAAAA AACGCAAAGT TGGCAGCGGC
AGCATGATAG CCGTGGAAGA GTGGCAGCCT GCGGACGGAC TGACCCTTGA GCCTAATGCA
AAGAGGGCTG CGAAGGCTAG AAAGAGGTGC CTGGCCCTGA CAGCGGGTCC CGGTGCCGGA
AAGACAGAGA TGCTCGCACA ACGCGCCGAC TTCTTGTTGA GGACCGGAAC CTGTCGGTAC
CCCAAGAGGA TACTGGCCAT CTCATTCAAA GTGGATGCAA GTAGAAACCT GAAGGACAGA
GTGGAGAGGA GGTGCGGCTA TGATTTGGCG TCAAGGTTTG ACAGTTATAC TTTCCACGCG
TTCGCCAAAA GGATCATCGA CCGCTTTAGG CCGGTGCTGA CAGGCAAGGA CGCCCTCGAC
GCAGGCTACA CCATCGTGGA TAAGAAGAAT GGCCCCTCTA GGACCCAGAT CGAGTTCGGC
GACCTTGTCC CCCTTGCCAT ACAAATCCTG CAATCAAGCA AAATTGCACG AAACGCGATC
CGCCAAACTT ACAGCGACAT CTTCCTGGAT GAGTTTCAGG ACTGTACAAA TCTGCAGTAC
GACTTGGTAA AACTTGCGTT CCAGGGTACG TCAATACGGC TGACGGCTGT TGGCGATACC
AAGCAGAAGA TAATGGCCTG GGCTGGAGCC CTGGACGGCA TTTTCCAGAC GTTTGCCAAC
GATTTCAACG CCGTGTCCCT GAACATGTAT AGGAATTTCA GAAGCAAGCC ACAACTGCTC
AGGGTTCAAA ATGAAATTAT CAGGAAGTTG GACCCCGATT CCGTGATGCC TGACGAACAA
CTTGACGGTG ATGAAGGCGA GGTCTATGCG TGGAGGTTCG AGGATAGCTG CAAGGAAGCC
GTGTATCTTG CGGACCTTAT CAATGGCTGG ATCAACACCG AACAGCTGCC CCCAGCGGAG
ATCGCCGTAC TGGTCAGCAA ACAGCTCGAC CTCTATGTCG ACCACTTGAT GACTGAGCTC

GAGGCTCGGG GAATCCCCTA CAGGAACGAG CAGCAGCTTC AAGACATCAC CATAGAGCCG
GCAGCTAGAC TCATTGTGGA CTACTTGAGT TGCCTCTACG GCAAGAGAGA GCCGAAAGCA
TGGATCCGGC TCATGAACCA GCTGATCCCA TTCGCGGACG AGGAGATCCA ATCTAGTGCT
CGAAAGGACC TCGACCAGTT GATAAAGAAG CAGAGAAAAA GGGTGAGCGA CGCGAAGCAC
ACCGATTCAC CTTTCAGCGA TTGGGCACAA CTCGCAATTG AATTCCTGAA GTACATAGGC
AGTAAGATGC TGGTGGCACT GAGTCCAGAT TACGAAACGC GCGAGAGGCT GAATGACGTG
ATCAGGGAAA CTTTCGCGAG GATCAAGGAA CTGCTCAAGA GCGAGCCCGA TCTGCCCAAG
GCGCTGGGCC GGTTTGCCGA TGACCAGGCG GTGCGAATAC TGACCATCCA CAAGAGCAAG
GGCCTGGAAT TCGACAGTGT GATCATCATG GCCGTCGAGA ACGAGATATT CTTCGGGAAC
CAGGACGAGA ATAGGTGCGC TTTCTTCGTA GGTGTGAGCC GAGCAAAAAG GAGGTTGATA
CTTACCCACG CCGACCAGAG GGAAAGGCCA GCGTCTGCCA AACGATGGAA TGTTAGTAGA
ACCGCTCAGA CTGAGTACAT TAGTTACGTC ACCCCTTTCG TGAGGCCACA GTAGTGA

103 Gene Name: gene-5976

Gene ID: BBF.2017.48.103

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAGA AAAAACGGAA GGTGGAGGAC CCCAAAAAGA AACGCAAAGT GGGTAGCGGC
TCAATGCTCG ACTTTAGCCT TACCCAGAAA GGTTGGGTGC TGCCCATCGT ACTGAACGCC
TTTCCGCTCA AGGTACCGGA CATGGAGCTC AAATTCGTGC AGATCCCCTA CGACAAGACG
ACCCTGGACT CACTGAGGTC AAGCCACAAG ATGACCCACG TATTCAGGAG GCAAGGCGAC
AGTATCCAGA TCTTTTCTAG CGACGGCACC TTTCCAAAGA GCGGCACCCC CCAGACCCTC
CAACTGAAGG ATAATCTGGG AATCTTTTTC TCTCTTGTAA AGGACGGCCT CCTCAAGCAC
TTCGCCGGTT TGGGCCGAAC CCCGTGCGGA TTCAACCCCA TTGAGGTCGT GTCAGCTCAG
GCCAAAGACA ATCTTCTGGC TAGCATCCTC GGAGAAGCCT ACCCGCTGAA AATTTGCGCC
AAGTACTCCA TCGACACCAG GACAGTGCAA GGTCAACCGT GTCTCATCAT CGACTGCAGC
ACTAGGAGAG TGGTTAAAGA GAACTGCCTC TTCTTCCTTA AGACCGGCTT TAACGTGATT
GGCCGCTATG TAGTGACCGA GCAGGACGAC GGGTTTCGGA AGCTGCTGGG TTTTGTGGAA
AACTGCCACG AAGGCAGGAC ACTGAGCGTT ATAAGGCCAG ATGGCCAAGC CGTGCATGCC
GAGGCCAAGG ACGTGTATCT CGAGGCATCT AGGGCCAACT TCGACGACTA CATCCTTTAT
ACGCACGGAA CTAAAAAGGA TAGCATCGTG GAGCGAATCA GACAAAGCGT GAGTATCTTC
AACGGCGGTA AGAACAAGAA AGATAGAATC GACGCGCTCA AAAAGTACAT CCAGGCCACC
AATATAAGCC TTTTGGATGG GACCAGGATC GAAATCGAGG AGCCCAGCGA CATTCAGAAG
GACTGCGCCC AGATGCAGAA GCCCGTGTTT GTGTTCAATG ACAATGGCGA GGCCGACTGG

ACCGAGAAGG GGCTGACTCA GAACGGCCCC TACACCAAGC GCACCTTCGA CCGAAACGAC
CCCAGCATCT GCGTGATCTG CGCACAACAC GACAGGGGGC GAGTGGAGCA GTTCGTTAGG
AAACTGCTGA AAGGCATGGC TAACAGCAAA TACTTCAGAA ACGGCCTTGA GGGCAAGTTC
GCGCTGGGAA CGTCCCGGGT AGAGGTGTTT GAAACCAGCA CAAATAGCGT GGACGCCTAT
AAGAGCGCGA TCGAAGCCGC CATCCGCAAG AAGGCCGATG ACGGCGGGAG GTGGGACCTG
GCATTGGTTC AAGTTAGGCA GAGCTTCAAG CAGCTGAAGG TGACTGACAA CCCCTACTAC
TTGGGAAAAA GCCTGTTCTA CATGCACCAG GTGCCAGTGC AGGATTTCAC TATCGAGCTC
CTGAGCCAGT CCGACTATTC ACTGGGCTAC AGCCTTAACA ACATGAGCCT CGCTTGCTAC
GCCAAAATGG GAGGAGTGCC CTGGCTGCTC AAGTCCTCTC CCACCCTTAG CCACGAGCTG
GTGATCGGCA TCGGCAGCGC CAACATTGTC CAGGAGAGGG GGGCACACAA CCAGAGGATC
ATGGGGATAA CCACCGTATT TAGTGGAGAT GGCAGCTACA TCGTCAGCAG CACGTCCAAA
GCTGTGGTTC CCGAAGCATA CTGCGAGGCG CTGACTAGCG TGCTGGGCGA GAATATCGAA
AAAATCCAAA GGAGAATGAA TTGGCAAAAG GGTGACTCAA TCCGACTGAT CTTCCACGCC
CAAGTGAAGA AGTTCAACAA GGAGGAGATT CAGGCAGTGC GAGCCGTGAT AGACAAGTAT
AGGGACTACC AGATCGAGTA CGCTTTTGTG AAAATCAGCG AGAACCACGG CCTGCACATG
TTTGACAGCT CAACCGCCAC CATGCCCAAG GGGAGGTTGG CCACACACAG GGGTAAGACC
TTTAAGCTGT CCAAAAACGA GATGTTGGTC TACCTGATCG GACAGAGGGA GCTGAGACAG
GAAACCGACG GCCACCCCAG GGGTGTCATC GTGAACGTAC ACAAGGACAG CACTTTCAAA
GATATCAAGT ACCTGAGCGC CCAACTGTAC TCTTTTGCGA GTCATTCTTG GAGGTCATAC
TTCCCCAACC CTATGCCCGT GACCATCACC TACAGCGACC TTATCGCCCA CAACCTCGGC
TGGCTGAACC AGCTGCCCGG GTGGTCTGAC AGCGTAATGA TAGGTAAAAT CGGTCATAGC
CAGTGGTTTC TGTAGTGA

104 Gene Name: gene-7520

Gene ID: BBF.2017.48.104

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAGAGGAA AGTTGAGGAT CCAAAAAAGA AACGAAAGGT AGGCAGCGGC
AGCGTAAAGC TTAATCACTT CCCCCTGAAT CCCGCTCTTG CAGTGTTCAA GACTACCTAC
AGGCACAGAA ACCCCAGGGG CTTCCTGGGA TTCGTTAGGT CACAAGGGTT GACCGCGGAG
AGAGTTGGCG AGGAAGTGTG TGTCTATCAC GGGCTTCCCC ACCCGGCTTT TAGAGGAGCC
ACCGCCCAAG GACACACCAG ACTGGCGCCT GGTGACACCG ATTACGACAG GGGCGTACTT
AGTCTGATCG GAGCCGCCCT GCTGAAAGCG GGTTACGTGC TTACTGAGCG CGAAAGGGCC
GCAGTGCACC CCACGCAGCA GAGAGTGCCC CTGCACACCC CTAGGAAACT CCCTGCCGAA
ATTGCGGTGA ATGCCCATCT TCGATGGGAA TGGGAACTGG AACGGCACAG CGGGAAGTCT
TGGCTTGTGC TTAGGCCCGG ACGCATGTTT TTGAGTGCGC TGAGCTGGCA CGATTTGGAC
CTGAGGGCAT GGGCACAGGA GTTGCCCCAG AGCGTACAGC AACTGCACGC GCTGTGCCTT
CGCTCCGGAC GACGAGAACG ACTGAGGCGC ATGGGTAACA CGTGGGCGTT CCAACGAGAG
GATAGGGAGC AAGAGGGAAG GTGGCACCTG AGCTTTAGCA CTAAGGCGCT TTCCGACCTG
AACCTGTCCG GGGATGCTCA CCATGCTGCT AGCCTGAGCA TGCCCGATGT GCAGAGGCTC
GTAAATCTGC CGGGTCTGTG GCAGCCCTTT GTGACAAGCC TTGAAGTCCT TGAGGTGCCT
GGTAAGGTGA TCGAGGGCAA AAGGCTGAGG TTCGGACGAG GAACAGGGCG CGACGTCACG
GATGTACACA AAAGGGGCAT CCTTCACCCT CCGCCGCAGC CAGTGCGCCT TGCGGTCGTG
CCCCCCATTC AGGCGGACGA AGAGGCGGAT GAGCAGCTCA GACGCGAGCT CCTTGCCCAC

CTCCTGCCAC GGGAAAAGGT GTTGGCCCAC CCCGAGGCTT CCCAGGGCCT CAAGAAGCAC
TTGAATCGAA GGGAAACCGA CGACACCTTC TACACCCTGT GGAGCGCTGG AGACTACTGC
AAACTGGGGC TGGAACCCTT TGATCTGGTG CGCGACCTCC ATAGGTACGA CCCCGGCACG
GGTCGCCTGC TGGCTCCAGA GAAGTTGCAT GGAGCAGCAG CCGCCGCGAG AGAGGCTGGC
AGGCAATTGA TTGGCCTCGT GATCCTGCCC GACACCATAG GGCGAGATGA GAGGGACGCA
CTGTCCGACG AACTGGCCAA GCTGGGTGTG AAGAAACTTC AGCACATCCG CAGGGACATG
CTGAACCGGC CCAGGACGCA GTATATGGCC TGGGTGAACG TGGCCGTGAA GCTCGCCCAG
AGGGCCGGAG CAGTCAGCTG GGACCTGGAA AAGTTGCCTG GAGTGTGCGA ACAGACCTTC
TTCGTTGGCG TGGATCTGGG CCATGACCAT CGGGAGAAGC AAAGCGTCCC GGCCTTCAGC
CTGCACGAGT TCCGAGGCAG GCCGGTCGAC TGCCTCACCC TTCCAAGGCG AGCCGGAAAT
GAAAGGTTGA GCCTGGCGGA GCTGAATCAA GGCCTGAGGA AGCTGCTTAA GGGTAAGAGG
CCAGCCCAAG TGATAGTGCA TAGGGACGGC AAGTACCTGG AGGGGGAGGT TGATGACTTC
ATAATCGCTT TGAACGACCT CGGCGTGCCG CGAGTCTCTC TTCTCGCCGT CAAAAAGTCC
AACCTCTCCA TGGTTGCCGG CGCTAAGGAG GGAGCGTTTT TGCCACTGGA CGAGCGGCGG
TGTCTGCTGG TTACCAATAC CCAAGCCGCG GTAGCTAGGC CGACAGAGCT GGAGGTGATG
CACTCAGATC ATCTGACTTT CGCCGAGCTG ACCGAGCAAG TGTTCTGGCT GACCCGAGTA
TTCATGAACA ACGCACAGCA TGCGGGTAGC GACCCTGCTA CCGTAGAGTG GGCGAACGGG
ATCGCTAGGA CCGGAAAGAG AATTGCCCTG TCTGGGTGGT CCGCCTAGTG A

105 Gene Name: gene-17978

Gene ID: BBF.2017.48.105

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGCGAAA AGTAGAGGAT CCAAAGAAGA AACGGAAGGT CGGCAGCGGA
AGTGTGAACC ATTACTATTT TTCCGAATGC AAGGCGGACG AGAAAGCCAG CGACATAGCC
ATCCACCTTT ACACCGTGCC CCTGTCCAAC CCCCATGAGA AATACAGCTA TGCGCACTCT
ATCGCCTATG AATTGAGAAA ACTCAACTCA TACATAACCG TGGCCGCGCA CGGTCAGTAT
ATCGCGTCTT TCGAGGAGAT ATGCCACTGG GGCGACCACA GGTACATACA GCACGAACAT
AGACCAATCC AGTGCAGCCT CCCGATGGAG AGGACCATAC TGGAAAGACT CCTCAAGAAA
GAGCTCGAGA ATAGGTGCAA AAGCAGCTAT AAGATGGACA ACGACCTTTT CCGGTTGGCT
AACGAGCAAA GCATGCACGT GGGCGAGATC AGCATACACC CAGCGATCTA CATCTCATTC
AGCGTGGAGG AAAATGGTGA CATATTTGTT GGCTTCGACT ACCAGCACCG GTTCGAGTAC
CGCAAAACAC TCCAAGACGT CATCAACAAC GATCCCTCCC TGCTTAAGGA AGGCATGGAA
GTGGTGGACC CCTTCAATAG AAGGGCCTAC TATTACACTT TTGTGGGCAT GGCCGATTAT
ACCGCCGGAC AGAAAAGCCC CTTCCTGCAG CAGTCTGTGA TCGACTATTA TCTCGAAAAG
AATGAGCTGT GGAAGCTCAA GGGTGTGCAC GAAAAAACCC CCGTGGTGCA CGTCAAGAGC
CGGGACGGTC ACTTGCTCCC GTATCTGCCG CATCTGCTCA AATTGACATG TTCATACGAA
CAGCTCTTGC CCAGCATGAC CAAGGAAGTC AATCGCCTGA TTAAGCTGAG CCCCAACGAG
AAGATGAGTA AGTTGTATAC GGAGATGTTT CGATTGCTCC GGCAGCAACA GGTGCTGACC
TTCAAGAAGG AAAACGTGCG AGCCGTCAAC CTCGGCTACG ATGTGAATGA ACTTGACAGC

CCGATCATGG AGTTCGGACA AGGCTACAAG ACAAACGAGA TCTATCGAGG CCTGAAGCAG
AGCGGAGTAT ACGAGCCCAG CTCAGTGGCC GTGAGCTTTT TTGTTGACCC CGAGCTTAAC
TACGACCCCC AGAAGCGGAA AGAAGTAGGT TGCTTCGTCA AAAAACTGGA GAGCATGAGC
GAGGCCCTGG GAGTAAAACT GAACATAAGC GACCAGCCCC GACAACTTTA TGGCCAGCTC
CCCAAGGACT TTTTCAAGCA GGACAACCTC TCATATCATT TGAAATCTAT CACCGACCAG
TTCAGGGGAA CGGTGGTGGT TGTTATCGGC ACTGAAGAGA ACATCGACCG GGCATACGTT
ACAATCAAAA AGGAATTCGG CGGCAAGGAG GATCTGATGA CCCAGTTTGT CGGCTTCACC
TCCTCCCTCG TCACGGAGAA CAACATTTTT CACTACTACA ACATCCTGCT CGGCATCTAT
GCGAAAGCTG GTGTTCAGCC CTGGATACTC GCCAGCCCAA TGCACTCAGA CTGTTTCATT
GGACTCGACG TAAGCCACGA GCACGGTAAG CACGCATCAG GGATAATACA AGTGATTGGA
CGGGACGGCA AGATTATCAA ACAAAAGAGC GTTGCGACAG CAGAGGCCGG AGAGACTATT
GCCAATAGCA CGATGGAAGA AATCGTCAAC GAAAGCATTT ATTCCTACGA GCAGATCTAC
GGGGCCAAAC CGCGCCACAT AACATTCCAT AGGGACGGGA TCTGTCGCGA GGACCTCGAT
TTTCTGCAAG CGTATTTGCG GAGTTTCCAA ATCCCATTCG ACTTCGTAGA AATCATAAAG
AAGCCGCGAC GCAGAATGGC GATATACTCT AATAAGAAGT GGGTCACGAA ACAGGGAATA
TACTACAGTA AGGGCAACAC CGCTTATCTG TGTGCCACGG ACCCCAGAGA ATCCGTGGGT
ATGGCGCAAC TTGTCAAGAT CGTACAAAAG ACTAACGGAT TGAGCGTTCA CGAGATAGTG
AGCGACGTGT ATAAGCTGTC CTTCATGCAC ATACACAGTA TGCTCAAGAC CAGGTTGCCT
ATCACGATAC ACTATAGCGA CCTCAGCTCA ACGTTCCACA ACCGGGGCTT GATCCATCCC
CGGTCCCAAC ATGAGAGAGC ACTCCCGTTC GTGTAGTGA

106 Gene Name: gene-18014

Gene ID: BBF.2017.48.106

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AAAAGCGGAA AGTTGAGGAC CCCAAAAAGA AACGAAAAGT CGGAAGCGGC
TCACTGGGGC TGAATAATGA GTCCAAAGAG TTCTTTAAGG GCATTAGCCG CATTTGGAGA
AATTACAAGG ACTACACCTA CCTTGACGGG ATTAAGCTGA GCCAGGCGCA GATCGATATC
ATCGAGAAGG AGGAGGACCA ATTGCTTATA GAGGGCTACG CCGGCACCGG TAAGTCCCTG
ACCCTTATAT ACAAGTTCAT TAACGTGCTG GTTCGGGAAG ATGGGAAGAG GGTGCTGTAT
GTGACTTTTA ACGATACGCT GATCGAGGAT ACGAAGAAAC GCCTTAGTTA TTGCAACGAG
TACAACGAGA ATAAAGAGAG GCACCACGTA GAGATTTGCA CATTCCATGA GATCGCCAGT
AATATCCTGA AGAAAAAGAA GATCATAGAC AGGGGTATTG AGAAACTGAC GGCTAAAAAG
ATAGAAGATT ACAAAGGTGC CGCTCTCCGC AGAATTGCGG GAATCCTGGC TAGGTACATC
GAGGGGGGAA AGTATTATAG CGAGTTGCCT AAAGAGGAAC GCCTCTACAA GACACATGAC
GAGAACTTTA TCAGGGAGGA GGTGGCCTGG ATCAAGGCCA TGGGCTTTAT AGAAAAGGAG
AAGTATTTCG AGAAAGATCG CATTGGGAGG TCCAAGAGTA TCAGGCTGAC GCGCTCACAA
CGCAAAACTA TATTCAAGAT ATTTGAAAAG TACTGCGAGG AGCAAGAAAA CAAATTCTTC
AAAAGCCTCG ACTTGGAGGA TTACGCCCTG AAGCTCATCC AGAACATAGA TAATTTCGAT
GACCTTAAGT TCGACTACAT TTTTGTGGAC GAGGTACAGG ATCTCGATCC CATGCAAATT
AAGGCGCTGT GTCTGCTGAC CAATACGAGC ATCGTGCTGT CAGGCGACGC GAATCAGCGG
ATTTACAAGA AATCTCCCGT GAAGTACGAG GAGCTCGGCC TCAGAATCAA AGAGAAGGGG

AAACGGAAAA TTCTGAACAA GAACTATCGG TCCACGGGTG AGATTGTCAA GCTCGCGAAC
TCAATCAAGT TCTTCGACGA GTCCATCAAT AAGTATAATG AAAAGCAGTT CGTAAAATCC
GGTGATCGCC CGATCATCCG GAAGGTGAAC GACAAAAAGG GTGCGGTGAA GTTCCTGATC
GGCGAGATCA AGAAAATCCA CGAAGAGGAC CCCTACAAAA CAATCGCCAT CATCCACCGA
GAGAAAAACG AGCTTATCGG CTTCCAAAAG TCCGAGTTCC GAAAGTACCT GGAAGGCCAG
CTGTACATGG AAAAATTCAG TGACATCAAG TCCTTTGAGT CAAAGTTTGA TTTGAGGGAA
AAGAACCAGG TGTTCTACAC CAACGGCTAC GATGTAAAGG GGCTGGAATT TGATGTGGTG
TTCATCATAA ACTTCAACAC GGCCAACTAC CCACTGAGTA AAGAGCTGAA GAAAATCAAG
GACGAAAACG ACGGCAAGGA AATGACGCTC ATTAAAGACG ATGTGCTCGA GTTTATCAAT
CGCGAGAAGA GGCTGCTGTA CGTAGCTATG ACCAGGGCCA AAGAAAAGCT GTATCTCGTG
GCCGACTGCA AAAACAGCAA CATCAGCAGC TTCATCTACG ACTTTAACAC CAAGTACTAT
GAGGCACAAA ATTTCAAGAA GAAAGAGATA GAGGAGAACT ACAACCGGTA CAAGATTAAC
ATGGAGCGCG AATACGGCAT CATCATTGAG GACGACGACT CCAACAACGT TAAGAACAAT
GACACGAAAC AAGAGAACAA GTTTAATACC GAATCTAAGG AAAAGGGCAA AGATGACATC
GACAAGATAA AGGTGTTTTT CATCAACAAG GGAATCGAGG TGGTGGACAA CCGAGATAAG
AGCGGGTGCT TGTGGATCGT CGCCGGGAAG GAAGCGATCC CTCTTATGAA GAAGTTCGGT
GTCCTGGGCT ATAACTTCAT ATTTATCGCA AACGGCGGTC GGGCATCTAA GAACCGGCCA
GCCTGGTACC TCAAGAATAG CTAGTGA

107 Gene Name: gene-18017

Gene ID: BBF.2017.48.107

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAAA AGAAGAGAAA GGTAGAGGAT CCCAAGAAGA AACGCAAGGT GGGGTCCGGC
AGTATGGACC GCGAGATCAT TGAAAACTTC AACCCCAGCG ACCCCAGGAC CGAGGGCGAG
AAGTATCTGA TGGATAACTT TTCAACCTCC CCCAGGTTTA ATGGCTGGAC AATATTTGAG
CAGCCCCACA TCAACTCAAT GAAGCCCGAC TTCATCTTGC TGCACCCCCA CAAGGGCATC
ATAATCATAG AAGTGAAGGA CTGGAACCTC AGCAGCGAGA CATATGAGAA CGGCGGTTAC
ATCTGGGGGG AAAACGGCGA GAGGATTAAG AAAAACCCCA TCAATCAAGT AGAAAACTAC
AAAAACTCTA TACTCAAGAT GGAACTTACA AACAGCATCG AATTTAGTGA AGTGTTCGGC
GACAAATACT TCGCGTGCAT AGAAACGGTG GTATACTTTC ACAAAGCCAA CAAAATTCAA
GCCGAGAACT TCTGCAGGAG GAACAATAAC TACACCAAGA TCTGGACCAA GGACGAGTTC
GACTACATAT GCAATATCAA TAACAAACTG AAGGGCAGTT GTCACACCTA TGCCCTGAGC
TACGAAAAAA GCACCCTTGA GGACAACAGA GGTATGCTGA GTAAACTGGT GGAGGAGCTC
AAGTGCAATC TCCAGTACAG TGACTACAAC TATGAACGAC GCCAACCGAT TAAGTTGACC
TATGAGCAAG AGAAGTTGGC GAGGCTGCAA AAGAATTCAA TCAGGAGGTG GAGCGGCGTG
GCAGGCGCTG GCAAGTCCCT GAGTCTGGCG CAAAAAGCCG TGAACGCCCT GAAGGAGGAC
CATAGCGTTC TGATCCTGAC CTACAACATA ACCCTGAGGC ACTATCTGCG CGATCTGTGC
TCTCAACAGT TCGGACCCGG CTCCTACAAA GGCGAGCGCA AGAAGCTGAG GAGCGACCTG
ACCATCTGTC ACTTTCATGA CTTTTTGAGA ATCATCATGG CCGAGTACGA GATCGAGGTC

GAACATGACG AGGACGACAA CTTCACCCAG CACTGGATAA ACAAGATCGA CAGTTGCATA
AAGGTGAACG GCATCAAGAG CCACCTCAAG TACGACTATA TCCTGATCGA CGAGGGCCAA
GACTTTGAAG GCGAATGGAT TAGGTTCCTG AAGCAGTTCT TCACCGAGGT GGGTGAGATC
TTTATCGTGT ACGACAAGGC CCAGGATCTC TACGAGCATG GCGTGTGGAT CGAGGACAGC
AACCAAATCA AAAACATCGG CTTTAAGGGC AAGCCCGGGA ACCTGAAAAT CAGTATGAGG
ATGCCTGAGA AGATGGTGTA CCTGGTGCAG GACATCAGAA ATGAGTTCAA GATAGATGAG
GAGGAGATCA CCCCAAACGT GAACAGCCAG CAGAGCTTCA TCGAGATAAC CAAGTGGATT
AACTGTATGC CCCTGACGCT CACTGAAAAG CTCGACCAGA TTGAAATACA GGTGGACTTT
CTGCGCCGAA ACAACAACAG CCTGGAGGAT ATCACGATCA TTACGACCAA CGAGGAAACC
GGAGTGGAGA TAGTGAATAG GTTCAAAAGC AGGGGTATCA AGACCAGCCA CGTCTACGAT
ATGGAGAAGC GGGGGAACCA GGCCAGGCGA AGGATGGAAA AATGGAAATT CCAGGGCGGC
ACCGGCAGAC TGAAGATTTG TAGCTATCAC AGCTATAAGG GCTGGGAGAC TCCGAACATC
ATCCTTGTGC TGGACGAGCC GAGCACAAAG TATGAGGACG GCATAATTAG TAAGGGGGAG
TATAACGAGA AGAACATTTT CGACGCTATC TTCATTAGCA TGTCCAGGGT GAAAAGGAAA
GCCCAAACCG GTGAGTTTAG CTTTACGTGC CTGAATTATC TTAGCGAATA CAATAAGATT
GAGGGCCTCT TCCACTAGTG A

108 Gene Name: gene-4715

Gene ID: BBF.2017.48.108

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGAGGAA GGTTGAGGAC CCCAAAAAGA AGCGCAAAGT AGGTAGCGGC
TCCATGCTGA CCAATAATCA GATTGTGCTG GAGCAGGAAC TTCTGGGAAG CATATTCAAA
AACAATAACC TGATGCTGAA AGCCCGAGAG AAGATAAAAC CGGAGATGTT CCTGTATAGC
AAACACATGA ACATTTACCT GGGCATCCTC GACATGGTGG CCAACAAGCT GGAGGTGGAC
CTGATCACCT TTCTCGAGCA CCATAAGAAA AGGGTGGGGG ATATGGATGG CGTAACTTAC
GTGACCGAGA TCTACACATG CAGCGCGTCC GACATTGGCT TCAATACAAA ACTTGACATG
CTGGTGAACA ACTACAAACG GCATCTGTAT GTGGAGATGA AGGACAAAAT CAACAGTGAT
ATGAGTCTTG AGGAGATCGA GAGCGAGGTT GAAGGGGTGA AGGTAAAGGT GCACAAATGC
AACATCAAGA AAGAACTGGA TATAGACAAG CAATATGACG ATTACATCAA CTGGCTTTAC
GACGAAAACA GAGACAAGGG GATGAAAAGC GGCCTGACCT ATCTGGACAA GTATCTCGGC
AACTTCCAGA AGGGCAGGCT CGTCACCGTG TTCGCCAGGA GCGGCGTCGG CAAGACCACG
TTCAGCTTGC AGCTGGCCGC CAATATGGCT CTGAAGGGCC ACAAGATATT CTACGGGAGC
GCAGAGATGA CCCGCAACCA GGTCTTTAAC AGGATCGTGG CCTCAGGTTT GAGCCTTAGC
GCGAAGGCGA TTGATGAGGA CACCATCCTG AAGGAGGACA AGGAGTCTAT CGCCAAGTTT
ATGACCAAGG TTATCAACAA CAAGTTCTAC GTGTCAACCG AAACCGACTT CGAAAAGTTC
ATCGACGAGA TAAAGGTTTA TAAGCTGCAG AACAGTCTGG ACGTGGTGTT CGTGGACTAC
ATTAACAAGT ACATCGACTT CACCGACAGG GACATGTTGA CCAACAAACT GGGGAAGATC

AGCGGCATGC TCAAGAGCCT GGCCATGGAA GAGGATATCT GCGTGGTGCT GATGGCCCAG
GCCAATAGAG TGATTGACAA GAAGGTGGGT GACAATGCCG TCGAAAAAAT CGACAGCAGC
GACATCCAGG ACAGCGCCAG AATCGAGCAA GACAGCGACC AAGTGATCGG CCTGTACCGG
AACGTGAAGC TCGATGATAA AATGTATAGG GAGAACCTGT TCAATCAGGG CAAGCTCAAG
TATAATTCCA AGAACGCCGA CGACAATCCG GAATGCATGA ACGCTGTGAT CATTAAGAAC
AGGCATGGCG ACCGAGGCAC GTGTGCACTG AGGTGGCACG GAAGGTACAG CAGGGTCAGC
GACTTCTAGT GA

109 Gene Name: gene-8494

Gene ID: BBF.2017.48.109

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAGCGGAA AGTCGAGGAT CCAAAGAAGA AGCGCAAGGT GGGTTCCGGG
AGCAAAGGGC GGCACCAGGC GAAACACTAC GCGGACGGCC TGGAAAAAAT GCACGGGCAA
AGGCCTGTGA TTTTCTACAC CAACGGCCAC GATATATGGA TATGGGATGA CCATCCGGCT
CAGCACTACC CGCCCAGACG GTTGTACGGA TTCTACGCGA AGTCCAGCCT GCAGTATTTG
ATAAGGCAGC GCAGTGAACG CAAGGCGCTG AATACGGTGA GCTCTAAAAC CGATATACTC
GGAGAAAGAC TCTACCAGCA CGAGGCACTG AAGCGGATCT GCGAACGCTT CGAAACCAAG
CAGAGGAAGG CACTCGCAGT CCAAGCGACC GGCACGGGGA AAACCCGCTT GAGTATCGCA
CTTACTGACT CTTGCATGAA GGCCGGGTGG GTGAAAAGGG TGCTTTTCCT GTGCGACCGA
AGGGAACTTA GAAAACAAGC TAAGAACGCC TTTAGCGAAT TCCTCAGCGC GCCTATTAGC
GTACTGACAA CGAAAAGTGC GCAGGATACC CACAATAGAA TCTTCGTGGC AACCTACCCC
GCAATGATGA AGGTGTACGA GCAACTGGAT ACGGGATTCT TCGACCTGAT CATAGCCGAC
GAGAGTCACC GAAGTATTTA CAACATCTAC GGCGACCTCT TTCGCTATTT TGACGCCCTT
CAAGTGGGCC TGACCGCAAC CCCCGTGGAG ATGGTATCTC GGAGCACATG CCAACTCTTC
GGGTGTGACT TTAAGCAACC AACTTCTAAT TACACACTCG AAACGGCTGT GGAGGAGGGT
TATTTGGTGC CCTACCAAGT CGTGAAACAT ACCACAAAGT TTCTGAGGGA TGGGATCAAG
GGCCACGCGC TTAGCGCGGA GGAACTGGCG GAGCTGGAGG ACAAGGGCAT CGATCCTAAC
ACTCTTGATT TCGACGCCGA GCAGATCGAC CGAGCGATCT ACAATAAAGA CACCAATCGG

AAAATCCTGC AGAACCTCAT GGAGAACGGT ATCCGGCAGG CCGATGGCCA GACCCTCGGT
AAGACGCTGG TATTTGCTAG GAACCACAAG CACGCCAAAC TCCTCGAACA GTTGTTCGAC
GAGCTGTACC CCCAGTACGG CGGTAAGTTC TGTCAGGTTA TAGACAACTA CGACCCCAGG
GCGGAAGAGT TGATAGACGA TTTTAAGGGC GAGGGCAGCA ACGAACAGCT CACTATAGCA
ATCTCAGTCG ACATGCTCGA CACCGGGATT GACGTCCCGG AGATCGTAAA CCTCGTATTC
GCACGGCCGG TTAAAAGCCC CGTGAAATTT TGGCAAATGG TTGGTCGGGG AACGCGACTC
TGTAAGAATT TGTTTGGACC CGGCAAGCAC AAGACGCACT TCCTTATTTT CGACCACTGG
GGAGTCGTGG AGTATCACGG CATGAAACAA CGCGAGGTAA CTGTGTCCCA GAGCAAGTCC
CTGATGCAGC AATTGTTTGA AAATAGATTG GAGCTCGCCA AGACCGCGTT GCACCACGCC
GAAGCCGACT TTTTTGAAAC GATGGCGGGG TGGCTGCACA AAACGATAAA TAGCCTGGAC
GATCGAACGA TTGCCGTTTG TGATAAGTGG AAAACTAAGC AGCAAATGTC CGACCTGGAA
ACGCTTAGAC AGTTCGGTGC AAACACCGTC ACGCTGCTTG AGTCAGAAAT CGCCCCGTTG
ATGCAATGGC TGGATGTCAG AGGGCATAGT GACGCATATC AGTGGGACCT CCTTGTCTCA
CAGATCCAAC AACAAAAATT GAAGCAGGCG GCAGCCTTCG ATGATCTCGC TGGGAGGGCA
ATCAATCAAC TGTGGCAGTT GCAGATGAAT TTGAATCAAG TTAAGGCAAA GTCCGAGTGG
ATTAAGCAGT GCCGAGAAAC GGAGTGGTGG CAGAAGGCGT CCCTGGATGA ACTGGAACAA
ATGCGACAAG AACTGCGGGG CATTATGCAG TACAGGAACA AGGGTGACAT TCCTAAGACA
GAGGCGCCCA TCATAGACAT AACGGACTCA GAGGAGGTGC GCGAGAAACA ATCCTCCTAC
CTGAACTCAG TTGACATGGT CGCGTATCGG GTCAAGGTTG AACAGGCGCT CCAGGAGCTC
TTTGAGAGAA ACCCCATCCT TCAGAAGATC CGGAACGGGG AGGCCGTGTC TGAGCGCGAG
CTTGAGAACT TGAACGCTCT CGTGCATACA CAACACCCGG ATATCGATCT CAACACACTT
AAAAAGTTCT ATGGGACCGC GGCTCCGATG GATCAAATCC TTCGGACAAT AGTAGGCATG
GACGGGAACA CGGTTAATCA GCGCTTTGCG GCGTTCATAC AACAGTACCC CTCACTGAGT
GCGCGCCAAG TTCAATTCCT GTCCCTGCTG AAACGACAAA TTGCTCAGAG TGGGGCCATA
GAGATTGACA ACTTGTACGA AATGCCATTC GCAGCTATCG GCGAACCCGA CAGCGTATTT
AGTAACGCGG AACAGATTGA TGACCTTCTG GCGATTGTGG AGAGCTTCGG GAAGCAGCCC
CAGCAGCAGT CTACGAGACA GGCCAATGAG ACATAGTGA

110 Gene Name: Archaea-2

Gene ID: BBF.2017.48.110

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAACGGAA GGTAGAGGAC CCCAAGAAAA AGCGGAAAGT TGGGAGTGGA
AGCATGCCGT TCAATAGCAA CCTGATCTTC GTGAAGCTCG ACGACCTCAA GAGAGCCTTT
CTCGAGGGCG TCCACAGTGG TCACGCCGTG GTGTATGAGG TGAGCGAGGG ACTGAGCACC
GAGGATCTGA AGAAAAGGCT TATCAAGGCC AGCGTGATGT ACCACTATAG GTATGGAAGG
AACGTGTTTG TGTTCGGCGT CAAGGAGGGC ACTAAGGTTG ACGATCTTGT ACCAGGCCGA
CGACTCGGCG AGCACGAGGT GAAGGAGGTT CTCAAGGGCA TCCCGTCTAA CAACCTGGTG
TCCATGATGA GCGCCATGCT CAATTACCAG CTCTCTGTGC TTCTCACCAG CAAGGGCTTC
CAGTATAGCT ACGAAGAGAT GCGGAGGGGC AAGTATCTGT GTGTCAGCAA CTATTACGGC
AAGCTGATAC GGAACCCCGT GAAGGTTTGC CTCAAGGTAA ATGTCATAAG GAGCCTCATT
GACGAGCAGG ATCAGTATCT GCCAATCGCG CTTAACTACA GGGTGAAAAA GAGCAGGCGG
CTTAGCCCCG AAGTAATGAA TGAGATCCAC GCGGAGTTCA TGGAGGCCTT CCCCAGCTAC
CTCAACGACC TGAAAATCAT AACTCGCGTC TTGAACGACG ATATGGTGAG GAACAGGGAA
CTGAAATTCC TGGAGATCGA GTACAAACCC CCTGCTATCA TTACGTTCCG GTTTCGAGGC
AACAGCACCG GCGAAAACGT GACCGACATT CTGAAGCTGG GCCCCTACTT CCTGCCTGGG
GAGGAGGAGA AGATCGATGT GGTCTTTGTG TACGAAAATG CTCTCGCTAG CCAGGCGAAG
AAACTCACCA AGGTTTTGGA GGATACCATC AAGGACGGGC TGGGCATAAA GCTGAACATA
GACGACGAAC ATAAGTTCAG CCACGACAAG CCGCTGGGCG ACGTTATTAA GCTGGTGCGC

GACCGATTCA TCAACAGCGG GAGTTGTCTG CTGGTCCTTA GCAAGGAGAA CCGCCTCGGT
CCTATCTTCA TGAGCATTAA ACCGCTCACG CTCAAGAAGA ACTTCTACTT CAAGTCTCAA
TTTATCACCA ACGAAACGAT TAGCAAACTG GACTCTTATG CGGTCAAAGC CAATATCGTG
AATAGCATCC TGTTCAGGGT TGAAGGTACC CCGTACATGC CCGTTCTGCG GGGCAATATA
GACGTACTGG CAAACAATTT GTTCGTGGGT ATCGCCCTGA GTAAGCCTCT GAGGAAGGGC
TACACCAAAG GAGGCATAGC CCTCATAGAC CCCTACAGCG CCCGAATTAT CACAAGGGCC
ATCGTGTTGA AGCGCAAGAT GAGGAGCGGC AAATTCGAAG CCTCAGACAT GCACGAGATC
GTGTCCAACA TCAAAGGCGT GCTGAAGGAC TACAAGGAGC TGTACAACGT CAACGAACTT
GTTATACATA TCTCCAAGTT TCTGTCTGAT GACGAATACG GCCTTTTTTA CGAGTACTTG
CAGGACCTTA ATGTCAACGT GCGACTCCTG AGCATCAGGA AGAGGGACGA CATTACACTG
GTTAGGGACG GGAGGATGGA CAGCCTGACC ATGATCAAGC GCGGCAAGAG TCATGTCGAG
GTCATGTATT GGCCTCACGA AAGGGCCTAC CACCCCCTTA CTATCAGGAT CTACGGCGAC
AATGTGGACA GGGACGTGAT GATGCGACAC CTGAGGTTTA TCGAGCTGCT CCGGCACATG
TACTACCCGG CCAGCAGCCG CTTCATAGTT GAGCCCGCGA CCATTAGCTA CAGCAGGAGG
GTCGCCAGAT TTGCCCCCTG GCTTTCAGAC AATACCTAGT GA

111 Gene Name: Archaea-4

Gene ID: BBF.2017.48.111

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAGA AAAAACGCAA GGTGGAGGAC CCAAAGAAGA AGAGGAAGGT CGGAAGCGGC
AGCATGGAGA AACAGACCTT CTACCAGGGC AACATGTACA GGCTGAAGGA TGAATTGATA
CAAGATATCC TCTCTGACAT TATCGTGGCG AGAGTAACTA ACATGCCAAG CAATCCCGAA
GAAGCCTACA GTGAAATACA GAAGATTGGC GGCATTATAC TCAATTACGA TGAGATGACC
AACAGCGCCT GGGTGGTGGG CAAGGAGTCT CTGCTGCAAA ATCACTATCC CGACGACATG
AAGGAGGTGC GAGCCTTCTC CTTTTCTGAG CTGTCCAAGG AAAACAAGAC GAAACTGGTC
CTTAATATCC TTAACGCCGA GGGCTATCTG CGCGACATTA GGGGGCACCG AGAAGTGGTG
AAGTCAATCA ACTCAGAGCG ATCAATCATT AGAAAATTCT TGGTGACGGT CGAGTACGAT
GGTCAACACT TCTATCTCGT AACCCTCCCA AAGTATAAGA TCATAGAGAA TCACACAATA
ATGGAACTCC TCATTGAGGG CAAGATCACC GTCAAAGAGC TCGTCCACAA CCTCCTCAAG
GACCCTAAGT GGAAAATCCA GACCAGTCGC AAAGATGTGC CCCTGCCTCC TGGGCACAGG
GTCGTGGAGA TCATTCTCAA GACTAAAGAT CCCGATCGAT ACCAGCAGGA ACTCGAACGC
ATCAACGAGT ATTTTACTAA AAAGACGGAA CTGGGGCCCA TTGACGATAG CAAGTATCCA
GATGATTATA ACATCATTTT CAGAAGCCAG ACGCGAGGCA AATACTTGAG CTATCACAGT
GCGCGGACCA AGCTCATCAG ACCGATTAAC AAAGAAATCC TCCGAGAAAT CTACAGGAGT
AACGAATTTA TCAAAGCACT GAATATCGCC AAAAAGCTGG TGGCCGACAT CATATACGAC
AGCACCAAAT ACCCGGGCAG GGCCATATTC CCCGCCTTTA AGATAGACGA ACGGACGATC

TCATACAAGG CCGTGTTCCT GAAGAATAAG ACGATAACTG AGAAAACCAT CCAACCCTAC
TACAATATCA AGGGTACCTT TAATTGGCTT TTCACCAACA CGCCGTTCGA CGATATTAGC
GAGCTGATAA TACCAATCCA GTCCCCCGAG TTCTTGAGGG ATAAGACCAT TGGAGTGTAC
ATCCTGTACC CTGCGAAGTA CAGAGAGAAC TCCGAAAGCC TGAAAGTGAT CCAGAATCTT
ATCAAGAGCG TAGATAGCAC GATCAAACGG CTGAGCGAGT ACTTTACATT CCTTCGAAAA
GTCAACGAAG GCCTGTCTCT CCCCTCTGCT ATAGATATCA TCTCTCGGAT CCCGGTTAAC
TATGAAAACT TGATAGAGAG TGCGTTTACC CGGATCCACA GCAAGAAGGG CGTTGAATAT
GACTACCACC TCGCGATAAC ACTGATACCT GACATGCGGC AGGAGCAGTT CGATAAAATC
AAAGGGTTCT TTTTCAATAA CGGGATTCTG CACAAGGCAA TAAACATCAA TAATCTGAGG
GACCCCAGCA AAGACCAAAA GAAGCTGATT GAGAGCATGA TCCTCCAGGC ACTGTACGCC
TTTGGCATCT ACTTCTACAG CCTTGACAAC CTGAACTACG ACTTTATCAT AGGACTCGAC
GTGACCAGGG AAATGGACAA GTCTGGTAGG TACTACGGTA TATCCGGAGC CGCGGTGGTC
CAAAATAAGA ACGGCCAGGT ATTGAAGATT ATACCGATCA CCAGCCCCCA GAGCAGCAGC
GAAACCGCAA ACATTAACTA CCTCATCGGC AATATCCAAC AGGAAGCCGC TGCAATCCTG
AATCGGAAGG GATACGCGGA CATATTGTTC CTCAGGGACG GCAAAGTGCC CGGTGGCGAA
CTGGAACAGT TTAAAGAGAT CAGCCGCAAG TACAACTACA GGTTTACTAT AATAGAGATC
CTCAAACGAC CCCTTGTCCG CTTTTTCTGG GAGAATTACA AGGAGCACAC CGTTAAGAGC
CCTAGGCATA ACTACTACTT CAAGATAGGC GACACGTATT ACTTGACCGC GCATTACTTC
ACGAATTACC TGAAGGTCCC ACTCAAATTG GGTAATACCT ATTTCGTGGC CCGAGGAAAG
ATAAGTAAAA ACGTGATTAG CCGCGAGGAC ATAATGACAA TCACAAAGCT CACTAAGCTC
AACTATAGCC AGCCCGAGAA CCCGGACAAA ATGAAGCTGC CTGCCCCCGT GCACCTGAGC
CACCGACTGA TCAATTATGA GAGGAGAGAG CTTAAGTTCA ACAGGTATGA GTTTCTTAAG
GAAGGAGCGC TTTATTTCCT GTAGTGA

112 Gene Name: Archaea-7

Gene ID: BBF.2017.48.112

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAACGGAA GGTGGAAGAT CCAAAAAAGA AGCGAAAGGT TGGTAGCGGC
TCAATGGCCT ATAGCCTTAA CGCTTTCGAA CTGGAAATTC CCGACATTGA CGCCGACCTC
TACAAAGTTG ACCCTCAACC CTCTGATGAC CCATATCGAA TCCTGGGGGG TTTGGAACGG
TCCTTCGAGC AACAACTGGA CGGCAAGGCC CAGAAATGGA AACAGGCGGA GGACGGAGAT
TGGTATATCG CCGTGATAGG CGCGTCAGAA AGGAAAACTA TCGAGTCCCC CTCCAGCGGT
ACGAGGGCAG GCTACACCAC CACGCATACG CTGGATCCGA GTAGCTTTTG GGACAGGATG
GTGTTGCAAA GGGCAATTAG CGACTCTGTA CGATGGTACA TGACCAACTA TCAGGACTTT
TGGTATCATG AGGATGCGGA TGCACTCTTT TATCCTTCTC CTAGAGGCAA AGTGGACGAG
TACGACGTCT ACACCGGATT TAGTCATAGG GTCGAGTTTT ATGACAGCCC ACAACTTGTC
GTGCGCAGCG TCACTAAGTT CATCTCCAGT GAAAGCCTGG CGGACCGGAT CAACCATCAG
GGCACAGAAG AAGCAACGGA AAAATACGGT GGTGAGAACT TTAGGCTGGA CAGGCCGGAA
CCAACCAAAT GTACTTTGCA CGGCATCTCA ACCGAGCGAA CGGTAAGTGA CAAGACGATA
GATTTTGGTG ACGAGATGCT GTCCGTGTTG GAGTTTGCAC AAAGAAAATA TGGCAGCGAG
TGGGCGGACA AAATCGATCC CGACGAACCA TTGGTGCAGA TACGCTTCGG GAACAGCGAC
CCCTACGACA CCGCTCCGAG CCTGCTGAAT GCGAGCCCTG AGGAGCTGAA TCGCAGGCTG
ACCAGCGAGG CAGCCCTCAG CGCACAAGAA AGGCAGAAGG CCATACAGAA CTTCATCGGC
AGGATACACT ACATCCAGGT TGAGGACGAG AAGGTGAGTG TCTCCGATGA CGGCGTACGG

CCCACCGAGC AGGGCGACTT CGACTACCCC GATCTTGCGT TTGGCAATGA CGAGGTGCTC
AGCACCGGCG TCCCGAACGC GGTAGATCCT AGCCAGGAGG TGCACCCGGG CAACTGGCGC
TGGATAATCA GGGACTACCT GGAGGAATAC GGCTTCTGGG AGTCACAACG AAAGCTGTCT
GAGATCGTGC TGGTGTACCC GAGAGGCGAA GAAAGACGGG CAGAGAACCT GTACCAGGAC
GTTAGGGAGA AGCTTTCAGA GATAGGAGGC GTTCAGATCA GGAGCGATCC ACACCGCGTG
TGTTACACCG ATCAGGTGGA GTTCGACGAA TGGGTGGCTG AATTCGGTGA CTCAATCGAC
GGTGTTCTTG GATTGATTGA GGGAGATGGG GACGAATACT ACGAAATCAT AGATGCATTT
GGCGGAGCAC CGACCCAGTA CGTCAACACT AGCACCTACT CAGAGCACAG AGGGGCGAGC
GACGACGTGA TCTTTAACAC TGCTTGCGGA CTGGCCGTGA AGTTGGGCGC ATATCCTTTT
GGCCTGGCCA ACGACCTGAA CAGTGACGTG TACCTCGGCC TTAGCGTGGC AGGGGATAGA
AGCACAACGG CCACCGCCGT TGCCATAGAC GGAAGAGATG GGAGGATTCT CTATCAAACA
GAGGAACCCC TGGGCCAGGG TAGCAGCACA GTAAGCGAGG GCTATCCCGC TAAGCGAATC
ATCCAGAGGA GCCTCAAGAC CGCCTCAAGC GCCTTTGATC GACCAATCGA GAGCTTCGAC
ATTCACAGGA ACGGAGACTT TGGCGACGCT GAGCTGGAAA CCCTTAGCAG TGAATTGCCT
GCACTCCAGG ACCAGGAATA TGTGCATACC GATGTTTCAT GGAGCGCCGT CGAGGTAATT
GAAAACCACC CTTACAGGCT CTTTAGTGAA CGGGGCAGCA GAGCTCCCGA TACCGGAGCC
TATGCTAAGC TGGACGACGA GCATGTACTG GTTACTACCT TTGGAGAGCC CCAGATCCAC
CAAGGTACGC CAAAACCGGT CCTGTGCAAG AGGAGAGCAA CGAGCCAAGA TCAAGACATC
ACCGCCATCG GAGAGGACGT GTTCAAACTC AGCTTCCTTA ACTGGGGTAG CCCAATGATG
AAGATGAAGC CACCTGTTAC CACTAAGATT CCGAAGGAAC TCAACGAGAT TTTCGAGAAG
TGCTCTAGGG TGAGATATCC CCCCTTCTAG TGA

113 Gene Name: Archaea-8

Gene ID: BBF.2017.48.113

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAAAGAAA GGTAGAGGAC CCGAAGAAGA AGCGCAAGGT CGGCTCCGGA
AGCATGAGTC AAGACTCTAG GAGCACCGAG GTGGAGAGGC AGGCCGAAAT ACAACCTGGT
ACCTACCTGT TGAACGGCCG GGGGGAAATT CAGTTGGATG AGGTTGACGC ATTCCAGTAC
GACCTCAAGG TGAGTGGAGG CGTGGAGCAG TATTGGGATC GGGAACAATT CACCAGCTCT
GCAGCCTACT ACCTGGACCA GGAACACGGG AGCCCTGTCG CTGAGATAGG CAAAATGAAC
GTGCTCAGCA AGACGGATTT GTCTAGATCA GTTAGAGTGT GGCAGAGAAA CGTGACTCCC
ATCAATAGGC AGAGCGTTAC ACTGACCGCA GCCCAACCCG AGGACCGAGA AAAGATCAAA
TCATTCGTGC AAAGCTGCTT CAAGAGGGCA GTGCCGACCG AAAAATACAG CTTTCGCTTT
CTCAACAAGA TTGTCAGGGA TGAGCCCGAG TTCACCACCG GCAGCGAAGG CTTTTCTGCA
CATCCGAAGC ACGACGTTAA GATACAGGTC ACCGCTGATG GCAATGTGCT TGTGCACGTG
GATAGCGGGT TCAGCATCAG GAGCAACAGC ACCCTGGACG AAATCTACTC TGAACAGGAT
AACCCTTACG GTAAGCGCGT TGCCCACGAC CCCGAGAGGT ATGGTACCCA GGGCCAAGGC
ACCCTTCGCG GTTGGAGCGA CTATCGGTAC ACAGACCATA TTTCTGATGC GGGTAGCTCT
GTGAACGAAA TGCACAAAGG GGTGGCGGAC GAAGAATGGC GGCAACGACT CGCAGAGGAG
AATCCCCGAC TTCTGAAAGT GGAGTATGGC AACAAAACTA GGAGGCAAGC CCCCCATTTC
CTGAGGCTCT CACCGCGGAT CGAACAGGTG CAGGATCAGG ATCGCGAGTT CTATTCAAGG
TTTAACAGCC GGAGCGCTAT GATGCCCGAC GAAAGATTTG AACTGTCTAA AGAGTTCCTG

CAGAACGTGA GCCGCTTGCC GGTATTGGAC ATGGAACTCG AGCCGGGTCC GGTGAACAGC
AGTTACGAGT TGCTGGAAAT GCGAGAGGAA AACAGGCTGG TTTTTGGAGG GAAGCAGAGG
GCTCGGGACC CGGGCAGCGG GCTTAGAGAG AATGGGGTGT ATCAAAGTCC CAGTCAGTAC
CGGCTGGGGG TGTTGACCCC CGAACGATGG GGAGAGAAGG CGAGCGAGCT GATCCCCCTG
ATTGTGTCCG GCCTGAACGA TCTGAGCGCA TCAGCAGGAG TTCGAGCATA TGGATACGAA
TTGGGGGACG TCAGCAATTA CACACCCGTG GTTCAGGACC TCCACGAGGA AACGGACGCT
GTGCTCGCCG TGGTCCCCAA TAAGGGTGTG GCCGAGGATT TTGGGATAGA CGATCCATAC
AAGGAGCTGA AAAGAACCCT CCTGCGGAAA GGGATACCCA CCCAAATGAT GCAAAAGTCC
ACGGTCGATG AAATCGTGGG TCAAAAGGCG GGAATCGGCA ATGACAAGTT TCTGAACGCA
CTTAGTGCAG TCGTGGCCAA AGTGGGCGGT ACCCCATGGC AGATCGATAG CCTCCCCGGG
AAAACCGACG CCTTCATGGG CTTGGACGTA ACTTACGACG AGAGTAGCGA GCAGCACGCA
GGCGCCAGTG CAAGCGTAGT ACTCGCGGAT GGGACGACTT TCGCAGCCGA GAGCACCACC
CAGCAAGGTG GCGAGAAGTT CAGTGCACGG CATGTAGAAC AGTTCGTGAG GGACCTCGTT
TTCGACTTTG CGGGGGAACA GGGCCGAGAC ATCGACAGAC TGTGCATAAT GAGAGATGGG
AAGATCAGCG AGGATATTGA CGCCGTAAGA GAGGGACTCA GTGGTATTGA GGCGGAGATC
GACATAGTTG GCATACGAAA ATCCGGGCAA CCTCGCATAG CTGAGTTTGA CGGTACTCGG
TTTCGGATCG CCGAAAAGGG CGTGGGCTTT GTGGACGCCG ACAGAAGCCA GTCTATCATC
CATGCATTCG GCAAACCCGA AATCCACGAC GACAATCCTG TGGGCACCCC ACGAACCTTT
CGACTGACCA AGGACTCTGG TCCCACAGAT GTGGAAACCC TGACCCGACA GGCATACTGG
TTGTCCGAGA TCCATTTTGG AAGCCCCGTT AGGTCCCCTA GGCTCCCCGT GCCAATAGAG
TACGCAGACA TGGCTGCTGA GTATGTTCGG GAGGAGTATG TCTCACCAGG GACTGTAATA
GAAGGGCCAG CATACATCTA GTGA

114 Gene Name: Archaea-6

Gene ID: BBF.2017.48.114

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAAAGGAA GGTTGAGGAC CCGAAGAAGA AACGCAAGGT CGGCAGCGGA
AGTATGAAAA CGCAGGATGA TATCGCGCAC AAGCAACCCA TTACCATCGA GGTCCAGATC
CTGAAGGAGC TCGACAAGCC AAGCCCAAAA ATGGCCACCC GGTTCCTCGT GGCCGATAGG
GACGGCAACA GGTTTAGCCT GGCTATCTGG AAGAACAACG CACTCAGCGA CTATGACTGG
ACGATTGGCC AGTGGTACAG GCTGGAAAAC GCCAGAGGAA ATGTCTTTAA CGGCAAACAG
TCCCTCAACG GTAGCAGCAA AATGCGCGCC ACTCCACTTG AGGCCAGCGA GGAGGACGAA
ACCAGCACGG ATGATGTGGG ACGGGTCGAC ACAATCCTGG GTAATATGAG CCCGGACCAG
GCTTACCTGA GCCTGTTTCC CATCAGTAGG TCTTTTGATA CCCTGTCTGT GTACGAGTAC
AGCATTGAGG CAGCCGAGGC ATTCGAGGAT GCGCCGGACA CCGTGACCTA CAGGTGCGCT
GGCAGGCTTC GGAGAATCAC GGGTGCGGGG GTCGCTTATG CTGGCTCAAT GAGGATCGTG
TCAACCCGCA AACTCCCGGA CAAGCTCGCG GACCCCTTTA GCTTGAGTGA ACCCACGGAG
AGGGAACTGA ACGCTACGGA CGCCAGGGAC AGGCATAGGA TAGAGCGGCT TCTCAAGAGC
CTCGTGAAGG CCGCCATCGA CGATAGCACC TACGACCCAT ACCAGATCAA CCGAATCAGG
GCCAGGACCC CGAGCATTAC CGCTGGCGAC GGGCTGTTCG AGGCGTGCTA TGAATTTGCA
GCAAGGGTCG ATGTGATGCC CTCCGGCGAC GCCTTCGTGG GAATTGAGGT AAGGTACCAC
ACGCGGAGCC AGGTCACTGC AGACGTTTAC GAGGACAAAA CCGCGGAACT GGTGGGCACC
ATCGTGGAGC ATGACCCAGA GAGGTACAAC ATTAGCGGTA CGGGCCGAGT AGTGGGTTTC

ACTGACCACC ACTTCACCGA CGCCCTCGAC GAATTGGGCG GTCTTAGTTT GGCGGACTGG
TACGCGCAGA AGGATCGCGT CCCAGAGGGG GTATTGGAGG CGCTGCGAGA GAAAAATCCT
AGGTTGGTTG ATATTCAGTA CCAGGAGGAC GAACCAGCCA GAATCCACGT CCCGGATTTG
CTCAGGGTAG CACCCCGCAA GGAAGTTGTC AAGGAGTTGG ATCCCGCCTT CCACAGAAGG
TGGGATCGAG AGGCCAAGAT GTTGCCCGAC AAAAGGTTCA GGCACGCCAT AGAGTTTGTG
GATCATCTCG GGTCCCTGCC GGATATAGAC GCCACGGTGG CACCCGAGCC TTTGGGGCCG
TCACTGTCTT ACATGAGCAC AGCAGTCGAC AGGGAGAAGA ATCTGCGCTT CAAAGATGGA
AGGACCGCCA CCACCCCGTC AAGCGGCATC CGGAGCGGCG TATACCAACA ACCGACGAGC
TTCGATATCG CCTATGTGTA CCCCACCGAG TCTGAACAGG AGAGCAAGCA ATTCATTTCT
AACTTCGAGA ACAAACTGTC CCAGTGCCAG TGCGAACCAA CTGCCGCTAG GCACGTTCCT
TATGAACTCG GCGGCGAGCT GAGTTACTTG GCTGTCATCA ATGAACTTGA GAGCGTGGAT
GCGGTGCTCG CTGTGGTGCC TCCCCGCGAC GATGACCGGA TAACGGCCGG AGACATAACT
GACCCCTATC CCGAATTCAA GAAGGGCCTC GGGAAGCAGA AAATACCCAG TCAAATGATC
GTGACCGAGA ACTTGGGCAC AAGATGGGTG ATGAACAATA CAGCCATGGG CCTGATCGCA
GGGGCAGGAG GCGTTCCGTG GAGGGTGGAT GAGATGCCGG GTGAGGCCGA TTGCTTCATA
GGACTGGATG TGACTCGCGA CCCGGAAACC GGCCAACACC TTGGCGCTAG TGCCAATGTC
GTTTATGCCG ACGGAACCGT TTTCGCCTCT AAAACGCAGA CCCTGCAGAG TGGGGAAACG
TTCGATGAGC AGAGCATAAT CGACGTGATC AAGGATGTAT TCCAGGAGTT CGTTAGGCGC
GAGGGGCGAT CCCCTGAACA CATTGTTATC CATAGGGATG GCCGGCTGTT TGAGGACGCC
GACGAAATCC AGGCCCCGTT CGCGGATAGC GGAGTGAGCA TAGACATTCT GGACATCAGG
AAATCTGGCG CTCCGAGGAT TGCCCAATAC GAGGACAACA GCTTCAAGAT TGACGAGAAA
GGCCGACTTT TCATCAGTCA AGATGACACG CATGGATTTA TCGCCACAAC GGGAAAGCCG
GAATTTGATG ATAGCGACAA CCTGGGCACT CCCAAGACTT TGAGGGTAGT GAGGCGGGCT
GGTGACACAC CGATGCTGAC TCTGCTGAAA CAGGTGTACT GGCTTAGCGA GGCACATGTT
GGCAGTGTGA GCCGAAGCGT TCGCCTGCCT ATCACAACTT ACTATGCAGA TCGCTGCGCC
GAACATGCGC GGGAGGGGTA TCTGCTCCAT GGCGAGTTGA TCGAGGGTGT GCCATATCTG
TAGTGA

115 Gene Name: Archaea-1

Gene ID: BBF.2017.48.115

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGCGCAA GGTAGAGGAC CCAAAGAAGA AACGGAAAGT GGGAAGCGGC TCAATGGAAG TGTCCCCCTT CTTCAACGAA CTGTTCAAGT ACTACATATT TCTGTTTTTT GGTTTCAAGG TGAACATCGT GAAATCACAT TACCAGAGCA TTAAGAAGCA CAAGATAATA TTCTATTCCG GTGGGATCAT GGACGAGTAT TACACTAACG CCTTCCCCAT CAACAAATAC TTTATCAACC GCATCATCTC TGAAAACTGC ATCCGCTGCC TGTGCAAAAT AACCAAGCTC GAGAAGAAAG AGAAGATCGA GGAGTTGCTT TACTCTATCA GCGCCACCCT GGGGGGCATT TACATCGACG ATTACAACCC AATGAAGAAT AAGTTCAGCT TCTACATTTG GAAGGGAATC CTGAATAAGA AGATTAAATC CTACGGGTCT GAATGGCTCA TTAACAAGAT GAAAAACATG GGCTTTAAGG ATCCGGAAAA CAAGACGCTG TTGAACTATG TGAAGAAAAA GTACGAGAAA GACATAAAGT TCGACATCAT AAAGAAAGAG AAGATAGAAT GGAGTAACCT CGACTGGGAG ATAAAGGAAA AGATAGTGCT GGGCGCCATA AAAACTCACC CTACCATTCG CAAACTGATT GAATACAAGA ATGAGAAATT CATTGACAAA ATTGGAAAGA AAATTCTGAC TTACTTTAGC ATCACAATCA CCAGCGACGA GAACGAGAAT TACTTTCTGA TCGTCAAGCC CAAGCATAAG ATCATCAGCT CAGAGACAAT TTACAACATG CTGAAGAACA ACAAAATCGA CTTTAAAACT CTTGAGAGGA AGCTGCTGAA CGGCAGCGCC CTGATAACCA CCAGTAGGGC AGTCGGCAGA CGGAAATACG TCAAAATCAA GAAAATCATA TCCCCCAAGG AGAAGGAGTA TTGGCAACAT ACCCAGGACA TCAATGAGCA CTACGAAAAG GAGGGCGTCC CGATCAGCGT CGGCGGTGAC

GACATCCACT GCTATATCTT CATCGGGGAG GACGATTACG CCTACCACAC GAAGAACTCC
TTGCTCTACG AGGGTGTGAC GGAGGACGTG CAGAAAATAC TCTTGGATAT GGGTAAGTTC
CTGGAGGAGC TGGAAACGGC AAAATCTATC CTCAAGCAGG GCAACCTCAT AGACTTCAGT
CGCGAATTCC TCAACATTAG CACGAAGGAC GACTACACCC TTACTCTCCT GAGCACACTG
TCCGATATCA AAGTGAAGCT TAAGACCGAG TCTGGTATCA TCACAGGCGA CTACCAGAAA
CTTAGGGAGA TCTTTGACTG GATCTTCGAC AAGAGCTTTA ACCCCTTGAA GCCTAAGAAT
TGCTACCTTC CGCTGAGTAT TCCCCCCATA CTGAATGACA AGAAAAAGAT CGGCGTGTAC
ATCTTCTATA GCAATATTAG CGACCCCGAG CTTAGGTTTA TCGAAGGGAT CTTTAAGAAA
CTGGGCCTGA TATGCGCCAT CAATAAGAGT GTGCCAAAAA TTGAGGTTAA ACTCAAGAAG
GAAGTGGACT TTGAGGACTA CGCCAACAGC AGGATCATAA TCACCCAGAC CGTACTGAGC
AATCTCGAGG ATGGCGAGCA GCCGTTCCTC ATATGTATAA GTCCCTTGCT GCCGAATAAC
GAGTTCGATG AACTCAAAAT GCATCTGTTC TCTCACCCGC AGCTGATATT TCACCAATTC
ATGTATCCGT TCAACCTTCG AAAGTGCCTT GAGAAAGAAT CATTCAAGAA ACCCTTCATC
AACTCAATCC TGTCTCAGTT CTTTCACAAA ATGGGCATGT ACCTCTTTAG TCTGTCTGAC
GAGCTGGGGA ACTACGACTT CATTATTGGT TACGACATAA GTAGGGAAAA GGATGACATC
GGGAAGATAA AAGGTATCGG CGGCTCCGCG ATCATCTACA ACAATTACGG CCATGTCAAG
TCAATCATAA CGTTCGACGA CGTAGGGTCT AGCGAGATAG GAAGGTACGA CCTCCTGTTC
GCACAGGTGC ACAGCGAACT GATACCCCAC CTGAATCTGA ACAATAAGCG GAAAATTAAG
ATTCTGCTTC TCAAAGACGG GCGGATTTTC AAAAAGGAAC TCGAAAAGCT CAGCCAAATC
AGCAAGAAGT ATAACTTCGA GATCACCTAC ATTGACGTTC GCAAGAGCAC GCTGCTCCGG
TTCTGGGGTG TGCGGAGGGG CAAAGTGGTG CCCGAGTATA AGAATAGCTA CGGGAAGTTC
GGACGCGCAT ACTATATTAG TAGCCATTAC TACAACCGCT TTTTCAAGCA ACCAATCGCA
ATCGTGGAGA AGTACCACAT AGACGAGGGC AATTACAAAC GCGTGGAAAT AGAGGAGAAT
GATATTAAGC AGCTGGTTCT GTTGACCAAG ATTAACTACA GCCAACTGAT GCCAGATAAG
ATGCGGCTGC CCGCACCCGT TCACTACGCA CACAAGCACG TGAACGCCGT GCGACGGGGC
TGGAAGATCA AGGATGTCTC TATACTGAGG AGCGGGTGCC TTCCTACGAT CTAGTGA

116 Gene Name: Archaea-3

Gene ID: BBF.2017.48.116

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAAAGAAA GGTGGAAGAT CCAAAGAAGA AACGCAAGGT GGGTAGCGGC
TCCATGACTA ACAAAACCAA ACAGAAAAGC AGGAAGCAGA GGTCCCTCAT AGAATTTCTT
AAGGTGAAGA AGATCAACAA GGAAGATGGT AAGAACCATA ACCTGATCAA GTATAGCACC
GAACGGATCG ATACAGGAGT GACCCAGAGC CTCATTGACA TCAATATATC CAGTAACATC
CTTAAGCTGC GGGGCAGCAT TGCTCAAGAG GTGTTCAAAC GGAAAATTGG CGTTTACTAC
GGGCTTGGGA AGTATTACGT TGCCGAAAAC AAGCTGAAGA ACACCGATCG AATGGATTTC
TTGAAGAGGG TCTACGAAAC CTTCCCCTAT AACTACCTCG ATAAACAGGA CCCGCACAGC
AAGATCAGCT TTTACGAGTA CTACACATTC CAGAAGTCCA TCGACAAAGA CGTGATAAAT
CTGCTTGAGC TGCAGAAGAT AAACGAGTAT AGTTGGGACA TACTGGACCC ACATATCGCC
ACGCGCCTTC TCACAAGCTA TGTGAAGCTT TACTTGGGCG ACTACTTGAA GCCAATCCTG
TCCTCTTTCG AGTACGTCCG GGCTCGAATC AAGACAAAGC AAAAGACCGT TCCAATCAAA
ATCCCCGTGA CCAAGAAGTT CGAGATCCGA ACTTTGGGGT ACGACCCGAC GCAGAGCGAA
ATTACTCTCG CCATAAAACG ACACGCCAGC ATGAACGCTG TGCTGTTGAG CAGCTTTCCC
CCCGACATCC TCGCGGTTGT GATAACTAAG CTCAAACGCC TCGTGAACGA GGCCGTGAAG
CAAGACTACC GAAAGGTCAG AATATACTCC GAAACCCAGC CGGGGAGCGG TACTGCCGCA
GTTGTTGAAA TCATCAGCGG CAGCCAAAAC GTGATGAAGT TTCTCGAGGA GCATCCGAAG
GGGGCCATCC ACGTTGAAAA GCGACTTAAA GAGCTGGGTA AATCACTGCA GGAGGTCCGG

TACCTTCTTA TCGGCGTCTA TGACAACAAC GTCAGCCTGG AGCGGGCGAA AAAAGACGAA
AGATACCACT ACTACTTCAC CGAGCATAAC GCTTACCTTG TACTTACGCC CGAGGTGCAA
AAGGCGCTCT TTGGCAAGTT GATCGACGAC TGGAAAACAA GCATTCTGAA TGAGTACCAA
AATAAGCTCC ACGAGATCAC GAGTCTTGGG ATGTTTAAGC ATTTGGAAAC CATACGGGGC
ATCCCGGTTT CCTTGAAAGA GAGGCTTGTG GTCCGCACCA GCGAGGGCTT GCAAACCGTA
GATGACATTA GGGACATTTT GACCAACCCC AAGATTCTTA GTAATATGTT GCCTATATCC
GAGGACGCGC TCAAGGAAAC GCGAAAGCAT AAACTGCGAA TCACCCTGTT CTGTCCGGAG
AAGTTTAGTG AGAGGATTCA CCGGACTATT TTCTACGACA AATTGAACCA GTTTCGCGAC
GGTCTGCTTA GCAACAGCTT CGCAAGCGTG GACGAAATCG AATTGTTCCA GGTCAAAGGC
GAAAACTCTA GCGATTATGA GGAGATCATG AAGGACGCTG GCCTTGATAA AATCCACGAT
TATACCCTGG CGGTCATCAT ATTTCCCGAA CATTATAGTA AGCGCAACCT TGAGTTGCGC
ATCTTTTACA ACTGGCTGAA AATGCGGTTC TACTCAGAGA ACAAGCCACT GGTTTTCCAG
GGCGCTCGGA TTGACAGCGT GTTCGGCCGG TATGCGAAGT ACGCATCATA CAACCTCATC
TTGCAGATCC CACCTAAATT GGGCATCTAC CCGTACTCAC TGGAGGAGCA CGAGGACTAT
GACTACATCA TCGGCATTGA TTACACCTAT TGGTACGAGA GAGATACGCC TAGTCTGGGC
GGTGGCGCCG TGTTGACCAG CCCGTCAGGG CTGATTGAGA GCATATACCC TATCGCACTC
CCGAGCCGCA CTGAATCCCT CAACATGTCC AAGATACTGA GCGAATGGTT CACGCGAACA
GTCAAAACGA ACCGGCATAT CATAGATAAG GGCCACGTGA CCGTGCTTAT CTCCAGGGAC
GGCATGATTC CTAAGTACGA ACGCCAGACA ATCCAGGAGT TCCTGAGTGA ATATAGCGGC
GACATGGGCA TGACCATAGA GGCAGTAGAA GTTAGGAAAC GGATCGCCGT GAGGACCTGG
GCTACACAAG AGCCCGTGGC CTACTACAGC CCGATAAAGG TTGGCGACTG TACCTACTAT
CTGGTCGACG CGCACACCGG ATACCCGCTG GGGGAGAAAG GGAACCGAAC CTTCTACAGC
TCACCCTATC TCATAGGAAG TTTTTACAGG TTCGAAAAGG GCAAATCCTC CCCCGTGCCA
GGTAGCGCAA AGAAGCACGT GATCGAAAGC CTGATAAGAC TTCAAAAAAT CAATTACGCC
ACCACCCGCA TGGATAACAT CAAGTTGCCC CTGCCCGTCG ACATCACCCA CAAACTCATT
AACTTTATCC GGGACACCAA GATGGAAATC AAGGGGGTCG GTATCCCAAA CAGTCTCTTT
ATGATATAGT GA

117 Gene Name: Archaea-4

Gene ID: BBF.2017.48.117

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCAAAGA AGAAGCGGAA AGTCGAGGAC CCTAAAAAGA AACGAAAGGT TGGCAGCGGT
AGCATGAAGA ACCTGAGATA CAAAATCAAC GCCTACAGAA TCAAAAAAGA CTATATTCCC AAG-
GAAGTTT ATAGATACAG GATCCGCTCC TTCATAGAGA ACATTAACAT ATATAGGTTC GTCG-
GTTTTT ACGGAGGCGT GGCCCTCAAT CAATCTGAGT TTATCCTTCC GTACCCGGTC GAAAATCTCG
TCCTGGAATA CGACGGAAAA GATGTAAAGC TTGAGCATAT CGACACACTG AACCTGGAGG
ACATCGAGAA TAAGGACAAG GAGAAAGCCG AGAAGCTGGT GAGGGGATAC CTGACCAGCA
TATACAAGTT GAAACCCATA CTCTACAAGA TCCTGCGGGA CGTTCGAGAG AGCAAGATCA
TTAACGATAT CAGAGTGGAT CCTATACCCG ACTTTACAGT AAAAAGGCAC AATAACGAAT
ACTACCTTGT CATCGATTTT AACCACACCG CGACCGTGTT GAAAAATCTT TGGGACTTCG
TGGGAAGGGA CAAGCTGAAA CTCGAGGATT ATATCGGTAA GAAAATCATA TTCAAGCCCA
ACCCGAAGAA GAGGTATACT ATAAAGAGCA TTGAAAAGCA GAACAAGAAG GACATTGATG
ACATTGTCGA GCACATCATC GAGTACTACA AGTGGACGGA GGAGGAAATT AAGAGCACCT
TCGGCGAAAT CGACTATACT CAGCCCATCA TCCATTGCGA GGGCATCCCC TACCCGTTCG
CACCGCAATT TTGCAATATC GTATTTACCA TGGAGGACTT GGATGAGAAT ACCCTCAAGG
ATCTGCAGAG CTACTGGAGG TTGCCCAACG AGATCAAAGG CAACATTATC AATCAGATCG
CTAAAAAACT GCGATTTGTG GAGAACGAGC CAATCGAATT GGAATTCATT AAGTTCAATA
ACACCCCCCT TATCGTGAAG GACGAAAATG GCAAACCAAC AAAGATATAC ACCACCAATC

GCCTCTTCCG ATGGAATTAC GATAGTAAAT CCAAACTGTA CTTGCCCTAC GACATCCCTG
ACATAATCAA GAACAAAACA CTGACAACGT TTGTGCTGAT CGACGAGAAT CTCAAAAACG
TGAGTGGTAA GATCAAGAGA AAGGTCTACC AAATGTTCAA GAATTACAAT AAGATCGCCA
GCAAGACTGA GCTCCCGAAA TTTGACTTCG CCAATAAATG GAAATACTTC TCTAACAACA
ACATCAGGGA CGTGATCCGA AAGATTAAGG ATGAGTTCAA CGAGGAGCTT GGCTTCGCGC
TCATTATCGG CAACCGATAC TATGAAAACG ATTATTACGA AACCCTGAAG ATGCAATTGT TCAAC-
CTGAA TATCATCTCC CAAAACATTC TCTGGGAGAA TTGGTCAAAA GACGATAATA ACTTCAT-
GAC AAACAATCTG CTCATACAAA TTATGGGCAA ACTCGGAATT AAGTACTTCG CACTGGACGC
AAAAGTGAAC TATGACTACA TCATGGGGTT GGACAGCGGC CTGGGCGCAT TCAAAAGCAA
CAGAGTGTCC GGGTGTACCG TGATCTATGA CAGCGAAGGG AAGATCCGAC GGATTCAACC
AATTGACGTG CCCAGCCCTG GGGAAAGGAT CCCCATTCAC CTGGTAGTGG AGTTCCTGGA
AACCAAGACC GACATCAATA TGGAAAACAA AAACATCCTG TTCCTTCGGG ACGGCTTTGT
GCAGAATAGT GAGAGGGAGG AGTTGAAGAA ACTGAGCAAA GAGCTGAATA GTAACATCGA
AGTGATCTCA ATCCGCAAGA ATAACAAGTA TAAAGTCTTT ACCAGCGACT ACGGTATCGG
CTCCATTTTT GGCAATGATG GCATATTCCT GCCACATAAA ACTACATTCG GAAGCAACCC
GGTGAAGCTC AGCACCTGGC TGCGCTTTAA CTCCGGGAAT GAGGAAAAAT TGAAGATAAA
TGAGTCTATA ATGCAACTTT TGTACGACCT TACCAAAATG AACTACAGCG CTCTGTACGG
GGAGGGTAGG AACCTTCGCA TCCCGGCACC GATTCACTAC GCCGACAAGT TTGTGAAGGC
CCTTGGAAAG AACTGGAAAA TAGACGAAGA GTTGCTGAAG CATGGCTTCC TCTACTTCAT
CTAGTGA

118 Gene Name: Archaea-5

Gene ID: BBF.2017.48.118

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAAAGGAA AGTGGAGGAT CCGAAGAAAA AGAGGAAGGT AGGCTCCGGG
AGCATGAAGC CAGTGAACTT GGATGAAAAC AGCCTCAACG ACGTCCCGGT AGGCGACACC
TATGCTGTCC GCTTCACTCT TGATGCAGTA TTCGAGAACG AAGGGCAGTA TCCCCGGAGG
AATCTGAAAT TCACAGACGG AGGCGGGGAT GACCGAACCA TCACTATTTG GAAAAACTCT
GCACCCGAGG AAATTTACGA GGCGGACTAT GAGCGCGGTG CGACGTATCT TATTACCGCC
GTCGAGTATG ACATCGACGA AGGTAATGAC GGCGAGCGAT ACCAGAATCT CACAGTCCAA
TCAGATGCTA CCTTGCTGGA GATGAGCGGT CCCCCTAGTA CCGAAGAGGC CTTGGAGGAC
GGCCTCGCCG AAACCCCAGA TACTAGCGCC GATTCAGGTG ACCACGGGTT GACAACCTTT
AGGACTACAG ACGATCTGCC GGATTATGAC GTCTATGAGT ACGAGCTGGT GCCGAAGCAA
GGATTCCGGC CGTCCGGAGA AAATGCCCTC CGAGCCACAT ACAGGGCACG ACGCAAGGTC
CGCCAGCAGT TGGACGTAAC ACCCGTCGTG GTCGGTGATG CGTTTAAGCT TGTGTCTCTG
GTCAAGCTGG CCCACGAGCG GGTCGAGCTT CCGCGATTCA AGATCAACGA GGTTGACGAG
AGGCCCATCG TCTACGCCGA TGAGGATGAC AGGGATGTGT TGGGGGAAAT GCTCGGTGAG
ATCCTCAAGG ACGCGAAACG GGACCAGTAC GACATCCATG GCATCGACAA AATACTGGAG
CCAGAGCCCG TCATAGAGAA AGAGGGCTTC AGGCTCCACG AACGGTACAA CCTGACCGTG
GAAGTTCTCC CTAGCAGGGC CGCTTATCTG CACGTGGACT ATCGACATCG GATATTGAGC
GACAGGACCC TGGATCAACT CGATGAGGAC GAAATCCACC CTGGCCTGCG CGTGACCCCC

TCATATAGGG ACATGGGTCT GTACGTTATA GGCGTTGGGC CGGAAACGGT GACCGATAAG
CTGCATATCG AGGGCAACAA GAGCCTGGTC CAATACCATC GGGAGGAGCC GTGGGTGGAC
CCGGCGAAGG TGCAAGAAAT CAAAGACGCA GATAGGGAAG TGATCTGGAC CGTGAGGCAA
CGGGGAGATG GCACCGAGAT GGCATTCCCG CCGGAGCTGC TCGCGCTTCA AGGGCACCCC
GAAAATTTGG CCCAGTTCGC CAGCGACTTT GCTGAACAAC AAAGGCTCAA CACGCGCCTT
TCCGCTGAGC AATGCATCAC CAAGGCTAAA AGGTTTGTGG AGCGACTCGG GCCCTTGCAA
TTCGACGGAC ACACTGTGGA ATTCGAAACC AACCCGCTGT TGGGCGATCG GAACATAGCC
ATAGATGGTC TGTTTCACCC GGAAGCAAAC GTGCTGCAGT TTAGCGGAGG CCAGACCGGC
ACCCACCCCT CAGATGTGAC ACAGCTGGGC GTGTACGAAG CCCCGGACCC CTTCAGGGTG
TGCCACATCA GGATGGAGAA GCGGGACAAA AGAATACAGA GGGGTTGGAG TACCTTGGAA
ACGAAGCTGG AGCAGATTGG AGCGCCTCCC GACAGTGTCG AGGAGGTCAC GTTCGACGCC
ACAATGAGCC CTGACCAGTT GGGTATGGAG ATAGCGGCCG AGATACCGGA CGACCATGAT
TACGACGCGG CCTTCTGCAC ATTGCCACCT AAAGACACCG GCTACTTTGA CACCGCAGAC
CCCGAGCGAG TTTACGATGA ACTTAAGAAA GTGTTGGCCA CCAAAGACCT TAACTCCCAA
TTCGCGTATG AAGCAACGCT GGACGAGCGC TTTACAATAA TCAATATAGC ACTGGGTCTT
GTCGCCGCAG CGGGAGGTAT TCCGTTCACA ATCGAGAGGG CGTTGCCAGG CGATAGCGAA
CTCCACCTGG GAATCGATGT AACCCACCAA TACGACGAGT CCGCGAATGG CAACCACATT
CACCTCGCTG CTGCGACGAC GGCTATCCAC GCTGATGGAG CTGTACTGGG CTACACCTCC
AGCCGCCCTC AGTCTGGGGA AAAGATTCCC CCCAAGGAGC TGAAAGAGAT CATCAAGCAA
GCGGTGATGG GCTTTCGCAC ACGCTACGAT CGCTACCCAA ATCATATAAC CATCCACAGG
GACGGGTTCG CAAACGAGGA CCTGTCCGAG GTAGAAAAGT TTCTGACGGA CCTCGACGTT
GAATATGATG TTGTCGAGAT CAGGAAGCAG GCCCCAGCGC GCGTCTTGAA ATACAGTGGT
GCCCACTTCG ACACGCCTCA AAAGGCGACC GCCGCAATCT ACGAGGACAT CCCGAAAGCG
ATTGTAGCGA CGTTTGGTGA ACCCGAGACT CTCGCTAGCC GGGAGTCAAC CGGGCTTCCC
CAACCAATCA CGGTGGAAAG GGTGCACGGA GAAACCCCCA TCGAGACACT TGCTGCGCAA
ACCTATCTGC TGAGCCAAGC CCACATAGGC GCCAGTAACG CTACAGCACG CTTGCCCATA
ACCACCATGT ATGCCGACTT GGCTAGTGCA GCGGCAGCCA GGCAACACCT TCCCCCGACC
AACAAGCTGA GGGATAAGAT CGGATTCATC TAGTGA

119 Gene Name: gene-3090

Gene ID: BBF.2017.48.119

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAGAGGAA GGTAGAGGAC CCTAAGAAGA AAAGAAAGGT AGGTTCCGGA TCCATGGAAG AAAATCTGTA TCTTGAATAC GACGCTTTCT TGAGGAGTGT GAAGCGCAAC GTGGACGTCC CTCATAGTTT CTTGCTTGGA GCCGGAGCTT CCATCTCCTC CGGAATTCAG TCTGCATACG ACTGTATATG GGAGTGGAAG AGAGATATCT ACATCACGAA GAATATAAAC GCCGCCGAGT ACTATAAAAA TCATAAAAAC GAAACGGTTC GCAAATCAAT ACAGAAGTGG CTGGACAACC ATGGCAACTA CCCCATCCTG GATGCAGCAG AAGAGTACAC ATTTTACGCC GAGAAAGCTC ATCCAATCGC TGACGATAGG AGAAAGTACT TCTTTAGTCT GATTGAGAAT AAAGAACCAT ATATCGGTTA CAAATTGCTG TGCTTTCTCG CTTCACAGGG GATTGTAAAG AGTGTATGGA CGACCAATTT TGACGGGCTG ATTGTACGAG CTGCTCACCA GAATAATTTG ACGCCTATAG AAATCACCTT GGATAACGCG GAGCGCATAT TCCGAAATCA GAGTACTAAG GAGCTTCTCT GCATAGCTCT GCACGGTGAC TACAAATATA GCACCTTGAA GAATACTGAT ACCGAACTGG ATAACCAACA CGAAATTTTT CAGGAGCACC TCGGAAATTA TCACGTAGAT AAAAATTTTA TAGTAGCTGG TTATAGTGGA CGCGACAAGT CTCTGATGGA TGCACTCAAG GCCGCTTATT CCAAGAAAGG ATCTGGTAGG TTGTATTGGT GTGGCTATGG TGAGAAGATA AATTCTGAAG TGAAAGATCT TCTTAAGTAT ATTAGAGCGA GTGGGAGGGA AGCATACTAT ATAGCTACGG ATGGGTTTGA CAAAATGCTC ATACACTTGT CAAAGGCAAT ATTTGAGGAT AGCCAAGAGC TGAGTGAGAA AATCCAGAAA ATACTCGAAA GCACGAATCA AACCGAAACC

TTCAACACAG AATTCAAGTT GGAGTTTAAG AAAACCGACA AATATATCAA ATCAAATCTG CAC-
CCTATTG TTTTTCCTAA GGAAGTATTT CAGTTGCAGA TCGAGTATGG CAATGAAAAA CCGTG-
GTCCT TCCTGAAAAC ACTGACAACT CAAACGAACA TTAGCGCCGT ACCGTTCAAA GGCAAT-
GTCT ACGCACTTGG TACGCTTAGC GAGATCAATT CCATCTTCAA GCCGTATCTT AAAAGC-
GAGG TCAAGAGGGA AGCGATCAGC CGATTCGACA TCGAAAACGT CACCGCATTC AAAAAC-
CTCA TGTTGACAGC CATATCCAAA TATTTTTGCT ACACGAAAGA AGTGAACTCT AACTA-
CAAAG ATAAGATTTG GTTGAAAAAC ATCCTGTCCA AGGTGGGGGA TATCACTGTT CACAAAG-
CAA TTTTCATATC CCTGTACTTT GACAAGAATT CCCATTTTGG TTATATGGCG TTCGCTCCTA
CCGTTTATTT GGATTCCGAC TGCGAAATTG AGAAGAGTCA AAAGCAATCC ATCAGTAAGA
ATTTGCTTGA GAAGTTGTAT AATAACAAAT ATAACGAGGA GCTCGAACTG TGGAATGGTA
TCTTGTTTAA TCATAAGAAA GTGAAATTTG AATATCCTCC CTTGTCTGGT ACGGGGTTCG
AATTTCAGAT ATCAAGCAAC ACTGCCTTCG GGGAGATAGA CGTGATTGAT AACAAGTACC
GCTCTTACGT CCCCCAGAAT TATGATAATA AGCAGACTCA GTTCCGGGGA ATCCAGTTTT
TGGAGCCGCA GCTGATATTT AAGAATATCG CAACGAACTC TGACTTCAAG GATTATCATC
CCATGCGAGG ACTGATTAAC AACCGACCAT ATGATGTAAA TCTCAACGGG ATTATCCACT
CCAATGAAAT TAACCTCTCA ATCATCTGTA GCCAAAAGTA TGGAGAAAGG TTGTTCGCAT
TCTTGACACA GCTCAATAGT AAGCACAGTA CAGAAAATAT CAACACTGAC TACCTGATAG AT-
TACCCCGG CTTCCTGTCC GCCTTTAATC TGCCCATCAA CATCCCAGCC ACCAACGATG ACGC-
TAGCTG GATGGACATC AACTTCGTAG CAGATAACTC TAAAGAAACA CACGAGAACG CTATAC-
GACT CGCGAGGGCA ATTACCAATA AGATCGAGAA GATTTCTGCT ATACAAAGCG CCAGCAC-
TAT AGTAATCTTT ATACCTTTCG AGTGGCAGCC CTTCGAAACA TATATTAACG AAATAGAAAC
GTTTGATTTG CACGACTACA TTAAAGCGTT TAGCGCCAGC AAGGGGATAT CAACGCAACT
TATTCGGGAG GACACCCTTG ACGATAAGCT CAAGTGCCAA ATATACTGGT GGTTGTCTCT
TTCTTTTTAC GTCAAGAGCC TCAGGACCCC ATGGATATTG AACAACCAGG AGCGGAAAAC
AGCTTATGCC GGAATTGGGT ACTCCATAAG CAAGGTAAAG AACAAGTCAG AGATCGTGAT
CGGATGTTCA CATATATATG ATTCAAATGG CCAAGGCCTT AAGTATCGCC TCTCAAAAAT
TGATAACTAC TTTCTCGATA AGCAAAATAA TCCGTACCTG TCTTATAAGG ACGCTTTTCA ATTTGGGGTT
AGTATCAGAG AACTCTTCTA TCAGTCACTC GATTCTCTGC CAGAAAGGGT CGTCATCCAT
AAAAGGACAA AATTCACCGA GGATGAGATC AATGGGATAA AGGCTTCACT CAACCAGGCT
GGTATTAAGA AGATTGATCT TATAGAGATC AACTACGATA TAGATGCAAA ATTCGTTGCC AT-
GAACGTGT TCGATAACAA ACTTCAGGTC GATAAATTCC CGATATCCAG AGGAACATGC ATTGT-
GACAA ATAAACGGAC GGCGTTGTTG TGGACGCATG GTATAGTACC TTCAGTTAAG CAGC-
CCAATT ATAAGTTCTA CCTGGGCGGG CGCTCTATCC CTGCGCCCAT AAAGATTACC AAG-
CATCACG GAGAAAGCAA CATTGATGTG ATAGCTAGTG AGATCCTCGG ACTCACAAAA AT-
GAATTGGA ATAGCCTGGA TCTCTACAGT AAACTTCCCT CTACGATAGA TTCTTCTAAC CA-

GATTGCTA AGATAGGAAA ACTTCTGTCT CGCTTTGAGG GCCGCTCATA TGACTACAGG CT-GTTTATTT AGTGA

120 Gene Name: gene-7321

Gene ID: BBF.2017.48.120

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCCAAAA AGAAACGAAA GGTCGAGGAC CCTAAGAAAA AGCGCAAGGT AGGTTCAGGC
TCTATGTCTG TGGACGCTAT GATCAGGAGT ATCGGGGTCG CACGGGACCG CCCGCTTCTC
GTTTTCCTCG GGGCCGGTGC CTCAATGAGC AGTGGTATGC CGTCCGCCAC TCAATGTATC
TGGGAGTGGA AACGAGAAAT CTTCTTGACA AACAACCCCG ACGTTGAAAA GACCCAGTTC
TCCGAGCTGA GCCTTCCCAG CGTCAGATTG CGCATCCAAG CATGGCTGGA TCGGCAACGA
CGCTATCCCG CTCTTGATCA TCCCGACGAG TATTCTACCT ACATAGGTGA GTGCTTTGCA
CGCTCTGACG ACCGCAGAAT CTACTTCGAG AAGTGGGTCA AACGCTGTAG TCCGCACCTT
GGATACCAAC TGCTTGCCGA ATTGGCACGG CAGGGGCTTG TGGCCAGCGT TTGGACTACT
AATTTCGATG CCTTGGCGGC TCGCGCAGCT ACGTCCATCA ATCTCACTGC AATCGAGATT
GGAATTGATT CACAGCAAAG ACTGTACCGG GCGCCGGGCG AGGCGGAACT GGCGTGTGTG
AGTCTGCATG GAGATTATCG GTATGATCCT TTGAAAAACA CCGCTCCAGA ACTCATAAAA
CAAGAGAAGG AGCTCAGAGA GTCACTTGTC CAAGCAATGA GAACTCACAC AGTCCTGGTT
TGCGGCTATA GTGGTCGGGA TGAGAGTGTC ATGGCAGCGT TTTCCGATGC CTATGACGCA
GCTCATTTTA AGGGTCATCA CCCCCTCTTC TGGACACAGT ACGGCGATTA TCCCGCCAGT
GAGCCCGTAG CTGGACTTCT TGCTTCACCG CTGGATCAGG AGCCTGCGAA GTTCCACGTG
CCTGGGGCAT CATTCGATGA TCTTATGCGC AGGATAGCAC TCCACGTGAG TGACGGTGAA
GCGCGCGAGC GGGTGCGGAA GATTCTTGAG AACTTCAAGA CGGCACCAGT TAACCAGAAG

CTCCCCTTTG CCTTGCCTAG CCTTCCTGTG ACGGGCCTCG TCAAGTCAAA CGCCATTCCG
TTGATACCGC CTGGAGAGCT TATAGAATTT GATCTTGTCC GGTGGCCGCC GTCCGGTGAA
GTTTGGAGCA CGCTCCGGGA AATAGGGGAT AGACACGGAT TCGTAGCTGC CCCTTTTCGC
GGGAAGGTGT ATGCTCTGGC TACGATAGAG CAACTGACAC AAGCCTTCGC GGACAATGTA
AAGGATGGCG CGTTCAACAG GGTGCCGCTG AATAATGATG ACCTCCGCTA CGAGGACGGA
ACCGCCAATC AGCTGATGCG ACGCGCTACT GTTCTGGCTT TGGCTGGGAA AGCTGGATGC
GCGAACGATG GGGATGCCAT TGTGTGGGAC ACATCTCGCT CAAAAACCGA AAGATTGGAT
AGGCAACTTT GGACTGTATA CGATGCAGTA CTTCTGCAGA TTCGGCCGCT GGGAACTAAG
CTCGCGCTCG TACTTAAGCC TACGCTGCGG GTTACGGATT CAACTGGCGA GGTAGCCCCG
AAAGAAATTG AACGGGCAGT CAAGGTGCGC GTATTGGGAT ACCAGCATAA CAAAGAGTTC
AACCAGGCGA CCGACTTTTG GAGGAAAAGG CTCCTGCCCT CAAGAGATCT CCTTGTCAGA
TTTCCTGATC TGGATGGTGG AATGACTTTC ACGATTTCAG GTCGGCCAAT ATTCGCCCGG
CTCACCGACG AAAGGACTGA AACTGTCACA CTGAACGATG CCCAAGAGCG ATCAGCATCT
CAAGTGGGGT TGCAGCTTGC AGAGCCTAAA CTGGTGTTTG CACGCACTGT AGGTACGGGT
CCCGCAACGG ACACCCTCCC GGTTAGAGGA TTGCTGCAAA ATAGACCTTT CGATGCTAAT
CTGACAGACT TGGGTATCGC GACGAACCTG AGGATCGCGG TTATTGCGCC CGCTCGGGAC
GCCAGAAGGG TACATGACTA TCTTGGGCAG CTGCATCAGC CTATAGATCC TACAAAGTGG
GATGCGGACT ATCTGATGAG GTTTCCCGGC TTCAGCTCCG CTTTTAAATG CCCTTTGGAC
ATTCCGCAGC CGGGCCAGGC AGCTTTTGTA ACACTTGACG AGCCACACGA TGAGAGTCCT
CAATCAGCGC GGACCCTTGC AGGCCGAATC ACAGCGGCAC TGTCTGCATT GAGGGCGACG
GAGAATCCCT CTGTTACAAT AATATATATT CCGGCGCGCT GGCACGCGCT GCGAGCATTC
GATCTCGAAT CAGAGCAATT CAATCTTCAT GACTTTGTTA AGGCCGCCGC AATTCCAGCG
GGCTGTTCCA CACAGTTTCT GGAGGAGTCA ACTCTTGCAA ATGGCCAACA GTGCAGAGTC
CGATGGTGGC TTAGCCTCGC TGTTTACGTA AAGGCAATGC GCACCCCGTG GGCTTTGACG
GGACTCGATA GGGACTCTGC CTTTGTAGGG CTGGGCTTCT CTGTAAGACG AAAGATCGAT
GGCGAAGGTC ACGTCGCGTT GGGTTGTTCT CATCTTTATA GCCCAAATGG TCATGGTTTG
CAGTTCCGCT TGAGTAAGAT TGATAATCCG ATAATGCTGC GAAAAAATCC TTTTATGTCC
TTTGACGACG CTAGAAAGTT GGGCGAAGGC ATCAGGGAAT TGTTTTTTGA CGCCCACCTC
CGGCTGCCGA ATCGCGTAGT TGTTCATAAA CAGACCCCGT TTCTTAAAGA GGAGCGGGAA
GGGCTCCAAG CAGGGCTCGA GGGAGTCGCG TGTGTGGAAC TCTTGCAAAT TTTTGTAGAC
GATACGTTGC GATATGTGGC TAGTCGACCA ATGCCGAATG GAGATTTCGA AATCCATGGC
TATCCTATCC GAAGGGGCAC CACAGTAGTG GTCGACGACC AGACCGCATT GTTGTGGGTA
CACGGCACAT CAACCGCGCT CAACCCGCGG CAGAGCTATT TTCAGGGCAA ACGCCGCATA
CCGGCCCCCC TTGTGATGAG GCGGCACGCG GGGACGTCTG ATCTGATGAT GTTGGCGGAC
GAAATATTGG GACTGTCCAA AATGAATTTT AACAGTTTTG ACCTGTATGG CCAACTCCCG

GCAACCATCG AAACGAGCCA AAGAGTCGCG AGGATAGGCG CTCTGCTGGA CCGCTATACG GAACGGTCAT ACGATTATCG ACTCTTTATG TAGTGA

121 Gene Name: gene-12619

Gene ID: BBF.2017.48.121

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCTAAGA AGAAAAGGAA AGTCGAAGAT CCGAAAAAGA AACGCAAAGT AGGGAGTGGT AGCATGATCA AACACCTCAA GTTCGACGAG TTCCTTCGCA GCGTGTCAAT TAGTAAGGAT AACACGTACT CCATGCTTAT CGGTGCCGGG TGCTCAATCA CTAGTGATAT CCAATCTGCC TATGACTGCA TATGGGAATG GAAGAAAATA ATTTACAAGT CCAATAACTT GAATACTCAG GACTGGATAG AGAATTACAA ATCCCCCAAA ACACAAGACG TGATACAAAA ATGGCTTGAC AACCAGGGAA ACAACCCTGA GAAAGATAAT ATCGAAGAGT ACTCATTCTA CGCAAAGAAA TGCTTTCCGA TAGATGAAAA TAGACGCCAG TACTTCCAGA AAATCTGCGC TAATAAGAAG CCCAGCGTCG GATATCGAGC CATTCCTCTC CTGGTGAAGC AAGGCATGCT CGACTCAATT TGGACAACCA ATTTTGATGA TCTTGTTAAT GTGGCGTGTA TAGGTGGTGG CGTTCAGGGG ATTGACATAT CCCTTCAGAC GGTAAACCGC ATAAATCAAC GCAATCAAAG CAAAAATGAA CTGCCTATTA TAAAGCTCCA CGGGGATTTC AAGTATGGCG ACCTTAAGAA CACGAGTGAG GAACTTCAGA ATCAAGACGA AACGCTTAGA TCAAAACTTT TGGACTACTT GAGCGATAAG AATCTCATAG TCATTGGCTA TAGTGGTCGG GACAACTCAC TCATGGAGAG CTTGAAAGAG ACTTATTCAA AACCTGGTGC GGGAATATTG TTTTGGTGTG GGTATGGGAA CAGTCCATCA AACCAAGTGA AGGAACTCCT TAAATTTATC AAGGATAAGG GGCGCAGCGC ATTCTATGTT TCCACTGAGG GATTCGATAA CACCATGCTG AACCTGACCA AGCATGTTAT TGAGGACGAT GATAACCTCA AAGAGGAATT CAGAGAACTC AAGAAGAGTA TCATTAATAA AAATACAACG ACCCCGTTTA

CGTTGAACCC GGAACGAATC AATAAGGTAC TGAAAAGTAA CCTCTTTCCT ATTACATTCC
CCAAAGAGAT CTTCGTATTC AATGCGACCT TCGATAAGAA ACCTTGGGAG CTTGTTAAGG
AGAAAACTCT GAGTGACTAT GAAATTTCAG CGATTCCATT TGAAAAAGAC ATATGGGCAT
TTGGGACTGC TAATAACGTC TACGAAAAGT TTGCAGATAT CATTAAGGGC GAGATCCAAC
GGAAGCCCCT GACCGATATC CGGCTTTATA ATCACAACAT AAAGTTCCTG CTCCTGTCAA
GCCTCTGCAA GCTGTTCTCA AAAACCTACA ATCTGAAAAC GGACTTTCGG TCTAAGATTT
GGGATGAGAG CTCATACAAA ACGGTTCACA ACCAAAAGGT CTATAACGCT ATAAAGATCG
ATCTCGTCAA AATACAAGAA CAGTCATATT TGTCACTCAA TCCAGACTTT CAATTGGCAG AT-
GATAACGT TCCCAATGAT ATCAACCAAC AGGTTGGACT GGAATTTTTT CATAAGATCT ATAAC-
GACAA ATTTAACGAC TATATAAACA TCTGGAGAAA GAAGATCCTC GAAACTACGT CATAC-
GAATT GCCACTGAAC TCCGGCACCG GGTTCGTATT TAAAATCTCT AAGAATCCAA TTTTCA-
CAAA TATAGATGAC CTTAATTCCA ACTATACGAA CGAGCACAAT ATACCCATAA ACATGATTAA
ACTTAAGGGG GTTCAATTCA AAGAAACGAA CCTCCTCTTT AGTTCACAAA ATGGAGATAA
AGTGGTTAAG GAAACCCACC CAATGAGAGG CCTCGTCAAT CATAGCCCGT TCGATAAGGG
ATTGAGTAGT CTTAAAAACA CTACGATCAA CCTGGGGATC GTATGCCCCC AACAGGATAG
CGAAAATTTT TATACTTTTT TGAATAAACA AAACCAAGAG ATTAAGAACG TTAATATTAA GGAT-
CAATAT GTAATCGATT ACAAAGGATT TCACAACACA TACGGTTTGA GTCTGAACAT ACCTAC-
TACG AGCAGTCCTA ATTGGGAAAT GACTAACGAG CCTGTCTCAA GGGACTCAAA GAAAATAATT
CATGAAATCA AGAATAATAT TTGCGACAAG ATAAATAAGC TTTGTAGTAT AGGCGGACAA AA-
GACAATAG TAATATTTAT CCCTAAACGC TGGGACAACT TCGTACACTA TAATGATGCC GTG-
GAAAGCT TTGATCTTCA CGATTATATC AAAGCGTTCT GTACCGAGAA AAAGGTTACC TCTCAGTTGA
TACGGGAAAA GACGATACTC GATAATAACC TCGAGTGCCA GATCAACTGG TGGTTGTCAC
TCAGTTATTT TGTAAAGTCC TTCCGAACAC CGTGGGTAAT CGACAACACC GACAATAAAA
CAGCTTTTGC GGGCATTGGT TATTCAGTAG AGTCCAAAAA AGAGGATAAG GGGCACATTA
TACTTGGCTG TTCCCATATT TACAGTAGTA ACGGGGAGGG ACTCAAGTAT AAGCTTTCCA
AGGTTAATGA TAAAATAGAA TGGATCAAGA AAAAGCCGCA TCTGTCCTAC GACGATGCTT
ACGAATTTGG TAAAAATGTG ATCAACCTGT TTTACGAAAG CATGAATGAG GTGCCAAAAC
GAGTGGTAAT TCACAAACGC ACCTTTTACA CTGAAGATGA GAAGCAAGGC ATACTTGACT
CTTTGCACGA TAACAAGAAA ATAGAAAACA TAGACTTGAT AGAAATAAAT TTCGAGGACA
ACATAAGGTA TGTCTCCTCT AAGATATATA ATCGGGAGGC AAAAATCGAC GGTTACTCAG
TATCACGCGG TACCTGTATC CTTCTTAACG AAAAAGAGGC ACTTTTGTAC GCCCATGGCG
TAATCCCGAG CGTGAAGAAT CCGAGTTATA ATTTTTATCC GGGAGGAAGG TACATACCGA
AGCCATTGAG GATAATAAAG CATTATGGAG TTGGTTCCCT GGAACAAATA GCAAATGAAA
TACTGGGCCT CACTAAGATG AACTGGAACT CTCTGAACAT GTATAGCCAA ATGCCTGCCA
CGATCGACTC AAGTAATAAG ATAGCCAAAA TAGGGAAACT CATAGAGAAT AGGGATAAAG

TAGAGTACGA TTATCGGTAT TTTATCTAGT GA

122 Gene Name: gene-12630

Gene ID: BBF.2017.48.122

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGAAAATTA TAGATAAGGA AACCTTCATC AGAAGTTTTA AAGTTTTGAG CAATCAATCC TTTGAC-
CTGT TCCTGGGCGC TGGCGCCTCC ATATCTAGCG GTATCCCTTC CGGAGGCGAC CTCGTCTGGC
ATTTTAAGCG CGAAATACTG AATTCCAACG GGAAGATAAA TATTAAAAAA TTTCAAGATC TTAA-
GATAGA AGATAATAAG AAGGTTATAC AAAGTTTCTT TGAGGAGACT GAGGAGAACA ACATTAT-
TAA TCCTTATTCC TATTATTTTA ACAAATGTTA TCCAGACCCC TTGATAAGAA AAGAATTCTT
GACGAATCTT GTGAGGGACA AGAAGCCTTC CATAGGATTT ATGTGCCTGT CTGCTCTCGT
GGAGCAGCAA AAAATCAACA CAGTATGGAC AACTAACTTC GATGACTTGA TTGAGAAGGC
GATTAACGGA TTGAATTACA AGTCCTGTCA AATTGTCTCA CCCGAGAATG CGGGCAGCGT
GAATAACTTT CGAACTGATA TCCCCACTGT TGTTAAGCTT CACGGAGATT TTAGGTATGA
CCCACTGCAG AATACTGACG AAGAGTTGCA GAAACTCGAA GAGTCCTTGC ATAAGTATTT
CGTAGAGGCA AGCACAAAGA GGGGACTTCT CGTAATGGGC TATTCTGGGT CAGATGAGTC
TGTGCTGCAA AGCCTTGAGA AGGCGCTGGA AGAGAACAAC GCGTTCCCTA AGGGACTCAT
TTGGTGCATC CCCAAAAGTG TCACCCCAAA CCAACGACTG GTCCGAATTA TATCTAAGGC
TAATGAGCAG AACCAGCGGT CCGGATTTAT GATTATCGAC AGTTTCGATT ATTTCTTGCA
TGAACTCTAC AAAATATGCG ACCTTACGAA TGACTATATC GACTCTATTA CCAAGGAGAG
ATTTGAAAAA AGGCAGTCAT TTAGGCTTAA CCAAACTCCG TCCTCTACTC TGCCAATCTT
GCTGAACGCA ATAAAAGCAA AGCACTTCCC GAAAAGTACC TTTCTGACTA AAACGAATAT

CTCAGGCATA GGTAAGTGGA AACGCTTGCG CGACGCTATA GGAAATAGCT CTATAGTCGG
ATCTTTCGGT AAGAACGATT CTCTCAGACT TTTTGGAAGT GAACAAGACA TTAATAATGT
ACTTAAGAAC TACTTGATTG ATGATTTGAA GATCAGTGAT ATCCCAGAGC ACTTGTTTTT
CCATTCTGAT TCATTCTACA TTGGCATGCT TTATGAACTG ATTGAAAAGT GTTTGATTAA
AGATTATGGG CTGTCAGTAT ATGCAAAGGG GAGAACTATC AGAAAGTTCT ATTCAATCAA
TAACCCGCTG CCGGAATCTG AAATCGCAGA TATTAAGAAG AGAAACAATA ATTTTAACAT CGA-
CAAAAAT ATAAATGTAT TTGAGGCGTT CGAGTTCTCC ATAGAATTCA TTAATAAGGA GCT-
GTTCCTG TTGCTGTGTC CCACCATACA TATTCAGACT AAACTCGGAG GTGAGGTCAA TCG-
CAATATC TCTCAGTACC TGTCAAACAC AATCATCAGC AATAGGTATA ATAACAAATA TGGGAAAAAG
CTGAATTGGT GGATTAACGA GCTCAAGAAG TATAACAAGG ACTTGGTTTT TAAATTGGGG
GACTTTGAGA TACGATTGAC AGATTATTAC TCCACGAGCG CTAAGCGCGT TAAAGATGAC
ATCTACTGTT TTGACGGATT TACTAAGTTG AGTGAGCCCA GTATATATTT CCACTATCAA GAC-
GAAGCAA AGCAGAGTAT CCATCCCATA AGTGGACTGA AGATACTCGG TCCATTGGAA GAAT-
CATTCG AGGCAAACGG TACATCTTCC ACAGTCAACC TTGCCATCAT TACTCCGGAC TTTG-
GCTTCT CCAAACTCAA GGCGCACCTC GAAAGTTTGC TTAATACAAT TTCCCCTATA TGGGA-
GAAGG AATACTTGAA GGAGTTCCCT GGTTTCGATA ACGTTTTTAA GAAGCACCTG ATAATAC-
CCA ATTCTATTCA AAGCGAGTAT GTAATCAGCA TACCTAATAA TGATGTAAAA CAGTTCTCAG
CAATTCAATT CTACGACTAC CTGAAGAGTA AGATCGACCG ACTCGCTCTG AAGTCCAATG
ACATTGATTG TCTTGTAATA TACATACCCG ACCAGTGGAA GAACTTCCGA GAGCTGAAAA
ATGAAAACAC ATATTATGAC CTTCACGACA GTCTTAAACT CTACTGCGTA AAAAAGGGGT
TGCGAATCCA GTTCATCGAA GATAAAAGCA TTAATTATAA AGACCAAGCC AAGATCCGGT
GGTGGCTGTC TCTGGGGCTC TACGTGAAGT CTAACGGCAC TCCCTGGAAG ATCAAAACAG
ATAATACAGA GACTGCCTTT GTGGGCCTCG GTTACGCTAT ACGACAAAAT GTTAAGAATA
AGGTTGTTCT CGGATCTTCA CAGATTTTCG ACGGTTATGG GAATGGCCTC AAGTTTCTTT
TGCAGCCCAT AGAGAAGCCA ATTTTTTACA ATAAAAACCC CTTCATGAGC AAAGAGGACT
CTTTTCGGCT TATCAGTAAT ATACGAAACA CATATCATAA GATCGATCCA GTTATCGGAC TTAA-
GAAACT CGTGTTGCAT AAGACAACTC ATTTTACTTC AGAGGAGATG GAGGGGATCT CTAAT-
GCTTT GGAAGGCATA GACAATATTG AACTCTTGCA GATTCAGCAA TTCTCATCAT GGAGGGCAAT
TAAGCTTATG AAAAATGCCA CAAAGCACGA TTTTAATGGT TATCCGATCG ATCGCGGAAC
TATAATTCAA CTCGACGACT TCTCTTTCCT TCTGTGGACA CACGGGCTTA TAGAGAACCA
AGAGCTGAAC GGTAAGTACT ACCAGGGAAA AAGAGGAATA CCGGCTCCGC TTCTTATTAA
GAGATTTAGA GGCACGGATC CAATAGAAAC GGTGGCAAAC GATATTCTTA AGCTGACCAA
GATGAATTGG AATGGTGCAG AGCTCTATAA AACCTTTCCT GTAACGATTG ATTTCAGTAA
AAAACTTTCA GTCATGGGGA AGTAGTGA

123 Gene Name: gene-7525

Gene ID: BBF.2017.48.123

Submitted by Dr. Stanley Qi on 2017.11.20. Contact at slqi@stanford.edu.

Description

Series of prokaryotic Argonaute proteins found with a bioinformatic search. Codon optimized for human cells.

Motivation

NA

References

NA

Sequence

ATGCCGAAAA AGAAAAGGAA GGTTGAGGAT CCTAAGAAGA AAAGAAAGGT CGGCAGCGGG
TCTATGTTCG ACATTGGATC AATGGTGAGA GTTAGGGGTC GAGACTGGGT CGTGTTGCCT
GGCAGTTCCG CAGACTTTCT CCTGCTTAAG CCACTCGGCG GATCAGATGC AGAAACGACA
GGGGTTTATG CCGGTCCCGG CGGCGAAGTT GTGAGATCAG CGACTTTTGC GCCACCCGAT
CCGCAAGCGT TTGGAACAGC CTCTGGCGCT CGGCTTCTCC TGAATGCAGC TAGATTGGCC
GTTAGGTCCG GCGCTGGACC GTTCCGCTCC CTTGGCAGGC TGGGGGTAGA ACCACGCCCA
TATCAACTTG TCCCCCTCCT TATGGCCCTG AGACAAAGTA CCGCCCGGCT CCTTATTGCC
GACGATGTAG GTATAGGAAA GACAGTTGAA GCGGCACTGA TCGCCAGGGA GCTGCTTGAC
CGCGGAGAGA TAGAGCGATT CGCTGTGCTT TGTCCGCCCC ATCTGGCTGG TCAGTGGGTA
GGTGAGCTGA GGAGCAAGTT TGGGATAGAT GCCGTCGCGG TCCTCCCCGG AACCGCGAGG
AGACTGGAGC GCGGCTGTAA CCCAGGCCAA TCTGTGTTCG CCAGATACCC TTTCGCAGTT
GTCTCTCTCG ACTTGGTCAA ATCAGACCGA TGGCGCCAGG ATTTTTTGCA GAACGCCCCC
GAGTTTGTTA TCGTCGACGA AGCGCACGCC AGTGCTGAGG GCGAGGGGTT GGGCGCGCGG
AGACATCAGA GATATCGCCT TTTGGAGGAC CTTGCGCGGG ACCCAGAGCG ACACTTGATA
CTCGTGACAG CTACGCCACA CAGCGGAAAG GAGGACGCAT TCAGATCCCT TTTGAGATTG
CTCAACCCTG AATTCGCCGC TCTGCCACTG GATCTCTCCG GCGCTCAAAA CGAAAGAGCT
CGGGCAGCTA TCGCTCGACA CTTGGTGCAG CGGAGGAGGG GTGACATCAC TGCATACCTT

CACGAGGACA CCCCATTTCC AGTCCGAAGG GACGCCGAGG TTAAGTATAC TCTGCACCCC
GATTATGCGG CATTGTTCGA GGACGTTCTG GCCTATGCAA GGGAGTCCGT GCACGTTCCA
GGCGAGGCGC ATAGTCGGAC GCGGATACGC TGGTGGGCCG CCCTGGGACT GCTTCGGGCT
TTGGCTTCTT CACCCCAAGC AGCCGCAGCC ACTCTCCGGG AAAGAGCAAG CACCGAAGGC
GAGACTGATG AAGCAGTTAT TGAAAGACTT GGCAGGGAAC TGGTGCTTGA CCCCGAGGAC
GGTGAACATG GGCTGCTGGA CGTCACCCCT GGAGCCCAGG TCGACGGTGA AGAAAGCGGG
ACCACGCGAC GCCTTCTCGC ACTCGCAGAG AGGGCCGACG CTCTGGCTGG GGCCAAAGAC
CGGAAGCTCG CACTCCTGAC CGCACAGGTC AGGGATCTTC TGCAGGAAGG TTTCGCGCCG
ATAGTTTTTT GTAGGTTCAT TGCGACCGCG GAGGCAGTAG CGGAGCACTT GAGGGGAGTT
CTGAAAGGAG CTGAAGTCGT GGCTGTCACA GGAAGGCTGA CGCCAGATGA GCGCGTCGCC
CGCATCGAGG AGCTTGCACC CCACGAGCGA CGGGTTCTTG TGGCAACGGA CTGCCTTAGT
GAGGGCATTA ATCTCCAAGC TGCCTTCAGC GCAGTAGTAC ACTATGATCT CCCCTGGAAC
CCTACCAGGC TCGATCAAAG GGAGGGCCGA ATTGACCGAT ATGGTCAACG ATCACCAGAG
GTCCGAGTGC TTACATTGTA TGGGGAGGAT AACAGGATAG ATACTCTGAT ACTGGATGTT
TTGATCCGAA AGCATCGGCT GATCCGGGCT ACCTTGGGAA TGGGTGTCCC CGCTCCCGAC
GAGGCAGAAG GATTGCTTGA CGTGCTGTTG GCGCGAGTAC TGGAACCCGA ACGAAGAGGT
TCTATTCAGC CATTGCTTCT GGATGAAGTG CAGGCTTTTG ATTTGAAATG GAGGGATGCG
GCTGAAAACG AAAAAAGGTC AAGGTCACGA TTCGCCCAGA ACTCTATAAG GCCCGAAGAA
GTAGCAGGGG AACTCGCAGC GGTACGGGAA GCGCTCGGTG ACGCTCGAGC CGCTCAGGAC
TTCGTTCTTG ATGCACTGCG AGGGGCCGGT GTTCAGGTGA CGCCGCGCCC CGACGGAAGC
TTCGAAGCGG ACCCCACCCA AGCCGATGTA GCACCGGAGG TCCGCGACTT TCTGCGGGGA
GCAAGGCGCT TCAGATTTGA CGCACGGGTA GAACGAGGTG TGACGCCCTT GGCGCGGAAC
CACCCATTGG TCGAGCAACT TGCAAGCACT GTACTGGGTC AGGCTCTGGA GTCTCCGCAG
GAGGCCGCAG CCAAGCGCGT AGGCGTCATT CGGACCTCTG GCGTAAGTAC TCAGACCACT
CTTTTGCTCC TTCGATGGAG ATTTCATCTT TCCGGACGAA AGGGAAACCG ATCTTGGCAA
ACTCTTGCTG AAGAACTTGA TCTTCTGGCT TACGCAGGAA GGGCAGAGGA TCCGCAGTGG
TTGGACGCTG AGGCCACCAG AGCTTTGCTC GATCTGACCC CTCAGGGTAA CTTGGATCCG
GTGCAGAAAG AGGAACGCCT TACTCGGACG CTTGAGGGAC TTAGCGCTTT GGAGGGGGTT
TTGGACCAGC GAGGAAGGGA TAGAGCCGCA GCTCTGCTTG ACGCTCACGA GAGAGTACGG
GGAGCAGCGC GAGGGCAAGG GGTGACCTAT TCTGCGGAGC CTCCTGGCCC CCCGGATCTG
CTTGGTGTCT ATCTCTTTCT CCCCGCACCA AGACTCGGAG GCCTCGCCTA GTGA

124 Gene Name: PIL253-1

Gene ID: BBF.2017.48.124

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707902

Sequence

GAAGACAAgt taactaactc aacgctagta gtggatttaa tcccaaatga gccaacagaa ccagagccag aaacagaatc agaacaagta acattggatt tagaaatgga agaagaaaaa agcaatgact tcgtgtgaat aatgcacgaa atcgttgctt attttttttt aaaagcggta tactagatat aacgaaacaa cgaactgaat agaaacgaaa aaagagccat gacacattta taaaatgttt gacgacattt tataaatgca tagcccgata agattgccaa accaacgctt atcagttagt cagatgaact cttcc-ctcgt aagaagttat ttaattaact ttgtttAaag acggtatata accgtactat cattatatag ggaaatcaga gagttttcaa gtatctaagc tactgaattt aagaattgtt aagcaatcaa tcggaaatcg tttgattgct ttttttgtat tcatttatag aaggtg-gagt ttgtatgaat catgatgaat gtaaaactta tataaaaaat agtttattgg agataagaaa attagcaaat atctatacac tagaaacgtt taagaaagag ttagaaaaga gaaatatcta cttagaaaca aaatcagata agtatttttc ttcggagggg gaa-gattata tatataagtt aatagaaaat aacaaaataa tttattcgat tagtggaaaa aaattgactt ataaaggaaa aaaatctttt tcaaaacatg caatattgaa acagttgaat gaaaaagcaa accaagttaa ttaaacaaTG GTCTTC

125 Gene Name: PIL253-2

Gene ID: BBF.2017.48.125

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707903

Sequence

GAAGACAAac aacctatttt ataggattta taggaaagga gaacagctga atgaatatcc cttttgttgt agaaactgtg cttcatgacg gcttgttaaa gtacaaattt aaaaatagta aaattcgctc aatcactacc aagccaggta aaagcaaagg ggctattttt gcgtatcgct caaaatcaag catgattggc ggtcgtggtg ttgttctgac ttccgaggaa gcgattcaag aaaatcaaga tacatttaca cattggacac ccaacgttta tcgttatgga acgtatgcag acgaaaaccg ttcatacacg aaaggacatt ctgaaaacaa tttaagacaa atcaatacct tctttattga ttttgatatt cacacggcaa aagaaactat ttcagcaagc gatattttaa caaccgctat tgatttaggt tttatgccta ctatgattat caaatctgat aaaggttatc aagcatattt tgttttagaa acgccagtct atgtgacttc aaaatcagaa tttaaatctg tcaaagcagc caaaataatt tcgcaaaata tccgagaata ttttggaaag tctttgccag ttgatctaac gtgtaatcat tttggtattg ctcgcatacc aagaacggac aatgtagaat tttttgatcc taattaccgt tattctttca aagaatggca agattggtct ttcaaacaaa cagataataa gggctttact cgttcaagtc taacggtttt aagcggtaca gaaggcaaaa aacaagtaga tgaaccctgg tttaatctTG GTCTTC

126 Gene Name: PIL253-3

Gene ID: BBF.2017.48.126

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707904

Sequence

GAAGACAAat ctcttattgc acgaaacgaa attttcagga gaaaagggtt taatagggcg taataacgtc atgtttaccc tctctttagc ctactttagt tcaggctatt caatcgaaac gtgcgaatat aatatgtttg agtttaataa tcgattagat caacccttag aagaaaaaga agtaatcaaa attgttagaa gtgcctattc agaaaactat caaggggcta atagggaata cattaccatt ctttgcaaag cttgggtatc aagtgattta accagtaaag atttatttgt ccgtcaaggg tggtttaaat tcaagaaaaa aagaagcgaa cgtcaacgtg ttcatttgtc agaatggaaa gaagatttaa tggcttatat tagcgaaaaa TCAgatgtat acaagcctta tttagtgacg accaaaaaag agattagaga agtgctaggc attcctgaac ggacattaga taaattgctg aaggtactga aggcgaatca ggaaattttc tttaagatta aaccaggaag aaatggtggc attcaacttg ctagtgttaa atcattgttg ctatcgatca ttaaagtaaa aaaagaagaa aaagaaagct atataaaggc gctgacaaat tcttttgact tagagcatac attcattcaa gagactttaa acaagctagc agaacgccct aaaacggaca cacaactcga tttgtttagc tatgatacag gctgaaaata aaacccgcac tatgccatta catttatatc tatgatacgt gtttgttttt tctttgctTG GTCTTC

127 Gene Name: PIL253-4

Gene ID: BBF.2017.48.127

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707905

Sequence

GAAGACAAtg ctgtttagcg aatgattagc agaaatatac agagtaagat tttaattaat tattaggggg agaaggagag agtagcccga aaacttttag ttggcttgga ctgaacgaag tgagggaaag gctactaaaa cgtcgagggg cagtgagagc gaagcgaaca cttgattttt taattttcta tcttttatag gtcattagag tatacttatt tgtcctataa actatttagc agcataatag atttattgaa taggtcattt aagttgagca tattagagga ggaaaatctt ggagaaatat ttgaagaacc cgattacatg gattggatta gttcttgtgg ttacgtggtt tttaactaaa agtagtgaat ttttgatttt tggtgtgtgt gtcttgttgt tagtatttgc tagtcaaagt gattaaatag aattgacgtc aaattctatc ataattgtgg tttcaaaatc ggctccgtcg atactatgtt atacgccaac tttgaaaaca actttgaaaa ggctgttttc tgtatttaag gttttagaat gcaaggaaca gtgaattgga gttcgtcttg ttataattag cttcttgggg tatctttaaa tactgtagaa aagaggaagg aaataataaa tggctaaaat gCGCATTtca ccggaattga aaaaactgat cgaagaatac cgctgcgtaa aagatacgga aggaatgtct cctgctaagg tatataagct ggtgggagaa aatgaaatcC TGtatttaaa aatgacggac agccggtata aagggaccTG GTCTTC

128 Gene Name: PIL253-5

Gene ID: BBF.2017.48.128

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707906

Sequence

GAAGACAAga ccacctatga tgtggaacgg gaaaaggaca tgatgCTGtg gctggaagga aagctgcctg ttccaaaggt cctgcacttt gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg gaagggtatg aagatgaaca aagccctgaa aagactatcg agctgtatgc ggagtgcatc CGCctctttc actccatcga cATTtcggat tgtcccctata cgaatagctt agacggccgc ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg gaagaagata ctccatttaa agatccgcgc gagctgtatg atttttttaaa gacggaaaag cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa gatggcaaag taagtggctt tattgatctt gggCGCagcg gcCGTgcgga caagtggtat gatattgcct tctgcgtccg gtcgatcCGC gaggatattg gggaagaaca gtatgtcgag CTGttttttg acttactggg gatcaagcct gattgggaga aaATTaaata ttatatttta ctggatgaat tgtttTAAga tacaattaaa ggctccttttt ggagcctttt tttttggaTT GAGATCGTTT TGGTCTGCGC GTAATCTCTT GCTCTGAAAA CGAAAAAACC GCCTTGCAGG GCGGTTTTTC GAAGGTTCTC TGAGCTACCA ACTCTTTGAA CCGAGGTAAC TGGCTTGGTG GTCTTC

129 Gene Name: PIL253-6

Gene ID: BBF.2017.48.129

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707907

Sequence

GAAGACAATT GGAGGAGCGC AGTCACCAAA ACTTGTCCTT TCAGTTTAGC CTTAACCGGC GCATGACTTC AAGACTAACT CCTCTAAATC AATTACCAGT GGCTGCTGCC AGTGGTGCTT TTGCATGTCT TTCCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG ACTGAACGGG GGGTTCGTGC ATACAGTCCA GCTTGGAGCG AACTGCCTAC CCGGAACTGA GTGTCAGGCG TGGAATGAGA CAAACGCGGC CATAACAGCG GAATGACACC GGTAAACCGA AAGGCAGGAA CAGGAGAGCG CACGAGGGAG CCGCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCACTGATTT GAGCGTCAGA TTTCGTGATG CTTGTCAGGG GGGCGGAGCC TATGGAAAgg cggtaatatt gttctggata ttaccagcaa ggccgatagt ttgagttctt ctactcaggc aagtgatgtt attactaatc aaagaagtat tgctacaacg gttaatttgc gtgatggaca gactctttta ctcggtggcc tcact-gatta GCAATGGCAC CTGCACCaAT CCGTAAAACG ACGGCCAGTa gtcaaaagcc tccgaccgga ggcttttgac tTGGTCAGGT GGAGTGGaat tcgcggccgc ttctagagtt GGAGcgagac ggCGCATCGC GGAGCGAGAC Ctgagagcgt tcaccgacaa acaacagata aaacgaaagg cccagtcttt cgactgagcc tttcgtttta tttgatgcct ggtcatttgt acagttcatc cataccatgc gtgatgcccg ctgcggttac gaactccagc agaaccatat gatcgcgttt ctcgttcgga tctttagaTG GTCTTC

130 Gene Name: PIL253-7

Gene ID: BBF.2017.48.130

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707908

Sequence

GAAGACAAta gacagaacgc tttgcgtgct cagatagtga ttgtctggca gcagaacagg accatcaccg attggagtgt tttgctggta gtgatcagcg agctgcacgc tgccatcctc cacgttgtgg cgaattttaa aattcgcttt aatgccattt ttttgtttat cggcggtgat gtaaacattg tggctgttaa aattgtattc cagcttatgg cccaggatat tgccgtcctc tttaaagtca atgcctttca gctcaatgcg gtttaccagg gtatcgcctt caaatttcac ttccgcacgc gttttgtacg tgccgtcatc cttaaaggaa atcgtgcgtt cctgcacata gccttccggc atggcggact tgaagaagtc atgctgcttc atatggtccg gataacgagc aaagcactga acaccataag tcagcgtcgt taccagagtc ggccaaggaa ccggcagttt accagtagta cagatgaact tcagcgtcag tttaccatta gttgcgtcac cttcaccctc gccacgcacg gaTGGTCTTC

131 Gene Name: PIL253-8

Gene ID: BBF.2017.48.131

Submitted by Keoni Gandall on 2017.11.20. Contact at koeng101@gmail.com.

Description

pIL253 plasmid resynthesis.

Motivation

Test pIL253 plasmid. Hypothetically this is a high copy theta-based replicating plasmid that functions in Bacillus subtilis.

References

PMID: 23707909

Sequence

GAAGACAAcg gaaaacttat gaccgttgac atcaccatcc agttccacca gaatagggac gacaccagtg aacagctcgt cgcctttacg catctagtat ttctcctctt tctctagtag ctaagatccc tagtatttct cctctttctc tagtatgtgt gaaattcgac aCATTATACG AGCCGATGAT TAATTGTCAA gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagGGT CTCTCGCTGC GATGCGccgt ctcACGCTac tagtCctgca gggctcGAGG TGTCAATCGT CGGaGCCGCT GAGCAATAAC TAGCATAACC CCTTGGGGCC TCTAAACGGG TCTTGAGGGG TTTTTTGCAT GGTCATAGCT GTTTCCTGTC GGCTTGGCAG GTGGCCATTG Cctggcgtac cgttcctgtc taaaatccct ttaatcggcc tcctgtttag ctcccgctct gattctaacg aggaaagcac gttatacgtg ctcgtcaagt taTGGTCTTC