

IDENTIFICATION

Species: *Sorghum bicolor*

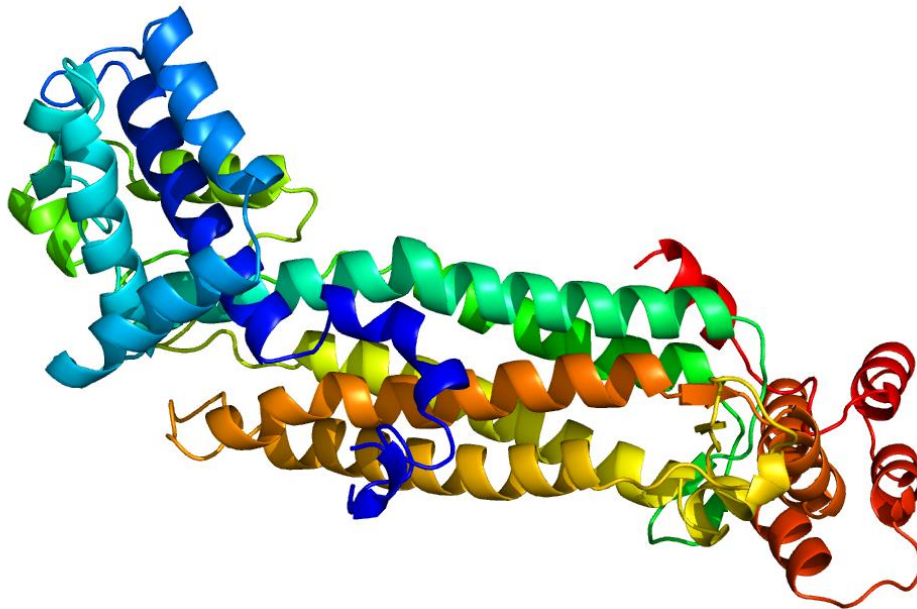
Locus: Sobic.006G157000

Gene Model: Sobic.006G157000.2.p

Description: SbEXPB-38

Family: Beta Expansin

3D structure:



GENOME DATABASES

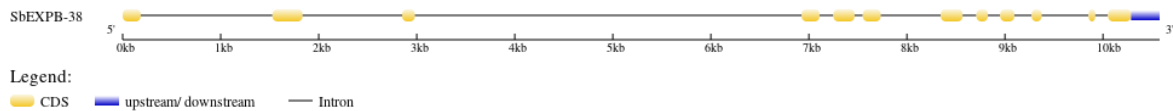
Phytozome: https://phytozome-next.jgi.doe.gov/info/Sbicolor_v3_1_1

KEGG: <https://www.genome.jp/entry/T01086>

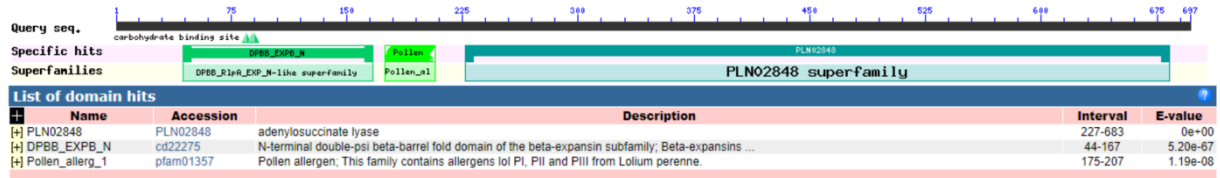
EXTERNAL RESOURCES

-

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>SbEXPB-38

MGSSRALALFLLCALLPAPPISVSAALLFGGGKSAKAAGADMWRPATATWYGEAE
GDGSDGGACGYGTLVDVPMKARVGSVSPVLFKDGEGCGACYKVKCLDRGICSR
AVTVIVTDECPGGLCAFGRTHFDLSGAAYSRMAVAGAGGRLRDRGQLNVVYRRTA
CKYGGKNIAFRVNEGSTNFWLSLLVEFEDGEEDIGSMQMKQMVGAPLDWAGRSLEE
LRSLPDHDTFCLMALSPLDGRYDRFVKELMPFFSEFGLIRYRVLIEIKWLLKLSQIPEIT
EVPQFSKEAQAFLNIIEDFCIDDAKEVKKIEKVTNHDVKA VEYFLKQRCSSNPEIAK
VSEFFHFGCTSEDINNLSHALALKEGVNRVMFPAMIDICRALCSLATQNSGYPLMART
HGQPASPTTVGKEMANFAARLSDIGKSFSEVKILGKFAGAVGNYNADVVAYPEVDW
PKVAEEFVRSGLQFNPHYVTQIEPHDYISKLFNLTQFNVLTFDRDMWSYISLGYF
KQIPKAGEVGSSTMPHKINPIDFENS DGNLCQANSILSGISMKLPISRMQRDLTDSTVL
RNLGMGLGHSLLAYKATIRGISKVQVNESRLAEDLEQTWEVLAEPIQTVMRRYGIPE
PYEKLKELTRGQAVTKDSMQQFIDGLDIPPEVRSKLSKLTPHSYTGLAEDLARDIEKW
VDLESGFQIK*

CDS (coding sequence)

>SbEXPB-38

ATGGGTTCCCTCGCGGCCCTCGCTCTCTTCCCTCCTCTGCGCTCTGCTGCCCCGCGCC
CCCCATCTCAGTCTCAGCCGCATTGCTGTTCCGGCGGCGGCAAGTCGGCCAAGGCG
GCGGGAGCCGACATGGAGTGGCGGCCGCGACCGCGACGTGGTACGGCGAGGCC
GAGGGCGACGGCAGCGACGGCGGCGCGTGC GGGTACGGGACGCTGGTGGACGTG
GTGCCGATGAAGGCGCGGGTCCGGTTCGGTGAGCCCCGTGCTGTTCAAGGACGGC
GAGGGCTGCGGCGCCTGCTACAAGGTCAAGTGCCTGGACCGCGGCATCTGCTCCC
GCCGGGCCGTCACGGTGATCGTCACGGACGAGTGCCCCGGCGGGCTCTGCGCCTT
CGGCCGCACCCACTTCGACCTCAGCGGCGCCGCCTACAGCAGGATGGCAGTCGCG
GGCGCCGGCGGCCGCTGCGTGACCGGGGCCAGCTGAACGTCGTGTACCGGAGG
ACGGCCTGCAAGTACGGCGGGAAGAACATAGCGTTCCGTGTGAACGAGGGCTCG
ACTAACTTCTGGCTCTCGCTCCTCGTCGAGTTCGAGGACGGCGAAGGCGACATCG
GATCCATGCAGATGAAGCAGATGGTTGGTGCTCCGCTGGACTGGGCGGGGAGGT
CGCTGGAGGAGTTGCGCAGCTTGCCGGACCACGACACGTTCTGCCTGATGGCGCT
CTCGCCGCTGGACGGGCGATACGATCGCTTCGTCAAGGAGCTGATGCCCTTCTTC

AGCGAGTTCGGCCTAATCAGATACCGTGTTCTCATCGAGATCAAGTGGCTACTGA
AACTTTCTCAAATTCCTGAGATCACTGAGGTGCCTCAGTTCAGCAAGGAAGCCCA
GGCTTTCTTGAATGCTATTATTGAGGATTTTTGCATTGACGATGCAAAAGAAGTA
AAGAAAATTGAGAAAGTAACCAACCATGATGTGAAAGCTGTGCAATACTTTCTG
AAGCAAAGGTGCAGCTCAAATCCAGAGATTGCAAAGGTGTCGGAGTTCCTCCATT
TTGGTTGTACCTCTGAAGATATCAACAATTTGTACATGCATTGGCTCTGAAAGA
GGGGGTAAATAGAGTTATGTTCCCTGCCATGATCGACATATGCAGAGCATTGTGT
TCCTTGGCAACACAAAATTCAGGCTATCCTATGTTGGCTCGAACTCATGGGCAGC
CAGCATCACCAACAACCGTGGGAAAGGAGATGGCAAACCTTTGCGGCCAGATTAT
CTGATATAGGAAAGAGTTTCTCAGAGGTGAAGATACTAGGGAAATTTGCTGGCGC
TGTTGGCAATTACAATGCTGATGTGGTTGCATATCCTGAAGTTGACTGGCCTAAG
GTGGCAGAAGAATTTGTTAGATCCTTGGGTTTGCAGTTTAATCCCTACGTTACTCA
GATTGAGCCTCATGACTACATATCAAAGCTCTTCAATCTATTACCCAGTTCAACA
ATGTGTTGACTGATTTTGATAGAGACATGTGGTCCTATATATCATTAGGCTACTTC
AAGCAGATACCAAAGGCTGGTGAAGTTGGTTCTTCCACTATGCCTCATAAAATCA
ACCCATTGATTTTGAAAACAGTGATGGCAATTTATGTCAGGCGAATTCATATT
ATCTGGTATAAGCATGAAACTGCCAATATCCCGGATGCAGCGTGACCTGACAGAC
TCAACTGTTTTGAGAACTTGGGTATGGGATTAGGTCATTCTCTATTGGCTTACAA
AGCTACCATACGTGGAATCAGCAAAGTTCAGGTGAATGAATCACGTTTAGCTGAA
GACTTGGAGCAAACCTTGGGAGGTCCTTGCTGAGCCAATACAGACAGTGATGCGA
AGATATGGAATACCTGAACCTTATGAGAAGCTGAAGGAACTGACGAGAGGTCAA
GCTGTGACCAAGGACAGCATGCAGCAATTCATTGATGGTCTAGACATAACCAGAG
GAGGTTGATCAAAGCTTTCGAAGCTAACCCCCATTCTTATACTGGGCTAGCGG
AGGATTTGGCCAGAGACATCGAGAAGTGGGTTGATCTGGAATCTGGATTTCAAAT
CAAGTGA

Nucleotide

>SbEXPB-38

ATGGGTTCCCTCGCGCGCCCTCGCTCTCTCCTCCTCTGCGCTCTGCTGCCCCGCGCC
CCCCATCTCAGTCTCAGCCGCAATTGCTGTTGCGGCGGCGGCAAGTCGGCCAAGGCG
GCGGGAGCCGACATGGAGTGGCGGCGGCGACCGCGACGTGGTACGGCGAGGCC
GAGGGCGACGGCAGCGACGGTGAGCACTCTCCTCCTGGGCTCCTCAACTCCGTAT
CGCACTGGTCTTATCCTAGCGTTTACTGGGGGGCATAGTAGATACGGAGTAGTTG
TAGTATAGGATACGAGCAAAATACTTGCAGTGTACAGCCATACTGTCATTCTTAA
TTCCTCGGTGAGTTGATTATTACCGCAGGATTCTTCAGCTGTTACTCTATCTATC
TACTGCTATCGCAGTGTACAATTTTCTTTACCGCGTATAGAAGAAAATGGTTTTGT
GGTGGCTGGTAGTGGTAGCAGTAAAAAAAATAGGAGAGGAATAAATAATTCCAC
CCGTTGCGACAATTTTTTTTTGGCGGGACCTCCCTGCGCATCCCATGAATCGCGGA
GCGAATTTGCGTGGGCCCCGGGCTGGGCCCCAGTACGCGTGTCAAGCGCCTAGGTG
CCCCGTTGCTTTAGGATTCGTGCTGCCGCCAGCCGTTCTCTCAGCCCCGGCGACC
TTCCAACATGCAACAGGTCCATGCAGTCCCCAGTCTTCGGTACCGTTTGCTGAAA
TCTTGACGAATTAATGCAGCATTGCACTCTCTGCCCCGTGTCACCGACACTGCG
GGTAATCTCCATGCTATCAAATCAGTAGAAGTTACAGTAATCCCTGCTCGGTTTTT
AACTGCTGCCATTTACGCCGTCTGAGCAGTACAGTACATACGTCGTGCACTGCA
TGACCTTCAGGTTTGTGTCCTATTGACGCGTCTCGCAACTGAACAGTACCCACGT
ATGAGCCCGTGACTTTTGTCCATCACAATTGGGCCTTGTTTAGTTCCAAAATTTTT

TGCAAAACAGGTACTGTAGCACTTTCGTTTGTGTTGTGACAAATATTGTCCAATCAT
GGACTAACTGGGCTCAAAGATTCATCTCATCAATTCCGACCAAATATGCAATT
AGTTTTTATTTTCGTTTATATTTAATACTTCATGCATGCGTCTAAAGATTCGATGTG
ACGGGGAATCTGAAAAATTTTGCAAAATTTTTTGGGAAGTAAACAAGGCCTTGAC
ACACACAGCCCACAGGCTCTGTTACAGTCAACCTGCATTTTCCATGCCCTGTACGT
CTCGGCTGTTCCCTTCCCTAATTTCTCTTCTTGTGAAAAAATTTTGAATTCGTTACT
GTAGCATTTTTATTTAGGATAAATATTATCCAATTATAGACTAACTAGGGTTAAA
AAAAAATCGTATTGTAATTTACAGTCAAACAATGTGATTAGTTTTTATTTTCGTTT
ATATTTAATGCTTAATGTATGTACCGCAAGATTCGATTTTTTTTGGTTTTTCGGTGTG
AAGTAAACAACGCCTTGTTTTTCTTTCAGGGCGGCGCGTGCGGGTACGGGACGCTG
GTGGACGTGGTGCCGATGAAGGCGCGGGTTCGGGTCGGTGAGCCCCGTGCTGTTCA
AGGACGGCGAGGGCTGCGGCGCCTGCTACAAGGTCAAGTGCCTGGACCGCGGCA
TCTGCTCCC GCCGGGCCGTACGGTGATCGTACGGACGAGTCCCCGGCGGGCT
CTGCGCCTTCGGCCGCACCCACTTCGACCTCAGCGGCGCCGCTACAGCAGGATG
GCAGTCGCGGGCGCCGGCGGGCCGCTGCGTGACCGGGGCCAGCTGAACGTCGTG
TACCGGAGGTGAGCACACACCGCACAAACCCCTCTCTAGGCCTGATTTAGTT
CTAAAATATTTTGCAAAATGGACACGGTAGCTTTTTTCGTTTGTATTTGACAAATAT
TATCTAATCATGGACTAACTAGGCTTAAAAGATTTGTCTCGTCAATTCGACCAA
ACTGTACAATTAGTTTTTATTTTTATCTATATTTAGTACTTCATGCATGTGTCTAAA
GATTCGATGTGACGAGAAATCTGATTTTTTTTTTTATTTTTTTGAACTAAACAAG
GGCAGTCTTGCTCTGAATGCTTCGTCAAGTTCAGAAATCAACTGCTGCAGGTGTCTCG
TAAATCGTGATGCTGTGCTGAAAAATGGCCACTTGGCGTCTGAAAATATCGCTGG
TGCTATGATTTACCCCTGTAGGAGTAGTAGTACTTAATGTACTACGATCCTGATTC
CTGCGGGGGTAGTGAATTCAATTGTCCAGGGTTAAATGGGAGGGAATGCAGCCC
AGCTGTCCGTCAGTCTGAGAAGTCTGCTGGTGTAAACCCTGGGTAAACATGCT
GCTGCCTGAAATAACTCAGTCAGAGTCGCATCTGATCACTTGTAATAACCTCCTG
ATCAATCTCTGCAGGGAAGGGGGTACTAGAGATTGCAGCAGATGTCACTACTAG
GGCCTTGTTTAGTTCCGAAAAATTTTGGAAAATCGACACTGTAGCACTTTCGTTTG
TATTTGACAAATATTATCCAATTATAGACTAACTAGGCTCAAAGATTCGTCTCGT
CAATTCGACCAAATCTGTGCAATTAATTTTTATTTTCATCTATATTTAATACTTCAT
GCATGCGTCTAAAGATTCGATGTGACGGGGAATGTGAAAAATTTTGCAAAATTTT
CTAGGAACTAAACAAGGCCTAGAAAAAAAACGATTAGTTCAGTTACTGGCTTTGG
CAGTTAGATTCCTGTTCGTTTTCTTCTGAACTGACCGGTGCAGTGTGTTTTG
CCTGTATGCATAAACGGATACCAGGACGGCCTGCAAGTACGGCGGGAAGAACAT
AGCGTTCCGTGTGAACGAGGGCTCGACTAACTTCTGGCTCTCGCTCCTCGTCGAG
TTCGAGGACGGCGAAGGCGACATCGGATCCATGCAGATGAAGCAGGTAAAGGAC
GATTCCTTCTGTGAATAATTATTTTAAATTACGGTAGTCAGGGATTTGAGCTCGAG
ACATATAATTCTAAAATCATATTAATAATTACACACACCGACTAATTTAACTTACAA
GCTTAATTCCATTGAGAAGACGGAGAATTCATTTTATATTTTAGGAGTAAGATAC
ATATTTACCCCTCCAAGCACCAACTGGCTAAATCGGTGGTTTTACGTTCTAGTATA
ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA
ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA
TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT
ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA
ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA
TAATATGTTTGACCTCAAGAACTCATAATTAATAATGTTATCTACATGTAAAGCA
TCAACACTTAAATAACATTTTTAGAATAAAGTTGGTACTAGTTTCGTGTTTTCTAA

TTTATCAAAGAAATTATTTATGAATTCTAAAACTAAACCAGATTTTTATTATTT
TTAGAAACCAACATAGGGACAAATTGACTAGTTTCAACAGTTGGAGGGTGTCT
GTCATCCCATTTCAGCGCTCAAAATTGCAGTGATATAACAACCTAGGTTTGGAAAA
ATATACTTTACCCTATGTATATTTTTACACATCGTAGAGTAGTGGCACACAGCAA
AGCCATCGGGAGGAAAAATAGACTATCCCCTCTCCATGGGCCATGTATAGGTCTC
GAGATGCACCGAACAAAAAAGCCACAGCCCATAGCTTAACCCAGAGCATGTT
GAAGACATCAGACGTGTGCGTGCTCCAGTGTTGCCGCTGTCGTTTGGCCAATGCT
ATCCTGCCCTGTCTGGTAGATCGAATCGATGTCTGTTCGATATGAAGAGGCGAGAG
GCCGGCCAAAAGATCTGGCAGCTAAAAATCTCTTGACGTTCTTTACCCATTCTTA
AGGTCACACAAGGTCACAAGGAAAGGAAGGCAGCACCTTATCTGGAACAGGGAA
AATCCAGCACCGCATGGAACAGACCACAGAGGCGTGCCGTTGCAGCGTGCTAGTG
CTGCTACTTTCTCGCGCTGCGCTGTCTACCACACAAGCCCCTAGTCAAAGCTTTGC
ATAAACCTAAATAGTTTAAGACATGCATGATCTTCTTTGCCGAATGATTCCCCTTC
TGAATGTGTTGGAAAAGCAGCACGTGAGCCCCCTGGCTGCTCATGCTGCTGTCTG
TTGCTCCATCGAGGACACTCATCTCGATCAGTCGCCTTTACGCCCCCTGTACTGTA
ACTGTAATGAACTGCTCTGTGCCGGCCGCCATGATGTTGTTGTTTCTTCATCAGC
TTCGGTCTGAGTCTAAAGACGGCCATGAAGCTGGAAGAGTAAAGAAACCCATGC
CTGCTTTGCTTTTGCCTTGCCTAGTGTGCTTTTACGGTGGCCGTTGAATCCGA
GACAGCCATGCCCGTGCCGTGTGCCCGGTGGCATGCACGCTAGCTTGTGGCCGAT
CGAGTCGTCGTCGTCCTGGCCTTTCCTTTGTTGGGCTAGGGAATCCAAAGATGCC
AAGGAATCCCATGCGACTTCTTGGTCCCTTGGAGACGTGCTGTGCTGCTGCCCTTG
TTTTGTTGCCCTTGTAAAAAGTTTCAGACGCCGCTGCGCTTTGCCTAGCCCTAGTG
GTAGTGGCCTGGGGTTTTAACTAGCAGTAACAACAGCTGTTCCCTGCTATTGCTTT
ACATTCATTCACCTTTGTGTTTGTCTTGCCAAACCATTTGTCATGCAATCGCTGCTC
TGCTATGCTGTGTATGTTGCAGGCCAACTCGGTGGAGTGGCTGGACATGAGGCAC
GTGTGGGGGGCCACGTGGTGCCTGGTGCGGGGCCCGCTCGTGGGCCCCTTCTCCG
TGAGGCTCACACGCTGTCCGCCCAGAACGCTCACGGCCCAGCAGTCATCCC
CAGGAATTGGACGCCCAAGGCCACCTACACCTCGCGCCTCAACTTCCAGCCGTCT
CTCTAGGGAGGAGCCCCGCGGAGACTGCCGGAGGGGAGAGGGGGCGCGGCCGA
CCGGGGGCCACGTTTTATCGGCCCGGACTCTGCCTTGTGGGCCCAGAAAGGCT
GGTGGCCCACTTGGCAGGCAGCCAGGCAGGAGGAATCCAGACTGTTTTTTTTTTT
TTTTTTTTTGTGTAGTGAAGCAGCTTGGTTTTATATGATTTTATTTTATCCTCGT
AGCGTAGCTAGCTAGTGTGAGGTGATTCATGATCATCATAGCAGGCATCGGAGGA
CATGCATGCGTGATGGCTGATGAGTTGAGTCCGTTAGGGGGCCAGAGTGTGTCAGG
TACAGGCAGCCGCCCTCTCAAAGGAGAGAGCTAGCGCTCGCTGCTTGTCTCGG
GCTCCAGCTCAAGCCCTTGGAGGCTCGGAATCCGATGTGGAAGCTGGTTGCGCCG
TGAGGGTTTTGGATCCCTCGTGGTGTCTGCTTCTGCTTCGCTCGTGTGTGGTGTGG
CTGTGTTTTTGGCGTCGATGTGTTTCTGTCTTGTGTTTGCCTGTTGATGAATGCA
GGGTTTCGTGACAAAAGTTGTTTATCTGTCCACAATGAGGAATGCGCGTCTGATGC
TGCTCAGAAAAGGAAAGAAAAAAGAGATGAACATTCGTGGCATAATATATGTAG
TGAAGATAGAACTCATCCAAAAACATACTTTGAACCCTTGAAAACTAAGCAT
ATATACAATATCCATAATGTGTTAAGAGATGACCTTTTTTTAAGCAAAGTTGAG
ATGCAAAGTCGTTGATGATTTTTATTTATTAACAAAAACGATTTTGTCTTCTTA
TTTGCACAAGCAGCGGCGTATATGGAAAACTAAGGATGTGTTTGTGTTTCTCGAT
AGACTGAACTGATTTTTTTTTCATGAGAAAACTAGGCTCTATTTTTTTTTTATTACAA
CAAATAGGTATGTCTTCGTTCAAAAACAAATAGGTATGCCATGAGCTGTTTTATA
TCCGTAATAACATACATGCACTCAACTAGTAAAAAATAAAAAAATAAAATCA

TCTTCAAAACTATTTATAATAGGGCTAGAGAGGTAAGATAGATGATTTTAAAAAT
TAAGAAGATGTTTTGTTTCGATTTTGAATTTAAGGAGGAAAAAAGACTTTTACCA
AAGTTGAGAGAGATAATGTTAATTTTTTTCCGTTCTCCTTGCTAAAGACCGAATA
GATTTTCGGTCTCCATGCTCTCCGGTACCAAAGCATCGAAGGTCCAACGCATCGA
GCCATGAACCAGCCCATTACTCCGTCTCCAGCGCAGGTCCGCGTCCTCCCTCGC
AACCACGCGGCCCTCCCCCGCGGTCCGCGGAGGTTCGCGGCCACAGGCGCT
CGGCCCTCCCCGCCGGAGGTTGCCGCCCCGCAGCTACGCCGTCCTTTCACGCCG
CATTTTCCCCGCACCCGCAGCCAAACCCACCGTATCGTTCGGTTCCCCTGAGCCTA
GGGTTTAACCAGCAAACGCCGCGCTTCTCGAAGCATGTCACCGCCGTCGTCGCTC
AAGCACCTGCCCCCGGCAACCGCCGCGGCTCTCCGCGCGCGGTTCGCCACCC
CGAGAGCCACCGGGATTCGCGGCCTCTGTGCACTCCCGACTTGGCGCCTCCGGTG
CTCAGCCTCCGCGCCGCTGGCCGCCACCGAGGTGATGGTGCACGCCTCCTCCACG
TGGAATTTCCCGGAATCCTGCTAGTGAGGTATTGGTTGCTTTGTGACTTACGGGCT
GTGCGGTGTTGTTTCAGATGGTTGGTGCTCCGCTGGACTGGGCGGGGAGGTTCGCTG
GAGGAGTTGCGCAGCTTGCCGGACCACGACACGTTCTGCCTGATGGCGCTCTCGC
CGCTGGACGGGCGATACGATCGCTTCGTCAAGGAGCTGATGCCCTTCTTCAGCGA
GTTCCGGCCTAATCAGATACCGTGTCTCATCGAGGTATCTGACATTTGGGACAGTT
TAAGCTCCTTGCAAACCTCTGCATCTGCCTACAATTTTGATTGGAGATTGGACGCTT
GAGAGCCTGTTTCTGGTGTATTGTCAATGCTGCTTAGCTGACGCACCTTTCTTGC
TCTTCAGATCAAGTGGCTACTGAAACTTTCTCAAATTCCTGAGATCACTGAGGTG
CCTCAGTTCAGCAAGGAAGCCCAGGCTTTCTTGAATGCTATTATTGAGGATTTTTG
CATTGACGATGCAAAAAGAAGTAAAGAAAATTGAGAAAAGTAACCAACCATGATGT
GAAAGCTGTGCAATACTTTCTGAAGCAAAGGTGCAGCTCAAATCCAGAGATTGCA
AAGGTTGCAAGTAGGATCTTGTTGTGTATAGGATTAATTATTTGTGGTTAGACTTT
AAGCTGAGTGGCATCTTGAACCTGTTTTTCAGGTGTCGGAGTTCTTCCATTTTGGTT
GTACCTCTGAAGATATCAACAATTTGTACATGCATTGGCTCTGAAAGAGGGGGT
AAATAGAGTTATGTTCCCTGCCATGATCGACATATGCAGAGCATTGTGTTCCCTG
GCAACACAAAATTCAGGCTATCCTATGTTGGCTCGAACTCATGGGCAGGTATGAC
TTATGAAGCCACCTTCACCATAATTTTTATTGCTGTTATACAGCAAGAACAAGA
AGTTTTTTGTACCCTTTGTTTGAAACATAGCTTATCCATAGTCACCTCTCGAAGGA
CTTCTATGTGAGATAATGATATAAAAATAATTAATTAATCGGGGCTTAATGTTTTGT
AAGTTTGCCTGGTTGATGCAGCAAGGTGATTTTGAGTGATAGTAAATTGATAAGT
TCATAATGAAAAGAATGGAATTGAATGACTTATCAACTCATTACAGCTTTTTCCAA
GGCAACATGGGTGGAGGAAATACAAACCTGCACTATTTATCAACATAGTCTAGG
AGATTTAATAAAAAGATTATTGTACTGTTCTGGCAGATAATCTTCAACCTTTCCAT
TTGTACAGGAAAACGGAAAAATATTTCTATATGTAAAAAAATTGTTCTTCAGCAG
TGTTGCTTTGTGGAGTTCAGCTCTGTAGCTCTTTTGATACAATAGTTATTAGTTAA
CATAATTGTCTTGCTCAATTGCTTGTATGTTAGTGGAATGCGCTTGCTTTGCCAAG
AAAGACATGATTTGGAACCCTGTCTTAAATTGTTTCCAATTCCTGCTTGTAGCCAG
CATCACCACAACCGTGGGAAAGGAGATGGCAAACCTTTGCGGCCAGATTATCTG
ATATAGGAAAGAGTTTCTCAGAGGTGAAGATACTAGGGAAATTTGCTGGCGCTGT
TGGCAATTACAATGCTGATGTGGTTGCATATCCTGAAGTTGACTGGCCTAAGGTG
GCAGAAGAATTTGTTAGATCCTTGGGTTTGCAGTTTAATCCCTACGTTACTCAGGT
GATCTATATTCTTTGTTTTGACATTTAGTTTCTTCTCTGGTCTTCTATGTTCTTGCA
CAACTAGGGAGTCTTCTCAAATTCAGGGTGATCATAACCATCAACGTAGTTGGT
TCTGATGCAGTTTTATCCCTTTTACAGATTGAGCCTCATGACTACATATCAAAGCT
CTTCAATCTATTCACCCAGTTCAACAATGTGTTGACTGATTTTGATAGAGACATGT

GGTCCTATATATCATTAGGCTACTTCAAGCAGGTAAACTAAAATGGCTTGACACT
ACAGTGTGGGAGGATTAACCTTGTACTGCAAGGGAACAAATTGATGGTATTGATT
TTTATCTTCCATATTAACCTTGCAACATCTTCCAAAACCTTTTTTTCAGATACCAA
GGCTGGTGAAGTTGGTTCTTCCACTATGCCTCATAAAATCAACCCCATGATTTTG
AAAACAGTGATGGCAATTTATGTCAGGCGAATTCTATATTATCTGGTATAAGCAT
GAAACTGCCAATATCCCGGATGCAGGTATTGTCACTAATCCAATTTTATGCCTCA
ATTTCAAGTTTTTCTAACATGAGGTGTCAGGATGAAAATCTGGTGCCTGTATTCATG
GCACCTATGCGAGATTGGATCTTAAATTACATTTACCATATCATTAGTATATACAC
ACACTAATGAACCACCATGTTGTATCATTGACAGCGTGACCTGACAGACTCAACT
GTTTTGAGAACTTGGGTATGGGATTAGGTCATTCTCTATTGGCTTACAAAGCTAC
CATACGTGGAATCAGCAAAGTTCAGGTAGGTGGCCCAGAGCTCGTCCTTTGTCTA
GCTTGTCTTTTTTCTCGTTACAGATGATATAGTTGTAGCTTCAACTTCAATATAGTT
GATGCATGTTTTAATTTAAGTTTATGATTATTATAAGTACTTGCAGAATGCTTGAT
GATCAAGCAGATAGACCTACCTCACCCTTGCTCTTTTCATAAGTTTTGGCTTGAC
GGGGGAAGATACACTATAAATTAGTGGTGACTGCTAGTGTAGGTGTTTACCCAAC
TAAAGTCCACACCCTCAAACCAGATGAGGATAGCTGGCTTAGTAGGATAAAGA
TCTTCCTTTTATGCAGTGCTCAACTGCACGCACTTCTTGCACCAATACATGGGAAT
ATCAGATGTTTACTTTTCTTTACTATTGAACTGAGATAACTAGGTAAATTTTTGC
ATGCATTTTTAGTAATTATGATCAGGCTAGAAATTTGTGCTGAATCGGTGCCTATA
CAGGTGAATGAATCACGTTTAGCTGAAGACTTGGAGCAAACCTGGGAGGTCCTTG
CTGAGCCAATACAGACAGTAAGCCATCTGATTTTTCATTTTCTGTCCATTTCCCCT
TAATTCTGTTCTGGTTGCTGCCACAAAGTGCTTGTTGTGTGCACATCAGGCATGTA
TCATGTTACTGTAATACTGTTATTACTCTTTGTAGGTGATGCGAAGATATGGAATA
CCTGAACCTTATGAGAAGCTGAAGGAACTGACGAGAGGTCAAGCTGTGACCAAG
GACAGCATGCAGCAATTCATTGATGGTCTAGACATACCAGAGGAGGTTTCGATCAA
AGCTTTCGAAGCTAACCCCCATTCTTATACTGGGCTAGCGGAGGATTTGGCCAG
AGACATCGAGAAGTGGGTTGATCTGGAATCTGGATTTCAAATCAAGTGAGCTTCC
ATTTAAAAATGGATAATCCTCCCCACTTTTTTCTCTTTGACAATGACAAAAAAA
TAACAGACGGGAGCCCTAGAAAACGAAAATTTTAATGGGTTGTGGCCTCTGAATT
GTCGACCAAGATAATTTACGTGGTTCTCATTTATGATAATTTTTGGCACATGTAAA
ATGCGCCTCAAGGAACTTATGCTAGCTGTATCTTGGTGAAAAATTATCGAGCTAA
AAAAGTTGCTTGCTCGATTTCTTATAAGTACGGAGTTGTTTTGCAACGCCTGGTGC
GGTAAGCT