

## IDENTIFICATION

**Species:** *Miscanthus sinensis*

**Locus:** Misin08G116100

**Gene Model:** Misin08G116100.1.p

**Description:** McsEXPA-44

**Family:** Alpha Expansin

**3D structure:**



## GENOME DATABASES

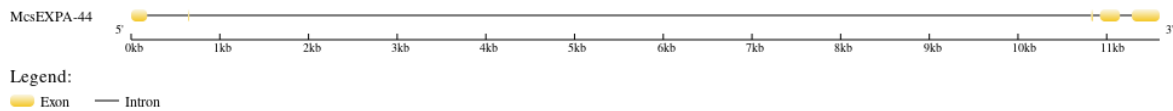
Phytozome: [https://phytozome-next.jgi.doe.gov/info/Msinensis\\_v7\\_1](https://phytozome-next.jgi.doe.gov/info/Msinensis_v7_1)

KEGG:-

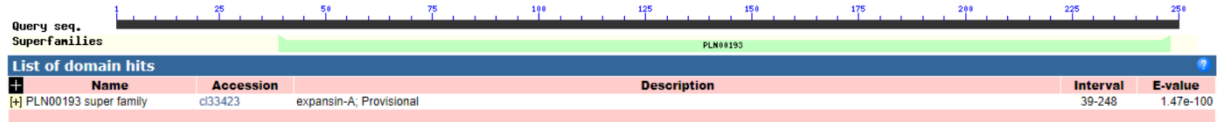
## EXTERNAL RESOURCES

<https://grass-genome-hub.southgreen.fr/Genomeassembly/47213>

## GENE STRUCTURE



## DOMAIN ARCHITECTURE



## SEQUENCES

### Peptide

>McsEXPA-44

MAAPARARLFTLLLLASAGWVSAMAGNDTVPSTTSSSPAPTGWLKAHATFYGGADA  
SDAMGGACGDINPITCSCGQCYKIACDRKRADPRFCKPGVRVTVTATNFCPPNLALP  
DGGWCNQQRPHFDMAQLAWEKIGVYSGGIIPVIVPCVKRGGVRFTINGHDYFNLVL  
VTNVAAGSIKSM DVKSSNSTDWTPMARNWGANWHSVAYFTGQMLSFRVTDTDG  
QTIEFTNVVPQG WKFGQTFTCKLQFK\*

### CDS (coding sequence)

>McsEXPA-44

ATGGCGGCTCCAGCTAGAGCTCGACTTTTCACACTCTTGCTGCTCGCATCTGCCGG  
TTGGGTGTTCGGCCATGGCTGGTAATGACACCGTCCCCTCCACCACGTCGTCGTCTC  
CGGCACCAACCGGGTGGCTAAAGGCGCATGCCACGTTCTATGGAGGCGCTGATG  
CCTCGGACGCCATGGGCGGCGCGTGCGGTGATATCAACCCTATCACTTGCTCGTG  
CGGGCAGTGCTACAAGATCGCGTGCGACCGCAAGAGAGCCGACCCGAGGTTCTG  
CAAGCCCGGCGTCAGGGTGACCGTCACGGCCACCAACTTCTGCCCGCCCAACTTG  
GCGCTGCCCGATGGCGGGTGGTGCAATCAGCAGCGCCACACTTCGACATGGCGC  
AGCTGGCCTGGGAGAAGATCGGCGTCTACAGCGGTGGCATCATCCCGTCATAGT  
TCCTTGCGTGAAGCGAGGTGGAGTCAGGTTACGATCAATGGACATGACTACTTC  
AATCTTGCTGCTTG TGACCAACGTTGCGGCTGCCGGCTCCATCAAGTCCATGGATG  
TCAAGAGTTCGAATTCGACCGACTGGACCCCAATGGCGCGCAACTGGGGTGCAA  
ACTGGCACTCTGTGGCCTATTTTACAGGGCAAATGCTCTCGTTTAGAGTCACCGA  
CACCGACGGACAGACTATTGAATTCACAAACGTGGTGCCACAAGGATGGAAGTT  
TGGCCAAACATTTACATGCAAGTTGCAGTTCAAGTGA

### Nucleotide

>McsEXPA-44

ATGGCGGCTCCAGCTAGAGCTCGACTTTTCACACTCTTGCTGCTCGCATCTGCCGG  
TTGGGTGTTCGGCCATGGCTGGTAATGACACCGTCCCCTCCACCACGTCGTCGTCTC  
CGGCACCAACCGGGTGGCTAAAGGCGCATGCCACGTTCTATGGAGGCGCTGATG  
CCTCGGACGCCATGGGTAATAGGGATGAAAACGGATCGGATACGGATGGATAT  
CATCCATATTATATTTGTTTTTCATATTTCTATTCGGATTTCGGATTTCGGATACGGAT  
AGTGTCAACTATGTCAGATAGGATACGATTGGATATCGACATCATAAATATACGA

TTTGAGTATTCGGATACGGATACGGTATCGGATGTTTAATATCCGGACTCAAATA  
CGAACAGATCTCAACCCCTCTAAACGGATTTCGGTTTCGAATACGGTCGAAAAATA  
TCCGTACCGTTTTTATCCCTATGGATAAACTCTAGCTAGCCGCTCAAGAACCATGC  
ATGGGACAGCTTTCAGCTTCTAGTTTCTATGCCATCTGCCATCTGCCATCAAGGC  
ATCAACATGATGAAGTGTGCCGAACGATATCACATCGCATCTGTCTGTCGTCAAAT  
CTAAGTATCTGACTCTCCTATATTTATGGTGGCAGGCGGCGCGTGGGGTAAGGC  
AACCTCTACTCCCAGCCTGTTTCGGTTGCTCGTAAACGATCGTAAATTTCCAGCTAG  
AAACAGTGTTTTTCTCTCATACTAAACCAGCCAGTAGTAAATAATTCACGATACG  
ATACGGCCTCCCGAACAGGCTGACGGCACGCGGACGGCGGGCGCTGAGCACGGTG  
CTCTTCAATGACAGCGCCTCCAATACGTCAGATTTGGACATCACTGTCGGCCTTAT  
CACTGTCGGATTTATGAGAACCGACAGTGATATTTTTTCACTGTCGGTTATGAGCT  
CGAAAATAATAAAATAATTGGGGCTGAGTTAAAACCGACAGTGAAGCACCACAT  
CACCGCCAGTTTGTGGTTTGAACCGACAGTATTTTGGCGGCCACCCATCACTGTC  
GGTTCAAACCACAAACCGGCAGTGATTGGGCTCTATCACTGTCGGTTCTAGCCAA  
GAACCGACGGTGATAAACCTACAGCACAAAAGGCTGCAGCCGTCCTTCTTCC  
TCCTCTCTCCATCCCGAGAACAGAGGCGCGTGTGGGGCGCTTCTCCATTGTT  
GTGGCTGCTCTTACCATGAAGCTAGCGCTTGGGTTTTGAAGAATGGCTTGATCTT  
TTACTTTAAGATTAGTAATAAACATCCACTCCTTTGATTTAGTTGCTTAATTAGC  
TCTGTTTTGATGGCTACATTGCTTGATTCTCATTCTAGTTCTTTCTCCATTCTAGAT  
AGCACCTTTTCAATTGGACATTTGTGCAAGGCTCAAATGTGGTGAATCCATCAT  
CCAAACGGTTGGTGTCTATGAACACATTTAAATTTGTAGCTAATTATTCCATGGTG  
GTAAAGATCATCATTGTTAGTATGTGATGCTATAATTAGTTCTTAGAAGTAGAGT  
AGAATAGTTGTTCTTCAATATTGTGCAATAATTGTTTTTACTATAATTTTCAATTTA  
CAAGGCCGGGGCTACCAATTTCAATTTTCCAATAATTAGTGTATAATAATGTTTTC  
CAAAAATGATACTTTCGAGCTATAATTGTTGTTTCATGAAGCTCAATTTCTTATTAA  
TTCTTTGTAATTACTGTCTCAATATTTTTTTTTGTGCAGAGATAGATCTAGAGTGG  
ATGCATCTGTCTCGAACGGACAAGCGGTACATGCATGGCGTCAGCCAGTTTATCA  
CCGATGCCAAAGCTCATGCTGGGAATAGGAACCCTGTCTTCTGCCCATGCAAAGA  
CTGCAAGAATCAAAGGAACTTTCGTCAAATTGAGCCTATACGATCGCACTTGATT  
ACTAGGGGATTCATGCTAAACTATACGATATGGACTATGCATGACGAGGTTGGTG  
TGAATGTTCTGCAGGAAAACGATGATGATGTGGACATGCCTAACGTAGCCATCCA  
CGATGGTGATGAGGAACTCGGTGTCAACACGGAACCTATGCCACAGTTAATAAT  
GTGTTTAGGAACACGCTAGTTGATGACTTACTTTGATATCTCATTGCATCTGCTCA  
TTCATCTTGTTGACCAAATTAGAGCCTTTGGCCTAATGTAACCTGCATCAGATGTTT  
TATTTGCAAAGATTGATGAAAGTTTTTCAGGAGGAATGTTAGGAACAGATTCAGGA  
CAGAAGGGGGCATGGTTGAAGGATGGTCAACGGGAGAAGGCCATTGAGTTCTGCA  
CATATTATCTGGATATCAAAGGTTTCGGAGTTCTAGAATCTCGTCATGAGGGAAG  
ACTACGTGGCAAAGGAACGATCGGGGAGAAATCTGTTACAGTAGACGACCCTGT  
TTTTTTTAGACAGGTACAGTTCGCTGTTCTTTAGCAAGTCGAGGGTGTGATGCCAT  
ACATTGACAAGCATAGGCAATCGCTGTAAACTCTATATCCAAGCAGGTCATAGGT  
TTTGCAGGATAAAAAATATAAGGAGGAATTTGTCAGCTGGTTGCGACGTCGCTTG  
CTTAGAATAAAGTTTGGTAATCAACTGGATACCTTAGCCAAGGGGCCCTTCGAGTA  
CATATCTTAAGTACCAAGGGTATAAGATCAATGGATTACATTTTACATAAAAAA  
GCAAGATGGAAAGAGCACATAACAAAATTGTGGTGTTCGTTTTAATGCTCACGAC  
AAAAACAGCAACGTGCAGGCGACATACTATGGTTTCATAGAGGAGATATGGGAG  
CTAGCCTACGGTCCGTTGAAGGCAGCTCTTTTTTGTCTGCCAGTGGGTCCGGCTCGA  
GGAAATCAACACTGACAGCAAAGGGTTCACTACCGTTGATCTACTAAGACCGCA

TATAGAGACGATCCCTTCGTCCTTGCAAGAGATGTTATGCAAGTCTTCTATGCAA  
GGGACAACAAGACAAAAGAAAGGCTAAAAGTAGTCCTAGAAGGGACAGGAAGA  
TTGTCCGGTGTTCGATGGAGTGACGGACGAAGAAGACTACAGGGGCTATCAAGAAA  
TGCATCCATTCCGGGGCAAACGTACCCCTACCTATCCTTCAAGTGGGTGACGAACC  
TGCTTACGTACGAATTAATCACAATGAGGCCCTCATTGTTGATGCACCTAAAGAT  
AGTTAGATTAATATTAATACTATGTAATCATTATGTAGACGTTGTATCAGTAATTCAC  
ACTGAATATTTAATTTATATTTTATGCGCATCTCTACAATTTAATTATTGACTATGT  
AATCAAATATTAAGATGATGTTCAAGACACTTGCTTTGAATGCTGGTGATATCCC  
CATCCCGTAAGCAAGCGAATGGTTTTCGAGCTACAACCAAGTGGATGTGGTAAA  
ATGGTCACTTATACTTGAGTCGTGGTTATTTGTTGTTTATTATTGTAATAACGTTCT  
CTATTTTTTTCTTTTTCTCTGCATGCAGATTGCGCTCCAGGACCTTACCTATGTTGC  
ATGTGGCTTATGGGCCGTTCTAGACCAAGCTACAGAGCAACTTGGGTACGATTGG  
GACTTCTACAATGGAAACCTCGGCCAACTCACCGTAGAAGGATGTCAGTACTGCA  
TAAGGATTACTAACAGGGGTAGTGGTCGTTTCGCTAGATTACATCTATGGCCGACC  
ATTCTTTCGGTATTGTGAGGCATGAGAGAGTGCACCTTACGGATATTGGACTTC  
ATCGATACAAGCGAATACATAATCCGTGACATCAATTACCATATACGGCTACATA  
GGCATGTTAGTGTACGTGGCCTGATTGATCCCGGCAGTGATTCGATAGTAATTA  
ATGTTTATTTAGCTAGGTTTTCAAGGTCGTAGTTTAATTGGGTTCTAATTTTAATA  
TGTATGGCTTTGTGTGTTAATTATTGGCATGCATGTAATTCGTGTAACATTTGTT  
GGGCAACTAGAAAATACATTCTGATGTCGTATTGGTCTCAAAATTATTCTTAGG  
CTTACATGTGCCACATGTGAAGGAAACACCAAGTTTGAGATTTTTTCGGAGATGGT  
TTGCTACTTTCTGAGGTTTAATTGAATTTCCGCGCGACGACACCGTTTAATTATAG  
GTTGAACTGAGTCTACCCACGAAAAGATCTGAAATGGTTCTAACTTTTACACAG  
GCTCTAATGTGCCCTAGGGATAAGAATTTTGTAGAGGTGGATGAAAAAATCTATT  
GGGTCCAAGATGAAATTCACCCCTCGTTTGTCTCATTCCGGCACATAGAAAAATATA  
ATAAATAAAAAAAGATCAGAAAAACCTCAGAACCTATGTGGGGTCTTAATTTGG  
GTGTACATGCTTGTCAAAAAAATTCAAGTCATTTAGACATAGGCGGAGTATACA  
TGATTCACAGAGGTACACGTGCCCTTCCATCGAACCATGACTATACACATAGGTT  
ATCAATGTTTCTGTGGTTCATAGGCATTCTAGTGTTCGTATTTCGTCTTAAATTTTTT  
CTCAGGCTTGCACGTGCCACGTGTGAAGGAAACACCAAGATTGGGATTTTTTGG  
GATGGTTTGCTACTTTCTAAGGTTTAATTGAATTTCCACGCGATGACATCGATTAA  
TTATAGGTTGAACTGAGTCTACCCATGAAAAGATCCGGAATGGTACCAAACCTTT  
ACACAGTCTCTAATGTACCCTGGGGAGAAGAATTTTGTGGAGGTGGATGAAAA  
ATCTATTGGGTCCAAGATGAAATTCATGCTCTTTTGTCTCATTTCAGCACACAGAAA  
AATATAAAAAAATAAAAAAATGATCAGAAAAACCTAAGATCCTGTGTGGAGTCT  
TAATTTGGGTGTACAAGCTTGCCAAAAAATTTCAAGCCATTTTCGACATAAGAAT  
ACATATGCTTCTATTTATTTAGTATATAAAAGAAATTTTTTAATATATATAAAC  
ATCACTGTCCGTTTCAAAAAACCGACATTGATGAGAATAAACCGATAGTGATACT  
GGTTATCACTGTCGGTTTATTTTTGAAAACCGACAGTGATATATAAAAAATAAAA  
AAATTATGCTAGTCTCAGGTTTCGAGCGTGGGTCTCGGGCTGCAGAGTTCAGAGA  
CTTAACCGCTGCGCTAGTATAGGATGTGTGCCAAGGGTATTTACTTTTTCTTT  
TGACCTTAAGGCCACTTAAAAAATTATAAAATAGAACAAAAAATTTAGGTCGCC  
TAGGACTCGAACATGGGTCTCGGGAGCTAAGTTGATGGGTCTTAACCGTTGCGCC  
AGTAGGAGAAGTTCGAAAGAAAATGAAAACCTTATCCTACTTAAAACCTAAGTTA  
ACACCTAACTGCTACATAACATCACTGTCGGTTTGGAGAATGACCCGACAGTTAT  
GTACCCCATCACTGTCGGTTTGAAGCAAAACCGACAGTGATGAGCTATTTCTTA  
AAATAAATTAATAATTTATAAAATAGAACAAAAAATTTGGGCCACCTGGGACTC

GAACACGGGTCTCGAGGGCTAAGTTGAGGGGTCTTAACCGTTGCGCCAGTAGGA  
GTTGAAAGAAAACGGAACTTATCCTACTTAAACACATAACTTAAACACCTAACTG  
CTACAGCACATCACTGCCAGTTTGGGGAGTGACCCGACAGTGATGTGCCCTATC  
ACTGTCGGTTTATAAGTAGAACCGACAGTGATGAGCTACTGCTATAAAGTGGCGC  
GGGTTGAAATTATCCCAATTCATTTCAACCCCGCGCCAACCCCTCCCCCACGTGC  
CCAGAGCACCTCCCCGCGCTGGAGCACCGCCCCACGCCGTCCAACGCCCATGCT  
AGAGCACCGCCCCACGTCGTCCATCGCCCCAAGCCGGAGCACCGCGTCGTCTGTC  
CCACCGCTGAGGCCTTCAGGAGTGCGCGGACAGCTGCTTTCCACCCCCACGGTG  
AGTGCGTGGCTCACTGTCCGCTACGTGTCTGATGAAATGTCCCACAGCTGCTTACT  
TGAATTAGTTTGATCATGCAGTGTGTGCTGTCATACAGTTTGATCATGCAGCTGTG  
AATGCTAGAACCTTGGTTTGGCACACAGTGTGTGCTGTCTGTTTCTTTGTGTTTCT  
GGACACATTGCTAGTAAGTAGTACATTGTTACTTAGTAGCACGTTGCTCAGCTCT  
ATTCATCAGTATACTGATAAGAACAAGCGTAATTTTATTTTTGGTTACCAGGA  
TAAACATAGCAGGGACTTGTAATACTTGACAGAAATTACTCTCTAAACATTAATG  
GTAGTTTGTGGAGTACTTGCATTTTCTCTATGAAACAACTTATGTTTTCTCT  
GTGAAACCATCTTGTGCATCATTGGAGCTTGGCTGGATACATATGTGAAACCAAT  
GGCTGGCCTCACCTCATAGCTTGAATTGCAACTGCGCTCCCATCAGGCAGATAC  
TCAATGTTTTTGGAACTGTGCCACCAATGCTCAATGTTTTTGGATGTGAGGATGGA  
GTATATATACTTTAGACTTATAGAATAATATATGTCCAACTTGTCATGGAATTC  
TCCATGCTATCAGCATAAAGATGCATAATTATTAACCTTATGCAAGATTACAAGTA  
ATACATAGTTATGTAGCCTTGTGCATTAACTGGGAATATTTTCCAACAATCCTAT  
CTTATTATGAGTTTTTGCATGTCACCTTCTAGGATAGTATGAGATGCATATTTGC  
CCCTACCAACCATGTAAGTTTCAGATCAAAGTATCACTAATTTGAAAAGATGCTT  
ACCTGCACTAATACTTGTGAAAGAACCACGCGTCGATGGTTGGGGGATTTATAGT  
GAGACTGAAGACTCTGAGACATCTCTAGAAGTATATCCAAAAGTCTTAAATTGGC  
TCTCTAAATCATCATTTGGTTTTTAACGAATAAAAATATTCTTTATATCTTTCCTAC  
ATCTAACAATACTAAGGCACGTGAATAGCTATTTTTTCAAACACTCCTATCG  
ATGCAATGGTTGTGTGTTATTCGTAAATATATTAATTGTATAAGAATACATCATG  
TTTGAATGACAACCTAATTTACTTAAATGTATAGGCAACCATAGCATGGTATATA  
GTGCATGTTGGTCACATGCCTGGGATTTATCGAACTTGGGAAGATTGCTATGCTC  
AAGTCAATCGCTACCCTAGTAATTTGCATAAAAAATACAACAAAGAAGCTGAAG  
CTTTAAGGGTCTACTATAGTCATTCGGCCTACCTTGCAAACCATAGGCAATCGGC  
CTACTATGGTCCAATGAATGCAAACCATGGCCATCCGCTAGCACTAGAGATCGAA  
GAGAACCCACCCGCCGAAGATGTGAAGATTATGAGAAATGCTCCTTAGTCTTGG  
AAGATGTCATGCGTTTATTTATGGCTATGGTCATCGCTTTTCTTATTTGAAAGTTG  
ATGTAGGCTTAGTTTTGTAGAACAGTAGGTGAACTTTGTTGTATATGGACTTATTA  
TTTACTTTGCATTAAACATATTATATGACTTTGTTGTATGTGGACTTATTATTAG  
ACTTTGCATTAAACATATTATTTATATATATATATATATATATATATATATATA  
TA  
TA  
TA  
AGTTGTCCACGGCCAAAATACTACGATCGTGTACAGCCATGACTCATCGTCG  
CCTGTTACCCCGTCGCCGAAGAACAGATCGGCAACAAGAAGCGGCTGATATCGCT  
GGGACGATAGAGCCGCATGATTCGACAAACGGTGACGGTGTCAATCAGGTTGGT  
GAGAACGAGCAGCCATGGACCCGTTCCGGCCTGCTCGATCTGGGTCAAATGACC  
AAGATGGCCAAGTAACTTTGGTATCATGTGACTATGTAGCCACACACGCTAATCA  
ATTGAGAGGGTCATAGCTTACTTGGTGTCTTAGCAGGAGCCTCCGCCGTCCGAG  
GAGCCAGAGGGTTCGGGTAGCAGGAAGGTCAGATCCAAGCGAGGCGTGTCAAAG

ATTTCTGTTGGTAGGAATGCACACTGGCACATTACTGAGTTCGATGAATTCAGGG  
ATCCACTCTCACCGCCTATAGTTTTGACCAAATACAAGACTATCCTCGGGTACTT  
GTTAGGGACTTCATCCCGATTAGGTATAAGAAGTGGATTGGGAAAGATGATGACC  
CGTGGAGGGTTCCTAAAAGTAAGAAGGATTACATATGGGATGTCAAGATTCCAG  
AGTATTTCACTTTCCAGCTAAGTACGACAGGGAGTTAGTCAAGAAAAAACA  
AAGAAATCATGGGGACATGCTTCAAGAATTTTAAGGGGACATTGTACAAGAATTT  
TGTCTCCAAAATAAAGAGCCAGATTTTCGATGGTGGACAGTTTAGCAAGCAGAA  
GGACTTCTGGCAGGATTTCAAGAAATAAAGGTTATCTGAGAAGTACTTAGCACTG  
AGCATGAAGAATAAGGAGAAGTTCGAGAAAGTGTGAATCCTCATCATCTCGGC  
TCTCGTGGCTACACCAAAAAGATGCCAAAATTTGAAGCAGAAGTTCAAAAGATG  
GATCGCCTTTCTGAGGAATGCGTTCAAGTTGAGACAGCCGATTGGGAGCCTAGAT  
TGGTATTATACTGTATGGGGAGGAATGTACGGCACGTCGAGGACGGAAGCTTTAG  
CTCCATAAATGAACCCATGAGCGAACTCATCCAGAGGATCTCTCAGGTAAGTGA  
GAGGTGAGGCAAGGGACTCGCACTTCTAATAGGGAGAAAGACGTGCTCACCCAA  
GCACTCGGAACCAAGGAGCACCCCTGGTCGCACTAGAGGAACCGGTGTTGTTCCAT  
AGAACTAGCATTCCCCTAAGAATCTAACACTTATCAGAGCCGCTCGAGGGGTAG  
AGCGAAACAGGAGGCAGAGTATTTGAGGCGACTAAAAGAGATGGAAGACAGAA  
TGGAAGCACGGATTAAGGCGACTGTTGAAGCGCGAGTCGAGCAAATCTTGCTGTC  
CAAGGGTTCAGGAGTACCACAAAACCCTACTTCTACTGCCTTTAGCCACAGTTT  
AGGGGTCGCAGCAGCTGCGGATCGACCCCGCTCGATGAAGAAGAGGCGAATGTG  
CCTCATCCGGTGGACGGCATCACTGAACCCGTCAATGTCAAGATGTACATCTGTC  
AGGAATGGACAAAGGTCAAGGTGGCGCTTGGCCAGGCCAGGCCTGCGGGAGACG  
GGACAATTAATGGCCGCCCAATTCACAGGGGTACGCTCGCGTCATCATTGACTG  
AATACTTGATAAGAAGTTTAACAAGATACCCATTGAGTACCCCGTAGCGGAAGAC  
AGGCCGAACTTGGTCAAATCAAGGGCTCTCAAGTGGCCTGGCGCAAGCGCTTCA  
TTAAGCTTGACCATCAATTGTCCTCTGATGATGAGGACGACTATGAGTCTTTGCC  
ACCTACGATGATCACTCACCATCTCCCAACAGAGATCACTCCCCTCCTCATCCGG  
AGCCATCACCTCCGAGAAGACAGAGGTCTCCTTTTATTCTTCTCGTCCGACTCCA  
TCTTCATCAGCCCCGAGAAAAGAGACTCCTCCTCCTCCCCCTCCTCCTCCATCTTC  
ATTAGCGCCGGGAAAACAGAATTCTCGTCTGTCATCCTCCTCCTCCACCACCTCAG  
CCCAAAACAAGATCAAGGACATCCTCGCAGTCGCAGAAAAGGTCCTTGGATTCTG  
ATCACTAATGCTCCCAAAGTGGACCAACAGAAAAGAAAAGGTCGTTCAAGTGGC  
AGCGTTACCACCTTGATGAAAGAAAATTTACAGCACACATCTCAAAGGAAGATT  
GTGTTAGTTTCTTCAATCTCTGTGCCTCACTGCCTTCGTATTTGTTGTAGTCCGAAT  
TCAAAGGCAAACACAGAATAAATACGCTACAACAAGAGACGAAGAAATACAC  
AATAAAGATTTAAGAACATACAGAAATACGTCTGAAGACAACCTAGGATTAACC  
ATGGAAGAGGCCACGAACATCTATTATGATGTACAAGGGACGGGAACACCGACA  
ATGAAGTATGAAATGAGCAATCTTTTGGTTCCCGATGACGAGTATAAAAATTTGA  
CAAGCCGGCAATAAAGATTATTGTAATAAGTCCGAGAAACATTGCACAATCTGTG  
AAGTCCGAGAACAATTTTTTTTTTATTTTGGAGAAATCGAGATCTGGATAAGAACAT  
TTGAACTGTAACCGTAAAATTTATAATGTTTAATATTGATGGATCTGTCTGCTATG  
TAGCATATATAAAACGAAACCTGATTCACGAAAAAATATAAATTAATAATCAA  
TTATAAAACAATAAAATGCGAATGGGCCATCATTGCCGGTTAGCAACGAATCGAC  
AATGTTGTCTTGCAAAACACTGCCGGTTATCAACAAACCGACAGTGTGTCTTGC  
AAAACACTGCCGGCTTGAGCCATGAACCGATAGTGTTGTTTTGCTCAACACTGTC  
GGCTTGAGCCATCAACCGATAGTGTTGGCTAGCTCAACACTGCCGGCTTGAGCCA  
TGAACCGACAGTGTAGGCTTGCTCAACAATGTTGGCTTGAGCCAAGAACCGACAC

TCTTGAAGCAACCATCACGATCGGTTGGGACTACAAACCGGCAGTGTTGTTCAAC  
TAGACACCGTCGGATGGAGCCAGAAACCGACACTGTTTCTTAGAGATCAGTGTCG  
GTTTTTAAAAATCAACAGTGATATCAACCCTATCACTGTTGGTATGTCAGTGTCAA  
TTTAAAATCCAACACCGAAGGGAGTTTTTAAACCGACGGTAATATTGAGATCTAA  
GGATAGTGCTCGTGCGGGCAGTGCTACAAGATCGCGTGCGACCGCAAGAGAGCC  
GACCCGAGGTTCTGCAAGCCC GGCGTCAGGGTGACCGTCACGGCCACCAACTTCT  
GCCCGCCCAACTTGGCGCTGCCCGATGGCGGGTGGTGCAATCAGCAGCGCCACA  
CTTCGACATGGCGCAGCTGGCCTGGGAGAAGATCGGCGTCTACAGCGGTGGCATC  
ATCCCCGTCATGTACCACAATACCAGAGGTCTTTCATTAATTTATTTCTACTTCT  
TCCTAGCTATATAGCCTCGATCTATCATCAAAAACCAAAAGTTTATTGGGATGCA  
ATGTCTTAATACTAGAGTTCCTTGCTTGCTTGTAGAGTTCCTTGCGTGAAGCGAGG  
TGGAGTCAGGTTACGATCAATGGACATGACTACTTCAATCTTGTGCTTGTGACC  
AACGTTGCGGCTGCCGGCTCCATCAAGTCCATGGATGTCAAGAGTTCGAATTGCA  
CCGACTGGACCCCAATGGCGCGCAACTGGGGTGCAAACCTGGCACTCTGTGGCCTA  
TTTTACAGGGCAAATGCTCTCGTTTAGAGTCACCGACACCGACGGACAGACTATT  
GAATTCACAAACGTGGTGCCACAAGGATGGAAGTTTGGCCAAACATTTACATGCA  
AGTTGCAGTTCAAGTGA