

Homework #6 (extra credit)

Due April 29th, 11:59pm

Each homework submission must include:

- An archive (.zip or .gz) file of the source code containing:
 - The makefile used to compile the code on Monsoon **(5pts)**
 - All .cpp and .h files **(5pts)**
- A full write-up (.pdf or .doc) file containing answers to homework's questions **(5pts)**, including the exact command line needed to execute every subproblem of the homework

The source code must follow the following guidelines:

- No external libraries that implement data structures discussed in class are allowed, unless specifically stated as part of the problem definition. Standard input/output and utilities libraries (e.g. math.h) are ok.
 - All external data sources (e.g. input data) must be passed in as a command line argument (no hardcoded paths within the source code **(5pts)**).
 - Solutions to sub-problems must be executable separately from each other. For example, via a special flag passed as command line argument **(5pts)**
-

Problem #1 (of 1): Suffix tree

Create a class called ***suffix_tree***. Use the **suffix tree (not trie)** data-structure to store a given genome (see Appendix A). You must construct the suffix tree without building the trie first (i.e. use the trick shown in class). At minimum, the class must contain:

- A default constructor
 - At least one custom constructor to build a tree from a given genome sequence
 - A function to traverse (*search*) the tree using a sequence fragment of a given length.
 - A destructor
- A. **(25pts) Perfect match only search:** Read in the SARS-COV2 genome sequence (Appendix A) and store it in the suffix tree. Implement a perfect match search function, which would take in a sequence fragment, search for presence of that fragment in the suffix tree and report presence/absence AND location of the fragment in the genome (finding a single hit is ok – no need to find all possible hits).
- What is the size of the tree (# of nodes)?
 - Generate 5K, 50K, and 100K random 36-mers from the SARS-COV2 genome sequence and use them to search the suffix tree. How many of your 36-mers had a match? Does it make sense? Explain why.
 - How long did it take (big O notation estimate)?

Appendix A: SARS-COV2 Genome

>NC_045512.2_SARS_COV2_complete_genome
ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTTCGATCTCTTGTAGATCTGTTCTCTAAA
CGAACTTTAAATCTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAAC
TAATTACTGTCGTTGACAGGACACGAGTAACCTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGTG
TTGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTC
CCTGGTTTTCAACGAGAAAACACACGTCCAACCTCAGTTTGCCTGTTTTACAGGTTTCGCGACGTGCTCGTAC
GTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGG
CTTAGTAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCAAACGTTTCGGAT
GCTCGAACTGCACCTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAACTCGAAGGCATTTCAGTACGGTC
GTAGTGGTGAGACACTTGGTGTCTTGTCCCTCATGTGGGCGAAATACCAGTGGCTTACCAGCAAGGTTCT
TCTTCGTAAGAACGGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTA
GGCGACGAGCTTGGCACTGATCCTTATGAAGATTTTCAAGAAAACCTGGAACACTAAACATAGCAGTGGTG
TTACCCGTGAACTCATGCGTGAGCTTAACGGAGGGGCATACACTCGCTATGTCGATAACAACCTTCTGTGG
CCCTGATGGCTACCCTCTTGAGTGCATTAAAGACCTTCTAGCACGTGCTGGTAAAGCTTCATGCACTTTG
TCCGAACAACCTGGACTTTATTGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTG
CTTGGTACACGGAACGTTCTGAAAAGAGCTATGAATTGCAGACACCTTTTGAAATTAAATTGGCAAAGAA
ATTTGACACCTTCAATGGGGAATGTCCAAATTTTGTATTTCCCTTAAATTCCATAATCAAGACTATTCAA
CCAAGGGTTGAAAAGAAAAGCTTGATGGCTTTATGGGTAGAATTTCGATCTGTCTATCCAGTTGCGTCAC
CAAATGAATGCAACCAAATGTGCCTTTCAACTCTCATGAAGTGTGATCATTGTGGTGAAACTTCATGGCA
GACGGGCGATTTTGTAAAGCCACTTGCGAATTTTGTGGCACTGAGAATTTGACTAAAGAAGGTGCCACT
ACTTGTGGTTACTTACCCCAAATGCTGTTGTTAAATTTTATTGTCCAGCATGTCACAATTCAGAAGTAG
GACCTGAGCATAGTCTTGCCGAATACCATAATGAATCTGGCTTGAAGAACATTCTTCGTAAGGGTGGTCG
CACTATTGCCTTTGGAGGCTGTGTGTTCTCTTATGTTGGTTGCCATAACAAGTGTGCCTATTGGGTTCCA
CGTGCTGCTACTAACATAGGTTGTAACCATAACAGTGTTGTTGGAGAAGGTTCCGAAGGCTTAATGACA
ACCTTCTTGAAATACTCCAAAAGAGAAAGTCAACATCAATATTGTTGGTGACTTTAAACTTAATGAAGA
GATCGCCATTATTTTGGCATCTTTTTCTGCTTCCACAAGTGCTTTTGTGGAACTGTGAAAGGTTTGGAT
TATAAAGCATTCAAACAAATTGTTGAATCCTGTGGTAATTTTAAAGTTACAAAAGGAAAAGCTAAAAAAG
GTGCCTGGAATATTGGTGAACAGAAATCAATACTGAGTCCTCTTTATGCATTTGCATCAGAGGCTGCTCG
TGTTGTACGATCAATTTTCTCCCGCACTCTTGAAACTGCTCAAAATTCTGTGCGTGTTTTACAGAAGGCC
GCTATAACAATACTAGATGGAATTTACAGTATTCACTGAGACTCATTGATGCTATGATGTTTACATCTG
ATTTGGCTACTAACAATCTAGTTGTAATGGCCTACATTACAGGTGGTGTGTTTTCAGTTGACTTCGCAGTG
GCTAACTAACATCTTTGGCACTGTTTATGAAAACTCAAACCCGTCCTTGATTGGCTTGAAGAGAAGTTT
AAGGAAGGTGTAGAGTTTCTTAGAGACGGTTGGGAAATTGTTAAATTTATCTCAACCTGTGCTTGTGAAA
TTGTGCGGTGGACAAATTGTCACCTGTGCAAAGGAAATTAAGGAGAGTGTTTCAACATTCTTTAAGCTTGT
AAATAAATTTTGGCTTTGTGTGCTGACTCTATCATTATTGGTGGAGCTAACTTAAAGCCTTGAATTTA
GGTGAAACATTTGTACGCACTCAAAGGGATTGTACAGAAAGTGTGTTAAATCCAGAGAAGAACTGGCC
TACTCATGCCTCTAAAAGCCCCAAAAGAAATTATCTTCTTAGAGGGAGAAACACTTCCCACAGAAGTGTT
AACAGAGGAAGTTGTCTTGAAAACCTGGTGATTTACAACCATTAGAACAACCTACTAGTGAAGCTGTTGAA
GCTCCATTGGTTGGTACACAGTTTGTATTAACGGGCTTATGTTGCTCGAAATCAAAGACACAGAAAAGT
ACTGTGCCCTTGCACCTAATATGATGGTAACAAACAATACCTTCACACTCAAAGGCGGTGCACCAACAAA
GGTTACTTTTGGTGATGACACTGTGATAGAAGTGCAAGGTTACAAGAGTGGAATATCACTTTTGAACCT
GATGAAAGGATTGATAAAGTACTTAATGAGAAGTGCTCTGCCTATACAGTTGAACTCGGTACAGAAGTAA
ATGAGTTTCGCTGTGTTGTGGCAGATGCTGTCTATAAAAACCTTGAACACAGTATCTGAATTACTTACACC
ACTGGGCATTGATTTAGATGAGTGGAGTATGGCTACATACTACTTATTTGATGAGTCTGGTGAGTTTAAA
TTGGCTTACATATGTATTGTTCTTTCTACCCTCCAGATGAGGATGAAGAAGAAGGTGATTGTGAAGAAG
AAGAGTTTGGCCATCAACTCAATATGAGTATGGTACTGAAGATGATTACCAAGGTAAACCTTTGGAATT
TGGTGCCACTTCTGCTGCTCTTCAACCTGAAGAAGAGCAAGAAGAAGATTGGTTAGATGATGATAGTCAA
CAAACCTGTTGGTCAACAAGACGGCAGTGAGGACAATCAGACAACCTACTATTCAAACAATTGTTGAGGTTT
AACCTCAATTAGAGATGGAACCTTACACAGTTGTTTACAGACTATTGAAGTGAATAGTTTTAGTGGTTATTT
AAAACCTTACTGACAATGTATACATTAAAAATGCAGACATTGTGGAAGAAGCTAAAAAGGTAAACCAACA
GTGGTTGTTAATGCAGCCAATGTTTACCTTAAACATGGAGGAGGTGTTGCAGGAGCCTTAAATAAGGCTA
CTAACAATGCCATGCAAGTTGAATCTGATGATTACATAGCTACTAATGGACCACTTAAAGTGGGTGGTAG
TTGTGTTTTAAGCGGACACAATCTTGCTAAACACTGTCTTCATGTTGTGCGCCCAAATGTTAACAAGGT
GAAGACATTCAACTTCTTAAGAGTGCTTATGAAAATTTTAAATCAGCACGAAGTTCTACTTGCACCATAT
TATCAGCTGGTATTTTTGGTGCTGACCCTATACATTCTTTAAGAGTTTGTGTAGATACTGTTTCGCACAAA

[illegible]

TCTTTACTTTGATAAAGCTGGTCAAAAGACTTATGAAAGACATTCTCTCTCTCATTGTTAACTTAGAC
AACCTGAGAGCTAATAACACTAAAGGTTCAATTGCCTATTAATGTTATAGTTTTTGTATGGTAAATCAAAAT
GTGAAGAATCATCTGCAAAATCAGCGTCTGTTTACTACAGTCAGCTTATGTGTCAACCTATACTGTTACT
AGATCAGGCATTAGTGTCTGATGTTGGTGATAGTGCAGGAAAGTTGCAGTTAAAATGTTTGATGCTTACGTT
AATACGTTTTTCATCAACTTTTAAACGTACCAATGGAAAACTCAAAACACTAGTTGCAACTGCAGAAGCTG
AACTTGCAAAGAATGTGTCCTTAGACAATGTCTTATCTACTTTTTATTTTCAGCAGCTCGGCAAGGGTTTTGT
TGATTGAGATGTAGAACTAAAGATGTTGTTGAATGTCTTAAATTGTCACATCAATCTGACATAGAAGTT
ACTGGCGATAGTTGTAATAACTATATGCTCACCTATAACAAAGTTGAAAACATGACACCCCGTGACCTTG
GTGCTTGTATTGACTGTAGTGCGCGTCATATTAATGCGCAGGTAGCAAAAAGTCACAACATTGCTTTGAT
ATGGAACGTTAAAGATTTTCATGTCAATTGTCTGAACAACTACGAAAACAAATACGTAGTGCTGTCAAAAAG
AATAACTTACCTTTTTAAGTTGACATGTGCAACTACTAGACAAGTTGTTAATGTTGTAACAACAAGATAG
CACTTAAGGGTGGTAAAATTGTTAATAATTGGTTGAAGCAGTTAATTAAGTTACACTTGTGTTCCTTTTT
TGTTGCTGCTATTTTTCTATTTAATAACACCTGTTTCATGTCTGCTAAACATACTGACTTTTTCAAGTGAA
ATCATAGGATACAAGGCTATTGATGGTGGTGTCACTCGTGACATAGCATCTACAGATACTTGTGTTTGCTA
ACAAACATGCTGATTTTTGACACATGGTTTAGCCAGCGTGGTGGTAGTTATACTAATGACAAAGCTTGCCC
ATTGATTGCTGCAGTCATAACAAGAGAAGTGGGTTTTGTCGTGCCTGGTTTTGCCTGGCAGCATATTACGC
ACAACATGTTGACTTTTTTGCACTTTCTTACCTAGAGTTTTTAGTGAGTTGGTAACATCTGTTACACAC
CATCAAACTTATAGAGTACACTGACTTTGCAACATCAGCTTGTGTTTTGGCTGCTGAATGTACAATTTT
TAAAGATGCTTCTGGTAAGCCAGTACCATATTGTTATGATAACCAATGTACTAGAAGGTTCTGTTGCTTAT
GAAAGTTTACGCCCTGACACACGTTATGTGCTCATGGATGGCTCTATTATTCAATTTCCCTAACACCTACC
TTGAAGGTTCTGTTAGAGTGGTAACAACCTTTTGATTCTGAGTACTGTAGGCACGGCACTTGTGAAAGATC
AGAAGCTGGTGGTTGTGTATCTACTAGTGGTAGATGGGTACTTAACAATGATTATTACAGATCTTTACCA
GGAGTTTTCTGTGGTGTAGATGCTGTAAATTTACTTACTAATATGTTTACACCACTAATTCAACCTATTG
GTGCTTTGGACATATCAGCATCTATAGTAGCTGGTGGTATTGTAGCTATCGTAGTAACATGCCTTGCCCTA
CTATTTTATGAGGTTTAGAAGAGCTTTTGGTGAATACAGTCATGTAGTTGCCTTTAATACTTTACTATTC
CTTATGTCATTCACTGTACTCTGTTTAAACACAGTTTACTCATTCTTACCTGGTGGTTTATTCTGTTATTT
ACTTGTACTTGACATTTTATCTTACTAATGATGTTTCTTTTTTAGCACATATTCAGTGGATGGTTATGTT
CACACCTTTAGTACCTTTCTGGATAACAATTGCTTATATCATTGTTGATTTCCACAAAGCATTTCATTTGG
TTCTTTAGTAATTAACCTAAGAGACGTGTAGTCTTTAATGGTGGTTTCTTTTAGTACTTTTGAAGAAGCTG
CGCTGTGACCTTTTTGTTAAATAAAGAAATGTATCTAAAGTTGCGTAGTGATGTGCTATTACCTCTTAC
GCAATATAATAGATACTTAGCTCTTTATAATAAGTACAAGTATTTTAGTGGAGCAATGGATACAACCTAGC
TACAGAGAAGCTGCTTGTGTGTCATCTCGCAAAGGCTCTCAATGACTTCAGTAACCTCAGGTTCTGATGTTT
TTTACCAACCACCACAAACCTCTATCACCTCAGCTGTTTTGCAGAGTGGTTTTAGAAAAATGGCATTCCC
ATCTGGTAAAGTTGAGGGTTGTATGGTACAAGTAACCTGTGGTACAACCTACACTTAACGGTCTTTGGCTT
GATGACGTAGTTTACTGTCCAAGACATGTGATCTGCACCTCTGAAGACATGCTTAACCTAATTATGAAG
ATTTACTCATTGTAAGTCTAATCATAATTTCTTGGTACAGGCTGGTAATGTTCAACTCAGGGTTATTGG
ACATTCTATGCAAAATTGTGTACTTAAGCTTAAGGTTGATACAGCCAATCCTAAGACACCTAAGTATAAG
TTTGTTCGCATTCAACCAGGACAGACTTTTTTCACTGTTAGCTTGTGTTACAATGGTTTACCCTCTGGTGGT
ACCAATGTGCTATGAGGCCCAATTTCACTATTAAGGGTTTCACTCTTAATGGTTTATGTGGTAGTGGTGG
TTTTAACATAGATTATGACTGTGTCTCTTTTTGTTACATGCACCATATGGAATTACCAACTGGAGTTTCA
GCTGGCACAGACTTAGAAGGTAACCTTTTATGGACCTTTTGTGACAGGCAAACAGCACAAAGCAGCTGGTA
CGGACACAACCTATTACAGTTAATGTTTTAGCTTGGTTGTACGCTGCTGTTATAAATGGAGACAGGTGGT
TCTCAATCGATTTACCACAACCTCTAATGACTTTAACCTTGTGGCTATGAAGTACAATTATGAACCTCTA
ACACAAGACCATGTTGACATACTAGGACCTCTTCTGCTCAAACCTGGAATTGCCGTTTTAGATATGTGTG
CTTCATTAAAAGAATTACTGCAAAATGGTATGAATGGACGTACCATATTGGGTAGTGCTTTATTAGAAGA
TGAATTTACACCTTTTGATGTTGTTAGACAATGCTCAGGTGTTACTTTCCAAAGTGCAGTGAAAAGAACA
ATCAAGGGTACACACCCTGGTTGTTACTCACAATTTTGAACCTTCACTTTTAGTTTTAGTCCAGAGTACTC
AATGGTCTTTGTTCTTTTTTTTTGTATGAAAATGCCTTTTTACCTTTTGCTATGGGTATTATTGCTATGTC
TGCTTTTGCAATGATGTTTGTCAAACATAAGCATGCATTTCTCTGTTTGTGTTTTGTTACCTTCTCTTGCC
ACTGTAGCTTATTTAATATGGTCTATATGCCTGCTAGTTGGGTGATGCGTATTATGACATGGTTGGATA
TGTTTGATACTAGTTTGTCTGGTTTTAAGCTAAAAGACTGTGTTATGTATGCATCAGCTGTAGTGTTACT
AATCCTTATGACAGCAAGAAGTGTGTATGATGATGGTGCTAGGAGAGTGTGGACACTTATGAATGTCTTG
ACACTCGTTTTATAAAGTTTATTATGGTAATGCTTTAGATCAAGCCATTTCCATGTGGGCTCTTATAATCT
CTGTTACTTCTAACTACTCAGGTGTAGTTACAACGTGCATGTTTTTGGCCAGAGGTATTGTTTTTATGTG
TGTTGAGTATTGCCCTATTTTCTTCATAACTGGTAATACACTTCAGTGTATAATGCTAGTTTTATTGTTTC
TTAGGCTATTTTTGTACTTGTACTTTGGCCTCTTTTGTGTTACTCAACCGCTACTTTAGACTGACTCTTG
GTGTTTATGATTACTTAGTTTCTACACAGGAGTTTAGATATATGAATTCACAGGGACTACTCCCACCCAA

GAATAGCATAGATGCCTTCAAACCTCAACATTAAATTGTTGGGTGTTGGTGGCAAACCTTGTATCAAAGTA
GCCACTGTACAGTCTAAAATGTGAGATGTAAAGTGCACATCAGTAGTCTTACTCTCAGTTTTTGCAACAAC
TCAGAGTAGAATCATCATCTAAATTGTGGGCTCAATGTGTCCAGTTACACAATGACATTCTCTTAGCTAA
AGATACTACTGAAGCCTTTGAAAAAATGGTTTTCACTACTTTCTGTTTTGCTTTCCATGCAGGGTGCTGTA
GACATAAACAAGCTTTGTGAAGAAATGCTGGACAACAGGGCAACCTTACAAGCTATAGCCTCAGAGTTTA
GTTCCCTTCCATCATATGCAGCTTTTGTACTGCTCAAGAAGCTTATGAGCAGGCTGTTGCTAATGGTGA
TTCTGAAGTTGTTCTTAAAAAGTTGAAGAAGTCTTTGAATGTGGCTAAATCTGAATTTGACCGTGATGCA
GCCATGCAACGTAAGTTGGAAAAGATGGCTGATCAAGCTATGACCCAAATGTATAAACAGGCTAGATCTG
AGGACAAGAGGGCAAAAGTTACTAGTGCTATGCAGACAATGCTTTTCACTATGCTTAGAAAAGTTGGATAA
TGATGCACTCAACAACATTATCAACAATGCAAGAGATGGTTGTGTTCCCTTGAACATAATACCTCTTACA
ACAGCAGCCAAACTAATGGTTGTGCATACCAGACTATAACACATATAAAAAATACGTGTGATGGTACAACAT
TTACTTATGCATCAGCATTGTGGGAAATCCAACAGGTTGTAGATGCAGATAGTAAAATTGTTCAACTTAG
TGAAATTAGTATGGACAATTCACCTAATTTAGCATGGCCTCTTATTGTAACAGCTTTAAGGGCCAATTCT
GCTGTCAAATTACAGAATAATGAGCTTAGTCCTGTTGCACTACGACAGATGTCTTGTGCTGCCGGTACTA
CACAACTGCTTGCCTGATGACAATGCGTTAGCTTACTACAACACAACAAAGGGAGGTAGGTTTGTACT
TGCCTGTTATCCGATTTACAGGATTTGAAATGGGCTAGATTCCCTAAGAGTGATGGAAGTGGTACTATC
TATACAGAAGTGAACACCTTGTAGGTTTGTACAGACACACCTAAAGGTCCTAAAGTGAAGTATTTAT
ACTTTATTAAAGGATTAAACAACCTAAATAGAGGTATGGTACTTGGTAGTTTAGCTGCCACAGTACGTCT
ACAAGCTGGTAATGCAACAGAAGTGCCTGCCAATTCACTGTATTATCTTTCTGTGCTTTTGTGCTAGAT
GCTGCTAAAGCTTACAAAGATTATCTAGCTAGTGGGGGACAACCAATCACTAATTGTGTTAAGATGTTGT
GTACACACACTGGTACTGGTCAGGCAATAACAGTTACACCGGAAGCCAATATGGATCAAGAATCCTTTGG
TGGTGCATCGTGTTGTCTGTACTGCCGTTGCCACATAGATCATCAAATCCTAAAGGATTTTGTGACTTA
AAAGGTAAGTATGTACAAATACCTACAACCTTGTGCTAATGACCCTGTGGGTTTTACACTTAAAAACACAG
TCTGTACCGTCTGCGGTATGTGGAAAGGTTATGGCTGTAGTTGTGATCAACTCCGCGAACCCATGCTTCA
GTCAGCTGATGCACAATCGTTTTTAAACGGGTTTGCCTGTAAAGTGCAGCCCGTCTTACACCGTGCAGCA
CAGGCACTAGTACTGATGTCGTATACAGGGCTTTTGACATCTACAATGATAAAGTAGCTGGTTTTGCTAA
ATTCCTAAAACTAATTGTTGTGCTTCCAAGAAAAGGACGAAGATGACAATTTAATTGATTCTTACTTT
GTAGTTAAGAGACACACTTTCTCTAACTACCAACTGAAGAAACAATTTATAATTTACTTAAAGGATTGTC
CAGCTTTGCTTAAACATGACTTCTTTAAGTTTGAAGATAGACGGTGACATGGTACCACATATATCAGTCA
ACGTCTTTACTAAATACACAATGGCAGACCTCGTCTATGCTTTAAGGCATTTTGTGATGAAGTAATTGTGAC
ACATTTAAAGAAATACTTGTACATACAATTGTTGTGATGATGATTATTTCAATAAAAAGGACTGGTATG
ATTTTGTAGAAAACCCAGATATATTACGCGTATACGCCAACTTAGGTGAACGTGTACGCCAAGCTTTGTT
AAAAACAGTACAATTCTGTGATGCCATGCGAAATGCTGGTATTGTTGGTGTACTGACATTAGATAATCAA
GATCTCAATGGTAAGTGGTATGATTTGCGGTGATTTACATACAAACCACGCCAGGTAGTGGAGTTCCTGTTG
TAGATTCTTATTATTATTGTTAATGCCTATATTAACCTTGACCAGGGCTTTAACTGCAGAGTCACATGT
TGACACTGACTTAAACAAAGCCTTACATTAAGTGGGATTTGTTAAATATGACTTCACGGAAGAGAGGTTA
AAACTCTTTGACCGTTATTTTAAATATTGGGATCAGACATACCACCCAAATTGTGTTAACTGTTTGGATG
ACAGATGCATTCTGCATTGTGCAAACCTTTAATGTTTTATTCTCTACAGTGTTCCACCTACAAGTTTTGG
ACCACTAGTGAGAAAAATATTTGTTGATGGTGTTCATTTGTAGTTTCACTGGATACCACTTCAGAGAG
CTAGGTGTTGTACATAATCAGGATGTAACTTACATAGCTCTAGACTTAGTTTTAAGGAATTACTTGTGT
ATGCTGCTGACCCTGCTATGCACGCTGCTTCTGGTAATCTATTACTAGATAAACGCACTACGTGCTTTTC
AGTAGCTGCACTTACTAACAATGTTGCTTTTCAAACCTGTCAAACCCGGTAATTTTAAACAAAGACTTCTAT
GACTTTGCTGTGTCTAAGGGTTTCTTTAAGGAAGGAAGTCTGTTGAATTTAAACACTTCTTCTTTGCTC
AGGATGGTAATGCTGCTATCAGCGATTATGACTACTATCGTTATAATCTACCAACAATGTGTGATATCAG
ACAACACTACTATTTGTAGTTGAAGTTGTTGATAAGTACTTTGATTGTTACGATGGTGGCTGTATTAATGCT
AACCAAGTCATCGTCAACAACCTAGACAAATCAGCTGGTTTTCCATTTAATAAATGGGGTAAGGCTAGAC
TTTATTATGATTCAATGAGTTATGAGGATCAAGATGCATTTTTCGATATACAAAACGTAATGTCTATCCC
TACTATAACTCAAATGAATCTTAAAGTATGCCATTAGTGCAAGAATAGAGCTCGCACCGTAGCTGGTGTC
TCTATCTGTAGTACTATGACCAATAGACAGTTTCATCAAAAATTATTGAAATCAATAGCCGCCACTAGAG
GAGCTACTGTAGTAATTGGAACAAGCAAATCTATGGTGGTTGGCACAACATGTTAAAACTGTTTATAG
TGATGTAGAAAACCTCACCTTATGGGTTGGGATTATCCTAAATGTGATAGAGCCATGCCTAACATGCTT
AGAATTATGGCCTCACTTGTCTTGTCTCGCAAACATACAACGTGTTGTAGCTTGTACACCGTTTCTATA
GATTAGCTAATGAGTGTGCTCAAGTATTGAGTGAAATGGTCATGTGTGGCGGTTCACTATATGTTAAACC
AGGTGGAACCTCATCAGGAGATGCCACAACCTGCTTATGCTAATAGTGTTTTTAACATTTGTCAAGCTGTC
ACGGCCAATGTTAATGCATTTTATCTACTGATGGTAACAAAATTGCCGATAAGTATGTCCGCAATTTAC
AACACAGACTTTTATGAGTGTCTCTATAGAAATAGAGATGTTGACACAGACTTTGTGAATGAGTTTTACGC
ATATTTGCGTAACATTTCTCAATGATGATACTCTCTGACGATGCTGTTGTGTGTTTCAATAGCACTTAT

GCATCTCAAGGTCTAGTGGCTAGCATAAAGAACTTTAAGTCAGTTCCTTTATTATCAAAACAATGTTTTTA
TGTCTGAAGCAAAATGTTGGACTGAGACTGACCTTACTAAAGGACCTCATGAATTTTGCTCTCAACATAC
AATGCTAGTTAAACAGGGTGATGATTATGTGTACCTTCCTTACCCAGATCCATCAAGAATCCTAGGGGCC
GGCTGTTTTGTAGATGATATCGTAAAAACAGATGGTACACTTATGATTGAACGGTTCGTGTCTTTAGCTA
TAGATGCTTACCCACTTACTAAACATCCTAATCAGGAGTATGCTGATGTCTTTCATTTGTACTTACAATA
CATAAGAAAGCTACATGATGAGTTAACAGGACACATGTTAGACATGTATTCTGTTATGCTTACTAATGAT
AACACTTCAAGGTATTGGGAACCTGAGTTTTATGAGGCTATGTACACACCGCATACAGTCTTACAGGCTG
TTGGGGCTTGTGTTCTTTGCAATTCACAGACTTCATTAAGATGTGGTGCTTGCATACGTAGACCATTCTT
ATGTTGTAAATGCTGTTACGACCATGTCATATCAACATCACATAAATTAGTCTTGTCTGTTAATCCGTAT
GTTTGCAATGCTCCAGGTTGTGATGTCACAGATGTGACTCAACTTTACTTAGGAGGTATGAGCTATTATT
GTAAATCACATAAACCCATTAGTTTTCCATTGTGTGCTAATGGACAAGTTTTTGGTTTTATATAAAAA
TACATGTGTTGGTAGCGATAATGTTACTGACTTTAATGCAATTGCAACATGTGACTGGACAAATGCTGGT
GATTACATTTTAGCTAACACCTGTACTGAAAGACTCAAGCTTTTTGCAGCAGAAACGCTCAAAGCTACTG
AGGAGACATTTAACTGTCTTATGGTATTGCTACTGTACGTGAAGTGCTGTCTGACAGAGAATTACATCT
TTCATGGGAAGTTGGTAAACCTAGACCACCTTAAACGAAATTATGTCTTTACTGGTTATCGTGTAACCT
AAAAACAGTAAAGTACAAATAGGAGAGTACACCTTTGAAAAAGGTGACTATGGTGATGCTGTTGTTTACC
GAGGTACAACAACCTTACAAATTAAATGTTGGTGATTATTTTGTGCTGACATCACATACAGTAATGCCATT
AAGTGCACCTACACTAGTGCCACAAGAGCACTATGTTAGAATTACTGGCTTATACCCAACACTCAATATC
TCAGATGAGTTTTCTAGCAATGTTGCAAATTATCAAAAGGTTGGTATGCAAAGTATTCTACACTCCAGG
GACCACCTGGTACTGGTAAGAGTCATTTTGCTATTGGCCTAGCTCTCTACTACCCTTCTGCTCGCATAGT
GTATACAGCTTGCTCTCATGCCGCTGTTGATGCACTATGTGAGAAGGCATTAAAATATTTGCCTATAGAT
AAATGTAGTAGAATTATACCTGCACGTGCTCGTGTAGAGTGTTTTGATAAATTCAAAGTGAATTC AACAT
TAGAACAGTATGTCTTTTGTACTGTAAATGCATTGCCTGAGACGACAGCAGATATAGTTGTCTTTGATGA
AATTTCAATGGCCACAAATTATGATTTGAGTGTTGTCAATGCCAGATTACGTGCTAAGCACTATGTGTAC
ATTGGCGACCCTGCTCAATTACCTGCACCACGCACATTGCTAACTAAGGGCACACTAGAACCAGAATATT
TCAATTCAGTGTGTAGACTTATGAAAACCTATAGGTCCAGACATGTTCCCTCGGAACCTTGTCGGCGTTGTCC
TGCTGAAATTGTTGACACTGTGAGTGCTTTGGTTTTATGATAATAAGCTTAAAGCACATAAAGACAAATCA
GCTCAATGCTTTAAATGTTTTATAAGGGTGTTATCACGCATGATGTTTCATCTGCAATTAACAGGCCAC
AAATAGGCGTGGTAAGAGAATTCCTTACACGTAACCCCTGCTTGGAGAAAAGCTGTCTTTATTTACCTTTA
TAATTCACAGAATGCTGTAGCCTCAAAGATTTTGGGACTACCAACTCAAAGTGTGATTATCACAGGGC
TCAGAATATGACTATGTCATATTCACTCAAACCACTGAAACAGCTCACTCTTGTAATGTAAACAGATTTA
ATGTTGCTATTACCAGAGCAAAAGTAGGCATACTTTGCATAATGTCTGATAGAGACCTTTATGACAAGTT
GCAATTTACAAGTCTTGAAATTCACGTAGGAATGTGGCAACTTTACAAGCTGAAAATGTAACAGGACTC
TTTAAAGATTGTAGTAAGGTAATCACTGGGTACATCCTACACAGGCACCTACACACCTCAGTGTTGACA
CTAAATTCAAACTGAAGGTTTTATGTGTTGACATACCTGGCATACCTAAGGACATGACCTATAGAAGACT
CATCTCTATGATGGGTTTTTAAATGAATTATCAAGTTAATGGTTACCCTAACATGTTTATCACCCGCGAA
GAAGCTATAAGACATGTACGTGCATGGATTGGCTTCGATGTGAGGGGTGTCATGCTACTAGAGAAGCTG
TTGGTACCAATTTACCTTTACAGCTAGGTTTTTCTACAGGTGTTAACCTAGTTGCTGTACCTACAGGTTA
TGTTGATACACCTAATAATACAGATTTTTCCAGAGTTAGTGCTAAACCACCGCCTGGAGATCAATTTAAA
CACCTCATACCCTTATGTACAAAGGACTTCCTTGGAATGTAGTGCGTATAAAGATTGTACAAATGTTAA
GTGACACACTTAAAAATCTCTCTGACAGAGTTCGATTTTGTCTTATGGGCACATGGCTTTGAGTTGACATC
TATGAAGTATTTTGTGAAAATAGGACCTGAGCGCACCTGTTGTCTATGTGATAGACGTGCCACATGCTTT
TCCACTGCTTCAGACACTTATGCCTGTTGGCATCATTCTATTGGATTTGATTACGTCTATAATCCGTTTA
TGATTGATGTTCAACAATGGGGTTTTACAGGTAACCTACAAAGCAACCATGATCTGTATTGTCAAGTCCA
TGGTAATGCACATGTAGCTAGTTGTGATGCAATCATGACTAGGTGTCTAGCTGTCCACGAGTGCTTTGTT
AAGCGTGTTGACTGGACTATTGAATATCCTATAATTGGTGATGAACTGAAGATTAATGCGGCTTGTAGAA
AGGTTCAACACATGGTTGTTAAAGCTGCATTATTAGCAGACAAATTCAGTTCTTTCACGACATTGGTAA
CCCTAAAGCTATTAAGTGTGTACCTCAAGCTGATGTAGAATGGAAGTTCTATGATGCACAGCCTTGTAGT
GACAAAGCTTATAAAATAGAAGAATTATTCTATTCTTATGCCACACATTCTGACAAATTCACAGATGGTG
TATGCCTATTTTGAATGCAATGTGATAGATATCCTGCTAATTCCATTGTTTGTAGATTTGACACTAG
AGTGCTATCTAACCTTAACCTGCCTGGTTGTGATGGTGGCAGTTTGTATGTAAATAAACATGCATTCCAC
ACACCAGCTTTTGATAAAAGTGCTTTTGTTAATTTAAACAATTACCATTTTTCTATTACTCTGACAGTC
CATGTGAGTCTCATGGAACAAGTAGTGTCAGATATAGATTATGTACCACTAAAGTCTGCTACGTGTAT
AACACGTTGCAATTTAGGTGGTGCTGTCTGTAGACATCATGCTAATGAGTACAGATTGTATCTCGATGCT
TATAACATGATGATCTCAGCTGGCTTTAGCTTGTGGGTTTTACAAACAATTTGATACTTATAACCTCTGGA
ACACTTTTACAAGACTTCAGAGTTTAGAAAATGTGGCTTTTAAATGTTGTAAATAAGGGACACTTTGATGG
ACAACAGGGTGAAGTACCAGTTTCTATCATTAATAACACTGTTTACACAAAAGTTGATGGTGTTGATGTA

GAATTGTTTGAAAATAAAACAACATTACCTGTTAATGTAGCATTTGAGCTTTGGGCTAAGCGCAACATTA
AACCAGTACCAGAGGTGAAAATACTCAATAATTTGGGTGTGGACATTGCTGCTAATACTGTGATCTGGGA
CTACAAAAGAGATGCTCCAGCACATATATCTACTATTGGTGTGTTTCTATGACTGACATAGCCAAGAAA
CCAAC TGAAACGATTTGTGCACCACTCACTGTCTTTTTTGATGGTAGAGTTGATGGTCAAGTAGACTTAT
TTAGAAATGCCCCGTAATGGTGTCTTATTACAGAAGGTAGTGTTAAAGGTTTACAACCATCTGTAGGTCC
CAAACAAGCTAGTCTTAATGGAGTCACATTAATTGGAGAAGCCGTAAAAACACAGTTCAATTATTATAAG
AAAGTTGATGGTGTGTCCAACAATTACCTGAAACTTACTTTACTCAGAGTAGAAATTTACAAGAATTTA
AACCCAGGAGTCAAATGGAAATTGATTTCTTAGAATTAGCTATGGATGAATTCATTGAACGGTATAAATT
AGAAGGCTATGCCTTCGAACATATCGTTTATGGAGATTTTAGTCATAGTCAGTTAGGTGGTTTACATCTA
CTGATTGGACTAGCTAAACGTTTTAAGGAATCACCTTTTGAATTAGAAGATTTTATTCCATATGGACAGTA
CAGTTAAAAACTATTTTCATAACAGATGCGCAACAGGTTTATCTAAGTGTGTGTGTTCTGTTATTGATTT
ATTACTTGATGATTTTTGTTGAAATAATAAAATCCCAAGATTTATCTGTAGTTTCTAAGGTTGTCAAAGTG
ACTATTGACTATACAGAAATTTTCAATTTATGCTTTGGTGTAAAGATGGCCATGTAGAAACATTTTACCCAA
AATTACAATCTAGTCAAGCGTGGCAACCGGGTGTGCTATGCCTAATCTTTACAAAATGCAAAGAATGCT
ATTAGAAAAGTGTGACCTTCAAATTATGGTGATAGTGCAACATTACCTAAAGGCATAATGATGAATGTC
GCAAAATATACTCAACTGTGTCAATATTTAAACACATTAACATTAGCTGTACCCTATAATATGAGAGTTA
TACATTTTGGTGCTGGTTCTGATAAAGGAGTTGCACCAGGTACAGCTGTTTTAAGACAGTGGTTGCCTAC
GGGTACGCTGCTTGTGATTAGATCTTAATGACTTTGTCTCTGATGCAGATTCAACTTTGATTGGTGAT
TGTGCAACTGTACATACAGCTAATAAATGGGATCTCATTATTAGTGATATGTACGACCCTAAGACTAAAA
ATGTTACAAAAGAAAATGACTCTAAAGAGGGTTTTTCACTTACATTTGTGGGTTTATACAACAAAAGCT
AGCTCTTGGAGGTTCCGTGGCTATAAAGATAACAGAACATTCTTGGAATGCTGATCTTTATAAGCTCATG
GGACACTTCGCATGGTGGACAGCCTTTGTTACTAATGTGAATGCGTCATCATCTGAAGCATTTTTAATTG
GATGTAATTATCTTGGCAAACACGCGAACAATAGATGGTTATGTCATGCATGCAAATTACATATTTTG
GAGGAATACAAATCCAATTCAGTTGTCTTCCTATTCTTTATTTGACATGAGTAAATTTCCCTTAAATTA
AGGGGTACTGCTGTTATGTCTTTAAAGAAGGTCAAATCAATGATATGATTTTATCTCTTCTTAGTAAAG
GTAGACTTATAATTAGAGAAAACAACAGAGTTGTTATTTCTAGTGATGTTCTTGTAAACAACTAAACGAA
CAATGTTTGTGTTTTCTTGTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCA
ATTACCCCTGCATACACTAATTCTTTCACACGTGGTGTGTTATTACCCTGACAAAGTTTTTCCAGATCCTCA
GTTTACATCTCAACTCAGGACTTGTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATACATG
TCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCCTGCTTACCATTAAATGATGGTGTGTTATTTTGC
TTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCTGAAGACCCAGTCC
CTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGAATGATCCAT
TTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGC
GAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTTC
AAAAATCTTAGGGAATTTGTGTTTAAAGAATATTGATGGTTATTTTAAATATATTCTAAGCACACGCCTA
TTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTTCGGCTTTAGAACCATTTGGTAGATTTGCCAATAGGTAT
TAACATCACTAGGTTTCAAACCTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCA
GGTTGGACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTAATAATATA
ATGAAAATGGAACCATTAACAGATGCTGTAGACTGTGCACCTTGACCCTCTCTCAGAAACAAAGTGACGTT
GAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACCTTCTAAGTTTGAAGTCCAACCAACAGAATCTATT
GTTAGATTTTCTAATATTACAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCTG
TTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCTTATATAATTCCGCATC
ATTTTCCACTTTTAAAGTGTTATGGAGTGTCTCCTACTAAATTAATGATCTCTGCTTTACTAATGTCTAT
GCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTG
ATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAAATCTAACAATCTTGATTCT
TAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAAGAG
GATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTGTAAGGTTTTAATTGTTACT
TTCTTTTACAATCATATGTTTTCCAACCCACTAATGGTGTGTTGTTACCAACCATAACAGATAGTAGTACT
TTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAAC
AAATGTGTCAATTTCAACTTCAATGGTTTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTC
TGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGA
GATTCTTGACATTACACCATGTTCTTTTGGTGGTGTGAGTGTATAACACCAGGAACAAATACTTCTAAC
CAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTTA
CTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTAAATAGGGGC
TGAACATGTCAACAACCTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACT
CAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACCTTG
GTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTAGTGTTAC

CACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCA
ACTGAATGCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTAACCGTGCTTTAACTGGAA
TAGCTGTTGAACAAGACAAAAACCCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACC
AATTAAAGATTTTGGTGGTTTTAATTTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCA
TTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAAACAATATGGTGATT
GCCTTGGTGATATTGCTGCTAGAGACCTCATTTGTGCACAAAAGTTTAACGGCCTTACTGTTTTGCCACC
TTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCTGGTTGG
ACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAAATGGTATTG
GAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAA
AATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACCTCAAGATGTGGTCAACCAAAATGCA
CAAGCTTTAAACACGCTTGTTTAAACAACCTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTAAATGATA
TCCTTTACCGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTTGATAGGTTGATCACAGGCAGACTTCAAAG
TTTGACAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTGCTGCT
ACTAAAATGTGAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTA
TGTCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAA
GAACTTCACAACCTGCTCCTGCCATTTGTCTATGATGGAAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTT
TCAAATGGCACACACTGGTTTGTAAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACA
CATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACC
TGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTTAAGAATCATACATCACCAGATGTTGATTTA
GGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTG
CCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACCTTGAAAGTATGAGCAGTATATAAAATGGCC
ATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGACAATTATGCTTTGCTGT
ATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAAGACG
ACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAAACGAACCTTATGGATTTGTTTATGAGA
ATCTTCACAATTGGAACCTGTAACCTTTGAAGCAAGGTGAAATCAAGGATGCTACTCCTTCAGATTTTGTTC
GCGCTACTGCAACGATACCGATACAAGCCTCACTCCCTTTCGGATGGCTTATTGTTGGCGTTGCACTTCT
TGCTGTTTTTTCAGAGCGCTTCCAAAATCATAACCCTCAAAAAGAGATGGCAACTAGCACTCTCCAAGGGT
GTTCACTTTGTTTGAACCTTGCTGTTGTTGTTGTTGTTGTAACAGTTTACTCACACCTTTTGCTCGTTGCTGCTG
GCCTTGAAGCCCCCTTTCTCTATCTTTATGCTTTTAGTCTACTTCTTGCAGAGTATAAACTTTGTAAGAAT
AATAATGAGGCTTTGGCTTTGCTGGAAATGCCGTTCCAAAACCCATTACTTTATGATGCCAATTTTTT
CTTTGCTGGCATACTAATTGTTACGACTATTGTATACCTTACAATAGTGTAACCTTCTTCAATTGTCATTA
CTTCAGGTGATGGCACAACAAGTCCTATTTCTGAACATGACTACCAGATTGGTGGTTATACTGAAAAATG
GGAATCTGGAGTAAAGACTGTGTTGTATTACACAGTTACTTCACCTCAGACTATTACCAGCTGTACTCA
ACTCAATTGAGTACAGACACTGGTGTGTAACATGTTACCTTCTTCATCTACAATAAAATTGTTGATGAGC
CTGAAGAACATGTCCAAATTCACACAATCGACGGTTCATCCGGAGTTGTTAATCCAGTAATGGAACCAAT
TTATGATGAACCGACGACGACTACTAGCGTGCCTTTGTAAGCACAAAGCTGATGAGTACGAACCTTATGTAC
TCATTGTTTTCGGAAGAGACAGGTACGTTAATAGTTAATAGCGTACTTCTTTTTCTTGCTTTCTGTTGAT
TCTTGCTAGTTACACTAGCCATCCTTACTGCGCTTCGATTGTGTGCGTACTGCTGCAATATTGTTAACGT
GAGTCTTGTAACCTTCTTTTTACGTTTACTCTCGTGTTAAAAATCTGAATTCTTCTAGAGTTCCTGAT
CTTCTGGTCTAAACGAACATAAATATTATATTAGTTTTTCTGTTTGGAACCTTAATTTTAGCCATGGCAGA
TTCCAACGGTACTATTACCGTTGAAGAGCTTAAAAAGCTCCTTGAACAATGGAACCTAGTAATAGGTTTC
CTATTCCTTACATGGATTTGTCTTCTACAATTTGCCTATGCCAACAGGAATAGGTTTTTGTATATAATTA
AGTTAATTTTCTCTGGCTGTTATGGCCAGTAACTTTAGCTTGTTTTGTGCTTGCTGCTGTTTACAGAAT
AAATTGGATCACCGGTGGAATTGCTATCGCAATGGCTTGTCTTGTAGGCTTGATGTGGCTCAGCTACTTC
ATTGCTTCTTTCAGACTGTTTGC GGTACGCGTTCATGTGGTCATTCAATCCAGAACTAACATTCTTC
TCAACGTGCCACTCCATGGCACTATTCTGACCAGACCGCTTCTAGAAAGTGAACCTCGTAATCGGAGCTGT
GATCCTTTCGTGGACATCTTCGTATTGCTGGACACCATCTAGGACGCTGTGACATCAAGGACCTGCCTAAA
GAAATCACTGTTGCTACATACGAACGCTTCTTATTACAAATTTGGGAGCTTCGACGCGTGTAGCAGGTG
ACTCAGGTTTTGCTGCATACAGTCGTACAGGATTGGCAACTATAAATTAACACAGACCATTCCAGTAG
CAGTGACAATATTGCTTTGCTTGTACAGTAAGTGACAACAGATGTTTCATCTCGTTGACTTTTCAGGTTAC
TATAGCAGAGATATTACTAATTATTATGAGGACTTTTAAAGTTTCCATTTGGAATCTTGATTACATCATA
AACCTCATAATTAATAAATTTATCTAAGTCACTAACTGAGAATAAATATTCTCAATTAGATGAAGAGCAAC
CAATGGAGATTGATTAAACGAACATGAAAATTATTCTTTTCTTGGCACTGATAACACTCGCTACTTGTGA
GCTTTATCACTACCAAGAGTGTGTTAGAGGTACAACAGTACTTTTAAAAGAACCTTGCTCTTCTGGAACA
TACGAGGGCAATTCACCATTTTCATCTCTAGCTGATAACAAATTTGCACTGACTTGCTTTAGCACTCAAT
TTGCTTTTGTCTTGTCTGACGGCGTAAAACACGTCTATCAGTTACGTGCCAGATCAGTTTTCACCTAACT
GTTTCATCAGACAAGAGGAAGTTCAAGAACCTTTACTCTCCAATTTTTCTTATTGTTGCGGCAATAGTGTTT

ATAACACTTTGCTTCACACTCAAAAGAAAGACAGAATGATTGAACTTTCATTAATTGACTTCTATTTGTG
CTTTTTAGCCTTTCTGCTATTCCTTGTTTTAATTATGCTTATTATCTTTTGGTTCTCACTTGAAGTGCAA
GATCATAATGAACTTGTACGCCTAAACGAACATGAAATTTCTTGTTTTCTTAGGAATCATCACAACGTG
TAGCTGCATTTACCAAGAATGTAGTTTACAGTCATGTACTCAACATCAACCATATGTAGTTGATGACCC
GTGTCCTATTCACCTTCTATTCTAAATGGTATATTAGAGTAGGAGCTAGAAAATCAGCACCTTTAATTGAA
TTGTGCGTGGATGAGGCTGGTTCTAAATCACCCATTTCAGTACATCGATATCGGTAATTATACAGTTTCCT
GTTTACCTTTTACAATTAATTGCCAGGAACCTAAATTGGGTAGTCTTGTAGTGCGTTGTTTCGTTCTATGA
AGACTTTTTAGAGTATCATGACGTTTCGTGTTGTTTTAGATTTTCATCTAAACGAACAAACTAAAATGTCTG
ATAATGGACCCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCTCAGATTCAACTGGCAG
TAACCAGAATGGAGAACGCAGTGGGGCGCGATCAAAACAACGTCGGCCCCAAGGTTTACCCAATAATACT
GCGTCTTGGTTCACCGCTCTCACTCAACATGGCAAGGAAGACCTTAAATTCCTCGAGGACAAGGCGTTC
CAATTAACACCAATAGCAGTCCAGATGACCAAAATTGGCTACTACCGAAGAGCTACCAGACGAATTTCGTGG
TGGTGACGGTAAAATGAAAGATCTCAGTCCAAGATGGTATTTCTACTACCTAGGAACTGGGCCAGAAGCT
GGACTTCCCTATGGTGCTAACAAGACGGCATCATATGGGTTGCAACTGAGGGAGCCTTGAATACACCAA
AAGATCACATTGGCACCCGCAATCCTGCTAACAATGCTGCAATCGTGCTACAACCTCCTCAAGGAACAAC
ATTGCCAAAAGGCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTTCCTCATCACGT
AGTCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACTTCTCCTGCTAGAATGGCTGGCA
ATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCTTGAGAGCAAAATGTCTGG
TAAAGGCCAACAACAACAAGGCCAACTGTCTACTAAGAAATCTGCTGCTGAGGCTTCTAAGAAGCCTCGG
CAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAGACGTGGTCCAGAACAACCC
AAGGAAATTTTGGGGACCAGGAATAATCAGACAAGGAAGTATTACAAACATTGGCCGCAAATTGCACA
ATTTGCCCCCAGCGCTTCAGCGTTCTTCGGAATGTGCGCATTGGCATGGAAGTCACACCTTCGGGAACG
TGGTTGACCTACACAGGTGCCATCAAATTGGATGACAAAGATCCAAATTTCAAAGATCAAGTCATTTTGC
TGAATAAGCATATTGACGCATACAAAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAGAAGGC
TGATGAACTCAAGCCTTACCGCAGAGACAGAAGAAACAGCAAAGTGTGACTCTTCTTCCTGCTGCAGAT
TTGGATGATTTCTCCAAACAATTGCAACAATCCATGAGCAGTGCTGACTCAACTCAGGCCTAACTCATG
CAGACCACACAAGGCAGATGGGCTATATAAACGTTTTTCGCTTTTCCGTTTACGATATATAGTCTACTCTT
GTGCAGAATGAATTCTCGTAACTACATAGCACAAAGTAGATGTAGTTAACTTTAATCTCACATAGCAATCT
TTAATCAGTGTGTAAACATTAGGGAGGACTTGAAAGAGCCACCACATTTTACCAGAGGCCACGCGGAGTAC
GATCGAGTGACAGTGAACAATGCTAGGGAGAGCTGCCTATATGGAAGAGCCCTAATGTGTAAAATTAAT
TTTAGTAGTGCTATCCCCATGTGATTTTAATAGCTTCTTAGGAGAATGACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA