

Homework #5

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): Prefix Trie

Create a class called ***prefix_trie***. The purpose of the class will be to contain a FASTA read set and all of the functions needed to operate on this set. Use the **prefix trie** data-structure to store the genomic sequences of the given read dataset. Here you will be performing fuzzy matching, tolerating up to 1 mismatch.

At minimum, the class must contain(25pts):

- A default constructor
 - At least one custom constructor to build a trie from a set of queries (of size n)
 - A function to traverse (*search*) the trie using a genome of size G. Note that you can assume that $G \gg n$. You will need to implement a ***fuzzy search tolerating up to 1 mismatch (substitutions only)***. Hint: use a stack to keep track of branches in the tree that need to be explored.
 - A destructor
 - A copy constructor
- A. **(25pts) Basic prefix trie:** Generate 5K, 50K, 100K, and 1M random 36-mers from the SARS-CoV2 genome sequence (Appendix A) and store them in the prefix trie. Hint: generate a random starting position somewhere in the genome and read 36 characters starting from that position.
- For each of the 36-mer datasets, what are the sizes of the trie (# of nodes)? Explain the pattern that you observed.
 - Iterate through all possible 36-mers in the SARS-CoV2 genome, using each to search / traverse the prefix trie with up to 1 mismatch. How many of your 36-mers had a match? Does it make sense? Explain why.
- B. **(25pts) Impact of error rate on trie structure:** Generate 5K, 50K, 100K, and 1M random 36-mers from the SARS-CoV2 genome sequence with **5% per-base error rate** and store them in the prefix trie. Hint: repeat the process from part A, except each base of 36-mer has a 5% chance of mutation/error.
- For each of the 36-mer datasets, what are the sizes of the trie (# of nodes)? Explain differences (if any) between the trie sizes in partA and part B.
 - Iterate through all possible 36-mers in the SARS-CoV2 genome, using each to search / traverse the prefix trie with up to 1 mismatch. How many of your 36-mers had a match? Does it make sense? Explain why.

Appendix A: SARS-COV2 Genome

>NC_045512.2_SARS_COV2_complete_genome

ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTTCGATCTCTTGTAGATCTGTTCTCTAAA
CGAACTTTAAATCTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAAC
TAATTACTGTCGTTGACAGGACACGAGTAACCTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGTG
TTGCAGCCGATCATCAGCACATCTAGGTTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTC
CCTGGTTTCAACGAGAAAAACACACGTCCAACCTCAGTTTGCCTGTTTTACAGGTTTCGCGACGTGCTCGTAC
GTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGG
CTTAGTAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCAAACGTTCCGAT
GCTCGAAGTGCACCTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAAGTGAAGGCATTAGTACGGTC
GTAGTGGTGAGACACTTGGTGTCTTGTCCCTCATGTGGGCGAAATACCAGTGGCTTACCGCAAGGTTCT
TCTTCGTAAGAACGGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTA
GGCGACGAGCTTGGCACTGATCCTTATGAAGATTTTTCAAGAAAACCTGGAACACTAAACATAGCAGTGGTG
TTACCCGTGAACCTCATGCGTGAGCTTAACGGAGGGGCATACACTCGCTATGTCGATAACAACCTTCTGTGG
CCCTGATGGCTACCCTCTTGAGTGCATTAAAGACCTTCTAGCACGTGCTGGTAAAGCTTCATGCACTTTG
TCCGAACAACCTGGACTTTATTGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTG
CTTGGTACACGGAACGTTCTGAAAAGAGCTATGAATTGCAGACACCTTTTGAAATTAAATTGGCAAAGAA
ATTTGACACCTTCAATGGGGAATGTCCAAATTTTGTATTTCCCTTAAATTCCATAATCAAGACTATTCAA
CCAAGGGTTGAAAAGAAAAAGCTTGATGGCTTTATGGGTAGAATTCGATCTGTCTATCCAGTTGCGTCAC
CAAATGAATGCAACCAATGTGCCTTTCAACTCTCATGAAGTGTGATCATTGTGGTGAAACTTCATGGCA
GACGGGCGATTTTGTAAAGCCACTTGCGAATTTTGTGGCACTGAGAATTTGACTAAAGAAGGTGCCACT
ACTTGTGGTTACTTACCCCAAAATGCTGTTGTTAAAATTTATTGTCCAGCATGTCACAATTCAGAAGTAG
GACCTGAGCATAGTCTTGCCGAATACCATAATGAATCTGGCTTGAACACCATTCTTCGTAAGGGTGGTCG
CACTATTGCCTTTGGAGGCTGTGTGTTCTCTTATGTTGGTTGCCATAACAAGTGTGCCTATTGGGTTCCA
CGTGCTAGCGCTAACATAGGTTGTAACCATAACAGGTGTTGTTGGAGAAGGTTCCGAAGGTCTTAATGACA
ACCTTCTTGAAATACTCCAAAAAGAGAAAGTCAACATCAATATTGTTGGTGACTTTAACTTAATGAAGA
GATCGCCATTATTTTGGCATCTTTTTCTGCTTCCACAAGTGCTTTTGTGGAACTGTGAAAGGTTTGGAT
TATAAAGCATTAACAATTTGTTGAATCCTGTGGTAATTTTAAAGTTACAAAAGGAAAAAGCTAAAAAG
GTGCCTGGAATATTGGTGAACAGAAATCAATAGTACTGAGTCTCTTTATGCATTTGCATCAGAGGCTGCTCG
TGTTGTACGATCAATTTTCTCCCGCACTCTTGAAACTGCTCAAAATTCTGTGCGTGTTTTACAGAAGGCC
GCTATAACAATACTAGATGGAATTTACAGTATTCACTGAGACTCATTGATGCTATGATGTTTACATCTG
ATTTGGCTACTAACAATCTAGTTGTAATGGCCTACATTACAGGTGGTGTGTTTCAAGTTGACTTCGCAGTG
GCTAACTAACATCTTTGGCACTGTTTATGAAAACTCAAACCCGTCCTTGATTGGCTTGAAGAGAAGTTT
AAGGAAGGTGTAGAGTTTCTTAGAGACGGTTGGGAAATTGTTAAATTTATCTCAACCTGTGCTTGTGAAA
TTGTGCGGTGGACAAATTGTCACCTGTGCAAAGGAAATTAAGGAGAGTGTTTCAAGACATTCTTTAAGCTTGT
AAATAAATTTTGGCTTTGTGTGCTGACTCTATCATTATTGGTGGAGCTAACTTAAAGCCTTGAATTTA
GGTGAACATTTTGTACGCACTCAAAGGGATTGTACAGAAAGTGTGTTAAATCCAGAGAAGAACTGGCC
TACTCATGCCTCTAAAAGCCCCAAAAGAAATTATCTTCTTAGAGGGAGAAACACTTCCACAGAAGTGTT
AACAGAGGAAGTTGTCTTGAAAACCTGGTGATTTACAACCATTAGAACAACCTACTAGTGAAGCTGTTGAA
GCTCCATTGGTTGGTACACCAAGTTTGTATTAACGGGCTTATGTTGCTCGAAATCAAAGACACAGAAAAGT
ACTGTGCCCTTGCACCTAATATGATGGTAACAAACAATACCTTCACACTCAAAGGCGGTGCACCAACAAA
GGTTACTTTTGGTGATGACACTGTGATAGAAGTGCAAGGTTACAAGAGTGTAATATCACTTTTGAACCTT
GATGAAAGGATTGATAAAGTACTTAATGAGAAGTGCTCTGCCTATACAGTTGAACTCGGTACAGAAGTAA
ATGAGTTTCGCTGTGTTGTGGCAGATGCTGTCTATAAAAACTTTGCAACCAGTATCTGAATTACTTACACC
ACTGGGCATTGATTTAGATGAGTGGAGTATGGCTACATACTACTTATTTGATGAGTCTGGTGAGTTTAA
TTGGCTTACATATGTATTGTTCTTTCTACCTCCAGATGAGGATGAAGAAGAAGGTGATTGTGAAGAAG
AAGAGTTTGAAGCATCAACTCAATATGAGTATGGTACTGAAGATGATTACCAAGGTAAACCTTTGGAATT
TGGTGCCACTTCTGCTGCTCTTCAACCTGAAGAAGAGCAAGAAGAAGATTGGTTAGATGATGATAGTCAA
CAAACCTGTTGGTCAACAAGACGGCAGTGAGGACAATCAGACAACCTACTATTCAAACAATTGTTGAGGTTT
AACCTCAATTAGAGATGGAACCTTACACCAAGTTGTTTCAAGTATTGAAGTGAATAGTTTTAGTGGTTATTT
AAAACCTTACTGACAATGTATACATTAAAAATGCAGACATTGTGGAAGAAGCTAAAAAGGTAAACCAACA
GTGGTTGTTAATGCAGCCAATGTTTACCTTAAACATGGAGGAGGTGTTGCAGGAGCCTTAAATAAGGCTA
CTAACAATGCCATGCAAGTTGAATCTGATGATTACATAGCTACTAATGGACCACTTAAAGTGGGTGGTAG
TTGTGTTTTAAGCGGACACAATCTTGCTAAACACTGTCTTCATGTTGTGCGCCCAAATGTTAACAAGGT
GAAGACATTCAACTTCTTAAGAGTGCTTATGAAAATTTTAAATCAGCACGAAGTTCTACTTGCACCATAT
TATCAGCTGGTATTTTTTGGTGCTGACCCTATACATTCTTTAAGAGTTTGTGTAGATACTGTTTCGACAAA

[illegible]

TCTTTACTTTGATAAAGCTGGTCAAAAGACTTATGAAAGACATTCTCTCTCTCATTGTTAACTTAGAC
AACCTGAGAGCTAATAACACTAAAGGTTCAATTGCCTATTAATGTTATAGTTTTTGATGGTAAATCAAAAT
GTGAAGAATCATCTGCAAAATCAGCGTCTGTTTACTACAGTCAGCTTATGTGTCAACCTATACTGTTACT
AGATCAGGCATTAGTGTCTGATGTTGGTGATAGTGCAGGAAAGTTGCAGTTAAAATGTTTGATGCTTACGTT
AATACGTTTTTCATCAACTTTTAACGTACCAATGGAAAACTCAAAACACTAGTTGCAACTGCAGAAGCTG
AACTTGCAAAGAATGTGTCCTTAGACAATGTCTTATCTACTTTTTATTTTCAGCAGCTCGGCAAGGGTTTTGT
TGATTGAGATGTAGAACTAAAGATGTTGTTGAATGTCTTAAATTGTCACATCAATCTGACATAGAAGTT
ACTGGCGATAGTTGTAATAACTATATGCTCACCTATAACAAAGTTGAAAACATGACACCCCGTGACCTTG
GTGCTTGTATTGACTGTAGTGCGCGTCATATTAATGCGCAGGTAGCAAAAAGTCACAACATTGCTTTGAT
ATGGAACGTTAAAGATTTTCATGTCAATTGTCTGAACAACTACGAAAACAAATACGTAGTGCTGTCAAAAAG
AATAACTTACCTTTTTAAGTTGACATGTGCAACTACTAGACAAGTTGTTAATGTTGTAACAACAAGATAG
CACTTAAGGGTGGTAAAATTGTTAATAATTGGTTGAAGCAGTTAATTAAGTTACACTTGTGTTCCTTTTT
TGTTGCTGCTATTTTTCTATTTAATAACACCTGTTTCATGTCTGCTAAACATACTGACTTTTTCAAGTGAA
ATCATAGGATACAAGGCTATTGATGGTGGTGTCACTCGTGACATAGCATCTACAGATACTTGTTTTTGCTA
ACAAACATGCTGATTTTTGACACATGGTTTAGCCAGCGTGGTGGTAGTTATACTAATGACAAAGCTTGCCC
ATTGATTGCTGCAGTCATAACAAGAGAAGTGGGTTTTGTCGTGCCTGGTTTTGCCTGGCAGCATATTACGC
ACAACATGTTGACTTTTTTGCACTTTCTTACCTAGAGTTTTTAGTGAGTTGGTAACATCTGTTACACAC
CATCAAACTTATAGAGTACACTGACTTTGCAACATCAGCTTGTGTTTTGGCTGCTGAATGTACAATTTTT
TAAAGATGCTTCTGGTAAGCCAGTACCATATTGTTATGATACCAATGTACTAGAAGGTTCTGTTGCTTAT
GAAAGTTTACGCCCTGACACACGTTATGTGCTCATGGATGGCTCTATTATTCAATTTCCCTAACACCTACC
TTGAAGGTTCTGTTAGAGTGGTAACAACCTTTTGATTCTGAGTACTGTAGGCACGGCACTTGTGAAAGATC
AGAAGCTGGTGGTTGTGTATCTACTAGTGGTAGATGGGTACTTAACAATGATTATTACAGATCTTTACCA
GGAGTTTTCTGTGGTGTAGATGCTGTAAATTTACTTACTAATATGTTTACACCACTAATTCAACCTATTG
GTGCTTTGGACATATCAGCATCTATAGTAGCTGGTGGTATTGTAGCTATCGTAGTAACATGCCTTGCCCTA
CTATTTTATGAGGTTTAGAAGAGCTTTTGGTGAATACAGTCATGTAGTTGCCTTTAATACTTTACTATTC
CTTATGTCATTCACTGTACTCTGTTTAAACACAGTTTACTCATTCTTACCTGGTGGTTTATTCTGTTATTT
ACTTGTACTTGACATTTTTATCTTACTAATGATGTTTCTTTTTTAGCACATATTCACTGGATGGTTATGTT
CACACCTTTAGTACCTTTCTGGATAACAATTGCTTATATCATTGTTGATTTCCACAAAGCATTTCCTATTGG
TTCTTTAGTAATTAACCTTAAAGAGACGTGTAGTCTTTAATGGTGGTTTCTTTTAGTACTTTTGAAGAAGCTG
CGCTGTGACCTTTTTGTTAAATAAAGAAATGTATCTAAAGTTGCGTAGTGATGTGCTATTACCTCTTAC
GCAATATAATAGATACTTAGCTCTTTATAATAAGTACAAGTATTTTAGTGGAGCAATGGATACAACCTAGC
TACAGAGAAGCTGCTTGTGTGTCATCTCGCAAAGGCTCTCAATGACTTCAGTAACCTCAGGTTCTGATGTTT
TTTACCAACCACCACAAACCTCTATCACCTCAGCTGTTTTGCAGAGTGGTTTTAGAAAAATGGCATTCCC
ATCTGGTAAAGTTGAGGGTTGTATGGTACAAGTAACCTGTGGTACAACCTACACTTAACGGTCTTTGGCTT
GATGACGTAGTTTACTGTCCAAGACATGTGATCTGCACCTCTGAAGACATGCTTAACCTAATTATGAAG
ATTTACTCATTGTAAGTCTAATCATAATTTCTTGGTACAGGCTGGTAATGTTCAACTCAGGGTTATTGG
ACATTCTATGCAAAATTGTGTACTTAAGCTTAAGGTTGATACAGCCAATCCTAAGACACCTAAGTATAAG
TTTGTTTCGCAATTAACAGGACAGACTTTTTTCACTGTTAGCTTGTACAATGGTTTACCCTCTGGTGGTT
ACCAATGTGCTATGAGGCCCAATTTCACTATTAAGGGTTTCACTTCTTAATGGTTTCACTGTTGTTGG
TTTTAACATAGATTATGACTGTGTCTCTTTTTGTTACATGCACCATATGGAATTACCAACTGGAGTTTCA
GCTGGCACAGACTTAGAAGGTAACCTTTTATGGACCTTTTGTGACAGGCAAACAGCACAAAGCAGCTGGTA
CGGACACAACCTATTACAGTTAATGTTTTAGCTTGGTTGTACGCTGCTGTTATAAATGGAGACAGGTGGTT
TCTCAATCGATTTACCACAACCTCTAATGACTTTAACCTTGTGGCTATGAAGTACAATTATGAACCTCTA
ACACAAGACCATGTTGACATACTAGGACCTCTTTCTGCTCAAACCTGGAATTGCCGTTTTAGATATGTGTG
CTTCATTAAAAGAATTACTGCAAAATGGTATGAATGGACGTACCATATTGGGTAGTGCTTTATTAGAAGA
TGAATTTACACCTTTTGATGTTGTTAGACAATGCTCAGGTGTTACTTTCCAAAGTGCAGTGAAAAGAACA
ATCAAGGGTACACACCCTGGTTGTTACTCACAATTTTGACTTCACTTTTAGTTTTAGTCCAGAGTACTC
AATGGTCTTTGTTCTTTTTTTTTGTATGAAAATGCCTTTTTACCTTTTGCTATGGGTATTATTGCTATGTC
TGCTTTTGCAATGATGTTTGTCAAACATAAGCATGCATTTCTCTGTTTGTGTTTTGTTACCTTCTCTTGCC
ACTGTAGCTTATTTAATATGGTCTATATGCCTGCTAGTTGGGTGATGCGTATTATGACATGGTTGGATA
TGTTTGATACTAGTTTGTCTGGTTTTAAGCTAAAAGACTGTGTTATGTATGCATCAGCTGTAGTGTTACT
AATCCTTATGACAGCAAGAAGTGTGTATGATGATGGTGCTAGGAGAGTGTGGACACTTATGAATGTCTTG
ACACTCGTTTTATAAAGTTTATTATGGTAATGCTTTAGATCAAGCCATTTCCATGTGGGCTCTTATAATCT
CTGTTACTTCTAACTACTCAGGTGTAGTTACAACGTGCATGTTTTTGGCCAGAGGTATTGTTTTTATGTG
TGTTGAGTATTGCCCTATTTTCTTCATAACTGGTAATACACTTCAGTGTATAATGCTAGTTTTATTGTTTC
TTAGGCTATTTTTGTACTTGTACTTTGGCCTCTTTTGTGTTACTCAACCGCTACTTTAGACTGACTCTTG
GTGTTTATGATTACTTAGTTTCTACACAGGAGTTTAGATATATGAATTCACAGGGACTACTCCCACCCAA

GAATAGCATAGATGCCTTCAAACCTCAACATTAAATTGTTGGGTGTTGGTGGCAAACCTTGTATCAAAGTA
GCCACTGTACAGTCTAAAATGTGAGATGTAAAGTGCACATCAGTAGTCTTACTCTCAGTTTTTGCAACAAC
TCAGAGTAGAATCATCATCTAAATTGTGGGCTCAATGTGTCCAGTTACACAATGACATTCTCTTAGCTAA
AGATACTACTGAAGCCTTTGAAAAAATGGTTTTCACTACTTTCTGTTTTGCTTTCCATGCAGGGTGTGTGA
GACATAAACAAGCTTTGTGAAGAAATGCTGGACAACAGGGCAACCTTACAAGCTATAGCCTCAGAGTTTA
GTTCCCTTCCATCATATGCAGCTTTTGTACTGCTCAAGAAGCTTATGAGCAGGCTGTTGCTAATGGTGA
TTCTGAAGTTGTTCTTAAAAAGTTGAAGAAGTCTTTGAATGTGGCTAAATCTGAATTTGACCGTGATGCA
GCCATGCAACGTAAGTTGGAAAAGATGGCTGATCAAGCTATGACCCAAATGTATAAACAGGCTAGATCTG
AGGACAAGAGGGCAAAAGTTACTAGTGCTATGCAGACAATGCTTTTCACTATGCTTAGAAAAGTTGGATAA
TGATGCACTCAACAACATTATCAACAATGCAAGAGATGGTTGTGTTCCCTTGAACATAATACCTCTTACA
ACAGCAGCCAAACTAATGGTTGTGCATACCAGACTATAACACATATAAAAAATACGTGTGATGGTACAACAT
TTACTTATGCATCAGCATTGTGGGAAATCCAACAGGTTGTAGATGCAGATAGTAAAATTGTTCAACTTAG
TGAAATTAGTATGGACAATTCACCTAATTTAGCATGGCCTCTTATTGTAACAGCTTTAAGGGCCAATTCT
GCTGTCAAATTACAGAATAATGAGCTTAGTCCTGTTGCACTACGACAGATGTCTTGTGCTGCCGGTACTA
CACAACTGCTTGCCTGATGACAATGCGTTAGCTTACTACAACACAACAAAGGGAGGTAGGTTTGTACT
TGCCTGTTATCCGATTTACAGGATTTGAAATGGGCTAGATTCCCTAAGAGTGATGGAAGTGGTACTATC
TATACAGAAGTGAACACCTTGTAGGTTTGTACAGACACACCTAAAGGTCCTAAAGTGAAGTATTTAT
ACTTTATTAAAGGATTAAACAACCTAAATAGAGGTATGGTACTTGGTAGTTTAGCTGCCACAGTACGTCT
ACAAGCTGGTAATGCAACAGAAGTGCCTGCCAATTCACTGTATTATCTTTCTGTGCTTTTGTGTAGAT
GCTGCTAAAGCTTACAAAGATTATCTAGCTAGTGGGGGACAACCAATCACTAATTGTGTTAAGATGTTGT
GTACACACACTGGTACTGGTCAGGCAATAACAGTTACACCGGAAGCCAATATGGATCAAGAATCCTTTGG
TGGTGCATCGTGTGTGTCTGTACTGCCGTTGCCACATAGATCATCAAATCCTAAAGGATTTTGTGACTTA
AAAGGTAAGTATGTACAAATACCTACAACCTTGTGCTAATGACCCTGTGGGTTTTACACTTAAAAACACAG
TCTGTACCGTCTGCGGTATGTGGAAAGGTTATGGCTGTAGTTGTGATCAACTCCGCGAAGCCATGCTTCA
GTCAGCTGATGCACAATCGTTTTTAAACGGGTTTGGCGGTGTAAGTGCAGCCCGTCTTACACCGTGGCGCA
CAGGCACTAGTACTGATGTGCTATACAGGGCTTTTGACATCTACAATGATAAAGTAGCTGGTTTTGCTAA
ATTCCTAAAACTAATTGTTGTGCTTCCAAGAAAAGGACGAAGATGACAATTTAATTGATTCTTACTTT
GTAGTTAAGAGACACACTTTCTCTAACTACCAACTGAAGAAACAATTTATAATTTACTTAAAGGATTGTC
CAGCTTTGCTTAAACATGACTTCTTTAAGTTTGAAGATAGACGGTGACATGGTACCACATATATACGTCA
ACGTCTTTACTAAATACACAATGGCAGACCTCGTCTATGCTTTAAGGCATTTTGTGATGAAGTAATTGTGAC
ACATTTAAAGAAATACTTGTACATACAATTGTTGTGATGATGATTATTTCAATAAAAAGGACTGGTATG
ATTTTGTAGAAAACCCAGATATATTACGCGTATACGCCAAGTTAGGTGAACGTGTACGCCAAGCTTTGTT
AAAAACAGTACAATTCTGTGATGCCATGCGAAATGCTGGTATTGTTGGTGTACTGACATTAGATAATCAA
GATCTCAATGGTAAGTGGTATGATTTGCGGTGATTTTACATAACACCGCCAGGTAGTGGAGTTCCTGTTG
TAGATTCTTATTATTATTGTTAATGCCTATATTAACCTTGACCAGGGCTTTAACTGCAGAGTCACATGT
TGACACTGACTTAAACAAAGCCTTACATTAAGTGGGATTTGTTAAATATGACTTCACGGAAGAGAGGTTA
AAACTCTTTGACCGTTATTTTAAATATTGGGATCAGACATACCACCCAAATTGTGTTAACTGTTTGGATG
ACAGATGCATTCTGCATTGTGCAAACCTTTAATGTTTTATTCTCTACAGTGTTCCACCTACAAGTTTTGG
ACCACTAGTGAGAAAAATATTTGTTGATGGTGTTCATTTGTAGTTTCACTGGATACCACTTCAGAGAG
CTAGGTGTTGTACATAATCAGGATGTAACTTACATAGCTCTAGACTTAGTTTTAAGGAATTACTTGTGT
ATGCTGCTGACCCTGCTATGCACGCTGCTTCTGGTAATCTATTACTAGATAAACGCACTACGTGCTTTTC
AGTAGCTGCACTTACTAACAATGTTGCTTTTCAAACCTGTCAAACCCGGTAATTTTAAACAAAGACTTCTAT
GACTTTGCTGTGTCTAAGGGTTTCTTTAAGGAAGGAAGTTCTGTTGAATTTAAACACTTCTTCTTTGCTC
AGGATGGTAATGCTGCTATCAGCGATTATGACTACTATCGTTATAATCTACCAACAATGTGTGATATCAG
ACAACACTATTTGTAGTTGAAGTTGTTGATAAGTACTTTGATTGTTACGATGGTGGCTGTATTAATGCT
AACCAAGTCATCGTCAACAACCTAGACAAATCAGCTGGTTTTCCATTTAATAAATGGGGTAAGGCTAGAC
TTTATTATGATTCAATGAGTTATGAGGATCAAGATGCATTTTTCGATATACAAAACGTAATGTCTATCCC
TACTATAACTCAAATGAATCTTAAAGTATGCCATTAGTGCAAGAATAGAGCTCGCACCGTAGCTGGTGTCT
TCTATCTGTAGTACTATGACCAATAGACAGTTTCATCAAAAATTATTGAAATCAATAGCCGCCACTAGAG
GAGCTACTGTAGTAATTGGAACAAGCAAATCTATGGTGGTTGGCACAACATGTTAAAACTGTTTATAG
TGATGTAGAAAACCTCACCTTATGGGTTGGGATTATCCTAAATGTGATAGAGCCATGCCTAACATGCTT
AGAATTATGGCCTCACTTGTCTTGTCTCGCAAACATACAACGTGTTGTAGCTTGTACACCGTTTCTATA
GATTAGCTAATGAGTGTGCTCAAGTATTGAGTGAAATGGTCATGTGTGGCGGTTCACTATATGTTAAACC
AGGTGGAACCTCATCAGGAGATGCCACAACCTGCTTATGCTAATAGTGTTTTTAACATTTGTCAAGCTGTC
ACGGCCAATGTTAATGCATTTTATCTACTGATGGTAACAAAATTGCCGATAAGTATGTCCGCAATTTAC
AACACAGACTTTTATGAGTGTCTCTATAGAAATAGAGATGTTGACACAGACTTTGTGAATGAGTTTTACGC
ATATTTGCGTAACATTTCTCAATGATGATACTCTCTGACGATGCTGTTGTGTGTTTCAATAGCACTTAT

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

GAATTGTTTGAAAATAAAACAACATTACCTGTTAATGTAGCATTTGAGCTTTGGGCTAAGCGCAACATTA
AACCAGTACCAGAGGTGAAAATACTCAATAATTTGGGTGTGGACATTGCTGCTAATACTGTGATCTGGGA
CTACAAAAGAGATGCTCCAGCACATATATCTACTATTGGTGTGTTTCTATGACTGACATAGCCAAGAAA
CCAAC TGAAACGATTTGTGCACCACACTCACTGTCTTTTTTGATGGTAGAGTTGATGGTCAAGTAGACTTAT
TTAGAAATGCCCCGTAATGGTGTCTTATTACAGAAGGTAGTGTTAAAGGTTTACAACCATCTGTAGGTCC
CAAACAAGCTAGTCTTAATGGAGTCACATTAATTGGAGAAGCCGTAAAAACACAGTTCAATTATTATAAG
AAAGTTGATGGTGTGTCCAACAATTACCTGAAACTTACTTTACTCAGAGTAGAAATTTACAAGAATTTA
AACCCAGGAGTCAAATGGAAATTGATTTCTTAGAATTAGCTATGGATGAATTCATTGAACGGTATAAATT
AGAAGGCTATGCCTTCGAACATATCGTTTATGGAGATTTTAGTCATAGTCAGTTAGGTGGTTTACATCTA
CTGATTGGACTAGCTAAACGTTTTAAGGAATCACCTTTTGAATTAGAAGATTTTATTCCATATGGACAGTA
CAGTTAAAAACTATTTTCATAACAGATGCGCAACAGGTTTATCTAAGTGTGTGTGTTCTGTTATTGATTT
ATTACTTGATGATTTTTGTTGAAATAATAAAATCCCAAGATTTATCTGTAGTTTCTAAGGTTGTCAAAGTG
ACTATTGACTATACAGAAATTTTCAATTTATGCTTTGGTGTAAAGATGGCCATGTAGAAACATTTTACCCAA
AATTACAATCTAGTCAAGCGTGGCAACCGGGTGTGCTATGCCTAATCTTTACAAAATGCAAAGAATGCT
ATTAGAAAAGTGTGACCTTCAAATTATGGTGATAGTGCAACATTACCTAAAGGCATAATGATGAATGTC
GCAAAATATACTCAACTGTGTCAATATTTAAACACATTAACATTAGCTGTACCCTATAATATGAGAGTTA
TACATTTTGGTGTGTTCTGATAAAGGAGTTGCACCAGGTACAGCTGTTTTAAGACAGTGGTTGCCTAC
GGGTACGCTGCTTGTGATTAGATCTTAATGACTTTGTCTCTGATGCAGATTCAACTTTGATTGGTGAT
TGTGCAACTGTACATACAGCTAATAAATGGGATCTCATTATTAGTGATATGTACGACCCTAAGACTAAAA
ATGTTACAAAAGAAAATGACTCTAAAGAGGGTTTTTCACTTACATTTGTGGGTTTATACAACAAAAGCT
AGCTCTTGGAGGTTCCGTGGCTATAAAGATAACAGAACATTCTTGGAATGCTGATCTTTATAAGCTCATG
GGACACTTCGCATGGTGGACAGCCTTTGTTACTAATGTGAATGCGTCATCATCTGAAGCATTTTTAATTG
GATGTAATTATCTTGGCAAACACGCGAACAATAGATGGTTATGTCATGCATGCAAATTACATATTTTG
GAGGAATACAAATCCAATTCAGTTGTCTTCCTATTCTTTATTTGACATGAGTAAATTTCCCTTAAATTA
AGGGGTACTGCTGTTATGTCTTTAAAGAAGGTCAAATCAATGATATGATTTTATCTCTTCTTAGTAAAG
GTAGACTTATAATTAGAGAAAACAACAGAGTTGTTATTTCTAGTGATGTTCTTGTTAACAACATAACGAA
CAATGTTTGTGTTTTCTTGTGTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCA
ATTACCCCTGCATACACTAATTCTTTCACACGTGGTGTGTTATTACCCTGACAAAGTTTTTCCAGATCCTCA
GTTTATACATCAACTCAGGACTTGTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATACATG
TCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCCTGCTTACCATTAAATGATGGTGTGTTATTTTGC
TTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCTGAAGACCCAGTCC
CTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGAATGATCCAT
TTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGC
GAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTTC
AAAAATCTTAGGGAATTTGTGTTTAAAGAATATTGATGGTTATTTTAAATATATTCTAAGCACACGCCTA
TTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTTCGGCTTTAGAACCATTTGGTAGATTTGCCAATAGGTAT
TAACATCACTAGGTTTCAAACCTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCA
GGTTGGACAGCTGGTGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTAAATATA
ATGAAAATGGAACCATTAACAGATGCTGTAGACTGTGCACCTTGACCCTCTCTCAGAAACAAAGTGACGTT
GAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACCTTCTAAGTTTGAAGTCCAAACCAACAGAATCTATT
GTTAGATTTTCTAATATTACAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCTG
TTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCTTATATAATTCCGCATC
ATTTTCCACTTTTAAAGTGTTATGGAGTGTCTCCTACTAAATTAATGATCTCTGCTTTACTAATGTCTAT
GCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTG
ATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAAATCTAACAATCTTGATTCT
TAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAAGAG
GATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTGGAAGGTTTTAATTGTTACT
TTCTTTTACAATCATATGTTTTCCAACCCACTAATGGTGTGTTGTTACCAACCATAACAGATAGTAGTACT
TTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAAC
AAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTC
TGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGA
GATTCTTGACATTACACCATGTTCTTTTGGTGGTGTGAGTGTATAACACCAGGAACAAATACTTCTAAC
CAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTTA
CTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTAAATAGGGGC
TGAACATGTCAACAACCTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACT
CAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACCTTG
GTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTAGTGTTAC

CACAGAAATTCTACCAAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCA
ACTGAATGCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTAACCGTGCTTTAACTGGAA
TAGCTGTTGAACAAGACAAAAACCCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACC
AATTAAAGATTTTGGTGGTTTTAATTTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCA
TTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAAACAATATGGTGATT
GCCTTGGTGATATTGCTGCTAGAGACCTCATTTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACC
TTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCTGGTTGG
ACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAAATGGTATTG
GAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAA
AATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACCTCAAGATGTGGTCAACCAAAATGCA
CAAGCTTTAAACACGCTTGTTTAAACAACCTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTAAATGATA
TCCTTTACCGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTTGATAGGTTGATCACAGGCAGACTTCAAAG
TTTGACAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTGCTGCT
ACTAAAATGTGAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTTGTGGAAAGGGCTATCATCTTA
TGTCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAA
GAACTTCACAACCTGCTCCTGCCATTTGTCTATGATGGAAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTT
TCAAATGGCACACACTGGTTTGTAAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACA
CATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACC
TGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTTAAAGATCATACATCACCAGATGTTGATTTA
GGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTG
CCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACCTTGAAAGTATGAGCAGTATATAAAATGGCC
ATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGACAATTATGCTTTGCTGT
ATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAAGACG
ACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAAACGAACCTTATGGATTTGTTTATGAGA
ATCTTCACAATTGGAACCTGTAACCTTTGAAGCAAGGTGAAATCAAGGATGCTACTCCTTCAGATTTTGTTC
GCGCTACTGCAACGATACCGATACAAGCCTCACTCCCTTTCGGATGGCTTATTGTTGGCGTTGCACTTCT
TGCTGTTTTTTCAGAGCGCTTCCAAAATCATAACCCTCAAAAAGAGATGGCAACTAGCACTCTCCAAGGGT
GTTCACTTTGTTTGAACCTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
GCCTTGAAGCCCCCTTTCTCTATCTTTATGCTTTTAGTCTACTTCTTGCAGAGTATAAACTTTGTAAGAAT
AATAATGAGGCTTTGGCTTTGCTGGAAATGCCGTTCCAAAACCCATTACTTTATGATGCCAATTTTTT
CTTTGCTGGCATACTAATTGTTACGACTATTGTATACCTTACAATAGTGTAACCTTCTTCAATTGTCATTA
CTTCAGGTGATGGCACAACAAGTCCTATTTCTGAACATGACTACCAGATTGGTGGTTATACTGAAAAATG
GGAATCTGGAGTAAAGACTGTGTTGTATTACACAGTTACTTCACCTCAGACTATTACCAGCTGTACTCA
ACTCAATTGAGTACAGACACTGGTGTGTAACATGTTACCTTCTTCATCTACAATAAAATTGTTGATGAGC
CTGAAGAACATGTCCAAATTCACACAATCGACGGTTCATCCGGAGTTGTTAATCCAGTAATGGAACCAAT
TTATGATGAACCGACGACGACTACTAGCGTGCCTTTGTAAGCACAAAGCTGATGAGTACGAACCTTATGTAC
TCATTGTTTTCGGAAGAGACAGGTACGTTAATAGTTAATAGCGTACTTCTTTTTCTTGCTTTCTGTTGAT
TCTTGCTAGTTACACTAGCCATCCTTACTGCGCTTCGATTGTGTGCGTACTGCTGCAATATTGTTAACGT
GAGTCTTGTAACCTTCTTTTTACGTTTACTCTCGTGTTAAAAATCTGAATTCTTCTAGAGTTCCTGAT
CTTCTGGTCTAAACGAACATAAATATTATATTAGTTTTTCTGTTTGGAACCTTAAATTTTAGCCATGGCAGA
TTCCAACGGTACTATTACCGTTGAAGAGCTTAAAAAGCTCCTTGAACAATGGAACCTAGTAATAGGTTTC
CTATTCCTTACATGGATTTGTCTTCTACAATTTGCCTATGCCAACAGGAATAGGTTTTTGTATATAATTA
AGTTAATTTTCTCTGGCTGTTATGGCCAGTAACTTTAGCTTGTTTTGTGCTTGCTGCTGTTTACAGAAT
AAATTGGATCACCGGTGGAATTGCTATCGCAATGGCTTGCTTGTAGGCTTGATGTGGCTCAGCTACTTC
ATTGCTTCTTTCAGACTGTTTGCAGCTACGCGTTCATGTGGTCATTCAATCCAGAACTAACATTCTTC
TCAACGTGCCACTCCATGGCACTATTCTGACCAGACCGCTTCTAGAAAGTGAACCTCGTAATCGGAGCTGT
GATCCTTTCGTGGACATCTTCGTATTGCTGGACACCATCTAGGACGCTGTGACATCAAGGACCTGCCTAAA
GAAATCACTGTTGCTACATACGAACGCTTCTTATTACAAATTTGGGAGCTTCGACGCTGTAGCAGGTG
ACTCAGGTTTTGCTGCATACAGTCGTACAGGATTGGCAACTATAAATTAACACAGACCATTCCAGTAG
CAGTGACAATATTGCTTTGCTTGTACAGTAAGTGACAACAGATGTTTCATCTCGTTGACTTTTCAGGTTAC
TATAGCAGAGATATTACTAATTATTATGAGGACTTTTAAAGTTTCCATTTGGAATCTTGATTACATCATA
AACCTCATAATTAATAAATTTATCTAAGTCACTAACTGAGAATAAATATTCTCAATTAGATGAAGAGCAAC
CAATGGAGATTGATTAAACGAACATGAAAATTATTCTTTTCTTGGCACTGATAACACTCGCTACTTGTGA
GCTTTATCACTACCAAGAGTGTGTTAGAGGTACAACAGTACTTTTAAAAGAACCTTGCTCTTCTGGAACA
TACGAGGGCAATTACCATTTTCATCTCTAGCTGATAACAAATTTGCACTGACTTGCTTTAGCACTCAAT
TTGCTTTTGTCTTGTCTGACGGCGTAAAACACGTCTATCAGTTACGTGCCAGATCAGTTTTCACCTAACT
GTTTCATCAGACAAGAGGAAGTTCAAGAACCTTTACTCTCCAATTTTTCTTATTGTTGCGGCAATAGTGTTT

ATAACACTTTGCTTCACACTCAAAAGAAAGACAGAATGATTGAACTTTCATTAATTGACTTCTATTTGTG
CTTTTTAGCCTTTCTGCTATTCCTTGTTTTAATTATGCTTATTATCTTTTGTTTCTCACTTGAAGTGCAA
GATCATAATGAACTTGTGACGCCTAAACGAACATGAAATTTCTTGTTTTCTTAGGAATCATCACAAGTG
TAGCTGCATTTACCAAGAATGTAGTTTACAGTCATGTACTCAACATCAACCATATGTAGTTGATGACCC
GTGTCCTATTCACCTTCTATTCTAAATGGTATATTAGAGTAGGAGCTAGAAAATCAGCACCTTTAATTGAA
TTGTGCGTGGATGAGGCTGGTTCTAAATCACCCATTTCAGTACATCGATATCGGTAATTATACAGTTTCCT
GTTTACCTTTTACAATTAATTGCCAGGAACCTAAATTGGGTAGTCTTGTAGTGCGTTGTTTCGTTCTATGA
AGACTTTTTAGAGTATCATGACGTTTCGTGTTGTTTTAGATTTTCATCTAAACGAACAAACTAAAATGTCTG
ATAATGGACCCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCTCAGATTCAACTGGCAG
TAACCAGAATGGAGAACGCAGTGGGGCGCGATCAAAACAACGTCGGCCCCAAGGTTTACCCAATAATACT
GCGTCTTGGTTCACCGCTCTCACTCAACATGGCAAGGAAGACCTTAAATTCCCTCGAGGACAAGGCGTTC
CAATTAACACCAATAGCAGTCCAGATGACCAAAATTGGCTACTACCGAAGAGCTACCAGACGAATTTCGTGG
TGGTGACGGTAAAATGAAAGATCTCAGTCCAAGATGGTATTTCTACTACCTAGGAACTGGGCCAGAAGCT
GGACTTCCCTATGGTGCTAACAAGACGGCATCATATGGGTTGCAACTGAGGGAGCCTTGAATACACCAA
AAGATCACATTGGCACCCGCAATCCTGCTAACAATGCTGCAATCGTGCTACAACCTTCTCAAGGAACAAC
ATTGCCAAAAGGCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTTCCTCATCACGT
AGTCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACTTCTCCTGCTAGAATGGCTGGCA
ATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCTTGAGAGCAAAATGTCTGG
TAAAGGCCAACAACAACAAGGCCAAACTGTCTACTAAGAAATCTGCTGCTGAGGCTTCTAAGAAGCCTCGG
CAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAGACGTGGTCCAGAACAACCC
AAGGAAATTTTGGGGACCAGGAATAATCAGACAAGGAAGTATTACAAACATTGGCCGCAAATTGCACA
ATTTGCCCCCAGCGCTTCAGCGTTCTTCGGAATGTGCGCATTGGCATGGAAGTCACACCTTCGGGAACG
TGGTTGACCTACACAGGTGCCATCAAATTGGATGACAAAGATCCAAATTTCAAAGATCAAGTCATTTTGC
TGAATAAGCATATTGACGCATACAAAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAGGC
TGATGAACTCAAGCCTTACCGCAGAGACAGAAGAAACAGCAAAGTGTGACTCTTCTTCCTGCTGCAGAT
TTGGATGATTTCTCCAAACAATTGCAACAATCCATGAGCAGTGCTGACTCAACTCAGGCCTAAACTCATG
CAGACCACACAAGGCAGATGGGCTATATAAACGTTTTTCGCTTTTCCGTTTACGATATATAGTCTACTCTT
GTGCAGAATGAATTCTCGTAACTACATAGCACAAAGTAGATGTAGTTAACTTTAATCTCACATAGCAATCT
TTAATCAGTGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTTTACCAGAGGCCACGCGGAGTAC
GATCGAGTGACAGTGAACAATGCTAGGGAGAGCTGCCTATATGGAAGAGCCCTAATGTGTAAAATTAAT
TTTAGTAGTGCTATCCCCATGTGATTTTAATAGCTTCTTAGGAGAATGACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA