

# Homework #5

Due April 13<sup>th</sup>, 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 readme.txt file outlining all modules (if any) needed for the execution of the code and the exact command lines needed to answer homework's questions **(5pts)**
- A full write-up (.pdf or .doc) file containing answers to homework's questions **(5pts)** – screenshots of code output are ok.

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, time.h, random.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).
  - 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:

- 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 (Hamming distance, 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. Generate 5K, 50K, and 100K 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.
- What are the sizes of the trie (# of nodes)? Does it make sense to you? Explain why.
  - 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. Generate 1K, 50K, and 100K 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.
- What are the sizes of the trie (# of nodes)? Does it make sense to you? Explain why.
  - 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  
GCTCGAACTGCACCTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAACTCGAAGGCATTACGTACGGTC  
GTAGTGGTGAGACACTTGGTGTCTTGTCCCTCATGTGGGCGAAATACCAGTGGCTTACCGCAAGGTTCT  
TCTTCGTAAGAACGGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTA  
GGCGACGAGCTTGGCACTGATCCTTATGAAGATTTTTCAAGAAAACCTGGAACACTAAACATAGCAGTGGTG  
TTACCCGTGAACTCATGCGTGAGCTTAACGGAGGGGCATACACTCGCTATGTCGATAACAACCTTCTGTGG  
CCCTGATGGCTACCCTCTTGAGTGCATTAAAGACCTTCTAGCACGTGCTGGTAAAGCTTCATGCACTTTG  
TCCGAACAACCTGGACTTTATTGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTG  
CTTGGTACACGGAACGTTCTGAAAAGAGCTATGAATTGCAGACACCTTTTGAAATTAAATTGGCAAAGAA  
ATTTGACACCTTCAATGGGGAATGTCCAAATTTTGTATTTCCCTTAAATTCCATAATCAAGACTATTCAA  
CCAAGGGTTGAAAAGAAAAAGCTTGATGGCTTTATGGGTAGAATTCGATCTGTCTATCCAGTTGCGTCAC  
CAAATGAATGCAACCAATGTGCCTTTCAACTCTCATGAAGTGTGATCATTGTGGTGAAACTTCATGGCA  
GACGGGCGATTTTGTAAAGCCACTTGCGAATTTTGTGGCACTGAGAATTTGACTAAAGAAGGTGCCACT  
ACTTGTGGTTACTTACCCCAAAATGCTGTTGTTAAAATTTATTGTCCAGCATGTCACAATTCAGAAGTAG  
GACCTGAGCATAGTCTTGCCGAATACCATAATGAATCTGGCTTGAAAACCATTTCTTCGTAAGGGTGGTCG  
CACTATTGCCTTTGGAGGCTGTGTGTTCTCTTATGTTGGTTGCCATAACAAGTGTGCCTATTGGGTTCCA  
CGTGCTAGCGCTAACATAGGTTGTAACCATAACAGGTGTTGTTGGAGAAGGTTCCGAAGGTCTTAATGACA  
ACCTTCTTGAAATACTCCAAAAAGAGAAAGTCAACATCAATATTGTTGGTGACTTTAACTTAATGAAGA  
GATCGCCATTATTTTGGCATCTTTTTCTGCTTCCACAAGTGCTTTTGTGGAACTGTGAAAGGTTTGGAT  
TATAAAGCATTAACAATTTGTTGAATCCTGTGGTAATTTTAAAGTTACAAAAGGAAAAGCTAAAAAG  
GTGCCTGGAATATTGGTGAACAGAAATCAATAGTACTGAGTCTCTTTATGCATTTGCATCAGAGGCTGCTCG  
TGTTGTACGATCAATTTTCTCCCGCACTCTTGAAACTGCTCAAAATTCTGTGCGTGTTTTACAGAAGGCC  
GCTATAACAATACTAGATGGAATTTACAGTATTCACTGAGACTCATTGATGCTATGATGTTTACATCTG  
ATTTGGCTACTAACAATCTAGTTGTAATGGCCTACATTACAGGTGGTGTGTTTCAAGTTGACTTCGCAGTG  
GCTAACTAACATCTTTGGCACTGTTTATGAAAACTCAAACCCGTCCTTGATTGGCTTGAAGAGAAGTTT  
AAGGAAGGTGTAGAGTTTCTTAGAGACGGTTGGGAAATTGTTAAATTTATCTCAACCTGTGCTTGTGAAA  
TTGTCGGTGGACAAATTGTCACCTGTGCAAAGGAAATTAAGGAGAGTGTTTCAAGACATTCTTTAAGCTTGT  
AAATAAATTTTGGCTTTGTGTGCTGACTCTATCATTATTGGTGGAGCTAACTTAAAGCCTTGAATTTA  
GGTGAACATTTGTACGCACTCAAAGGGATTGTACAGAAAGTGTGTTAAATCCAGAGAAGAACTGGCC  
TACTCATGCCTCTAAAAGCCCCAAAAGAAATTATCTTCTTAGAGGGAGAAACACTTCCCACAGAAGTGTT  
AACAGAGGAAGTTGTCTTGAAAACCTGGTGATTTACAACCATTAGAACAACCTACTAGTGAAGCTGTTGAA  
GCTCCATTGGTTGGTACACCAAGTTTGTATTAACGGGCTTATGTTGCTCGAAATCAAAGACACAGAAAAGT  
ACTGTGCCCTTGCACCTAATATGATGGTAACAAACAATACCTTCACACTCAAAGGCGGTGCACCAACAAA  
GGTTACTTTTGGTGATGACACTGTGATAGAAGTGCAAGGTTACAAGAGTGTAATATCACTTTTGAACCTT  
GATGAAAGGATTGATAAAGTACTTAATGAGAAGTGCTCTGCCTATACAGTTGAACTCGGTACAGAAGTAA  
ATGAGTTTCGCTGTGTTGTGGCAGATGCTGTCTATAAAAACTTTGCAACCAGTATCTGAATTACTTACACC  
ACTGGGCATTGATTTAGATGAGTGGAGTATGGCTACATACTACTTATTTGATGAGTCTGGTGAGTTTAA  
TTGGCTTACATATGTATTGTTCTTTCTACCTCCAGATGAGGATGAAGAAGAAGGTGATTGTGAAGAAG  
AAGAGTTTGAAGCATCAACTCAATATGAGTATGGTACTGAAGATGATTACCAAGGTAAACCTTTGGAATT  
TGGTGCCACTTCTGCTGCTCTTCAACCTGAAGAAGAGCAAGAAGAAGATTGGTTAGATGATGATAGTCAA  
CAAACCTGTTGGTCAACAAGACGGCAGTGAGGACAATCAGACAACCTACTATTCAAACAATTGTTGAGGTTT  
AACCTCAATTAGAGATGGAACCTTACACCAAGTTGTTTCAAGTATTGAAGTGAATAGTTTTAGTGGTTATTT  
AAAACCTTACTGACAATGTATACATTAAAAATGCAGACATTGTGGAAGAAGCTAAAAAGGTAAACCAACA  
GTGGTTGTTAATGCAGCCAATGTTTACCTTAAACATGGAGGAGGTGTTGCAGGAGCCTTAAATAAGGCTA  
CTAACAATGCCATGCAAGTTGAATCTGATGATTACATAGCTACTAATGGACCACTTAAAGTGGGTGGTAG  
TTGTGTTTTAAGCGGACACAATCTTGCTAAACACTGTCTTCATGTTGTGCGCCCAAATGTTAACAAGGT  
GAAGACATTCAACTTCTTAAGAGTGCTTATGAAAATTTTAAATCAGCACGAAGTTCTACTTGCACCATAT  
TATCAGCTGGTATTTTTGGTGCTGACCCTATACATTCTTTAAGAGTTTGTGTAGATACTGTTTCGACAAA

[illegible]

TCTTTACTTTGATAAAGCTGGTCAAAAGACTTATGAAAGACATTCTCTCTCTCATTGTTAACTTAGAC  
AACCTGAGAGCTAATAACACTAAAGGTTTCATTGCCTATTAATGTTATAGTTTTTGTATGGTAAATCAAAAT  
GTGAAGAATCATCTGCAAAATCAGCGTCTGTTTACTACAGTCAGCTTATGTGTCAACCTATACTGTTACT  
AGATCAGGCATTAGTGTCTGATGTTGGTGATAGTGCAGGAAAGTTGCAGTTAAAATGTTTGATGCTTACGTT  
AATACGTTTTTCATCAACTTTTAAACGTACCAATGGAAAACTCAAAACACTAGTTGCAACTGCAGAAGCTG  
AACTTGCAAAGAATGTGTCCTTAGACAATGTCTTATCTACTTTTTATTTTCAGCAGCTCGGCAAGGGTTTTGT  
TGATTGAGATGTAGAACTAAAGATGTTGTTGAATGTCTTAAATTGTCACATCAATCTGACATAGAAGTT  
ACTGGCGATAGTTGTAATAACTATATGCTCACCTATAACAAAGTTGAAAACATGACACCCCGTGACCTTG  
GTGCTTGTATTGACTGTAGTGCGCGTCATATTAATGCGCAGGTAGCAAAAAGTCACAACATTGCTTTGAT  
ATGGAACGTTAAAGATTTTCATGTCAATTGTCTGAACAACTACGAAAACAAATACGTAGTGCTGTCAAAAAG  
AATAACTTACCTTTTTAAGTTGACATGTGCAACTACTAGACAAGTTGTTAATGTTGTAACAACAAGATAG  
CACTTAAGGGTGGTAAAATTGTTAATAATTGGTTGAAGCAGTTAATTAAGTTACACTTGTGTTCCTTTTT  
TGTTGCTGCTATTTTTCTATTTAATAACACCTGTTTCATGTCTGCTAAACATACTGACTTTTTCAAGTGAA  
ATCATAGGATACAAGGCTATTGATGGTGGTGTCACTCGTGACATAGCATCTACAGATACTTGTTTTTGCTA  
ACAAACATGCTGATTTTTGACACATGGTTTAGCCAGCGTGGTGGTAGTTATACTAATGACAAAGCTTGCCC  
ATTGATTGCTGCAGTCATAACAAGAGAAGTGGGTTTTGTGCTGCCTGGTTTTGCCTGGCAGCATATTACGC  
ACAATAATGGTGACTTTTTGCATTTCTTACCTAGAGTTTTTAGTGAGTTGGTAACATCTGTTACACAC  
CATCAAACTTATAGAGTACACTGACTTTGCAACATCAGCTTGTGTTTTGGCTGCTGAATGTACAATTTT  
TAAAGATGCTTCTGGTAAGCCAGTACCATATTGTTATGATAACCAATGTACTAGAAGGTTCTGTTGCTTAT  
GAAAGTTTACGCCCTGACACACGTTATGTGCTCATGGATGGCTCTATTATTCAATTTCCCTAACACCTACC  
TTGAAGGTTCTGTTAGAGTGGTAACAACCTTTTGATTCTGAGTACTGTAGGCACGGCACTTGTGAAAGATC  
AGAAGCTGGTGGTTGTGTATCTACTAGTGGTAGATGGGTACTTAACAATGATTATTACAGATCTTTACCA  
GGAGTTTTCTGTGGTGTAGATGCTGTAAATTTACTTACTAATATGTTTACACCACTAATTCAACCTATTG  
GTGCTTTGGACATATCAGCATCTATAGTAGCTGGTGGTATTGTAGCTATCGTAGTAACATGCCTTGCCCTA  
CTATTTTATGAGGTTTAGAAGAGCTTTTGGTGAATACAGTCATGTAGTTGCCTTTAATACTTTACTATTC  
CTTATGTCATTCACTGTACTCTGTTTAAACACAGTTTACTCATTCTTACCTGGTGGTTTATTCTGTTATTT  
ACTTGTACTTGACATTTTATCTTACTAATGATGTTTCTTTTTTAGCACATATTCAGTGGATGGTTATGTT  
CACACCTTTAGTACCTTTCTGGATAACAATTGCTTATATCATTGTTATTTCCACAAAGCATTCTATTGG  
TTCTTTAGTAATTAACCTTAAAGAGACGTGTAGTCTTTAATGGTGGTTTCTTTTAGTACTTTTGAAGAAGTG  
CGCTGTGCACCTTTTTGTTAAATAAAGAAATGTATCTAAAGTTGCGTAGTGATGTGCTATTACCTCTTAC  
GCAATATAATAGATACTTAGCTCTTTATAATAAGTACAAGTATTTTAGTGGAGCAATGGATACAACCTAGC  
TACAGAGAAGCTGCTTGTGTGTCATCTCGCAAAGGCTCTCAATGACTTCAGTAACCTCAGGTTCTGATGTTT  
TTTACCAACCACCACAAACCTCTATCACCTCAGCTGTTTTGCAGAGTGGTTTTAGAAAAATGGCATTCCC  
ATCTGGTAAAGTTGAGGGTTGTATGGTACAAGTAACCTGTGGTACAACCTACACTTAACGGTCTTTGGCTT  
GATGACGTAGTTTACTGTCCAAGACATGTGATCTGCACCTCTGAAGACATGCTTAACCTAATTATGAAG  
ATTTACTCATTTCGTAAGTCTAATCATAATTTCTTGGTACAGGCTGGTAATGTTCAACTCAGGGTTATTGG  
ACATTCTATGCAAAATTGTGTACTTAAGCTTAAGGTTGATACAGCCAATCCTAAGACACCTAAGTATAAG  
TTTGTTTCGCAATTAACAGGACAGACTTTTTTCACTGTTAGCTTGTACAATGGTTTACCCTCTGGTGGTT  
ACCAATGTGCTATGAGGCCCAATTTCACTATTAAGGGTTTCACTTCTTAATGGTTTCACTGTTGGTGGT  
TTTTAACATAGATTATGACTGTGTCTCTTTTTGTTACATGCACCATATGGAATTACCAACTGGAGTTTCA  
GCTGGCACAGACTTAGAAGGTAACCTTTTATGGACCTTTTGTGACAGGCAAACAGCACAAAGCAGCTGGTA  
CGGACACAACCTATTACAGTTAATGTTTTAGCTTGGTTGTACGCTGCTGTTATAAATGGAGACAGGTGGTT  
TCTCAATCGATTTACCACAACCTCTAATGACTTTAACCTTGTGGCTATGAAGTACAATTATGAACCTCTA  
ACACAAGACCATGTTGACATACTAGGACCTCTTTCTGCTCAAACCTGGAATTGCCGTTTTAGATATGTGTG  
CTTCATTAAAAGAATTACTGCAAAATGGTATGAATGGACGTACCATATTGGGTAGTGCTTTATTAGAAGA  
TGAATTTACACCTTTTGATGTTGTTAGACAATGCTCAGGTGTTACTTTCCAAAGTGCAGTGAAAAGAACA  
ATCAAGGGTACACACCCTGGTTGTTACTCACAATTTTGAACCTTCACTTTTAGTTTTAGTCCAGAGTACTC  
AATGGTCTTTGTTCTTTTTTTTTGTATGAAAATGCCTTTTTACCTTTTGCTATGGGTATTATTGCTATGTC  
TGCTTTTGCAATGATGTTTGTCAAACATAAGCATGCATTTCTCTGTTTGTGTTTTGTTACCTTCTCTTGCC  
ACTGTAGCTTATTTAATATGGTCTATATGCCTGCTAGTTGGGTGATGCGTATTATGACATGGTTGGATA  
TGTTTGATACTAGTTTGTCTGGTTTTAAGCTAAAAGACTGTGTTATGTATGCATCAGCTGTAGTGTTACT  
AATCCTTATGACAGCAAGAAGTGTGTATGATGATGGTGCTAGGAGAGTGTGGACACTTATGAATGTCTTG  
ACACTCGTTTTATAAAGTTTATTATGGTAATGCTTTAGATCAAGCCATTTCCATGTGGGCTCTTATAATCT  
CTGTTACTTCTAACTACTCAGGTGTAGTTACAACGTGCATGTTTTTGGCCAGAGGTATTGTTTTTATGTG  
TGTTGAGTATTGCCCTATTTTCTTCATAACTGGTAATACACTTCAGTGTATAATGCTAGTTTTATTGTTTC  
TTAGGCTATTTTTGTACTTGTACTTTGGCCTCTTTTGTGTTACTCAACCGCTACTTTAGACTGACTCTTG  
GTGTTTATGATTACTTAGTTTCTACACAGGAGTTTAGATATATGAATTCACAGGGACTACTCCCACCCAA

GAATAGCATAGATGCCTTCAAACCTCAACATTAATTTGTTGGGTGTTGGTGGCAAACCTTGTATCAAAGTA  
GCCACTGTACAGTCTAAAATGTGAGATGTAAAGTGCACATCAGTAGTCTTACTCTCAGTTTTTGCAACAAC  
TCAGAGTAGAATCATCATCTAAATTTGTGGGCTCAATGTGTCCAGTTACACAATGACATTCTCTTAGCTAA  
AGATACTACTGAAGCCTTTGAAAAATGGTTTTCACTACTTTCTGTTTTGCTTTCCATGCAGGGTGCTGTA  
GACATAAACAAGCTTTGTGAAGAAATGCTGGACAACAGGGCAACCTTACAAGCTATAGCCTCAGAGTTTA  
GTTCCCTTCCATCATATGCAGCTTTTGTACTGCTCAAGAAGCTTATGAGCAGGCTGTTGCTAATGGTGA  
TTCTGAAGTTGTTCTTAAAAAGTTGAAGAAGTCTTTGAATGTGGCTAAATCTGAATTTGACCGTGATGCA  
GCCATGCAACGTAAGTTGGAAAAGATGGCTGATCAAGCTATGACCCAAATGTATAAACAGGCTAGATCTG  
AGGACAAGAGGGCAAAAGTTACTAGTGCTATGCAGACAATGCTTTTCACTATGCTTAGAAAAGTTGGATAA  
TGATGCACTCAACAACATTATCAACAATGCAAGAGATGGTTGTGTTCCCTTGAACATAATACCTCTTACA  
ACAGCAGCCAAACTAATGGTTGTGCATACCAGACTATAACACATATAAAAAATACGTGTGATGGTACAACAT  
TTACTTATGCATCAGCATTGTGGGAAATCCAACAGGTTGTAGATGCAGATAGTAAAATTGTTCAACTTAG  
TGAAATTAGTATGGACAATTCACCTAATTTAGCATGGCCTCTTATTGTAACAGCTTTAAGGGCCAATTCT  
GCTGTCAAATTACAGAATAATGAGCTTAGTCCTGTTGCACTACGACAGATGTCTTGTGCTGCCGGTACTA  
CACAACTGCTTGCCTGATGACAATGCGTTAGCTTACTACAACACAACAAAGGGAGGTAGGTTTGTACT  
TGCCTGTTATCCGATTTACAGGATTTGAAATGGGCTAGATTCCCTAAGAGTGATGGAAGTGGTACTATC  
TATACAGAAGTGAACACCTTGTAGGTTTGTACAGACACACCTAAAGGTCCTAAAGTGAAGTATTTAT  
ACTTTATTAAGGATTAAACAACCTAAATAGAGGTATGGTACTTGGTAGTTTAGCTGCCACAGTACGTCT  
ACAAGCTGGTAATGCAACAGAAGTGCCTGCCAATTCACTGTATTATCTTTCTGTGCTTTTGTGCTAGAT  
GCTGCTAAAGCTTACAAAGATTATCTAGCTAGTGGGGGACAACCAATCACTAATTGTGTTAAGATGTTGT  
GTACACACACTGGTACTGGTCAGGCAATAACAGTTACACCGGAAGCCAATATGGATCAAGAATCCTTTGG  
TGGTGCATCGTGTTGTCTGTACTGCCGTTGCCACATAGATCATCAAATCCTAAAGGATTTTGTGACTTA  
AAAGGTAAGTATGTACAAATACCTACAACCTTGTGCTAATGACCCTGTGGGTTTTACACTTAAAAACACAG  
TCTGTACCGTCTGCGGTATGTGGAAAGGTTATGGCTGTAGTTGTGATCAACTCCGCGAAGCCATGCTTCA  
GTCAGCTGATGCACAATCGTTTTTAAACGGGTTTGCCTGTAGTGAAGTGCAGCCCGTCTTACACCGTGC  
CAGGCACTAGTACTGATGTCGTATACAGGGCTTTTGACATCTACAATGATAAAGTAGCTGGTTTTGCTAA  
ATTCCTAAAACTAATTGTTGTGCTTCCAAGAAAAGGACGAAGATGACAATTTAATTGATTCTTACTTT  
GTAGTTAAGAGACACACTTTCTCTAACTACCAACTGAAGAAACAATTTATAATTTACTTAAAGGATTGTC  
CAGCTGTTGCTAAACATGACTTCTTTAAGTTTGAAGATAGACGGTGACATGGTACCACATATATACGTC  
ACGTCTTACTAAATACACAATGGCAGACCTCGTCTATGCTTTAAGGCATTTTGTGATGAAGTAATTTGTGAC  
ACATTAAGAAAATACTTGTACATACAATTGTTGTGATGATGATTATTTCAATAAAAAGGACTGGTATG  
ATTTTGTAGAAAACCCAGATATATTACGCGTATACGCCAAGTTAGGTGAACGTGTACGCCAAGCTTTGTT  
AAAAACAGTACAATTCTGTGATGCCATGCGAAATGCTGGTATTGTTGGTGTACTGACATTAGATAATCAA  
GATCTCAATGGTAAGTGGTATGATTTGCGGTGATTTTACATAACACACGCCAGGTAGTGGAGTTTCTGTTG  
TAGATTCTTATTATTATTGTTAATGCCTATATTAACCTTGACCAGGGCTTTAACTGCAGAGTCACATGT  
TGACACTGACTTAAACAAGCCTTACATTAAGTGGGATTTGTTAAATATGACTTCACGGAAGAGAGGTTA  
AAACTCTTTGACCGTTATTTTAAATATTGGGATCAGACATACCACCCAAATTGTGTTAACTGTTTGGATG  
ACAGATGCATTCTGCATTGTGCAAACCTTTAATGTTTTATTCTCTACAGTGTTCCACCTACAAGTTTTGG  
ACCACTAGTGAGAAAAATATTTGTTGATGGTGTTCATTTGTAGTTTCACTGGATACCACTTCAGAGAG  
CTAGGTGTTGTACATAATCAGGATGTAACTTACATAGCTCTAGACTTAGTTTTAAGGAATTACTTGTGT  
ATGCTGCTGACCCTGCTATGCACGCTGCTTCTGGTAATCTATTACTAGATAAACGCACTACGTGCTTTTC  
AGTAGCTGCACTTACTAACAATGTTGCTTTTCAAACCTGTCAAACCCGGTAATTTTAAACAAGACTTCTAT  
GACTTTGCTGTGTCTAAGGGTTTCTTTAAGGAAGGAAGTTCTGTTGAATTAACAACACTTCTTCTTTGCTC  
AGGATGGTAATGCTGCTATCAGCGATTATGACTACTATCGTTATAATCTACCAACAATGTGTGATATCAG  
ACAACACTACTATTTGTAGTTGAAGTTGTTGATAAGTACTTTGATTGTTACGATGGTGGCTGTATTAATGCT  
AACCAAGTCATCGTCAACAACCTAGACAAATCAGCTGGTTTTCCATTTAATAAATGGGGTAAGGCTAGAC  
TTTATTATGATTCAATGAGTTATGAGGATCAAGATGCACCTTTTCGCATATACAAAACGTAATGTCTATCCC  
TACTATAACTCAAATGAATCTTAAGTATGCCATTAGTGCAAGAATAGAGCTCGCACCGTAGCTGGTGTCT  
TCTATCTGTAGTACTATGACCAATAGACAGTTTCATCAAAAATTATTGAAATCAATAGCCGCCACTAGAG  
GAGCTACTGTAGTAATTGGAACAAGCAAATCTATGGTGGTTGGCACAACATGTTAAAACTGTTTATAG  
TGATGTAGAAAACCTCACCTTATGGGTTGGGATTATCCTAAATGTGATAGAGCCATGCCTAACATGCTT  
AGAATTATGGCCTCACTTGTCTTGTCTCGCAAACATACAACGTGTTGTAGCTTGTACACCGTTTCTATA  
GATTAGCTAATGAGTGTGCTCAAGTATTGAGTGAAATGGTCATGTGTGGCGGTTCACTATATGTTAAACC  
AGGTGGAACCTCATCAGGAGATGCCACAACCTGCTTATGCTAATAGTGTTTTTAACATTTGTCAAGCTGTC  
ACGGCCAATGTTAATGCACCTTTTATCTACTGATGGTAACAAAATTGCCGATAAGTATGTCCGCAATTTAC  
AACACAGACTTTTATGAGTGTCTCTATAGAAATAGAGATGTTGACACAGACTTTGTGAATGAGTTTTACGC  
ATATTTGCGTAACATTTCTCAATGATGATACTCTCTGACGATGCTGTTGTGTGTTTCAATAGCACTTAT

GCATCTCAAGGTCTAGTGGCTAGCATAAAGAACTTTAAGTCAGTTCTTTATTATCAAAACAATGTTTTTA  
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  
TCAATTCAAGTGCTAGACTTATGAAAACCTATAGGTCCAGACATGTTCCCTCGGAACCTTGTCGGCGTTGTCC  
TGCTGAAATTGTTGACACTGTGAGTGCTTTGGTTTTATGATAATAAGCTTAAAGCACATAAAGACAAATCA  
GCTCAATGCTTTAAATGTTTTATAAGGGTGTTATCACGCATGATGTTTCATCTGCAATTAACAGGCCAC  
AAATAGGCGTGGTAAGAAATTCCTTACACGTAACCCCTGCTTGGAGAAAAGCTGTCTTTATTTACCTTTA  
TAATTCACAGAATGCTGTAGCCTCAAAGATTTTGGGACTACCAACTCAAAGTGTGATTATCACAGGGC  
TCAGAATATGACTATGTCATATTCACTCAAACCACTGAAACAGCTCACTCTTGTAATGTAAACAGATTTA  
ATGTTGCTATTACCAGAGCAAAAGTAGGCATACTTTGCATAATGTCTGATAGAGACCTTTATGACAAGTT  
GCAATTTACAAGTCTTGAAATTCACGTAGGAATGTGGCAACTTTACAAGCTGAAAATGTAACAGGACTC  
TTTAAAGATTGTAGTAAGGTAATCACTGGGTACATCCTACACAGGCACCTACACACCTCAGTGTTGACA  
CTAAATTCAAACTGAAGGTTTTATGTGTTGACATACCTGGCATACCTAAGGACATGACCTATAGAAGACT  
CATCTCTATGATGGGTTTTAAATGAATTATCAAGTTAATGGTTACCCTAACATGTTTATCACCCGCGAA  
GAAGCTATAAGACATGTACGTGCATGGATTGGCTTCGATGTCGAGGGGTGTCATGCTACTAGAGAAGCTG  
TTGGTACCAATTTACCTTTACAGCTAGGTTTTTCTACAGGTGTTAACCTAGTTGCTGTACCTACAGGTTA  
TGTTGATACACCTAATAATACAGATTTTTCCAGAGTTAGTGCTAAACCACCGCCTGGAGATCAATTTAAA  
CACCTCATACCCTTATGTACAAAGGACTTCCTTGGAATGTAGTGCGTATAAAGATTGTACAAATGTTAA  
GTGACACACTTAAAAATCTCTCTGACAGAGTCGTATTTGTCTTATGGGCACATGGCTTTGAGTTGACATC  
TATGAAGTATTTTGTGAAAATAGGACCTGAGCGCACCTGTTGTCTATGTGATAGACGTGCCACATGCTTT  
TCCACTGCTTCAGACACTTATGCCTGTTGGCATCATTCTATTGGATTTGATTACGTCTATAATCCGTTTA  
TGATTGATGTTCAACAATGGGGTTTTACAGGTAACCTACAAAGCAACCATGATCTGTATTGTCAAGTCCA  
TGGTAATGCACATGTAGCTAGTTGTGATGCAATCATGACTAGGTGTCTAGCTGTCCACGAGTGCTTTGTT  
AAGCGTGTTGACTGGACTATTGAATATCCTATAATTGGTGATGAAGTGAAGATTAAATGCGGCTTGTAGAA  
AGGTTCAACACATGGTTGTTAAAGCTGCATTATTAGCAGACAAATTCAGTTCTTTCACGACATTGGTAA  
CCCTAAAGCTATTAAGTGTGTACCTCAAGCTGATGTAGAATGGAAGTTCTATGATGCACAGCCTTGTAGT  
GACAAAGCTTATAAAATAGAAGAATTATTCTATTCTTATGCCACACATTCTGACAAATTCACAGATGGTG  
TATGCCTATTTTGAATGCAATGTGATAGATATCCTGCTAATTCCATTGTTTGTAGATTTGACACTAG  
AGTGCTATCTAACCTTAACCTTGCTGTTGTGATGGTGGCAGTTTGTATGTAAATAAACATGCATTCCAC  
ACACCAGCTTTTGATAAAAGTGCTTTTGTTAATTTAAACAATTACCATTTTTCTATTACTCTGACAGTC  
CATGTGAGTCTCATGGAACAAGTAGTGTCAGATATAGATTATGTACCACTAAAGTCTGCTACGTGTAT  
AACACGTTGCAATTTAGGTGGTGCTGTCTGTAGACATCATGCTAATGAGTACAGATTGTATCTCGATGCT  
TATAACATGATGATCTCAGCTGGCTTTAGCTTGTGGGTTTTACAAACAATTTGATACTTATAACCTCTGGA  
ACACTTTTACAAGACTTCAGAGTTTAGAAAATGTGGCTTTTAAATGTTGTAAATAAGGGACACTTTGATGG  
ACAACAGGGTGAAGTACCAGTTTCTATCATTAATAACACTGTTTACACAAAAGTTGATGGTGTTGATGTA

GAATTGTTTGAAAATAAAACAACATTACCTGTTAATGTAGCATTTGAGCTTTGGGCTAAGCGCAACATTA  
AACCAGTACCAGAGGTGAAAATACTCAATAATTTGGGTGTGGACATTGCTGCTAATACTGTGATCTGGGA  
CTACAAAAGAGATGCTCCAGCACATATATCTACTATTGGTGTGTTTCTATGACTGACATAGCCAAGAAA  
CCAAC TGAAACGATTTGTGCACCACACTCACTGTCTTTTTTGATGGTAGAGTTGATGGTCAAGTAGACTTAT  
TTAGAAATGCCCCGTAATGGTGTCTTATTACAGAAGGTAGTGTTAAAGGTTTACAACCATCTGTAGGTCC  
CAAACAAGCTAGTCTTAATGGAGTCACATTAATTGGAGAAGCCGTAAAAACACAGTTCAATTATTATAAG  
AAAGTTGATGGTGTGTCCAACAATTACCTGAAACTTACTTTACTCAGAGTAGAAATTTACAAGAATTTA  
AACCCAGGAGTCAAATGGAAATTGATTTCTTAGAATTAGCTATGGATGAATTCATTGAACGGTATAAATT  
AGAAGGCTATGCCTTCGAACATATCGTTTATGGAGATTTTAGTCATAGTCAGTTAGGTGGTTTACATCTA  
CTGATTGGACTAGCTAAACGTTTTAAGGAATCACCTTTTGAATTAGAAGATTTTATTCCATATGGACAGTA  
CAGTTAAAAACTATTTTCATAACAGATGCGCAACAGGTTTATCTAAGTGTGTGTGTTCTGTTATTGATTT  
ATTACTTGATGATTTTTGTTGAAATAATAAAATCCCAAGATTTATCTGTAGTTTCTAAGGTTGTCAAAGTG  
ACTATTGACTATACAGAAATTTTCAATTTATGCTTTGGTGTAAGATGGCCATGTAGAAACATTTTACCCAA  
AATTACAATCTAGTCAAGCGTGGCAACCGGGTGTGCTATGCCTAATCTTTACAAAATGCAAAGAATGCT  
ATTAGAAAAGTGTGACCTTCAAATTATGGTGATAGTGCAACATTACCTAAAGGCATAATGATGAATGTC  
GCAAAATATACTCAACTGTGTCAATATTTAAACACATTAACATTAGCTGTACCCTATAATATGAGAGTTA  
TACATTTTGGTGCTGGTTCTGATAAAGGAGTTGCACCAGGTACAGCTGTTTTAAGACAGTGGTTGCCTAC  
GGGTACGCTGCTTGTGATTGATCTTAATGACTTTGTCTCTGATGCAGATTCAACTTTGATTGGTGAT  
TGTGCAACTGTACATACAGCTAATAAATGGGATCTCATTATTAGTGATATGTACGACCCTAAGACTAAAA  
ATGTTACAAAAGAAAATGACTCTAAAGAGGGTTTTTCACTTACATTTGTGGGTTTATACAACAAAAGCT  
AGCTCTTGGAGGTTCCGTGGCTATAAAGATAACAGAACATTCTTGGAATGCTGATCTTTATAAGCTCATG  
GGACACTTCGCATGGTGGACAGCCTTTGTTACTAATGTGAATGCGTCATCATCTGAAGCATTTTTAATTG  
GATGTAATTATCTTGGCAAACACGCGAACAATAGATGGTTATGTCATGCATGCAAATTACATATTTTG  
GAGGAATACAAATCCAATTCAGTTGTCTTCCTATTCTTTATTTGACATGAGTAAATTTCCCTTAAATTA  
AGGGGTACTGCTGTTATGTCTTTAAAGAAGGTCAAATCAATGATATGATTTTATCTCTTCTTAGTAAAG  
GTAGACTTATAATTAGAGAAAACAACAGAGTTGTTATTTCTAGTGATGTTCTTGTTAACAACTAAACGAA  
CAATGTTTGTGTTTTCTTGTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCA  
ATTACCCCTGCATACACTAATTCTTTCACACGTGGTGTGTTATTACCCTGACAAAGTTTTTCCAGATCCTCA  
GTTTACATTTCAACTCAGGACTTGTCTTACCTTTCTTTTCCAATGTTACTTGGTTCATGCTATACATG  
TCTCTGGGACCAATGGTACTAAGAGGTTTGATAAACCTGTCCCTACCATTTAATGATGGTGTGTTTTTGC  
TTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTTGGTACTACTTTAGATTGGAAGACCCAGTCC  
CTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGAATGATCCAT  
TTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGC  
GAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTTC  
AAAAATCTTAGGGAATTTGTGTTTAAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACACGCCTA  
TTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTTCGGCTTTAGAACCATTTGGTAGATTTGCCAATAGGTAT  
TAACATCACTAGGTTTCAAACCTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCA  
GGTTGGACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTAAATATA  
ATGAAAATGGAACCATTAACAGATGCTGTAGACTGTGCACCTTGACCCTCTCTCAGAAACAAAGTGACGTT  
GAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACCTTCTAAGTTTGAAGTCCAAACCAACAGAATCTATT  
GTTAGATTTTCTAATATTACAACTTGTGCCCTTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCTG  
TTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCTATATAATTCCGCATC  
ATTTTCCACTTTTAAAGTGTTATGGAGTGTCTCCTACTAAATTAATGATCTCTGCTTTACTAATGTCTAT  
GCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTG  
ATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAAATCTAACAATCTTGATTCT  
TAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGA  
GATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTGGAAGGTTTTAATTGTTACT  
TTCTTTTACAATCATATGTTTTCAACCCACTAATGGTGTGTTGTTACCAACCATACAGAGTAGTAGTACT  
TTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAAC  
AAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTC  
TGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGA  
GATTCTTGACATTACACCATGTTCTTTTGGTGGTGTGAGTGTATAACACCAGGAACAAATACTTCTAAC  
CAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTTA  
CTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTAAATAGGGGC  
TGAACATGTCAACAACCTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACT  
CAGACTAATTCTCTCGGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACCTTG  
GTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTAGTGTTAC



CACAGAAATTCTACCAAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCA  
ACTGAATGCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTAACCGTGCTTTAACTGGAA  
TAGCTGTTGAACAAGACAAAAACCCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACC  
AATTAAAGATTTTGGTGGTTTTAATTTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCA  
TTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAAACAATATGGTGATT  
GCCTTGGTGATATTGCTGCTAGAGACCTCATTTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACC  
TTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCTGGTTGG  
ACCTTTGGTGCAAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAAATGGTATTG  
GAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAA  
AATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACCTCAAGATGTGGTCAACCAAAATGCA  
CAAGCTTTAAACACGCTTGTGTTAAACAACCTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTAAATGATA  
TCCTTTACCGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTTGATAGGTTGATCACAGGCAGACTTCAAAG  
TTTGACAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTGCTGCT  
ACTAAAATGTGAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTTGTGGAAAGGGCTATCATCTTA  
TGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCACTGTGACTTATGTCCCTGCACAAGAAAA  
GAACTTCACAACCTGCTCCTGCCATTTGTCTATGATGGAAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTT  
TCAAATGGCACACACTGGTTTGTAAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACA  
CATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACC  
TGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTTAAAGATCATAATCACCAGATGTTGATTTA  
GGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTG  
CCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACCTTGAAAGTATGAGCAGTATATAAAATGGCC  
ATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGACAATTATGCTTTGCTGT  
ATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAAGACG  
ACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAAACGAACCTTATGGATTTGTTTATGAGA  
ATCTTCACAATTGGAACCTGTAACCTTTGAAGCAAGGTGAAATCAAGGATGCTACTCCTTCAGATTTTGTTC  
GCGCTACTGCAACGATACCGATACAAGCCTCACTCCCTTTCGGATGGCTTATTGTTGGCGTTGCACTTCT  
TGCTGTTTTTTCAGAGCGCTTCCAAAATCATAACCCTCAAAAAGAGATGGCAACTAGCACTCTCCAAGGGT  
GTTCACTTTGTTTGAACCTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG  
GCCTTGAAGCCCCCTTTCTCTATCTTTATGCTTTTAGTCTACTTCTTGCAGAGTATAAACTTTGTAAGAAT  
AATAATGAGGCTTTGGCTTTGCTGGAAATGCCGTTCCAAAACCCATTACTTTATGATGCCAATGATTTTT  
CTTTGCTGGCATACTAATTGTTACGACTATTGTATACCTTACAATAGTGTAACCTTCTTCAATTGTCATTA  
CTTCAGGTGATGGCACAACAAGTCCTATTTCTGAACATGACTACCAGATTGGTGGTTATACTGAAAAATG  
GGAATCTGGAGTAAAGACTGTGTTGTATTACACAGTTACTTCACCTCAGACTATTACCAGCTGTACTCA  
ACTCAATTGAGTACAGACACTGGTGTGTAACATGTTACCTTCTTCATCTACAATAAAATTGTTGATGAGC  
CTGAAGAACATGTCCAAATTCACACAATCGACGGTTCATCCGGAGTTGTTAATCCAGTAATGGAACCAAT  
TTATGATGAACCGACGACGACTACTAGCGTGCCTTTGTAAGCACAAAGCTGATGAGTACGAACCTTATGTAC  
TCATTGTTTTCGGAAGAGACAGGTACGTTAATAGTTAATAGCGTACTTCTTTTTCTTGCTTTCTGTTGAT  
TCTTGCTAGTTACACTAGCCATCCTTACTGCGCTTCGATTGTGTGCGTACTGCTGCAATATTGTTAACGT  
GAGTCTTGTAACCTTCTTTTTACGTTTACTCTCGTGTTAAAAATCTGAATTCTTCTAGAGTTCCTGAT  
CTTCTGGTCTAAACGAACATAATATTATATTAGTTTTTCTGTTTGGAACTTTAATTTTAGCCATGGCAGA  
TTCCAACGGTACTATTACCGTTGAAGAGCTTAAAAAGCTCCTTGAACAATGGAACCTAGTAATAGGTTTC  
CTATTCCTTACATGGATTTGTCTTCTACAATTTGCCTATGCCAACAGGAATAGGTTTTTGTATATAATTA  
AGTTAATTTTCTCTGGCTGTTATGGCCAGTAACCTTTAGCTTGTTTTGTGCTTGCTGCTGTTTACAGAAT  
AAATTGGATCACCGGTGGAATTGCTATCGCAATGGCTTGTCTTGTAGGCTTGATGTGGCTCAGCTACTTC  
ATTGCTTCTTTTCACTGTTTGGCGGTACGCGTTCATGTGGTCATTCAATCCAGAACTAACATTCTTC  
TCAACGTGCCACTCCATGGCACTATTCTGACCAGACCGCTTCTAGAAAGTGAACCTCGTAATCGGAGCTGT  
GATCCTTTCGTGGACATCTTCGTATTGCTGGACACCATCTAGGACGCTGTGACATCAAGGACCTGCCTAAA  
GAAATCACTGTTGCTACATACGAACGCTTTCTTATTACAAATTTGGGAGCTTCGACGCGTGTAGCAGGTG  
ACTCAGGTTTTGCTGCATACAGTCGTACAGGATTGGCAACTATAAATTAACACAGACCATTCCAGTAG  
CAGTGACAATATTGCTTTGCTTGTACAGTAAGTGACAACAGATGTTTCATCTCGTTGACTTTTCAAGTTAC  
TATAGCAGAGATATTACTAATTATTATGAGGACTTTTAAAGTTTCCATTTGGAATCTTGATTACATCATA  
AACCTCATAATTAATAAATTTATCTAAGTCACTAAGTGAATAAATATTCTCAATTAGATGAAGAGCAAC  
CAATGGAGATTGATTAAACGAACATGAAAATTATTCTTTTCTTGGCACTGATAACACTCGCTACTTGTGA  
GCTTTTATCACTACCAAGAGTGTGTTAGAGGTACAACAGTACTTTTAAAAGAACCTTGCTCTTCTGGAACA  
TACGAGGGCAATTACCATTTTATCCTCTAGCTGATAACAAATTTGCACTGACTTGCTTTAGCACTCAAT  
TTGCTTTTGTCTTGTCTGACGGCGTAAAACACGTCTATCAGTTACGTGCCAGATCAGTTTACCTAACT  
GTTTCATCAGACAAGAGGAAGTTCAAGAACCTTTACTCTCCAATTTTTCTTATTGTTGCGGCAATAGTGTTT

ATAACACTTTGCTTCACACTCAAAAGAAAGACAGAATGATTGAACTTTCATTAATTGACTTCTATTTGTG  
CTTTTTAGCCTTTCTGCTATTCCTTGTTTTAATTATGCTTATTATCTTTTGTTTCTCACTTGAAGTGCAA  
GATCATAATGAACTTGTGACGCCTAAACGAACATGAAATTTCTTGTTTTCTTAGGAATCATCACAAGTG  
TAGCTGCATTTACCAAGAATGTAGTTTACAGTCATGTACTCAACATCAACCATATGTAGTTGATGACCC  
GTGTCCTATTCACCTTCTATTCTAAATGGTATATTAGAGTAGGAGCTAGAAAATCAGCACCTTTAATTGAA  
TTGTGCGTGGATGAGGCTGGTTCTAAATCACCCATTTCAGTACATCGATATCGGTAATTATACAGTTTCCT  
GTTTACCTTTTACAATTAATTGCCAGGAACCTAAATTGGGTAGTCTTGTAGTGCGTTGTTTCGTTCTATGA  
AGACTTTTTAGAGTATCATGACGTTTCGTGTTGTTTTAGATTTTCATCTAAACGAACAAACTAAAATGTCTG  
ATAATGGACCCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCCTCAGATTCAACTGGCAG  
TAACCAGAATGGAGAACGCAGTGGGGCGCGATCAAAACAACGTCGGCCCCAAGGTTTACCCAATAATACT  
GCGTCTTGGTTCACCGCTCTCACTCAACATGGCAAGGAAGACCTTAAATTCCTCGAGGACAAGGCGTTC  
CAATTAACACCAATAGCAGTCCAGATGACCAAAATTGGCTACTACCGAAGAGCTACCAGACGAATTTCGTGG  
TGGTGACGGTAAAATGAAAGATCTCAGTCCAAGATGGTATTTCTACTACCTAGGAACTGGGCCAGAAGCT  
GGACTTCCCTATGGTGCTAACAAGACGGCATCATATGGGTTGCAACTGAGGGAGCCTTGAATACACCAA  
AAGATCACATTGGCACCCGCAATCCTGCTAACAATGCTGCAATCGTGCTACAACCTTCTCAAGGAACAAC  
ATTGCCAAAAGGCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTTCCTCATCACGT  
AGTCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACTTCTCCTGCTAGAATGGCTGGCA  
ATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCTTGAGAGCAAAATGTCTGG  
TAAAGGCCAACAACAACAAGGCCAAACTGTCTACTAAGAAATCTGCTGCTGAGGCTTCTAAGAAGCCTCGG  
CAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAGACGTGGTCCAGAACAACCC  
AAGGAAATTTTGGGGACCAGGAATAATCAGACAAGGAAGTATTACAAACATTGGCCGCAAATTGCACA  
ATTTGCCCCCAGCGCTTCAGCGTTCTTCGGAATGTGCGCATTGGCATGGAAGTCACACCTTCGGGAACG  
TGGTTGACCTACACAGGTGCCATCAAATTGGATGACAAAGATCCAAATTTCAAAGATCAAGTCATTTTGC  
TGAATAAGCATATTGACGCATACAAAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAGAAGGC  
TGATGAAACTCAAGCCTTACCGCAGAGACAGAAGAAACAGCAAAGTGTGACTCTTCTTCCTGCTGCAGAT  
TTGGATGATTTCTCCAAACAATTGCAACAATCCATGAGCAGTGCTGACTCAACTCAGGCCTAAACTCATG  
CAGACCACACAAGGCAGATGGGCTATATAAACGTTTTTCGCTTTTCCGTTTACGATATATAGTCTACTCTT  
GTGCAGAATGAATTCTCGTAACTACATAGCACAAAGTAGATGTAGTTAACTTTAATCTCACATAGCAATCT  
TTAATCAGTGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTTTACCAGAGGCCACGCGGAGTAC  
GATCGAGTGACAGTGAACAATGCTAGGGAGAGCTGCCTATATGGAAGAGCCCTAATGTGTAAAATTAAT  
TTTAGTAGTGCTATCCCATGTGATTTTAATAGCTTCTTAGGAGAATGACAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAA