

BIOL/CMPU 353 – BIOINFORMATICS Spring 2020
Translation homework

1. Imagine your collaborator gave you the following DNA sequence from their project sequencing the dsDNA genome of a bacterium. They want to know if this segment of the genome might possibly represent a small protein-coding segment of a gene. You decide to translate this sequence into amino acid sequence, and then determine if the amino acid sequence consists of an unbroken string of amino acids. There are six possible translations that could be derived from this DNA sequence.

GCATATGAATGAAGCCGGGATAGAGTAAGGAGTTATCCCATTTCTATTATAG

Report all six possible translations:

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.

Which of these likely represents the true reading frame? Why?

CODON TABLE:

		Second base				
		U	C	A	G	
First base	U	UUU } Phenyl-alanine F UUC } UUA } Leucine L UUG }	UCU } UCC } Serine S UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	U C A G
	C	CUU } CUC } Leucine L CUA } CUG }	CCU } CCC } Proline P CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } CGC } Arginine R CGA } CGG }	U C A G
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } ACC } Threonine T ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }	U C A G
	G	GUU } GUC } Valine V GUA } GUG }	GCU } GCC } Alanine A GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } GGC } Glycine G GGA } GGG }	U C A G

Now, visit <https://web.expasy.org/translate/>, and paste the DNA sequence into the translate window. Select “verbose” and run the tool. Scroll down to see the translations. Does this match your prediction?

2. Another collaborator who has been doing a gene expression study provides you with a DNA sequence that they derived from an mRNA transcript. They think that the DNA sequence from the mRNA is the coding strand, and that it includes a small portion of the upstream untranslated region of the gene along with the beginning of the protein coding segment. Help them identify where the protein-coding portion of the gene begins.

>AccessionNumber

... .GCCTGCTTTCCACTCGCTAGCCCCGCCGGGGTCCGTGTCCTGTCTCGGTGGCCGGACCCGGGCCCCGA
GCCCGAGCAGTAGCCGGCGCCATGTCGGTGGTGGGCATAGACCTGGGCTTCCAGAGCTGCTACGTCGCTG
TGGCCCGCGCCGGCGGCATCGAGACTATCGCTAATGAGTATAGCGACCGCTGCACGCCGGCTTGCATTTC
TTTTGGTCCTAAGAATCGTTCAATTGGAGCAGCAGCTAAAAGCCAGGTAATTTCTAATGCAAAGAACACA
GTCCAAGGATTTAAAAGATTCCATGGCCGAGCATTCTCTGATCCATTTGTGGAGGCAGAAAAATCTAACC
TTGCATATGATATTGTGCAGTTGCCTACAGGATTAACAGGTATAAAGGTGACATATATGGAGGAAGAGCG
AAATTTTACCACTGAGCAAGTGACTGCCATGCTTTTGTCCAAACTGAAGGAGACAGCCGAAA... .