Lab 05: Central Dogma I: Integrated pipeline for Building Web-based DNA Translation Tool

Your Task

To receive full credit for this lab, you must **answer** all following questions.

Task 1: Get protein sequence for DNA sequence (https://www.ncbi.nlm.nih.gov/nuccore/NM_207618.2), and compare it with sequence on NCBI website. Do you get same results?

```
GGTCAGAAAAAGCCCTCTCCATGTCTACTCACGATACATCCCTGAAAACCACTGAGGAAGTGGC
TTTTCAGATCATCTTGCTTTGCCAGTTTGGGGTTGGGACTTTTGCCAATGTATTTCTCTTTGTC
TATAATTTCTCTCCAATCTCGACTGGTTCTAAACAGAGGCCCAGACAAGTGATTTTAAGACACA
TGGCTGTGGCCAATGCCTTAACTCTCTTCCTCACTATATTTCCAAACAACATGATGACTTTTGC
TCCAATTATTCCTCAAACTGACCTCAAATGTAAATTAGAATTCTTCACTCGCCTCGTGGCAAGA
AGCACAAACTTGTGTTCAACTTGTGTTCTGAGTATCCATCAGTTTGTCACACTTGTTCCTGTTA
ATTCAGGTAAAGGAATACTCAGAGCAAGTGTCACAAACATGGCAAGTTATTCTTGTTACAGTTG
TTGGTTCTTCAGTGTCTTAAATAACATCTACATTCCAATTAAGGTCACTGGTCCACAGTTAACA
GACAATAACAATAACTCTAAAAGCAAGTTGTTCTGTTCCACTTCTGATTTCAGTGTAGGCATTG
TCTTCTTGAGGTTTGCCCATGATGCCACATTCATGAGCATCATGGTCTGGACCAGTGTCTCCAT
GGTACTTCTCCTCCATAGACATTGTCAGAGAATGCAGTACATATTCACTCTCAATCAGGACCCC
AGGGGCCAAGCAGAGACCACAGCAACCCATACTATCCTGATGCTGGTAGTCACATTTGTTGGCT
TTTATCTTCTAAGTCTTATTTGTATCATCTTTTACACCTATTTTATATATTCTCATCATTCCCT
GAGGCATTGCAATGACATTTTGGTTTCGGGTTTCCCTACAATTTCTCCTTTACTGTTGACCTTC
AGAGACCCTAAGGGTCCTTGTTCTGTGTTCTTCAACTGTTGAAAGCCAGAGTCACTAAAAATGC
CAAACACAGAAGACAGCTTTGCTAATACCATTAAATACTTTATTCCATAAATATGTTTTTAAAA
GCTTGTATGAACAAGGTATGGTGCTCACTGCTATACTTATAAAGAGTAAGGTTATAATCACTT
GTTGATATGAAAGATTTCTGGTTGGAATCTGATTGAAACAGTGAGTTATTCACCACCCTCCAT
TCTCT
```

- Your Protein Sequence:

MSTHDTSLKTTEEVAFQIILLCQFGVGTFANVFLFVYNFSPISTGSKQRPRQVILRHMAV
ANALTLFLTIFPNNMMTFAPIIPQTDLKCKLEFFTRLVARSTNLCSTCVLSIHQFVTLVPV
NSGKGILRASVTNMASYSCYSCWFFSVLNNIYIPIKVTGPQLTDNNNNSKSKLFCSTSDFS
VGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRMQYIFTLNQDPRGQAETTATHTIL
MLVVTFVGFYLLSLICIIFYTYFIYSHHSLRHCNDILVSGFPTISPLLLTFRDPKGPCSVFFN
C

- NCBI's Protein Sequence:

```
MSTHDTSLKTTEEVAFQIILLCQFGVGTFANVFLFVYNFSPISTGSKQRPRQVILRHMAVANALTLFLTIFPNNMMT
FAPIIPQTDLKCKLEFFTRLVARSTNLCSTCVLSIHQFVTLVPVNSGKGILRASVTNMASYSCYSCWFFSVLNNIYI
PIKVTGPQLTDNNNNSKSKLFCSTSDFSVGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRMQYIFTLNQDPRGQ
AETTATHTILMLVVTFVGFYLLSLICIIFYTYFIYSHHSLRHCNDILVSGFPTISPLLLTFRDPKGPCSVFFNC
```

Task 2: Get protein sequence for DNA sequence (https://www.ncbi.nlm.nih.gov/nuccore/NM_008220.5), and compare it with sequence on NCBI website. Do you get same results?

```
TTACGTTTGCTTCTGATTCTGTTGTGTTGACTTGCAACCTCAGAAACAGACATCATGGTGCACC
TGACTGATGCTGAGAAGGCTGCTGTCTCTGGCCTGTGGGGAAAGGTGAACGCCGATGAAGTTGG
TGGTGAGGCCCTGGGCAGGCTGCTGGTTGTCTACCCTTGGACCCAGCGGTACTTTGATAGCTTT
GGAGACCTATCCTCTGCCTCTGCTATCATGGGTAATGCCAAAGTGAAGGCCCATGGCAAGAAG
TGATAACTGCCTTTAACGATGGCCTGAATCACTTGGACAGCCTCAAGGGCACCTTTGCCAGCCT
CAGTGAGCTCCACTGTGACAAGCTGCATGTGGATCCTGAGAACTTCAGGCTCCTGGGCAATATG
ATCGTGATTGTGCTGGGCCACCACCTGGGCAAGGATTTCACCCCCGCTGCACAGGCTGCCTTCC
AGAAGGTGGTGGCTGGAGTGGCTGCTGCCCTGGCTCACAAGTACCACTAAGCCCCTTTTCTGCT
ATTGTCTATGCACAAAGGTTATATGTCCCCTAGAGAAAACTGTCAATTGTGGGGAAATGATGA
AGACCTTTGGGCATCTAGCTTTTATCTAATAAATGATATTTACTGTCATCTCAAAAAAAAAAAA
AAAAAA
```

- Your Protein Sequence:

MVHLTDAEKAAVSGLWGKVNADEVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMG
NAKVKAHGKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLG
HHLGKDFTPAAQAAFQKVVAGVAAALAHKYH

- NCBI's Protein Sequence:

```
MVHLTDAEKAAVSGLWGKVNADEVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMGNAKVKAHGKKVITAFNDGLN
HLDSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLGHHLGKDFTPAAQAAFQKVVAGVAAALAHKYH
```

Task 3: Get protein sequence for DNA sequence (https://www.ncbi.nlm.nih.gov/nuccore/NM_025580.2), and compare it with sequence on NCBI website. Do you get same results?

```
CGACGGGGAGCGCAGTGAAGGGGGGGTGGGACCCGAACATGGCGGCGGTGGTAGCTGCTACGGCG
CTGAAGGGCCGGGGGGCGAGAAATGCCCGCGTCCTCCGGGGGATCCTCTCTGGAGCCACAGCTA
ACAAGGCTTCCCAGAACAGGACCAGAGCACTGCAGAGCCACAGCTCACCAGAATGCAAGGAGGA
GCCCGAGCCCCTCTCCCCTGAGCTAGAATACATTCCCAGAAAGAGGGGCAAGAACCCCATGAAA
GCCGTGGGGCTAGCCTGGGCCATCGGCTTCCCCTGTGGTATCCTCTTCTTCGTCCTCACCAAGC
AGGAAGTGGACAAGGACCGATTGAAGCAGATGAAAGCTCGTCAGAACATGCGGGTGTCCAACAC
GGGCGAGTATGAGAGCCAGAGGTTCAGGGCTTCACCCCAACAAGCCCAGTTCCCTGAGGTTGGG
TCTGGGGCCCAGACCTGAGTGAGGATCACTGCAACCTTGTGCTAGACTATTGACTGCCGTCTTG
TGTGGCCATGGCTCCTGAGGCCCACCAGAGGGCGCAGGATGCCCAAGCTGCTGCCAGCACCCAA
CTCTTCCACACCGTGGAAGGAGCCAAATGCAAATAAAGTTATTAAGTACTGACACGGAAAAAAA
AA
```

- Your Protein Sequence:

MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSHSSPECKEEPEPLSPE
LEYIPRKRGKNPMKAVGLAWAIGFPCGILFFVLTKQEVDKDRLKQMKARQNMRVSNTG
EYESQRFRASPQQAQFPEVGSGAQT

- NCBI's Protein Sequence:

```
MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSHSSPECKEEPEPLSPELEYIPRKRGKNPMKAVGL
AWAIGFPCGILFFVLTKQEVDKDRLKQMKARQNMRVSNTGEYESQRFRASPQQAQFPEVGSGAQT
```

Task 4: Get protein sequence for DNA sequence (https://www.ncbi.nlm.nih.gov/nuccore/NM_009735.3), and compare it with sequence on NCBI website. Do you get same results?

```
TTTCAGTGGCTGCTACTCGGCGCTTCAGTCGCGGTCGCTTCAGTCGTCAGCATGGCTCGCTCGG
TGACCCTGGTCTTTCTGGTGCTTGTCTCACTGACCGGCCTGTATGCTATCCAGAAAACCCCTCA
AATTCAAGTATACTCACGCCACCCACCGGAGAATGGGAAGCCGAACATACTGAACTGCTACGTA
ACACAGTTCCACCCGCCTCACATTGAAATCCAAATGCTGAAGAACGGGAAAAAAATTCCTAAAG
TAGAGATGTCAGATATGTCCTTCAGCAAGGACTGGTCTTTCTATATCCTGGCTCACACTGAATT
CACCCCCACTGAGACTGATACATACGCCTGCAGAGTTAAGCATGCCAGTATGGCCGAGCCCAAG
ACCGTCTACTGGGATCGAGACATGTGATCAAGCATCATGATGCTCTGAAGATTCATTTGAACCT
GCTTAATTACAAATCCAGTTTCTAATATGCTATACAATTTATGCACGCAGAAAGAAATAGCAAT
GTACACATCACCTTCTTTATATCTTACTTTAAATATTTTATGCATGTTTTCAAAAAATTGGAAA
TATCCTAGATAGCTGAGCAATAAATCTTCAATAAGTATTTTGATCAGAATAATAAATATAATTT
TAAGAACAATAGTTGATCATATGCCAAACCCTCTGTACTTCTCATTACTTGGATGCAGTTACTC
ATCTTTGGTCTATCACAACATAAGTGACATACTTTCCTTTTGGTAAAGCAAAGAGGCCTAATTG
AAGTCTGTCACTGTGCCCAATGCTTAGCAATTCTCACCCCCAACCCTGTGGCTACTTCTGCTTT
TGTTACTTTTACTAAAAATAAAAAAC
```

- Your Protein Sequence:

MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILNCYVTQFHPPHIEIQML
KNGKKIPKVEMSDMSFSKDWSFYILAHTEFTPTETDTYACRVKHASMAEPKTVYWDRD
M

- NCBI's Protein Sequence:

```
MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILNCYVTQFHPPHIEIQMLKNGKKIPKVEMSDMSFS
KDWSFYILAHTEFTPTETDTYACRVKHASMAEPKTVYWDRDM
```

Task 5: Get protein sequence for DNA sequence (https://www.ncbi.nlm.nih.gov/nuccore/NM_011662), and compare it with sequence on NCBI website. Do you get same results?

```
GTCACACCAGGTCCCACCAGCCCCTGGACTGTGGTGTCCAGTGCATATCTGGCCACCATGGGGG
CTCTGGAGCCCTCCTGGTGCCTTCTGTTCCTTCCTGTCCTCCTGACTGTGGGAGGATTAAGTCC
CGTACAGGCCCAGAGTGACACTTTCCCAAGATGCGACTGTTCTTCCGTGAGCCCTGGTGTACTG
GCTGGGATTGTTCTGGGTGACTTGGTGTTGACTCTGCTGATTGCCCTGGCTGTGTACTCTCTGG
GCCGCCTGGTCTCCCGAGGTCAAGGGACAGCGGAAGGGACCCGGAAACAACACATTGCTGAGAC
TGAGTCGCCTTATCAGGAGCTTCAGGGTCAGAGACCAGAAGTATACAGTGACCTCAACACACAG
AGGCAATATTACAGATGAGCCCACTCTATGCCCATCAGCGGCCTGATGCCCGGATCCGGTCATT
CCAGATGCCTACTCAACAAGCCCTCTCTGGGATCAGGACTCCCGTTGGAATACAGATCCACAGG
GTACCTCCCTGAGATATCTGACATTGTACCATTTCTGTCCCCAAATAAAAGACAGAGCAAAAAA
AAAAAAAAAAAAAAAA
```

- Your Protein Sequence:

MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGVLAGIVLGDLVLTLLIA
LAVYSLGRLVSRGQGTAEGTRKQHIAETESPYQELQGQRPEVYSDLNTQRQYYR

- NCBI's Protein Sequence:

MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGVLAGIVLGDLVLTLLIALAVYSLGRLVSRGQGTA
EGTRKQHIAETESPYQELQGQRPEVYSDLNTQRQYYR

Task 6: Give a brief overview of your work, discussing any thoughts, successes or challenges you faced.

First, I used my code from lab04 and manipulated the function I created to mold to the new variable and the new test DNA sequence. Once I verified that the function was running smoothly, I copied the HTML code into my notebook and opened the second link to view the webpage. Upon doing so, I pasted the DNA sequences to be translated and outputted the translated amino acids. After this, I verified with the NCBI publication to make sure the amino acid sequences matched up.