

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
TTTTTCAGATCATCTTGCTTTGCCAGTTTGGGGTTGGGACTTTTGCCAATGTATTTCTCTTTGTC
TATAATTTCTCTCCAATCTCGACTGGTTCTAAACAGAGGCCAGACAAGTGATTTTAAGACACA
TGGCTGTGGCCAATGCCTTAACTCTCTTCCTCACTATATTTCCAAACAACATGATGACTTTTGC
TCCAATTATTCCTCAAACCTGACCTCAAATGTAAATTAGAATTCTTCACTCGCCTCGTGGCAAGA
AGCACAAACTTGTGTTCAACTTGTGTTCTGAGTATCCATCAGTTTGTCACTTGTTCCTGTTA
ATTCAGGTAAAGGAATACTCAGAGCAAGTGTCAAAACATGGCAAGTTATTCTTGTTACAGTTG
TTGTTTCTTCAGTGTCTTAAATAACATCTACATTCCAATTAAGGTCACCTGGTCCACAGTTAACA
GACAATAACAATAACTCTAAAAGCAAGTTGTTCTGTTCCACTTCTGATTTTCACTGTAGGCATTG
TCTTCTTGAGGTTTGCCCATGATGCCACATTCATGAGCATCATGGTCTGGACCAGTGTCTCCAT
GGTACTTCTCCTCCATAGACATTGTCAGAGAATGCAGTACATATTCACTCTCAATCAGGACCCC
AGGGGCCAAGCAGAGACCACAGCAACCCATACTATCCTGATGCTGGTAGTCACATTTGTTGGCT
TTTATCTTCTAAGTCTTATTTGTATCATCTTTTACACCTATTTTATATATTCTCATCATTCCTT
GAGGCATTGCAATGACATTTTGGTTTTCGGGTTTCCCTACAATTTCTCCTTTACTGTTGACCTTC
AGAGACCCTAAGGGTCCTTGTTCTGTGTTCTTCAACTGTTGAAAGCCAGAGTCACTAAAAATGC
CAAACACAGAAGACAGCTTTGCTAATACCATTAAATACTTTATTCCATAAATATGTTTTTAAAA
GCTTGTATGAACAAGGTATGGTGCTCACTGCTATACTTATAAAAGAGTAAGGTTATAATCACTT
GTTGATATGAAAAGATTTCTGGTTGGAATCTGATTGAAACAGTGAGTTATTCACCACCCTCCAT
TCTCT
```

- Your Protein Sequence:

```
MSTHDTSLKTTEEVAFAQIILLCQFGVGTAFANVFLFVYNFSPISTGSKQRPRQVILRHMAV
ANALTLFLTIFPNNMMTFAPIIPQTDLKCKLEFFTRLVARSTNLCSTCVLSIHQFVTLVPV
NSGKGILRASVTNMA SYSCYSCWFFSVLNNIYIPIKVTGPQLTDNNNNSKSKLFCSTSDFS
VGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRMQYIFTLNQDPRGQAETTATHTIL
MLVVTFVGFYLLSLICIIFYTYFIYSHSLRHCNDILVSGFPTISPLLLTFRDPKGPCSVFFN
C
```

- NCBI's Protein Sequence:

```
"MSTHDTSLKTTEEVAFAQIILLCQFGVGTAFANVFLFVYNFSPIST
GSKQRPRQVILRHMAVANALTLFLTIFPNNMMTFAPIIPQTDLKCKLEFFTRLVARST
NLCSTCVLSIHQFVTLVPVNSGKGILRASVTNMA SYSCYSCWFFSVLNNIYIPIKVTG
PQLTDNNNNSKSKLFCSTSDFSVGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRM
QYIFTLNQDPRGQAETTATHTILMLVVTFVGFYLLSLICIIFYTYFIYSHSLRHCND
ILVSGFPTISPLLLTFRDPKGPCSVFFNC"
```


Task 2: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nucore/NM_008220.5), and compare it with sequence on NCBI website. Do you get same results?

```
TTACGTTTGCTTCTGATTCTGTTGTGTTGACTTGCAACCTCAGAAACAGACATCATGGTGCACC
TGACTGATGCTGAGAAGGCTGCTGTCTCTGGCCTGTGGGGAAAGGTGAACGCCGATGAAGTTGG
TGGTGAGGCCCTGGGCAGGCTGCTGGTTGTCTACCCTTGGACCCAGCGGTACTTTGATAGCTTT
GGAGACCTATCCTCTGCCTCTGCTATCATGGGTAATGCCAAAGTGAAGGCCCATGGCAAGAAAG
TGATAACTGCCTTTAACGATGGCCTGAATCACTTGGACAGCCTCAAGGGCACCTTTGCCAGCCT
CAGTGAGCTCCACTGTGACAAGCTGCATGTGGATCCTGAGAACTTCAGGCTCCTGGGCAATATG
ATCGTGATTGTGCTGGGCCACCACCTGGGCAAGGATTTACCCCCGCTGCACAGGCTGCCTTCC
AGAAGGTGGTGGCTGGAGTGGCTGCTGCCCTGGCTCACAAGTACCACTAAGCCCCCTTTTCTGCT
ATTGTCTATGCACAAAGGTTATATGTCCCCTAGAGAAAACTGTCAATTGTGGGGAAATGATGA
AGACCTTTGGGCATCTAGCTTTTATCTAATAAATGATATTTACTGTCATCTCAAAAAAAAAAAAA
AAAAAA
```

- Your Protein Sequence:

```
MVHLTDAEKAAVSLWGKVNADDEVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMG
NAKVKAHGKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLG
HHLGKDFTPAAQAAFQKVVAGVAAALAHKYH
```

- NCBI's Protein Sequence:

```
MVHLTDAEKAAVSLWGKVNADDEVGGEALGRLLVVYPWTQRYFD
SFGDLSSASAIMGNAKVKAHGKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPE
NFRLLGNMIVIVLGHHLGKDFTPAAQAAFQKVVAGVAAALAHKYH
```

Task 3: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nucore/NM_025580.2), and compare it with sequence on NCBI website. Do you get same results?

```
CGACGGGGAGCGCAGTGAAGGGGGGTGGGACCCGAACATGGCGGCGGTGGTAGCTGCTACGGCG
CTGAAGGGCCGGGGGGCGAGAAATGCCCGCGTCCTCCGGGGGATCCTCTCTGGAGCCACAGCTA
ACAAGGCTTCCCAGAACAGGACCAGAGCACTGCAGAGCCACAGCTCACCAGAATGCAAGGAGGA
GCCCCAGCCCCTCTCCCCTGAGCTAGAATACATTCCCAGAAAGAGGGGCAAGAACCCCATGAAA
GCCGTGGGGCTAGCCTGGGCCATCGGCTTCCCCTGTGGTATCCTCTTCTTCGTCCTCACCAAGC
AGGAAGTGGACAAGGACCGATTGAAGCAGATGAAAGCTCGTCAGAACATGCGGGTGTCCAACAC
GGGCGAGTATGAGAGCCAGAGGTTTCAAGGCTTCAACCCCAACAAGCCCAGTTCCCTGAGGTTGGG
TCTGGGGCCCAGACCTGAGTGAGGATCACTGCAACCTTGTGCTAGACTATTGACTGCCGTCTTG
TGTGGCCATGGCTCCTGAGGCCACCAGAGGGCGCAGGATGCCCAAGCTGCTGCCAGCACCCAA
CTCTTCCACACCGTGGAAGGAGCCAAATGCAAATAAAGTTATTAAGTACTGACACGGAAAAAAA
AA
```

- Your Protein Sequence:

MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSHSSPECKEEPEPLSPE
LEYIPRKRKGKNPMKAVGLAWAIGFPCGILFFVLTKQEVDKDRLKQMKARQNM RVSNTG
EYESQRF RASPQQAQFPEVGSGAQT

- NCBI's Protein Sequence:

MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSH
SSPECKEEPEPLSPELEYIPRKRKGKNPMKAVGLAWAIGFPCGILFFVLTKQEVDKDRL
KQMKARQNM RVSNTGEYESQRF RASPQQAQFPEVGSGAQT

Task 4: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nucore/NM_009735.3), and compare it with sequence on NCBI website. Do you get same results?

```
TTTCAGTGGCTGCTACTCGGCGCTTCAGTCGCGGTTCGCTTCAGTCGTCAGCATGGCTCGCTCGG
TGACCCTGGTCTTTCTGGTGCTTGTCTCACTGACCGGCCTGTATGCTATCCAGAAAACCCCTCA
AATTCAAGTATACTCACGCCACCCACCGGAGAATGGGAAGCCGAACATACTGAACTGCTACGTA
ACACAGTTCCACCCGCTCACATTGAAATCCAAATGCTGAAGAACGGGAAAAAAATTCCTAAAG
TAGAGATGTCAGATATGTCCTTCAGCAAGGACTGGTCTTTCTATATCCTGGCTCACACTGAATT
CACCCCCACTGAGACTGATACATACGCCTGCAGAGTTAAGCATGCCAGTATGGCCGAGCCCAAG
ACCGTCTACTGGGATCGAGACATGTGATCAAGCATCATGATGCTCTGAAGATTCATTTGAACCT
GCTTAATTACAAATCCAGTTTCTAATATGCTATACAATTTATGCACGCAGAAAGAAATAGCAAT
GTACACATCACCTTCTTTATATCTTACTTTAAATATTTTATGCATGTTTTCAAAAAATTGAAA
TATCCTAGATAGCTGAGCAATAAATCTTCAATAAGTATTTTGATCAGAATAATAAATATAATTT
TAAGAACAATAGTTGATCATATGCCAAACCCTCTGTACTTCTCATTACTTGGATGCAGTTACTC
ATCTTTGGTCTATCACAACATAAGTGACATACTTTCCCTTTTGGTAAAGCAAAGAGGCCTAATTG
AAGTCTGTCACTGTGCCCAATGCTTAGCAATTCTCACCCCCAACCCCTGTGGCTACTTCTGCTTT
TGTTACTTTTACTAAAAATAAAAAAC
```

- Your Protein Sequence:

MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILNCYVTQFHPPHIEIQML
KNGKKIPKVEMSDMSFSKDWSFYILAHTFTPTETDTYACRVKHASMAEPKTVYWDRD
M

- NCBI's Protein Sequence:

MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILN
CYVTQFHPPHIEIQMLKNGKKIPKVEMSDMSFSKDWSFYILAHTFTPTETDTYACRV
KHASMAEPKTVYWDRDM

Task 5: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nucore/NM_011662), and compare it with sequence on NCBI website. Do you get same results?

```
GTCACACCAGGTCCCACCAGCCCCTGGACTGTGGTGTCCAGTGCATATCTGGCCACCATGGGGG
CTCTGGAGCCCTCCTGGTGCCTTCTGTTCCCTTCCTGTCCTCCTGACTGTGGGAGGATTAAGTCC
CGTACAGGCCCAGAGTGACACTTTCCTCAAGATGCGACTGTTCTTCCGTGAGCCCTGGTGTACTG
GCTGGGATTGTTCTGGGTGACTTGGTGTGACTCTGCTGATTGCCCTGGCTGTGTACTCTCTGG
GCCGCCTGGTCTCCCGAGGTCAAGGGACAGCGGAAGGGACCCGAAACAACACATTGCTGAGAC
TGAGTCGCCTTATCAGGAGCTTCAGGGTCAGAGACCAGAAGTATACAGTGACCTCAACACACAG
AGGCAATATTACAGATGAGCCCACTCTATGCCCATCAGCGGCCTGATGCCCCGGATCCGGTCATT
CCAGATGCCTACTCAACAAGCCCTCTCTGGGATCAGGACTCCCGTTGGAATACAGATCCACAGG
GTACCTCCCTGAGATATCTGACATTGTACCATTCTGTCCCCAAATAAAAGACAGAGCAAAAAA
AAAAAAAAAAAAAAAAAAAA
```

- Your Protein Sequence:

MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGVLAGIVLGDLVLTLLIA
LAVYSLGRLVSRGQGTAEGTRKQHIAETESPYQELQGQRPEVYSDLNTQRQYYR

- NCBI's Protein Sequence:
- MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGV
LAGIVLGDLVLTLLIALAVYSLGRLVSRGQGTAEGTRKQHIAETESPYQELQGQRPEV
YSDLNTQRQYYR

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

Overall I feel like I did a good job on this assignment. The function I wrote worked but it had a few issues that we were able to fix. Once we fixed the function applying the HTML codes was easy and everything worked well.