

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
TCCAATTATTCTCAAACTGACCTCAAATGTAAATTAGAATTCTTCACTCGCCTCGTGGAAGA
AGCACAAACTTGTGTTCAACTTGTGTTCTGAGTATCCATCAGTTTGTCACTTGTTCCTGTTA
ATTGAGGTAAAGGAATACTCAGAGCAAGTGTCAAAACATGGCAAGTTATTCTTGTTACAGTTG
TTGTTTCTTCAGTGTCTTAAATAACATCTACATTCCAATTAAGGTCAGTGGTCCACAGTTAACA
GACAATAACAATACTCTAAAAGCAAGTTGTTCTGTTCCACTTCTGATTTCAAGTGTAGGCATTG
TCTTCTTGAGGTTTGCCCATGATGCCACATTCATGAGCATCATGGTCTGGACCAGTGTCTCCAT
GGTACTTCTCCTCCATAGACATTGTGAGAGAATGCAGTACATATTCCTCTCAATCAGGACCCC
AGGGGCCAAGCAGAGACCACAGCAACCCATACTATCCTGATGCTGGTAGTCACATTTGTTGGCT
TTTATCTTCTAAGTCTTATTTGTATCATCTTTTACACCTATTTTATATATTCTCATCATTCCTT
GAGGCATTGCAATGACATTTTGGTTTTCGGGTTTCCCTACAATTTCTCCTTTACTGTTGACCTTC
AGAGACCCTAAGGGTCCTTGTTCTGTGTTCTTCAACTGTTGAAAGCCAGAGTCACTAAAAATGC
CAAACACAGAAGACAGCTTTGCTAATACCATTAAATACTTTATTCCATAAATATGTTTTTAAAA
GCTTGTATGAACAAGGTATGGTGCTCACTGCTATACTTATAAAAGAGTAAGGTTATAATCACTT
GTTGATATGAAAAGATTTCTGGTTGGAATCTGATTGAAACAGTGAGTTATTCACCACCCTCCAT
TCTCT
```

- Your Protein Sequence:

```
MSTHDTSLKTTEEVAFAQIILLCQFGVGTFANVFLFVYNFSPISTGSKQRPRQVILRHMAV
ANALTLFLTIFPNNMMTFAPIIPQTDLKCKLEFFTRLVARSTNLCSTCVLSIHQFVTLVPV
NSGKGILRASVTNMA SYSCYSCWFFSVLNNIYIPIKVTGPQLTDNNNNNSKSKLFCSTSDFS
VGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRMQYIFTLNQDPRGQAETTATHIL
MLVVTFVGFYLLSLICIFYTYFIYSHSLRHCNDILVSGFPTISPLLLTFRDPKGPSCSVFFN
C
```

- NCBI's Protein Sequence:

```
MSTHDTSLKTTEEVAFAQIILLCQFGVGTFANVFLFVYNFSPISTGSKQRPRQVILRHMAVANAL
TLFLTIFPNNMMTFAPIIPQTDLKCKLEFFTRLVARSTNLCSTCVLSIHQFVTLVPVNSGKGIL
RASVTNMA SYSCYSCWFFSVLNNIYIPIKVTGPQLTDNNNNNSKSKLFCSTSDFS VGIVFLRFAH
```

DATFMSIMVWTSVSMVLLLHRHCQRMQYIFTLNQDPRGQAETTATHTILMLVVTFVGFYLLSLI
CIIFYTYFIYSHSLRHCNDILVSGFPTISPLLLTFRDPKGPCSVFFNC

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
GGAGACCTATCCTCTGCCTCTGCTATCATGGGTAATGCCAAAGTGAAGGCCCATGGCAAGAAAG
TGATAACTGCCTTTAACGATGGCCTGAATCACTTGGACAGCCTCAAGGGCACCTTTGCCAGCCT
CAGTGAGCTCCACTGTGACAAGCTGCATGTGGATCCTGAGAACTTCAGGCTCCTGGGCAATATG
ATCGTGATTGTGCTGGGCCACCACCTGGGCAAGGATTTCACCCCGCTGCACAGGCTGCCTTCC
AGAAGGTGGTGGCTGGAGTGGCTGCTGCCCTGGCTCACAAGTACCACTAAGCCCCTTTTCTGCT
ATTGTCTATGCACAAAGGTTATATGTCCCCTAGAGAAAACTGTCAATTGTGGGGAAATGATGA
AGACCTTTGGGCATCTAGCTTTTATCTAATAAATGATATTTACTGTCATCTCAAAAAAAAAAAAA
AAAAAA

- Your Protein Sequence:

MVHLTDAEKAAVSGLWGKVNADGVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMG
NAKVKAHGKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLG
HHLGKDFTPAQAQAFQKVAGVAAALAHKYH

- NCBI's Protein Sequence:

MVHLTDAEKAAVSGLWGKVNADGVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMGNAKVKAH
GKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLGHHLGKDFTPAQ
AAFQKVAGVAAALAHKYH

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?

CGACGGGGAGCGCAGTGAAGGGGGGTGGGACCCGAACATGGCGGCGGTGGTAGCTGCTACGGCG
CTGAAGGGCCGGGGGGCGAGAAATGCCCCGCGTCCTCCGGGGGATCCTCTCTGGAGCCACAGCTA
ACAAGGCTTCCCAGAACAGGACCAGAGCACTGCAGAGCCACAGCTCACCAGAATGCAAGGAGGA
GCCCCAGCCCCTCTCCCCTGAGCTAGAATACATTCCCAGAAAGAGGGGCAAGAACCCCATGAAA
GCCGTGGGGCTAGCCTGGGCCATCGGCTTCCCCTGTGGTATCCTCTTCTTCGTCCTCACCAAGC
AGGAAGTGGACAAGGACCGATTGAAGCAGATGAAAGCTCGTCAGAACATGCGGGTGTCCAACAC
GGGCGAGTATGAGAGCCAGAGGTTCAAGGGCTTCACCCCAACAAGCCCAGTTCCCTGAGGTTGGG
TCTGGGGCCCAGACCTGAGTGAGGATCACTGCAACCTTGTGCTAGACTATTGACTGCCGTCTTG
TGTGGCCATGGCTCCTGAGGCCACAGAGGGCGCAGGATGCCCAAGCTGCTGCCAGCACCCAA
CTCTTCCACACCGTGGAAGGAGCCAAATGCAAATAAAGTTATTAAGTACTGACACGGAAAAAAA
AA

- Your Protein Sequence:

MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSHSSPECKEEPEPLSPE
LEYIPRKRGNPMKAVGLAWAIGFPCGILFFVLTKQEVDKDRLKQMKARQNM RVSNTG
EYESQRF RASPQQAQFPEVGSGAQT

- NCBI's Protein Sequence:

MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSHSSPECKEEPEPLSPELEYIP
RKRGNPMKAVGLAWAIGFPCGILFFVLTKQEVDKDRLKQMKARQNM RVSNTGEYESQRF RASP
QQAQFPEVGSGAQT

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?

```
TTTCAGTGGCTGCTACTCGGCGCTTCAGTCGCGGTCGCTTCAGTCGTCAGCATGGCTCGCTCGG
TGACCCTGGTCTTTCTGGTGCTTGTCTCACTGACCGGCCTGTATGCTATCCAGAAAACCCCTCA
AATTCAAGTATACTCACGCCACCCACCGGAGAATGGGAAGCCGAACATACTGAACTGCTACGTA
ACACAGTTCCACCCGCTCACATTGAAATCCAAATGCTGAAGAACGGGAAAAAAATTCCTAAAG
TAGAGATGTCAGATATGTCCTTCAGCAAGGACTGGTCTTTCTATATCCTGGCTCACACTGAATT
CACCCCCACTGAGACTGATACATACGCCTGCAGAGTTAAGCATGCCAGTATGGCCGAGCCCAAG
ACCGTCTACTGGGATCGAGACATGTGATCAAGCATCATGATGCTCTGAAGATTCATTTGAACCT
GCTTAATTACAAATCCAGTTTCTAATATGCTATACAATTTATGCACGCAGAAAGAAATAGCAAT
GTACACATCACCTTCTTTATATCTTACTTTAAATATTTTATGCATGTTTTCAAAAAATTGGAAA
TATCCTAGATAGCTGAGCAATAAATCTTCAATAAGTATTTTGATCAGAATAATAAATATAATTT
TAAGAACAATAGTTGATCATATGCCAAACCCTCTGTACTTCTCATTACTTGGATGCAGTTACTC
ATCTTTGGTCTATCACAACATAAGTGACATACTTTCCTTTTGGTAAAGCAAAGAGGCCTAATTG
AAGTCTGTCACTGTGCCCAATGCTTAGCAATTCTCACCCCCAACCCCTGTGGCTACTTCTGCTTT
TGTTACTTTTACTAAAAATAAAAAAC
```

- Your Protein Sequence:

```
MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILNCYVTQFHPPHIEIQML
KNGKKIPKVEMSDMSFSKDWSFYILAHTFTPTETDTYACRVKHASMAEPKTVYWDRD
M
```

- NCBI's Protein Sequence:

```
MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILNCYVTQFHPPHIEIQMLKNGK
KIPKVEMSDMSFSKDWSFYILAHTFTPTETDTYACRVKHASMAEPKTVYWDRDM
```

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
CTCTGGAGCCCCTCCTGGTGCCTTCTGTTCCCTTCCTGTCCTCCTGACTGTGGGAGGATTAAGTCC
CGTACAGGCCCAGAGTGACACTTTCCCAAGATGCGACTGTTCTTCCGTGAGCCCTGGTGTACTG
GCTGGGATTGTTCTGGGTGACTTGGTGTGACTCTGCTGATTGCCCTGGCTGTGTACTCTCTGG
GCCGCCTGGTCTCCCAGAGTCAAGGGACAGCGGAAGGGACCCGAAACAACACATTGCTGAGAC
TGAGTCGCCTTATCAGGAGCTTCAGGGTCAGAGACCAGAAGTATACAGTGACCTCAACACACAG
AGGCAATATTACAGATGAGCCCACTCTATGCCCATCAGCGGCCTGATGCCCGGATCCGGTCATT
CCAGATGCCTACTCAACAAGCCCTCTCTGGGATCAGGACTCCCGTTGGAATACAGATCCACAGG
GTACCTCCCTGAGATATCTGACATTGTACCATTCTGTCCCCAAATAAAAGACAGAGCAAAAAA
AAAAAAAAAAAAAAAAAAAA
```

- Your Protein Sequence:

```
MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGVLAGIVLGDLVLTLLIA
LAVYSLGRLVSRGQGTAEGTRKQHIAETESPYQELQGQRPEVYSDLNTQRQYYR
```

- NCBI's Protein Sequence:

```
MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGVLAGIVLGDLVLTLLIALAVY
SLGRLVSRGQGTAEGTRKQHIAETESPYQELQGQRPEVYSDLNTQRQYYR
```

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

I think this assignment was easy until question 6. I spent two hours working on it and I was able to add on 3 frame and somewhat start the complementary frame (still have a lot to be done) . I was trying to work more on it, but I was not able to because I have other homework that I have to attend to (exam on Monday9/14) I hope that I will have some time next week to talk over with you about question 6, I think I manage to set up the new reading frame 5' to 3' but the 3' to 5' is quite difficult.