

## 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](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
GACAATAACAATAACTCTAAAAGCAAGTTGTTCTGTTCCACTTCTGATTTCAAGTGTAGGCATTG
TCTTCTTGAGGTTTGCCCATGATGCCACATTCATGAGCATCATGGTCTGGACCAGTGTCTCCAT
GGTACTTCTCCTCCATAGACATTGTCAGAGAATGCAGTACATATTCACTCTCAATCAGGACCCC
AGGGGCCAAGCAGAGACCACAGCAACCCATACTATCCTGATGCTGGTAGTCACATTTGTTGGCT
TTTATCTTCTAAGTCTTATTTGTATCATCTTTTACACCTATTTTATATATTCTCATCATTCCTT
GAGGCATTGCAATGACATTTTGGTTTTCGGGTTTCCCTACAATTTCTCCTTTACTGTTGACCTTC
AGAGACCCTAAGGGTCCTTGTTCTGTGTTCTTCAACTGTTGAAAGCCAGAGTCACTAAAAATGC
CAAACACAGAAGACAGCTTTGCTAATACCATTAAATACTTTATTCCATAAATATGTTTTTAAAA
GCTTGTATGAACAAGGTATGGTGCTCACTGCTATACTTATAAAAGAGTAAGGTTATAATCACTT
GTTGATATGAAAAGATTTCTGGTTGGAATCTGATTGAAACAGTGAGTTATTCACCACCCTCCAT
TCTCT
```

- Your Protein Sequence:

```
MSTHDTSLKTTEEVAFAQIILLCQFGVGTAFANVFLFVYNFSPISTGSKQRPRQVILRHMAV
ANALTLFLTIFPNNMMTFAPIIPQTDLKCKLEFFTRLVARSTNLCSTCVLSIHQFVTLVPV
NSGKGILRASVTNMASYSCYSCWFFSVLNNIYIPIKVTGPQLTDNNNNNSKSKLFCSTSDFS
VGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRMQYIFTLNQDPRGQAETTATHTIL
MLVVTFVGFYLLSLICIIFYTYFIYSHSLRHCNDILVSGFPTISPLLLTFRDPKGPCSVFFN
C
```

- NCBI's Protein Sequence:

```
"MSTHDTSLKTTEEVAFAQIILLCQFGVGTAFANVFLFVYNFSPISTGSKQRPRQVILRHMAVANALTLFLTIFPNNMM
TFAPIIPQTDLKCKLEFFTRLVARSTNLCSTCVLSIHQFVTLVPVNSGKGILRASVTNMASYSCYSCWFFSVLNNIY
IPIKVTGPQLTDNNNNNSKSKLFCSTSDFSVGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRMQYIFTLNQDPRG
QAETTATHTILMLVVTFVGFYLLSLICIIFYTYFIYSHSLRHCNDILVSGFPTISPLLLTFRDPKGPCSVFFNC"
```

### Task 2: Get protein sequence for DNA sequence

([https://www.ncbi.nlm.nih.gov/nucore/NM\\_008220.5](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:

```
MVHLTDAEKA AVSGLWGKVN ADEVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMG
NAKVKAHGKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLG
HHLGKDFTPAAQAAFQKVVAGVAAALAHKYH
```

- NCBI's Protein Sequence:

```
"MVHLTDAEKA AVSGLWGKVN ADEVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMGNAKVKAHGKKVITAFNDGL
NHLDSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLGHHLGKDFTPAAQAAFQKVVAGVAAALAHKYH"
```

### Task 3: Get protein sequence for DNA sequence

([https://www.ncbi.nlm.nih.gov/nucore/NM\\_025580.2](https://www.ncbi.nlm.nih.gov/nucore/NM_025580.2)), and compare it with sequence on NCBI website. Do you get same results?

```
CGACGGGGAGCGCAGTGAAGGGGGGTGGGACCCGAACATGGCGGCGGTGGTAGCTGCTACGGCG
CTGAAGGGCCGGGGGGCGAGAAATGCCCGCGTCTCCGGGGGATCCTCTCTGGAGCCACAGCTA
ACAAGGCTTCCCAGAACAGGACCAGAGCACTGCAGAGCCACAGCTCACCAGAATGCAAGGAGGA
GCCCCAGCCCCTCTCCCCTGAGCTAGAATACATTCCCAGAAAGAGGGGCAAGAACCCCATGAAA
GCCGTGGGGCTAGCCTGGGCCATCGGCTTCCCCTGTGGTATCCTCTTCTTCGTCCTCACCAAGC
AGGAAGTGGAACAAGGACCGATTGAAGCAGATGAAAGCTCGTCAGAACATGCGGGTGTCCAACAC
GGGCGAGTATGAGAGCCAGAGGTTTCAAGGCTTCAACCCCAACAAGCCCAGTTCCCTGAGGTTGGG
TCTGGGGCCAGACCTGAGTGAGGATCACTGCAACCTTGTGCTAGACTATTGACTGCCGTCTTG
TGTGGCCATGGCTCCTGAGGCCCACCAGAGGGCGCAGGATGCCCAAGCTGCTGCCAGCACCCAA
CTCTTCCACACCGTGGAAGGAGCCAAATGCAAATAAAGTTATTAAGTACTGACACGGAAAAAAA
AA
```

- Your Protein Sequence:

MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSHSSPECKEEPEPLSPE  
LEYIPRKRKGKNPMKAVGLAWAIGFPCGILFFVLTKQEVDKDRLKQMKARQNMVRSNTG  
EYESQRFRA SPQQAQFPEVGSQAQT

- NCBI's Protein Sequence:

"MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSHSSPECKEEPEPLSPELEYIPRKRKGKNPMKAVG  
LAWAIGFPCGILFFVLTKQEVDKDRLKQMKARQNMVRSNTGEYESQRFRA SPQQAQFPEVGSQAQT"

#### Task 4: Get protein sequence for DNA sequence

([https://www.ncbi.nlm.nih.gov/nuccore/NM\\_009735.3](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  
ACACAGTTCCACCCGCTCACATTGAAATCCAAATGCTGAAGAACGGGAAAAAATTCCTAAAG  
TAGAGATGTCAGATATGTCCTTCAGCAAGGACTGGTCTTTCTATATCCTGGCTCACACTGAATT  
CACCCCCACTGAGACTGATACATACGCCTGCAGAGTTAAGCATGCCAGTATGGCCGAGCCCAAG  
ACCGTCTACTGGGATCGAGACATGTGATCAAGCATCATGATGCTCTGAAGATTCATTTGAACCT  
GCTTAATTACAAATCCAGTTTCTAATATGCTATACAATTTATGCACGCAGAAAGAAATAGCAAT  
GTACACATCACCTTCTTTATATCTTACTTTAAATATTTTATGCATGTTTTCAAAAAATGGAAA  
TATCCTAGATAGCTGAGCAATAAATCTTCAATAAGTATTTTGATCAGAATAATAAATATAATTT  
TAAGAACAATAGTTGATCATATGCCAAACCCTCTGTACTTCTCATTACTTGGATGCAGTTACTC  
ATCTTTGGTCTATCACAACATAAGTGACATACTTTCCTTTTGGTAAAGCAAAGAGGCCTAATTG  
AAGTCTGTCACTGTGCCCAATGCTTAGCAATTCTCACCCCCAACCTGTGGCTACTTCTGCTTT  
TGTTACTTTTACTAAAAATAAAAAAC

- Your Protein Sequence:

MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILNCYVTQFHPPHIEIQML  
KNGKKIPKVEMSDMSFSKDWSFYILAHTFTPTETDTYACRVKHASMAEPKTVYWDRD  
M

- NCBI's Protein Sequence:

"MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILNCYVTQFHPPHIEIQMLKNGKKIPKVEMSDMSF  
SKDWSFYILAHTFTPTETDTYACRVKHASMAEPKTVYWDRDM"

#### Task 5: Get protein sequence for DNA sequence

([https://www.ncbi.nlm.nih.gov/nuccore/NM\\_011662](https://www.ncbi.nlm.nih.gov/nuccore/NM_011662)), and compare it with sequence on NCBI website. Do you get same results?

GTCACACCAGGTCCCACCAGCCCCTGGACTGTGGTGTCCAGTGCATATCTGGCCACCATGGGGG  
CTCTGGAGCCCTCCTGGTGCCTTCTGTTCTTCTGTCTCTCTGACTGTGGGAGGATTAAGTCC

CGTACAGGCCCAGAGTGACACTTTCCCAAGATGCGACTGTTCTTCCGTGAGCCCTGGTGTACTG  
GCTGGGATTGTTCTGGGTGACTTGGTGTGACTCTGCTGATTGCCCTGGCTGTGTACTCTCTGG  
GCCGCCTGGTCTCCCGAGGTCAAGGGACAGCGGAAGGGACCCGAAACAACACATTGCTGAGAC  
TGAGTCGCCTTATCAGGAGCTTCAGGGTCAGAGACCAGAAGTATACAGTGACCTCAACACACAG  
AGGCAATATTACAGATGAGCCCACTCTATGCCCATCAGCGGCCTGATGCCCCGGATCCGGTCATT  
CCAGATGCCTACTCAACAAGCCCTCTCTGGGATCAGGACTCCCGTTGGAATACAGATCCACAGG  
GTACCTCCCTGAGATATCTGACATTGTACCATTTCTGTCCCCAAATAAAAGACAGAGCAAAAAA  
AAAAAAAAAAAAAAAAAAAA

- Your Protein Sequence:

MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGVLAGIVLGDVLVLTLLIA  
LAVYSLGRLVSRGQGTAEGRKQHIAETESPYQELQGQRPEVYSDLNTQRQYYR

- NCBI's Protein Sequence:

"MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGVLAGIVLGDVLVLTLLIALAVYSLGRLVSRGQGT  
AEGTRKQHIAETESPYQELQGQRPEVYSDLNTQRQYYR"

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

Overall, I didn't face any significant challenges. The most time I had to spend on was organizing all of my codes from Lab 04 and putting them into a function. I was successfully able to complete this lab without any major complications.