

Transcription4

May 15, 2023

```
[1]: # Prompt the user to enter the input fasta file name

input_file_name = input("Enter the name of the input fasta file: ")
```

Enter the name of the input fasta file: CFTR_sequence.txt

```
[2]: # Open the input fasta file and read the DNA sequence

with open(input_file_name, "r") as input_file:
    dna_sequence = ""
    for line in input_file:
        if line.startswith(">"):
            continue
        dna_sequence += line.strip()
```

```
[4]: # Transcribe the DNA to RNA

rna_sequence = ""
for nucleotide in dna_sequence:
    if nucleotide == "T":
        rna_sequence += "U"
    else:
        rna_sequence += nucleotide
```

```
[5]: # Prompt the user to enter the output file name

out_file_name = input("Enter the name of the output file: ")
```

Enter the name of the output file: CFTR_RNA.txt

```
[9]: # Save the RNA sequence to a text file

with open(out_file_name, "w") as output_file:
    output_file.write(rna_sequence)
    print(f"The RNA sequence has been saved to (out_file_name)")
```

The RNA sequence has been saved to (out_file_name)

```
[7]: print(rna_sequence)
```

```
AUGCAGAGGUCGCCUCUGGAAAAGGCCAGCGUUGUCUCCAAACUUUUUUUCAGCUGGACCAGACCAAUUUUGAGGAAAGG
AUACAGACAGCGCCUGGAAUUGUCAGACAUUAUACCAAUCCCUUCUGUUGAUUCUGCUGACAAUCUAUCUGAAAAAUUGG
AAAGAGAAUGGGAUAGAGAGCUGGCUUCAAGAAAAAUCCUAAACUCAUUAUAGCCUUCGGCGAUGUUUUUCUGGAGA
UUUAUGUUCUAUGGAAUCUUUUUAUUAUUAGGGGAAGUCACCAAAGCAGUACAGCCUCUCUACUGGGAAGAAUCAUAGC
UUCCUAUGACCCGGAUAAACAAGGAGGAACGCUCUAUCGCGAUUUUAUCUAGGCAUAGGCUUAUGCCUUCUCUUAUUGUGA
GGACACUGCUCCUACACCCAGCCAUUUJGGCCUUAUCAUAUUGGAAUGCAGAUAGAGAAUAGCUAUGUUUAGUUUGAUU
UAUAAGAAGACUUUAAAGCUGUCAAGCCGUGUUCUAGAUAUUUAAGUAUUGGACAACUUGUUAUGUCUCCUUCCAACAA
CCUGAACAAAUUGAUGAAGGACUUGCAUUGGCACAUUUCGUGUGGAUCGCUCCUUGCAAGUGGCACUCCUCAUGGGGC
UAAUCUGGGAGUUGUUAACAGGCGUCUGCCUUCUGUGGACUUGGUUCCUGAUAGUCCUUGCCUUUUUCAGGCUGGGCUA
GGGAGAAUGAUGAUGAAGUACAGAGAUCAAGAGCUGGGAAGAUCAGUGAAAGACUUGUGAUUACCUCAGAAAUGAUUGA
AAAUAUCCAAUCUGUUAAGGCAUACUGCUGGGAAGAAGCAAUGGAAAAAUGAUUGAAAACUUAAGACAAACAGAACUGA
AACUGACUCGGAAGGCAGCCUUAUGUGAGAUACUUAUAGCUCAGCCUUCUUCUUCAGGGUUCUUGUGGUGUUUUUA
UCUGUGCUUCCCUAUGCACUAAUCAAAGGAUCAUCCUCCGGAUUUAUUCACCACCAUCUCAUUCUGCAUUGUUCUGCG
CAUGGCGGUCACUCGGCAAUUUCCUGGGCUGUACAAAACUUGUAUGACUCUCUUGGAGCAAUAAAACAAAUACAGGAUU
UCUUAACAAAAGCAAGAAUAUAAGACAUUGGAAUAUAACUUAACGACUACAGAAGUAGUGAUGGAGAAUGUAACAGCCUUC
UGGGAGGAGGAUUGGGGAAUUAUUGAGAAAGCAAAACAAUAAACAAUAGAAAAACUUCUAAUGGUGAUGACAG
CCUCUUCUUCAGUAAUUCUCACUUCUUGGUACUCCUGUCCUGAAAGAUUAUUAUUCUAAAGAUAGAAAGAGGACAGUUGU
UGGCGGUUGCUGGAUCCACUGGAGCAGGCAAGACUUCACUUCUAAUGGUGAUUAUGGGAGAACUGGAGCCUUCAGAGGGU
AAAAUUAAGCACAGUGGAAGAAUUCUUCUGUUCUAGUUUCCUGGAUUAUGCCUGGCACCAUUAAGAAAAUAUCAU
CUUUGGUGUUUCCUAUGAUGAAUAUAGAUACAGAAGCGUCAUCAAGCAUGCCAAUAGAAGAGGACAUCCUAAAGUUUG
CAGAGAAAGACAAUAUAGUUCUUGGAGAAGGUGGAAUACACUGAGUGGAGGUCAACGAGCAAGAAUUCUUAAGCAAGA
GCAGUAUACAAAGAUUGCUGAUUUGUAUUUAUAGACUCUCCUUUUGGAUACCUAGAUGUUUUAACAGAAAAAGAAUAUU
UGAAAGCUGUGUCUGUAAACUGAUGGCUAACAAAAACUAGGAUUUUGGUCACUUCUAAAAUGGAACAUUUAAGAAAGCUG
ACAAAUUAUUAAUUAUUGCAUGAAGGUAGCAGCUAUUUUUAUGGGACAUUUUCAGAACUCCAAAAUCUACAGCCAGACUUU
AGCUCAAAAUCUAGGGAUGUGAUUCUUCGACCAAUUUAGUGCAGAAAGAAGAAAUUCAAUCCUAAACUGAGACCUUACA
CCGUUUCUCAUUAAGAAGGAGAUGCUCUCCUGUCUCCUGGACAGAAACAAAAAAACAAUCUUUUAACAGACUGGAGAGUUUG
GGGAAAAAAGGAAGAAUUCUUAUUCUCAAUCCAAUCAAUCUUAUACGAAAAUUUCCAUUGUGCAAAAGACUCCCUUACAA
AUGAUGGCAUCGAAGAGGAUUCUGAUGAGCCUUUAGAGAGAAGGCUGUCCUAGUACCAGAUUCUGAGCAGGGAGAGGC
GAUACUGCCUCGCAUCAGCGUGAUCAGCACUGGCCCCACGCUUCAGGCACGAAGGAGGCAGUCUGUCCUGAACCUGAUGA
CACACUCAGUUAACCAAGGUCAGAACAUUCACCGAAAGACAACAGCAUCCACACGAAAAGUGUCACUGGCCCCUCAGGCA
AACUUGACUGAACUGGAUAUAUAUUAAGAAGGUUAUCUCAAGAAACUGGCUUGGAAUAAGUGAAGAAAUUAACGAAGA
AGACUUAAGGAGUGCUUUUUUGAUGAUUUGGAGAGCAUACCAGCAGUGACUACAUGGAACACAUACCUUGCAUAUAUA
CUGUCCACAAGAGCUUAAUUUUUGUGCUAAUUUGGUGCUUAGUAAUUUUUCUGGCAGAGGUGGCUGCUUCUUGGUUGUG
CUGUGGCUCUUGGAAACACUCCUCUUAAGACAAAGGGAUAGUACUCAUAGUAGAAUAACAGCUAUGCAGUGAUUAU
CACCAGCACCAGUUCGUUAUAUGUGUUUACAUAUACGUGGGAGUAGCCGACACUUGCUUGCUAUGGGAUUCUUCAGAG
GUCUACCACUGGUGCAUACUCUAAUCACAGUGUCGAAAAUUUACACCACAAAAUGUUAUUAUUCUGUUCUUAAGCACCUC
AUGUCAACCCUCAACACGUUGAAAGCAGGUGGGAUUCUUAUAUAGAUUUCUCCAAAGAUUAAGCAAUUUUGGAUGACCUUCU
GCCUCUUAACCAUAUUUGACUUAUCCAGUUGUUAUUAUUGUGAUUGGAGCUAUAAGCAGUUGUCGCAGUUUUACAACCCU
ACAUCUUUGUUGCAACAGUGCCAGUGAUAGUGGCUUUUAUUAUGUUGAGAGCAUAUUCUCCAAACCUCACAGCAACUC
AAACAACUGGAAUCUGAAGGCAGGAGUCCAAUUUUCACUCAUCUUGUUAACAAGCUUAAAAGGACUAUGGACACUUCGUGC
CUUCGGACGGCAGCCUUAUUUGAAACUCUGUCCACAAAGCUCUGAAUUUAUACUACUGCCAACUGGUUCUUGUACCUGU
CAACACUGCGCUGGUUCCAAAUAGAGAAUAGAAUAGAUUUUUGUCAUCUUCUUAUUGCUGUUAACCUUAUUUCCAUUUUA
ACAACAGGAGAAGGAGAAGGAAGAGUUGGUUAUAUCCUGACUUUAGCCAUGAAUAUCAUGAGUACAUUGCAGUGGGCUGU
AAACUCCAGCAUAGAUGGGAUAGCUUGAUGCGAUCUGUGAGCCGAGUCUUUAAGUUAUUGACAUGCCAACAGAAGGUA
AACCUACCAAGUCAACCAAAACCAUACAAGAAUGGCCAACUCUCGAAAGUUAUGAUUAUUGAGAAUUCACACGUGAAGAAA
```

GAUGACAUCUGGCCUCAGGGGGCCAAUGACUGUCAAAAGAUCUCACAGCAAAAUACACAGAAGGUGGAAAUGCCAUUU
AGAGAACAUUUCCUUCUCAAUAAGUCCUGGCCAGAGGGUGGGCCUCUUGGGAAGAACUGGAUCAGGGAAGAGUACUUUGU
UAUCAGCUUUUUUGAGACUACUGAACACUGAAGGAGAAAUCCAGAUCAUGGUGUGUCUUGGGAUUCAAUAACUUUGCAA
CAGUGGAGGAAAAGCCUUUGGAGUGAUACACAGAAAGUAUUUAUUUUUUCUGGAACAUUUAGAAAAAACUUGGAUCCCUA
UGAACAGUGGAGUGAUCAAGAAAUAUGGAAAGUUGCAGAUAGGUUGGGCUCAGAUUCUGUGAUAGAACAGUUUCCUGGGA
AGCUUGACUUUGUCCUUGUGGAUGGGGGCUGUGUCCUAAGCCAUGGCCACAAGCAGUUGAUGUGCUUGGCUAGAUCUGUU
CUCAGUAAGGCGAAGAUCUUGCUGCUUGAUGAACCCAGUGCUCAUUUUGGAUCCAGUAACAUAACCAAUAUUAGAAGAAC
UCUAAAACAAGCAUUUGCUGAUUGCACAGUAAUUCUCUGUGAACACAGGAUAGAAGCAAUGCUGGAAUGCCAACAAUUUU
UGGUCAUAGAAGAGAACAAAGUGCGGCAGUACGAUUCUCCAGAAACUGCUGAACGAGAGGAGCCUCUCCGGCAAGCC
AUCAGCCCCUCCGACAGGGUGAAGCUCUUUCCCCACCGGAACUCAAGCAAGUGCAAGUCUAAGCCCCAGAUUGCUGCUCU
GAAAGAGGAGACAGAAGAAGAGGUGCAAGAUACAAGGCUUAG

[]: