1. **Work with this unknown nucleotide sequence:**

 901 ctggagagca cgggtttgtt gtcttcactc tgggctccat ggtgtcacag ctacccgaag

 961 ccaaagccag agagttcttt gaggcattta gacagatacc tcagagagtg ttgtggagat

 1021 acactggacc agtcccagaa aatgcaccaa agaatgtcaa attaatgaaa tggctgccac

 1081 agaatgacct cttgggccat cctaaggtta gggcttttgt tacacatggt ggatcacatg

 1141 gaatctatga aggaatctgt aatggagtgc ctatggtgat gcttcctctg tttggagacc

 1201 aaggggataa tgctcagcgt ttggtgtctc gaggagtcgc agaaagcctg actatctatg

1. Rewrite this sequence in FASTA format.

>filtered DNA sequence consisting of 360 bases.

ctggagagcacgggtttgttgtcttcactctgggctccatggtgtcacagctacccgaag

ccaaagccagagagttctttgaggcatttagacagatacctcagagagtgttgtggagat

acactggaccagtcccagaaaatgcaccaaagaatgtcaaattaatgaaatggctgccac

agaatgacctcttgggccatcctaaggttagggcttttgttacacatggtggatcacatg

gaatctatgaaggaatctgtaatggagtgcctatggtgatgcttcctctgtttggagacc

aaggggataatgctcagcgtttggtgtctcgaggagtcgcagaaagcctgactatctatg



1. Identify what it encodes and which organism it originates from.



1. Manually design primers that will amplify the CDS of the identified sequence, ensuring that their Tm does not exceed 60°C.

CDS: 

F: ATG GCT TGC TTG TGG TTG CT



R: TTA TGC CTT CTT CTT CTT CTG AG

Reverse complement sequence:





1. Translate the CDS into protein and write down the sequece.



1. Does the corresponding protein contain any transmembrane regions?

Yes, one



**B)** **Download the protein sequence Q0VC44.**

1. What does this sequence encode, and from which organism does it come?



1. Downlaod the FASTA format of the protein sequence.

>sp|Q0VC44|SPXN\_BOVIN Spexin OS=Bos taurus OX=9913 GN=SPX PE=3 SV=1

MKGFKSLVVMTLTLFLVFSFMGNCNSAPQRLFERRNWTPQAMLYLKGAQGRRFLSDQSRR

KDLSDRPPLERRSPNSQQLTLPEAAAVLLAFLQKPQEAGDENLDQTRFLEDSLLNW

1. Compare the protein sequence with the corresponding human homolog—how similar are these proteins?

Human sequence:

>sp|Q9BT56|SPXN\_HUMAN Spexin OS=Homo sapiens OX=9606 GN=SPX PE=1 SV=1

MKGLRSLAATTLALFLVFVFLGNSSCAPQRLLERRNWTPQAMLYLKGAQGRRFISDQSRR

KDLSDRPLPERRSPNPQLLTIPEAATILLASLQKSPEDEEKNFDQTRFLEDSLLNW





1. What is the length of the coding sequence (CDS) of the corresponding gene?







**CDS length: 420-70+1=351**

1. Will the corresponding nucleotide sequence (CDS only) be cleaved by the restriction enzymes BamHI, EcoRI, or KpnI?







