**Version A1**

1. **Work with following sequence:**

GCCTTCATCGAGTTTAAGAATTTCAAAAAGGAAAAGGAAACCCCCTTTAACCAAGTCCCC

ATACTCGAAATGGACAACATCATATTTGCCCAAAGTCAAGCCATAGTAAGGTACCTCTCC

AAGAAATACAAAATAAGTGGGAACAGCGAGCTGAACGAATTTTACGCCGACATGATTTTT

TGTGGAGTACAAGACATACATTATAAATTTAACAATACAAATTTATTTAAGCAGAATGAG

ACCACCTTTCTTAATGAGGAGTTGCCGAAGTGGTCAGGTTACTTTGAGAATATTTTGAAA

CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG

GCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGA

TAACAATTTCACACAGGAAA

* Find out if there is any vector contamination.



* If yes, get the “purified” sequence and write it down in FASTA format



>purified sequence

GCCTTCATCGAGTTTAAGAATTTCAAAAAGGAAAAGGAAACCCCCTTTAACCAAGTCCCC

ATACTCGAAATGGACAACATCATATTTGCCCAAAGTCAAGCCATAGTAAGGTACCTCTCC

AAGAAATACAAAATAAGTGGGAACAGCGAGCTGAACGAATTTTACGCCGACATGATTTTT

TGTGGAGTACAAGACATACATTATAAATTTAACAATACAAATTTATTTAAGCAGAATGAG

ACCACCTTTCTTAATGAGGAGTTGCCGAAGTGGTCAGGTTACTTTGAGAATATTTTGAAA

* Would be the purified sequence cut by restriction endonucleases KpnI, HpaII or RsaI?





* How long is the longest peptide after translation? In which open reading frame (ORF) is it encoded?



Or if you select “any sense codon” (which is better if we don´t know if the sequence is full length.



* From which organism does this sequence likely come from?

=identification of unknown sequence-BLASTn



1. **Download the sequence of the accession number: NM\_000771.4**
* How long is the coding sequence and how long will the translated protein be?



* Design primers for the detection of this gene for amplification of the product no longer than 500nt.



* How many prolines does corresponding protein have?



* How many times would be the sequence cut by pepsin (pH1.3)?



* Compare the corresponding protein with protein with accession number: P10632. What is the level of identity?

