

# BIOINFORMATIKA V PRAXI

## CVIČENÍ 3 – DRUHÁ ČÁST

### IDENTIFIKACE GENŮ, PROTEINŮ A JEJICH FUNKCE

#### STUDIJNÍ MATERIÁLY

Studijní materiály předmětu C2130 Úvod do chemoinformatiky a bioinformatiky, přednáška **Predikce genu, Sequence-evolution-function: Computational Approaches in Comparative Genomics**.

#### CHYBY PŘI PREDIKCI GENŮ

Velmi zjednodušený přístup k predikci prokaryotických genů (genem je nejdelší ORF) vede k chybám, ale jejich množství je poměrně malé. Chyby mohou také vznikat při sekvencování DNA. Přidání/odstranění startovního a/nebo stop kodonu může vést ke zkrácení, prodloužení nebo úplnému vynechání genu.

#### ÚKOL 1 – příklady chyb vzniklých při sekvenaci

Pomocí predikčního programu **GeneMark** (<http://exon.gatech.edu/GeneMark>) identifikujte geny v sekvenačních výstupech a porovnejte je s původní sekvencí z databáze (část genomu *E. coli*). Určete, k jaké chybě došlo. Pro porovnání rozdílně predikovaných genů využijte program **Align** (<http://www.ebi.ac.uk/Tools/emboss/align/>).

#### SEKVENCE *E. COLI*

```
TTAAGAAATTCGGTATCAACTTCGAGGCCCTTTCAGGTACCGTGGCGTATTCGGGGCGATAGATTTGTATTTGTAGCCGCAACTG
CGATCGCCCCGGCGAAAAACGCAGATCCGGCGGTACTTTGTGAAATCACCGTTTCAATCCAGCCAGCGTGCCGCGTGAACGAAACGAT
GGCGATCGTACAGTGTGCGGATCCAGACGTTTTCGCGCTGGAACCGGAGAAATCGACATCTGAGCCCCAGGCCGTGTGGAGTAGTCGATAGA
GTAGCGTTGCGAGTGCGCCCTGGTTGGCATCAGGCCACCACGAGAACCGGTGCGGCTAATCATACCCCAGGATAAAACAGCATCGTGGTGTG
GTAATTTACCCCTGAGTAAAGTGGTCGAGACTCCAGCGCAGGTTAATGGCAGCGTGCCAGCCGCTGGAGAGATCCAGTAGCGAGAAGCCACCA
CGGTAGTGGAGTCAGATTCGGTATCGTTTCAGGTTCAGTGCCTTAAACCGCCCTGCACCAATAATATTGTTCCAGTGGATTCTTCAGCAGCGG
CATTTTATAGCTGAAGTCGAGGGTCTGTTCCGGCGCGAAATACTGGTACTGGTGGTTCAGACTGTGACCATAAGAGTTTCCACGGCTTTTTC
CAGTAGCTTTACGCGCGGTCCACGTCGGTAGAGTAACCGACCCCGGTTTCGATGGTGTTCAGTTCGCGCGGAAACCACGCCCGTCAAGG
GTAATACTTTTCGTTTCGCGCGCTTTATCAAATGTGGAGCCACTACGACCGAGTTAAACCAGCCGGTAGCAGAAAGTCGACGGTTAAGTTCGCG
CAGATCTTTCGATTCGATCATCGCCCTTTTAAACGGCACCAGATTTGACGTTATTCATCGCGGATTTGTGATCCTTCAAAGGTCACATGC
CCAAAGCGGTAACGTTCCGCACTGTTATAATCAATATCCAGAAGGCTTTATGCAGGCCGAGCGCAATGCCAGCTGCGCTTTGGTAAATTCGC
TATCGAAATAACCTTTAGCAAGCAATGCTGGTTAAGGACTTTTGAATTTTCATAATCGCCCTGGTTCAGTACCGTGCCAATAGCCGGGCG
AGTATCGAGCAATTTCAATAGTCTTTATCGGTCGCGCGCGCGCGCAATACCACATCGGTGCCGCCAATTAACACCGGCACGCCCTGGCGTG
ACTTTGGCGATCAATACCTGCCGCCCTTTCTTTGGCGGTGGACGGAGATCAAATTCATGGTTCGGTGGTAATAACCCAGCGCTTCAGACCTT
CGCGGATGGCATCATCGACGCTGCGCGAAAGCGACGGTCTGGCGTCACTTTCATCACTTTCATCGTAGAAAGTGCAGCAGAACGTTCTTTTC
CAGCTGTCCCATAACCCCTCGACTGTAGACGGACGTTTCGCGCGGACGGCAGATCCGTTAAGCAGAGTAAGCTTACACAGCATAACTGTCCG
ATATAGCGCATATTTCTCCTGAATATCCTTTTTCTTCCGCCCCGGAACCGGTTAAACCGGCTTAAACAAAAATCCAGTAATATGGATTAAGAA
CAGACTAAACCCAAATATTTCTATGTTTTACTTTAGACCTATTCACGGTGGTATTGTTGTGCAATACGCCCTTGTGTACAACCTTAACCCC
AATGACCGAATTTTCGGGAGAGCGACACCATGAGTTTATTTGATAAAAAGCATCTGGTTTTCCCGCCGATGCCCTGGACGTGAACACCCCG
ATGCCCGTAGCCAGCTGCATGCGGTCAACGGTCACTCAATGACCAATGTACCTGACGGAATGGAGATTGCCATTTTGGCATGGGTTGTTTCT
GGGGTGTGGAGCGTCTGTCTTGGCAGTTACCCGCGGTTTACAGCACCGCCGACGGCTATACCGCGGCTATACGCCAAATCCGACTTATCGGGA
AGTGTGCTCCGGTATGCGGATGACCGAAGCGGTACGCATTGTTTACGATCCTTCTGTGCATCAGTATGAGCAGTTGCTACAGGATTTTGG
GAGAATCAGATCCCGCCAGGGAATGCGTCAGGGCAATGACCGCAATGACCGCATCGAGTATCGTTTCAGCGATTATCCGCTGACCCAGAACAGGATG
CCGACGCTCGCGCAGTCTGGAACGTTTTCAGGCGCGATGCTTCCGCGCATGACGATCGTCACATCACCCAGGAAATCGTAAACGCCACACC
GTTTTATATGCGGAAGATGACCACCAGCAATATCTGCATAAAAACCCGATATGGTTACTGTGGAATTGGCGGAATTGGCGTCTGTGCCACCG
GAAGCATAGCGTTACGGTACAAATGTAGATTGTTGATAAAGTGCCTTTGTTTATGCCGAATGCAGCGTGAATGCCTTACCAGCCTACAAAA
TCGTCCAAATCAATATATGCGAGGACTGCGTAGGCCATGACCGCAATGACCGCATAGGCAATTTGCAATTTACGAGTATGCAATTTTAAAC
CTCTGGCGACTTTACAGTACCTTACGCTATAC TAGCCACTGAAAATGCCGGATCACTTTCTTCGAATCGGCTTTCAATGTGATTTACACAAA
TTAATCAACTTCCCTTCCGAGGATCTGGCCTGAAAGTTCGGATAAGATATGTTAAACAGTATTTTAGTCATACTCTGCTTGATCGCTGTAAAGT
CGGTTCTTCTCGATGCTGAGATCTCACTTGCCTCCTCAGCAAAATCAAATTAACCTGCTGGCTGATGAAGCAATATAAATGCCAACCGG
TTCGTAATATGCAAGAAATATGCGGATGATGCGGATGAGTGAAGCCGCTGCTGCTGAAACCGAGTTGCGATTTCGCGCGGATCGTCGGTATGC
GGCATTCTTCCAGCTTTTACAGCCTGTTCTCCGCTATATGTCGGCAGAACCTCTGAGCAACTGAGCTTTATCTCTCTTTCTCGTTAGTG
ACTGGCATGTTTATCCTGTTTGGGATTTAACCCGAAACGCATCGGTATGATTGCGCCAGAAGCGGTGGCTTTGCGTATCATCAACCCGATGC
GCTTTCGCTGTACGTTGCAACCCCGCTGGTGTGGTCTTCAACGGCTGGCGAACATGATCTCCGATTTTCAAACCTGCCAATGGTACGTA
AGATGACATCACTCTGATGACATCTACCGGATGAGTGAAGCCGCTGCTGCTGAAACCGAGTTGCGATTACGTAACAGGAAACAGCAGGATTTAAAC
GTCTTTGAGCTGGAATCCCGTACCGTCCGCTTCAATGACACCGCTGAAAACGTGATTTGGTTTGTATCTCCACGAAGATGAGCAAAGCCTGA
AGAATAAGGTGGCGGAACATCCGCACTTAAGTTCCTCGTCTGTAATGAAGATATTGACCACATCATCGGTTATGTCGATTCTAAAGACCTGCT
```

GAACCCGGTGTGGCTAACCAAAGCCTGGCACTGAACAGCGGGCTACAAATTCGCAACACGCTGATTGTGCCGGATACGTTAACCCCTTCAGAG  
GCGTTGGAAAGTTTAAACCAGCTGAAAGACTTCGCGGTGATCATGAACAGTACCGCTGGTGGTGGGGATCATCACCCCTCAACGACGTGA  
TGACCACGCTGATGGGCTCGTTCGGTTCAGGGCTGGAAGAGCAGATTTGTCGCCCGTGTAGAAACATCATGGCTGATGACGGCGGCCACCC  
AATTGACGACGTCATGCGCGTGTGGATATTGACGAGTTCGCCGAGTTCGGGCACTACGAAACCATCGGAGGCTTTATGATGTTTATGCTGCGT  
AAGATCCCGAAACGCACCGATTTCGGTGAATTCGCCGGCTACAAATTTGAAGTGGTGGATATCGATAACTACCGCATCGACCAGCTGCTGGTGA  
CCCGGATCGACAGCAAGGCCACCGCCCTTTCCGCAAACTGCCTGACGCTAAAGATAAAGAAGAAAGCGTTCGCGTAACCCCCAGAAACATCAAC  
GGCTCCTGAATCAGGACCGCTTATTACTGCATAGCACTTTGGTTAAGCCATCTGTTTGCAGATAAAGCATTCAATTTTTCATATGATTTGTCGGCTGA  
TCACTGACAAATGCAGCTTATCCTTCACTTTTGGGATAAAGATCTTACCCTTATCAAATTCAAAAGCGCCAACGTCCTTGATGTATAACCGTCC  
ACGGAACAGGATCTTTACGTACTTCGCCACCTGAAGTGGGTTTACCGTTGAAAATTTTCATTTGTTATCTCCTGCTGAGTATTACGCCCTGTGC  
GGTGCCACAATCGGCCAACTATTATGAGGCGCAATTTAATGCCTAGTGACTATAGACTATCCGGGCAATGTTTCCACCGCGTATAACTTTT  
TTTACCTAAAGGTTAACCAATATTACAGAAATATCTTTTACCGGAGCGCGGCTTCAGTATAAAGCATTCAATTTTTCATATGATTTGTCGGCTGA  
CCGCAAACTGGCATCACACTTGGCGGAAATTCGATAAATAGCACATATGATTAATAACTCAGACCCAAGTGGTTCGGATCACCTGCATATCATAAG  
AAGGAAACACCATGACCTACGCAAGATTTCTGGCACTACCTGCTGCTGTTGCGGATGATGGCTTCCGCACATCAGTTGAAAACCGGTCAGCG  
AGTGCCGCGGATTGGCATCACCGATCGTGGCGAGTGGTGTGATAAAGATCAGTTTAGCTACAAAACCTGGAACAGCGCGCAGTTAGTGGGA  
AAAGTCGCACTGACCAACTGATTTGCTGGTTCGCACTCTGCAAAAAGAGAAAACCGCAGCGCTGATTGAAGCAGTAAACTCAGCGAACTCCGC  
ACGATCGTTACCAGACCACCACCTTTGTTAACACCGACGACGCAATTCGGGGTTCGGCATGTTTGTGCGCAGCAGTCTGGAGAGTAATAAAAA  
GCTTTATCCCTGGTGCAGTTTTATTGTTGATAGCAATGGCGTTCGACGCGGCTGCTGGCAGCTGGATGAAAAAGTTCGGCTGTAGTGGTGTG  
GATAAAGACGGTTCGCTGCAATGGGCCAAAGACAGGGCGCTTACTCAGGAAGAGGTGCAGCAAGTGTGACCTGCTGCATAAATTAATTAATA  
AATGACACTCTGACCAAGGATTCAGGAACGACTCGCGCGGATTCGCGGAGTAACTCCCGAGCTTAACTCCGAGTTCAGCGGTTTACCGGATG  
GCGGCCACAGCATGTCCAGCGCGGTGTCCCAATATTCGTTGGTCCGAAAGCGCGGGTACAGCTGCGCCTGTCTTCCGCCACCAGGCAGAAAT  
TCAGCGAAGAGCGGATGACGTTGCTGATGTTCCGCAAGCTGTTGCAGATACCTTTTTCAGCTCCGCATCTGCATGGGAACGGCTGATAACCAC  
CAGCGCGGGCGCGCATCGCGACCTGAATCTGCTTGGCAGCACCGGACTCTTCTTCCAGGCTTGCCTTTCGCCGCGCTGTACATTAACGTTT  
ATTACCGCGCGTACACCACTAAATCGGTTTGCATGTTGCAATGTTAAGCGTGAATTCGCGGTAATTCGCGGTAATTCGCAATAAATCACT  
TAGTACCATCCAGCGGGTCTACCAGCCAGTAAACGCTGCCAGTGTGACGGACTTCCCAACCGGGAGGATCTTCTTCAGAAAGGACCGGAATATC  
CGGTGTACGCTACGTAACCCGTCATGATAACGGTGTGAGCGGCAATATCCGCTGCCGTTACCGGAGAAATGTCGCTTTGCTGACGAGCTCC  
ATCGGTTTCGTTCCGTCGTAGACCTGCATAATGGCATCGCTGCATTCGCTGCAAGCTGGCATACTTGATCTAACATTTCTCCACCTCGTCTCT  
GTGACGCGTCTTAACTTATGTTTACTTATAACCTATCGTTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG  
GCATTTCTTCACTGTTCTATAAGTAAGACGTTTATTTCTCTTTTCTTTTCGATTTCCCGATGATAAAGGATGTCCTGATGATTAAGTTTAGCG  
CAACGCTCCTGGCCACGCTGATTGCCCGCAGTGTGAATGCAGCGACGGTGCATCTGCGTATCATGAAACCACTGATCTGCATAGCAACATGAT  
GGATTTGATTTTACAAAGACACCGCCACGAAAAATTCGGACTGGTACGTACGGCAAGCCTGATTAACGATGCCCGCAATGAAGTGAAAAAAC  
AGCTACTGGTCGATAACCGCGATTTGATTCAGGGGATCCGCTGGCGGATACATGTCGGCGAAAGGATTAAGACAGGTGATGTTTTCATCCGG  
TTTATAAGGGCGTGAATACGCTGGATTATACGGTTCGGTACACTCGGCAATCATGAATTTAACTACGGTCTGGATTACCTGAAAAATGCGCTGGC  
GGGAGCGAAATTCCTTATGTAATGCCAACGTCATTGACGCCAGAACCAAGCCAAATGTTTACACCGTATTTAATTAAGATACCGAAGTG  
GTGCGATAAAGACGGAATAAAGACAGCGCTGAAGATTTGGCTATATTTGGCTGCTGCCCGCAAAATCATGGGCTGGGATAAAGCTAATTTATTCGG  
GAAAAGTGACGGTGAATGATATTACCGAAACCTGCGCAAAATACGTTGCTGAAATGCGCGAGAAAGGTCGGATGTCGTTGCTGTTTGGCGCA  
TTCCGGGCTGTCTGCCGATCCGTTATAAAGTAATGGCGGAAAACTCAGTTTATTACCTCAGTGAATTCGGGCGTTAACGCCATTATGTTTGGT  
CATGCTACGCGGTTTTCCCGGTAAGATTTTGTGATATCGAAGGGGTGATATCACCAAAGGCACGCTGAATGGTGTTCGGGCGGTAATGC  
CGGGCATGTGGGCGCATCATCTTGGGGTGGTGCACCTTACACTCAGTAATGACAGCGGTAATGGCAGGTGACCGAGCGGAAAGCGGAAAGCTCG  
GCCGATTTACGACATCGCAATAAATAAATCCCTCGCGGGGAGACAGCAAGCTGGTAGAAACACTCAAAGCCGATCAGATGCCACAGCCGAC  
TTCGTCAGCAAGCAATCGGTAATCTGCCGACAATATGTATAGCTATCTGGCGCTGGTGCAGGACGATCCGACCGTGCAGTGTGAACAACG  
CGCAAAAAGCGTATGTCGAACATTACATTCAGGGCGATCCGGATTCGGCAAACTGCCGCTGCTTTCAGCTGCCGCACCGTTTAAAGTTGGTGG  
TCGCAAAAATGACCCCGCAAGCTATGTGGAGGTGGAAGAAAGCGACTGACCTTCCGTAATGCCCGCATCTTTATCTTACCCCAATACGCTG  
ATTGTGGTGAAGCCAGCGTAAAGAGTGAAGAGTGGTGGTGGTTCGCGCGGACAGTTTAAACGATGATCTTAAACAGCAGCAAGAAACCGC  
AATCACTCATCAACTGGGATGGTTTCCGCACCTATAAATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG  
TGACGGGAGTGCAGATGATTAATGCCAATGCGGAAAGGATTAAGAACCTGACCTTTAATGGCAAGCCGATTGATCCGAAACGCGATGTTCCCTC  
GTTGCCACCAATAACTATCGCGCTTACGGCGGCAAAATTTGCCGCTACGGGCGACAGCCATATCGCTTTTGCCTTACCCGGATGAGAACCGCTCGG  
TGCTGGCAGCTGGATTGCTGATGAGTGAACCGTTCGGGGGAAATTCACCCGCGCGAGATAACAACCTGGCGTTTAGCACCGATAGCTGGCGA  
TAAGAACTGGATATCCGTTTTGCAAACTTCCGTCAGATAAAGCCGCGGCTTATTAAAGAGAAAGGCAATATCCGATGAATGAATGAGTGGCG  
ACCGATGATATCGGGTTTGGCATTTATCAGGTGGATTTGAGTAAGTAAACACTTCTTTTTCGGCTATAAATCATCAACCGCATCCGGCATT  
ATTGGCGGATGCGATGCTGGCGCATCTTATCCGCCCTACAAGCCATGACCGTAGACCCAGATAAGCTCAGCGCATCCGGCAGTTATGCCGACG  
TTCATCCCGCACCGCAACCTCGGGCAATTCACACTCACTCAGCTTCGCGAGTGCAGCAACCTTTTCGCTCACCTGCTCCGCCAGCGCGGT  
AGCTATATTTCCACATCGCGCGCACCCAGGATACCGGATACCGTTACGGAAGAAACCATCTGTTTCAACGCTGTAAACGATGCGCAAGCATCT  
TTTCGCTACCCCGCGGATTTTGGCGCGAGGTGCTAAAGCGATGAGTACCTTCCGCTAGCGCCACCAGAATCAACACCCCCCAACGGCTGGT  
GACGTGTTTCAACACCTCGCGGACGGGCACTGTTCCGCAAGAGATTAACCTCTTTCAGTTGTTGCGACAGGCTAACCTGACTCATTTTATAC  
TTACTTTTGTGCGCACTTACTAAAGTAAGTTTAGTGTGTTAGCGTATTTAAACACAAGCAAAAACGATGGAGACTTCCCATGATCGCTATTA  
CTGGTCCCACTGGCAACTTGGTCACTATGTTTGAATCCTTGAAGAAACCGTTCCTGTGACGCAAAATGAGGCTATGATGCTGATTAATCGCT  
AAAAGCCAGGCTCTGGCAGCACAGGCATTACCGTGCCTCAGGCTGACTACGGCGATGAAGCCGCATGACATCTGCATCTCAGGGAGTGGAA  
AAACTACTGCTGATCTTCCAGCGAAGTGGGTCAACGTGCCCGCAGCATCGTAATGTTATTAATGCCGCAAGGGCGGTGGTGTGAATTTA  
TCGCTTATACCAGCTGCTACATGCAGATACCTCCCGCTCGGCTCCGCAATGAGCAGATCGAGACGGAGAAAAATGTTGGTGTATTCTGGCAT  
AGCCGATTTCCCGCGGCACTGAAAAGCGTTCGGACTGCCCGGCACTGGCGGATATGCTGGCGGATTTGACGTTGGCGCATGAAAAGTGGT  
CTGTTTGAATAACAGCAAAACGTTAGCAAAATGATTGGCCGCCAACAGCAACGTTAGCCGAAAGCGTAAAGCACTTTTTTAAAGTAAATAACT  
AGTTAATTAAGTGGCATCTCCCGCATCTCTCTAATAATGACGGGATGCCGGGAGCAATCATGCTGCTTCTTGAACCTTCTTCTGACAGAC  
CAATGGATGCCAGTAATGATTAGCGCGTGTGCTGACGCCCTGTAGCAGGGTGTGATGTTGGGGCTTATTTTGTGCGGGCGTGTGATCGTACCGG  
AATACCCGGCGATGTTGCGAGTCGATGGGGCTTATCTGGCGTTAGGTTAATTCGCTGCCATTTGCCTGGCTGGGACGCGTGGCTGCTGCA  
GTTGGCGCTGGGCACTGGCTTACCGCTTATGCTCACTATGATGGCAACCTTACTTATTACTTCTGCTTCCGCTGCGGAGTCACTCAACCTACT  
GGCGCGCTGTTTCCAGATGATTATCGGCACCTGCCGTTGGTCACTTCTGCTTTGCAATCTGCTTTATAGCCAGCGCGACGGCAAACTCG  
CTGGGGAAAACTCGCCCGGCACTGATTTGATTTGGCATCGGCTGGCGGATGTTGAATATTGCTGAGTTAAACCACGGACTCCCGATTTTGA  
CTGGGACGTTTAACTCTGGCATCTGCTAGCGTTAGTTTCCGTTGCTGCTGGCATGTTGATGCTGCGCAACCGCTGGCTGGGAA  
AATCCCGCAACACATCCGATGATGTGGCGCAGCGCGCAGGCGTGGTTCAGCTACCGCTGGCTTCTCATCGGCTATCTGCTGCTGTTACTGGC  
TGAATATAACAACCGCGACTTCTCCTTACCTTTTGGCCCGCTCGCTGGTGTATTATGCTGATGTTGCGATAGCCGCTGCTTGTCTCATG  
GTTGGCGCACTGCTGGAACCTGCCAGCCAGCGATTACCGCAGTGGATTTCCGGGCGCTGATGTTTTCGAAACGCTGGCAGGTTTGTCTG  
TACACTTTTTTACTCCGACGAAATGCCCGCGCTAATGACGCTGAGCGGATTCGCGCTGTTAGTGTATGGCTGGTGTGCTTGGCTGAGAGCAA  
AACAGAAAAACCTTTAATGAATCTGCTCAGAAAAGTTGACAGCTGCGCAGTGAAGTAAATAAGCCTCTGCTACGTAAGGGTTTAGCTTTT  
CCTTAAAGATGCATTTAAAAATACATCTTATCTTATTAAGAATGAGGTATCAGCTATGGCTTATCGCGACCAACCTTTAGTGAACGCGCTCT

CTATCCCTCGCGCTTCGGCTCTGTTTCGTAAATATGATATGGATTACTGCTGTGGCGGTAAGCAGACCGCTGGCGCGCGGGCGGCACGTAAGA  
ACTGGATGTTGAGGTCATTGAAGCTGAACTGGCAAAGCTTGCTGAACAACCGATTGAGAAAGACTGGCGTAGCGCCCCGCTGGCAGAAATCATT  
GACCATATCATCGTGCGCTAC

## PŘÍKLAD 1

TTAAGAAATCCGGTTATCAACTTCGAGGCCCTTTCAGGTCACCGTTGGCGTATTTCCGGGCGATAGATTTGTATTTGTAGCCGGAATACTG  
CGATCGCCCCCGGCGAAAAACGCAGATCCGGCGGTACTTTGTGCGAAATCACCGGTTTCAATCCAGCCCAGCGTGCCCGGTGTAACGAAACGAT  
GGCGATCGTACAGTGTGCGGATCCAGACGTTCTGCGCCTGGAAAACGGAGAAATCGACATCTGAGCCCCAGGCGGTGTTGGAGTAGTCGATAGA  
GTAGCGTTGCGAGTTCGCCCTGGTTGGCATCAGGCCACCACGAGAACCGGTGCGGCTAATCATACCCCAGGATAAAAACAGCATCGTGGTGTG  
GTAATTTACCCCTGAGTAAAGTGGTCGAGACTCCAGCGCAGGTTAATGGCAGCGTCCAGCCGCTGGAGAGATCCCAAGTAGCCGAAAGCCACCA  
CGGTAGTGGAGTCAGATTCGGTATCGTTTCAGGTCAGTCCGCTTAAAACCGCCCTGCACCAAATAATATTGTTCCAGTGGATTCTTCAGCAGCGG  
CATTTTATAGCTGAAGTCGAGGGTCTGTTCCGGCGCGGAAATACTGGTACTGGTGGTCAGACTGTGACCATAAGAGTTTACACCGGCTTTTTT  
CAGGTAGCTTTTACGCGCGGTCCACGTCCTGAGAGTAAACCGCCGTTTCGATGGTGTTCAGTTTCGCGCGGAAACCACGCCCCGTAAGG  
GTAATACTTTTCGTTTCGCGCGCTTTATCAAATTTGGAGCCACTACGACCGAGTTAAACCAGCCGGTAGCAGAAAGTCGACGGTTAAGTTCCGC  
CAGATCTTTTCGATTCGATTAATTCGCCCCCTTTAAACCGCACAGATTTTGCAGGTATTTCATCGCGGATTTGTGATCCAAAGGTCACATGC  
CCAAAGCGGTAACGTTCCGCCACTGTTATAATCAATATCCAGAAAGGCTTTATGCAGGCCGAGCGCAATGCCAGCTGCGCTTTGGTAAATTCGC  
TATCGAAATAACCTTTAGCAACGCAATGCTGGTAAAGGACTTTTGAATTTTCATAATCGCCCTGGTTCAGTACCGTCCAAATAGCCGGGG  
AGTATCGAGCAATTTCAAATAGTCTTTATCGGTCGCGCGCGCGGCAATACCACATCGGTGCCGCAATTAACACCCGACGCGCTTGGCGGTG  
ACTTTGGCGATCAATACTTCGCCCCCTTTTGGCGGTCCGCGGAGCAATAACGCAATCGGTTGCGCGCTGGTAAATACCAGCGCTTTCAGACTT  
CGCGGATGGCATCATCGACGCGTGGCGGAAAGCGACGGTCTGGCGTCACTTTCATCACTTTCAATCGTAGAAAGCTGGCAGCAACGTTCTTTT  
CAGCTGTCCCGATAACCCCTCGACCTGTAGACGGACGTTTCGCGCGCAGCGCAGATCCGCTTAAGCAGAGTAAGCTTACACAGCATAACTGTCCG  
ATATAGCGCATATTTCTCTGAAATCTTTTCTTCCGCCCCGGAACCGCGCTTAAACCGGCTTAAACAAAATCCAGTAAATATGGATTAAGAA  
CAGACTAAACCTTACGATTAATTCGCTTATGTTTACTTTAGACTTTACCGTTCAGTATTTGTGTGCAAAATACGCCCCCTTACAGGATTTTTC  
AATGACCGATTTTCGGGAGAGCGACACCATGAGTTTATTTGATAAAAAGCATCTGGTTTCCCCCGCCGATGCCCTGCCTGGACGTAACACCCCG  
ATGCCCGTAGCCACGCTGCATGCGGTCAACGGTCACTCAATGACCAATGTACCTGACGGAATGGAGATTGCCATTTTGGGATGGGTTGTTTCT  
GGGGTGTGGAGCGTCTGTTTGGCAGTTACCCGGCGTTTACAGCACCCGCGCAGGCTATACCCGGCGCTATACGCCAAATCCGACTTATCGGGA  
ACTGTGCTCCGGTGATACGGGTATGCCGAAGCGGTACGCACTTGTTTACGATCCCTTCTGTCATCAGCTATGAGCAGTTGCTACAGGATTTTGG  
GAGAATCAGATCCCGCCAGGGAATGCGTCAGGGCAATGACCAGGCACGCAGTATCGTTTCAGCGATTTATCCGCTGACCCAGAACAGGATG  
CCGCGACTCGCGCCAGTCTGGAACGTTTTCAGCGCGGATGCTTGCAGCGGATGACGATCGTTCACATCACCACGGAATCGTAAACGCCACACC  
GTTTTATATGCGGAAGATGACCACCAGCAATATCTGCATAAAAACCCGATGGTTACTGTGGAATGGCGGAATGGCGGTCTGTCTGCCACCG  
GAAGCATACCGTTACGGGTACAATGTAGATTGTTGATAAAGTCCGTTTGTGTTATGCCGAATGACGCTGAATGCCCTTACCAGCCATCAAAA  
TCGTCCAAATTAATATATTGACAGGACTGCGTAGGCCATGATCGGCATAGCGCATCAGGCAGTATTGCATTTATCAGCAGTCTGAATTTAAC  
CTCTGGCGACTTTACAGTACCTTACGCTATACAGCCACTGAAATGCCGGATCACTTTCTTCGAATCGGCTTTCAATGPTATTTCACACAAA  
TTAATCAACTCCCCTTCGAGGATCTGGCCTGAAAGGTCGGATAAGATATGTTAAACAGTATTTTAGTCATACTCTGCTGTATCGCTGTAAGT  
GCGTTCTCTCGATTCAGATTCGAGACTCTCACTTGGCCCTCACGCAAAATCAAACTTAAACTGCTGGCTGATGAAGGCAATATAAAATGCCAACCG  
TTCTGAATATGCAGGAAATCCCGCATGTTCTTTACGGTGGTCCAAATCGGTCGTAACGCGAGTTGCGATTCTCGGGGATCGTCCGGTATGCG  
GGCATTCTCCAGCTTTTCACAGCCTGTTCTCCCGCTATATGTCGGCAGAACCTCTGAGCAACTGAGCTTTATCTCTCTTTCTCGTTAGTG  
ACTGGCATGTTTATCTGTTTGGCGGATTTAACCCGAAACCGATCCGGTATGATGGCCAGAACGCGGTGGCTTTGGCTATCATCAACCCGATGC  
GCTTCTGCTCGTGTGATTCAGACTTGCACCCGTTGCTGTTTCTTCAACGCAATGATCTTCCGATATTTTCCAACTTACAGCTTACAGCTAA  
AGATGACATCACTTCTGATGACATCTACCGGTTAGTGAAGCCGGTGGCTGGCGGGCGTGTACGTAACAGGAAACAGAGCTGATTGAAAAC  
GTCTTTGAGCTGGAATCCCGTACCGTTCGCTTCAATGACACCGCGTGAACAGCTGATTGGTTGATCTCCAGGAAGATGAGCAAAGCCTGA  
AGAATAAGGTGGCGGAACATCCGCACTTAAGTTCCCTCGTCTGTAATGAAGATATTGACCACATCATCGGTTATGTCGATTCTAAAGACCTGT  
AAGCAACACCATGACCTACCGAAGATTTGCGCACTGCAACAGCGGCTTACCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT  
GCGTTGGAAAAGTTTTAAACCGCAGGTGAAGACTTTCGCGGTGATCATGAACGAGTACGCGCTGGTGGTGGGATCATCACCCCTCAACGACGTGA  
TGACCACGCTGATGGGCGATCTGGTCCGTCAGGGGCTGGAAGAGCAGATTGTCGCCCGTATGAGAATCATGGCTGATTGACGCGCGCACCCC  
AATGACGAGCTCATGCGCGTGTGGATTTGACGAGTTCCCGCAGTCCGGCACTACGAAACCATCGGAGGCTTTATGATGTTTATGCTGCGT  
AAGTCCCGAAGACGCACCGATTCCGGTAAATTCGCCCGTCAAAATTCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
CCCGGATCGACAGCAAGGCCACCGCCCTTTCGCCAAAACCTGCCTGACGCTAAAGATAAAGAAAGAAAGCGTCCGTAACCCCCAGAAACATCAAC  
GGCTCCTGAATCAGGAGCGCTTATTACTGCATAGCACTTTGGTTAAGCCATCTCTGTTTGCAGACGCATAACCTGACGGTTAACTTCGGACA  
TCACTGACAAATCGAGCTTATCCTTCACTTTGGGATAAAGATCTTACCTTATCAAAATTCAAAAGCGCCACAGCTCTTGATGATAACCGCTG  
ACGGAACAGGATTTTACTGATTTTCGCAACTGAGAGTGGGTTGTACCTTGGAAATTTTCAATTTGTTTATCTCCTGCTGATTACGCTGCTGC  
GGTGCCCAATCGGCCAACTATTATGAGGCGCAAAATTTAATGCCTAGTGAATATAGACTATCCGGGCAATGTTTCCACCGCGTATAACTTTT  
TTTACCTAAAGGTTACAATTATTCAGAAATATCTTTTTACCGAAGCGCGGCTTTCAGTATAAGCATTCATTTTTCATATGATTTGTGCGCTTGA  
CCGCAAACTGGCATCACACTTGGCGGAAATTCGATAAATAGCAGATATGATTAAAACTCAGACCAAGTGGTCCGATCAGCTGCATATCATAAG  
AAGGAAACACCATGACCTACCGAAGATTTTCGGCACTCACCTGCCTGCTGTTGGCAGATGATGGCTTCCGCACATCAGTTTCGAAACCGGTCAGC  
AGTGGCCCGGATTGGCATCACCGATCGTGGCGAGTTGGTGTGATAAAGATCAGTTTACTGATAAAAACCTGGAACAGCGCGCAGTTAGTGGGA  
AAAGTGGAGTACTGCAACATATGCTGGTCCGACCTTGCAAAAGAGAAAACGCGCAGCTGATTGAAGCGATTAAATCAGCGAAGTTACCGC  
ACGATCGTTACCAGACCACCCATTTAATACCCGACGACGCAATTCGGGTTCCGGCATGTTTGTGCGCAGAGCTTGAGAGTAATAAAAA  
GCTTTATCCCTGGTGCAGTTTATTGTTGATAGCAATGGCGTGCACCGGTCAGCTGGCAGCTGGATGAAAAGTTCGCTGTAGTGGTGTGCTG  
GATAAAGACCGTTCGCTGATTCGGCCAAAGACAGGGCGCTTACTCAGGAAGAGGTTGCAGCAAGTATGAGACTGCTGCATATAAATTAATA  
AATAGACACTCTGAACCCAGGATTCAGGAACGACTCGCGCGGAGTGAATTCAGCGGTTTACCTGCCAGTCTGTAACGTTGCGCTCCGGCAGCT  
GCGGCCACAGACTGTCCAGCGCGGTTGCCAAATATTCGTTGGTCCGAAGCGCGGTTACAGCTGCGCCTGTCTTCCGCCACCGGAGCAAT  
TCAGCAAGAGCCGATGGACGGTCTGATGTTCCGCAAGCTGTTGCAGATCTTTTACGCTCCGATCTGCATGGGAACCTGATAAACC  
CAGCGCGGGCGCATCGCGGCAATCTGCTTGGCCACACCGCACTTCTTCTTCCAGGCTTGCCTTCTGCGCGCTGTACATTACGTTT  
ATTACCGCGCATACACCACCTAAAATCGGTTTGCATGGTCAATGAGCGCAATGTTAACGGTGAATTCGCCATTGCGTTTAAATAAATCTT  
TAGTACCATCCAGCGGGTCTACAGCCAGTAACGCTGCCAGTGTGACGGACTTCCCAACCGGGAGGATCTTCTTCAGAAAGGACCGGAAATATC  
CGGTGTACGCTACGTTAAACCTCATGATAACCGTGTGAGCGGCAATACTCCGTTGCCGTTACCGGAGAAATTTGCTGCTGACGACCTCC  
ATCGGTTTCTGCTCCGCTAGACTGATTAATGGCATTCGCTGACCTGCTGCAAGTCTCCGTCAGAGCTGGCATACTTGATCTAACATTTCTCCACTCGTCT  
GTGAGCGGTGTTAACTTATGTTTACTTATACCCATCGTTAATGAATGCGCAACTGTGATAGTGTATCATTTTCAAAGCGTAAAATTTGTG  
GCATTTCTCACTGTTCTATAAGTAAGACGTTTATTTCTCTTTTCTTTCGATTTCCCGATGATAAAGGATGTCCTGTGATGATTAAAGTTAGCG  
CAACGCTCCTGGCCACGCTGATGGCCGAGTGAATGCAGCAGCGGTCGATCTGCTGATCATGGAACCACTGATCTGCATAGCAACATGAT  
GGATTTGATTAATCAAAAGACACCCGCCAGGAAAATTCGGACTGTTGACGCTGATTAACGATGATTAACGATGATTAACGATGATTAAGTGA  
AGCGTACTGGTCGATAACGGCGATTGATTACGGGAGTCCGCTGGCCGATTACATGTCGGCGAAAGGATTAAGGACGAGTGTATCTCCGG  
TTTATAAGGCGCTGAATACGCTGGATTATACGGTCCGTACACTCGGCAATCATGAATTTAACTACGGTCTGGATTACCTGAAAATGCGCTGGC

GGGAGCGAAATTCCTTTATGTAATGCCAACGTCATTGACGCCAGAACCAACAGCCAAATGTTTACACCGTATTTAATTAAGATACCGAAGTG  
GTCGATAAAGACGGAAAAAACAACAGCGTGAAGATTTGGCTATATTTGGCGTCGTGCCCGCAAATCATGGGCTGGGATAAAGCTAATTTATCCG  
GAAAAGTGACGGTGAATGATATACCGAAACCGTGCAGCAATACGTGCCGTGAAATGCGCGGAGAAAAGGTGCCATGTCGTTGTCGTTCTGGCGCA  
TTCCGGGCTGCTGCGGATCCGTATAAAGTAATGGCGGAAAACCTAGTTTATTACCTCAGTGAATTCGCGGCGTTAACGCCATTATGTTTGGT  
CATGCTCACGCGGTTTTCCCGGTAAAGATTTTGCTGATATCGAAGGGGCTGATATCACCAAAGGCACCGTGAATGGTGTTCGCGCGGTAAATGC  
CGGGCATGTGGGGCGATCATCTTGGGGTGGTCGACTTACAACCTAGTAAATGACAGCGGTAAATGGCAGGTGACGCAGGCGAAAAGCGGAAGCTCG  
GCCGATTTACGCACATCGCAATAAAAAAATCCCTCGCGCGGAAAGACAGCAAGCTGGTAGAAAACACTCAAAGCCGATCACGATGCCACAGCCAG  
TTCGTCAGCAAGCCAATCGGTAATCTGCCGACAATATGTATAGCTATCTGGCGCTGGTGCAGGACGATCCGACCGTGCAAGTAGTGAACAACG  
CGCAAAAAGCGTATGTCGAACATACATTCAGGGCGATCCGGATCTGGCAAAACTGCCGGTGCCTTCAGCTGCCGACCGCTTTAAAGTTGGTGG  
TCGCAAAAATGACCCGGCAAGCTATGTGGAGGTGGAAAAGGCCAGTGGACCTCCGTAATGCCGCGGATCTTTATCTTACCCCAATACCGTGT  
ATGTGGTGAAAAGCCAGCGTAAAGAGGTGAAAAGGTGGCTGGAGTGCTCCGCGGACAGTTTAAACAGATTGATCTAACGACCGAAAACCGC  
AATCACTCATCAACTGGGATGGTTTCCGCACCTATAACTTTGATGTGATTGATGGTGTGAATTAFCAGATTGATGTACCCAACTGCCCGTTA  
TGACGGCGAGTGCCAGATGATTAAATGCCAATGGCGAAAAGGATTAAGAACTGACCTTTAATGGCAAGCCGATTGATCCGAAACCGCATGTTCTCT  
GTTGCCACCAATAACTATCGCGCTTACGGCGGCAAATTTGCCGGTACGGCGCAGCCATATCGCTTTTGCTTACCAGGATGAGAACCCTGCG  
TGCTGGACCGTGGATTGCTGATGAGTGAAGCTGCGGGGAAATTCACCAGCGGCGAGATAACAACCTGGCGTTTAGCACCGATGATCGGGCGA  
TAAGAACTGGATATCCGTTTCGAAACTTCTCCGTCAGATAAAGCCGACGCTTTATTAAGAGAAAAGGGCAATATCCGATGAATAAAGTTCGCG  
ACCGATGATATCGGGTTTGCAGTTTATCAGGTGGATTTGAGTAAGTAAACACTTCTTTTTCGGCTATAAATCATCAACCGCATCCGGGATTT  
ATTGGCGGATGCGATGCTGGCGCATCTTATCCGCCCTACAAGCCATGCACCGTAGACCAGATAAGCTCAGCGCATCCGGCAGTTATGCCGCAC  
ATTCCTCCGCAACCGTGGCAACCTCGGCAAAATTCACCTCAACTCACTCCAGCTCCGCGGAAATTCACCAGCGATGAGTACCTTCGCGTAGCCAC  
AGGCTATATTCCACATGCGGCGGCACCCCGGATACCGGATACGGTAAAGAAAACCTCCTGTCTAACGCCTGTAACGACTGCGCAAGCATCT  
TTTCGCTCACCCCGCGGATTTTCGCGCCGAGGTCGCTAAAGCGATGAGTACCTTCGCGTAGCCACCGAATCAACACCCCAACGGCTGGT  
GACGCTGTTTAAACACTCGCGCACGGGCACTGTTCCGCAAGAGATTAACCTCTTTCAGTTGTGGCACAGGCTAACCTGACTCATTTATCATACT  
TTACTTTTGTAGCTTACTAAAAGTAAAGTTAGTTAGCTTAATTAACCAAGCAAAACGATGGAGACTCCCATGATCCGCTATTA  
CTGGTGGCACTGGCAACTTGGTCACTATGTTATGAATCCTTGAAGAAAACGGTTCCTGTGACGCAAAATAGTGGCTATCGTTGCTAAATCCGGC  
AAAAGCCAGGCTTCGGCAGCAAGGATTTACCGTGGTGCAGGCTGACTACGGCGATGAAGCCGCACCTGACATCTGCATCTCAGGGAGTGGAA  
AAACTACTGCTGATCTCTCCAGCGAAGTGGTCAACGTCGCCGCGAGCATCGTAATGTTAATTAATGCCGCAAAAGGGCTGGTGTGAATAATTA  
TGCTTATACCCTGGATTGCTGATGAGTACCTCCCGGCTGCGCAATCGCAATCGAGACGGAGAAAATGTTGGCTGATGCTGGCAT  
CGTTTACACCCCTGCTGCGCAACGGCTGGTACAGCGAAAACCTCCTCCGACCGCCCGCGCAGCACTGGAACACGGCGTATTTATCGGTGCGGCG  
GGCGATGGCAAAATCGCCTCAGCAACCGCGGCGAGATTATGCGGCGAGCTGCGGCACCGGATTAGCGAAGCCGGTACGAAAGGCAAGGTTTACG  
AACTGGCGGGCGATAGTGCCTGGACATTGACACAGTTAGCGGCGAGAGCTGACCAAAACAGAGCGGCAAAACAGGTTACCTATCAAAAATCTGAGCGA  
AGCCGATTTCCGCGCGGCACTGAAAAGCGTCGGACTGCCGAGCGGATATGCTGGCGGATTCGACGTTGGCGCATCGAAAGGTTGGT  
CTGTTTTGATAACAGCAAAACCGCTTAGCAAAATGATTGGCCGCCAAGCACAACGTTAGCGCAAAAGCGTAAGCCATCTTTTTAATGTTAATAACT  
AGTTAATTAAGTGGCATCTCCCGCATCCTCTCTAATAATGACGGGATGCCGGGAGCAATCATGCTGCTTCTGAACCTTCTTCTGACAGAC  
CAATGGATGGCCAGTAATGATAGCGGCTGCTGACCGCCGTTTAGCAGGTTGATGGGGGGCTATTTTTGTGCGGGCGTGGATCGTACCCGG  
AATACCCGGCGATGTTGAGTCGATGGGGGCTTATCTGGCTTAGGGTTAATTCGCTGCCCATGFCCTGGCTGGGACCGTGGCTGCTGCTCA  
GTTGGCGCGTGGGACTGGCTTACCOCCTGATGCTCACTATGATGGGCAACCTCATTATTACTTCTGCTTGGCAGTGCCATTCAACGTACT  
GGCGGCTGTTTTCCACGATGATTATCGGCACCCCTGCCGGTGGTCATTCTGTGTTTGGCAATCTGCTTTATAGCCAGCGCGACGGCAAACTCG  
CGTGGGAAAACCTCGCCCGGCACTGATTTGTAATTTGGCATCGCGCTGGCGAGTGTGAATAATGCTGAGTTAAACCACGGACTCCCCGATTTTGA  
CTGGGCACTTTATACCTCTGGCATCGTGTAGGTTAGTTTTCCGTTGCTGCTGCGCATGGTATGCCCCTGGCAACCGCCGCTGGCTGGCGGAA  
AATCCCGCAAAACATCCGATGATGTGGGCGACGGCGAGGCGTGGTCAAGCTGCCGGTTCCTCTCATCGGCTATCTCGTCGCTGTTACTGGC  
TGAATATAACAAAGCCGACTTCTCCTTACCTTTTGGCCCGCTCCGCTGGTGTATTATAGTCTGATGGTTGCGATAGCCGTGCTTTGCTCATG  
GGTGGGCGCATCTGCTGGAACGCTCGCCAGCCAGCATTACCGACAGTATTCTGGGGCGCTGATTGTTTTGCAAAACCGTGGCAGGTTTGTCT  
TACACTTTTTACTCCGCGCAAAATGCCCGCCGTAATGACGCTGAGCGGATTCGCGCTGTTAGTATTGGCTGGTCAATGAGGTTTAAAGGTTT  
AACCAGAAAACCTTTAACTGAATCTGTCTCAGAAAAGTTGACACGCTGCGCAGTAGTTAAATAAGCCCTCTGCTACGTAAGGTTTATAGCTTTT  
CCTTAAAGATGCAATTTAAATACATCTTATCTTATTAAGAATGAGGTATCAGCTATGGCTTATCGCGAACCACTTTAGTGAACTGGCGCTCT  
CTATCCCTCGCGCTTCGCTCTGTTTCGTAATAATGATATGGATTACTGCTGTTGGCGGTAAGCAGACGCTGGCGCGCGCGCGGACGTAAGA  
ACTGGATGTTGAGGTCATTGAAGCTGAACTGGCAAAGCTTGTGAAACAACCGATTGAGAAAAGCTGGCGTAGCGCCCGCTGGCAGAAATCATT  
GACCATATCATCGTGGCGTAC

## PRÍKLAD 2

TTAAGAAATTCGCGTTATCAACTTCGAGCCCTTTTACAGTACCGGTTGGCGTATTTCCGGGGCGATAGATTTGTAATTTGTAGCCGGAATAC  
CGATCGCCCCGCGGAAAAAACGCAGATCCGGCGGTACTTTGTCGAAATCACCAGGTTTCAATCCAGCCCAGCGTGCCCGGTGTAACGAAACGAT  
GGCGATCGTACAGTGTGGGATCCAGAGCTTCTGCGCTGGAAAACGGAGAAATCGACATCTGAGCCCCAGGCGGTGTTGGAGTAGTCGATAGA  
GTAGCGTTGCGAGTGCGCCCTGGTTGGCATCAGGCCACCACGAGAACCGGTGCGGCTAATCATACCCAGGATAAAAACAGCATCGTGGTGTG  
GTAATTTACCCTGAGTAAAGTGGTGCAGACTCCAGCGAGTTAATGGCAGCTGCCAGCCGTTGGAGAGATCCGACTAGCGAGAAGCCACCA  
GGCTAGTGGAGTCAGATTCGGTATCGTTCAGGTCAGTGCCTTAAAACCGCCCTGCACCAAAATAATATTGTTCCAGTGGATTCTTCAGCAGCGG  
CATTTTTATAGCTGAAGTGCAGGGTCTGTTCCGGCGCGAAATACTGGTACTGGTGGTCAGACTGTGACCATAAGAGTTATCCAGGCTTTTTT  
CACGTAGCTTTACAGCGCGGTTCCACGTCCTGAGGTAACCGACCCCGGTTTTGATGGTGTTTTTAGTTTCGCGGCAAAACAGCCCGCTCAAG  
GTAATACTTTCCGCGCGCTTTATCAAATTTGGAGCCACTACGACCGGTTAAACCAGCCGGTAGCAGAAAGTCAGCCGTTAAGTTAAGTCCGC  
CAGATCTTTCGATTTCGATCTCACTCCCTTTTAAACCGCACCAAGTTTTCAGGTTATTCATCCCGGATTTGTTGATCTTCAAAGGTCACATG  
CCAAAGCGGTAACGTTCCGCACGTATAATCAATATCCAGAAGGCTTTATGACGGCGAGCGCAATGCCAGCTGCCTTGGTAAATTCGC  
TATCGAAATAACCTTACGCAACGCAATGCTGGTTAAGGACTTTTTGAAATTTTCATAATCGCCCTGGTTTCAGTACCGTCCCAATAGCCGGG  
AGTATCGAGCAATTTCAAATAGCTTTTATCGGTCGCGCGCGCGCAATACCACATCGGTGCCCAATTAACACCGCGACGCTTACGTTGCGGTG  
ACTTTGGCGATCAATACTGCCCGCTTTCTTTGGCGGTGGACGGAGATCAAATTCAAATGGTTCGCTGTTAAATACCAGCGCTTTCAGACTT  
CGCGGATGGCATCATGACGCGTGGCGGAAAGCGACGGTCTGGCGTCACTTCACTCTTCAATCGTAGAAAGCTGCGCAGAACGTTCTTTTC  
CAGCTGTCGCGATAAACCCTCGACTGTAGACGGACGTTGCGCGCGACGGCAGATCCGCTTAAAGCAGAGTAAGCTTACACAGCATAAAGTTCGG  
ATATAGCGCATAATTTCTCTGAATATCTTTTCTGCCCTGGAACCGCGTTAAACCGCTTAACAAAAATCCAGTAATATGATTAATAAAAG  
CAGACTAAACCCCAAAATTTCTTATGTTTACTTTAGCCTTACGGGTTTACGGTGGTTATTGTTGTCGAAATACGCCCTTGTACAACCTTAAACCC  
AATGACCGATTTTCCGGAGAGCGACACCATGAGTTTATTTGATAAAAAGCATCTGGTTTTCCCGCCGATGCCCTGCCTGGACGTAACACCCG  
ATGCCCGTAGCCAGCTGCATGGGTCAACGGTCACTCAATGACCAATGTACCTGACGGAATGGAGATTGCCATTTTGGCATGGGTGTTTCT  
GGGGTGTGGAGCGTCTGTTGCGAGTTACCCGGCTTTACAGCACCGCCGAGCTATACGGCGGCTATACGCCAAATCCGACTTATCCGGA  
ACTGTGCTCCGGTATACGGGTATGCCGAAGCGGTACGCATTGTTACGACTTCTGTATCATGACTTCTGTCTCATGACTATGAGGATTTTGG  
GAGAATCAGATCCCGCCAGGGAATGCGTCAGGGCAATGACCAGGCACGCAGTATCGTTCAGGATTTATCCGCTGACCCAGAACAGGATG  
CCGAGCTCGCGGCGAGTCTGGAACGTTTTACGGCGCGATGCTTGCAGCGATGACGATCGTCACTACCCAGGAAATCGTAAACGCCACAC

GTTTTATTATGCCGAAGATGACCACCAGCAATATCTGCATAAAAAACCCGTATGGTTACTGTGGAATTGGCGGAATTGGCGTCTGTCTGCCACC  
GAAGCATAGCGTTACGGGTACAAATGTAGATTTGTGATAAAGTGCCTTTTGTATGCCGAATGCAGCGTGAATGCCTTACCAGGCCTACAAAA  
TCGTCCAAATTCATATATTCAGGGACTGCGTAGCCGTGATCGGCATAGCGCATCAGGCAGTATTGCATTTATCAGCAGTCTGAAATTTAAC  
CTCTGGCGACTTTACAGTACCTTACGCTATACTAGCCACTGAAAATGCCGGATCACTTTCTTCGAATCGGCTTTCAATGTGTATTTCACACAAA  
TTAATCAACTTCCCTTCCGAGGATCTGGCCTGAAAGTTCGGATAAAGATATGTTAAACAGTATTTTAGTCATACCTCTGCTTGATCGCTGTAAAGT  
GCGTCTCTCGATGTCTGAGATCTCACTTGGCCCTCAGCAGCAAAATCAAACTAAACTGCTGGCTGATGAAGGAATATAAATGCCCAACGG  
TCGTGAATATGCAGGAAATCCCGCATGTTCTTACGGTGGTCCAAATTCAGCTGAAACGCAGTTGCGATTTCTCGGCGTATCGCTGGTATGC  
GGCATTCTTCCAGCTTTTCACAGCCTGTTCTCCCGCTATATGTCGGCAGAACTCTCTGAGCAACTGAGCTTTATTTCTCTTTCTCTGTTAGTG  
ACTGGCATGTTTATCCTGTTTGGCGATTAAACCCGAAACGCATCGGTATGATTGCGCCAGAAGCGGTGGCTTTGCGTATCATCAACCCGATGC  
GCTTCTGCCTGTACGTTTGCACCCCGCTGGTGTGGTCTTCAACGGCTGGCGAACATGATCTTCCGTATTTTCAAACCTGCCAATGGTACGTAA  
AGATGACATCACCTTGTAGTACATCTACCGGTAGTGAAGACCGGTGGCTGGCTGGCGGGCTGTTACGTAAACAGGAACCCGAGCTGAATGAAAAC  
GTCTTTGAGCTGGAATCCCGTACCGTTCCTGCTTCAATGACACCGGTGAAAACGTGATTTGGTTTGTATCTCCACGAAGATGAGCAAAGCCTGA  
AGAATAAGGTGGCGGAACATCCGCACTTAAGTTCCCTGCTGTGAATGAAGATATTGACCACATCATCGGTTATGTCGATTTCAAAGACCTGCT  
GAACCCGCTGCTGGTAAACCAAGCCTGGCACTGAACAGCGGGTACAAAATTCGCAACAGCTGATTGTGCCGGATACGTTAACCTTTTCAGAG  
GCGTTGGAAGATTTTAAACCCGAGTGAAGACTTCGCGGTGATCATGAACGATCAGCGCTGGTGGGGATCATCCCTCAACGATCAACGAGTGA  
TGACCAGCTGATGGCGATCTGGTGGTCCAGGGCTGGAAGAGCAGATTGTCGCCCCTGATGAGAACTCATGGCTGATTGACGGCGGCACCC  
AATTGACGACGTATGCGGTGCTGGATATTGACGAGTTCCCGGAGTCCGGCACTACGAAACCATCGGAGGCTTTATGATGTTTATGCTGCGT  
AAGATCCCGAAACGCACCGATTCCGTGAAATTCGCGGGCTACAAATTTGAAGTGGTGGATATCGATAACTCCGCATCGACAGCTGCTGGTGA  
CCGGATCGACAGCGCAAGTCCCTTTCGCCAAAACCTGACCTGACGTGAAAGTAAAGAAAGCCTGCGCTAACCCGACCGCGAATGAAAAC  
GGCTCTGAAATCAGGAGCCGCTTATTACTGCATAGCACTTTGGTTAAGCCATCTCTGTTTGCAGACGCATAACCTGACGGTTAACTTCCGACA  
TCACTGACAAAATCGAGCTTATCCTTCACTTTTGGGATAAAGACTTACCTTATCAAATTCAAAAGCGCAACGCTCTTATGATATAACCGTCC  
ACGGAACAGGATCTTACGTACTTCCGCACTGAAGTGGTGTACCTTGGAAAATTTTCATTTGTTATCTCTGCTGAGTATTACGCGCTGTGC  
GGTGGCACAGTCCGCCAACTTATGAGGCGCAAAATTTAATGCGCTAGTGAATTTAGACTATACCGGGCAATTTCCAGCTGATACTAACTTTT  
TTTACCTAAAGTTACAATTATTCAGAATTATCTTTTACCAGCGCGGTCTTCACTATAAGCATTCATTTTTCATATGATTTGTGCGCTTGA  
CCGCAAACTGGCATCACACTTGGGGAAATTCGATAAATAGCACATATGATTAAGAACTCAGACCCAGTGGTCCGATCACTGCATATCATAAG  
AAGGAAACCACTACCCATACGCAAGATTCTGGCACTCACTGCTGCTGTTGGCCGATGATGGCTTCCGCACATCAGTTCGAAACCCGCTCAGCG  
AGTGGCCCGGATTGGCATCACCGTGGCGAGTTGGT  
AAAGTGGAGTACTGCAACATATTGCTGGTCCGACCTCTGCAAAAGAGAAAAACGCGACGCTGATTGAAGCGATTAAATCAGCGAAGTTACCGC  
ACGATCGTTACAGACACCACCATTGTTAACACCGACGACGCAATTCGCGGTTCCGGCATGTTTGTGGCGAGCAGTCTGGAGAGTAATAAAAA  
GCTTTATCCCTGGTCCGAGTTTATTGTTGATAGCAATGGCGTCCGACCGGGTGGCTGGCAGCTGGATGAAAAAAGTTCCGCTGTAGTGGTGGCTG  
GATAAAGACGGTCCGCAATGGGCCAAAAGACAGGGCGCTTACTCAGGAAGAGGTGCAGCAAGTGGATGGACCTGCTGCATAAAATTAATTAATA  
AATAGACTCTGAACCCAGGATTCAGGAACGACTCGCGCGGAGTGAATCCAGCGGTTTACCCTGCCAGTCTGAAACGTCGCTCCGGCAGCT  
CGGCCACAGCATGTCCAGCGGGGTGTCCAAATATTGTTGGTCCGAAAGCGGGTACAGCTGCGCTGTCTTCCGCCACCAGGCAGAAAT  
TCAGCGAAGAGCCGATGGACGTGGTCTGATGTTCCGCAAGCTGTTGCAGATCTCTTTCAGCTCCGATCTGCATGGGAACGGCTGATAACCC  
CAGCGCGGGCGCATGCGCGGACTGAATCTGCTTGGCACACCGCATCTTCTTCCAGGCTTGCCTTTCGCGGCTGTACATCCAGCTTCC  
ATTACCGCGCATACACCACCTAAAATCGGTTTGCATGGTCAATGAGCGCAATGTTAACGGTGAATTCCGCCATTGCGTTTAAATAAATCTTT  
TAGTACCATCCAGCGGGTCTACAGCCAGTAACGCTGCCAGTGGTGCAGGACTTCCCAACCGGGAGGATCTTCTTCAGAAAGGACCGGAATATC  
CGGTGTACGCTACGTAACCTGCTTATGATAACGGTGTGAGCGGCAATATCCGCTCCGCTTACCAGGAAATTTGTCGCTTGTGACGACGCTC  
ATCGGTTTCGTCGCTGAGCTGACATGCAATAAGGCATCGCCGCTTACTTCGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  
GTGAGCGGTGTTAACTTATGTTTACTTATACCCATCGTGAATGAATGCGCAACTGTGATAGTGTATCATTTTTCAAAGCGTAAAATTTGTG  
GCATTTCTACTGTTCTATAAGTAAAGCCTTTATCTTCTTTCTTTCTGATTTCCCGATGATAAAAGGATGTCCCTGATGATTAAGTTTAGCG  
CAACGCTCTGCGACCGTATGTTGCGCGCAGTGTGAATGCAAGCGGCTGATGCTGCGTATCATGGAACCACTGATCTGCATGCAACATGAT  
GGATTTTCGATTATTAACAAGACCCGACCGAAAAATTCGGACTGTTACGTCAGTACGCAAGCCTGATTAAACGATCCCGCAATGAAGTGA  
AGCGTACTGGTCCGATAACCGCGATTGATTTCAGGGGATTCGCTTGGCCGATTACGTCGCGGAAAGGATTAAAGCGAGTGTATGTTTATCCCG  
TTTATAAGCGCTGAATACGCTGGATTATACGCTGGTACACTCGGCAATCATGAATTTAACTACGCTGGATTACCTGAAAAATGCGCTGGC  
GGGAGCGAAATTCCTTTATGTAATGCCAAGCTCATTGACGCCAGAACCAAAACAGCCAATGTTTACACCGTATTTAATTAAGATACCGAAGT  
GTCGATAAAGACGGAAAAAACAGACGCTGAAGATTGGCTATATTGGCTGCTGCGCCGCAAAATCATGGCTGGGATAAAGCTAATTTATCCG  
GAAAAGTGCAGGTGAATGATATTACCGAACCCTGCGCAAAATACGTCGCTGAAATGCGCGAGCAAAAGGTGCGCATGTCGTTGTCGTTTCCGCGCA  
TTCGCGCTGTCTGCCGATCCGTATAAAGTAAAGTGGCGAAAACTCAGTTTATTACCTCAGTGAATTCGCGGCTTAAACCCATTATGTTTGGT  
CATGCTCAGCGGTTTTCCCGGTAAGATTTTGTGATATCGAAGGGGCTGATATCACCAAAAGCGCAGCTGAATGGTGTTCGCGCGTAAATG  
CGGCATGTGGGGCATCATCTGGGGTGGTGCATTAACAATCAGTAATGACAGCGTAAATGGCAGGTGACCGCAGGTAAGTGGCAGGTGACCGCAG  
GCCGATTTACGACATCGCAATAAAAAATCCCTGCGCGGGAAGACAGCAAGCTGGTGAAGAACTCAAAAGCAGTACAGTGCACACCGCCAG  
TTCGTCAGCAAGCAATCGTAAATCTGCCACAATATGTATAGTATCTGGCGTGGTGCAGGACGATCCGACCGTGCAGTGTGAACACG  
CGCAAAAGCGTATGTCGAACATTCATTCAGGGCATCCGGATCTGGCAAACTGCGCGTGTTCAGCTGCCGACCGTTTAAAGTTGGTGG  
TCGCAAAATGACCCGGACAGCTATGTGGAGGTGGAAGAGGCTGACTTCCGTAATGCCGCGATCTTTATCTACCCCAATACCGCTG  
ATTGTTGGTGAAGCGGTAAGAGGTGAAAGAGGTGAGGATGCTTCCCGCGACAGTTTAACAGGATGATCTTAAACAGCAGTAAACCGCCG  
AATCACTCATCAACTGGGATGGTTTCCGCACCTATAACTTTGATGTGATGATGGTGTGAATTTATCAGATTGATGTTACCAACTGCCCGTTA  
TGACGGCGAGTGCAGATGATTAATGCCAATGCGGAAAGGATTAAGAAGCTGACCTTTAATGGCAAGCGGATTGATCCGAACGCCATGTTCTC  
GTTGCCACCAATAACTATCGCGCTTACGGCGGCAAAATTTGCCGTTACGGGCGACGCCATATCGCTTTTGGCTTACCAGGATGAGAACCCTCGG  
TCTGGCAGCTGGATTGCTGATGAGTGAAGACGTCGGGGGAAATTCACCCGCGCAGATAAACAAGTGGCGTTTACGACCCGATGCTGGCGA  
TAAGAAACTGGATATCCGTTTCGAAACTTCTCCGTCAGATAAAGCCGACGCTTTAATTAAGAGAAAGGGCAATATCCGATGAATAAAGTCCGG  
ACCGATGATATCGGGTTTGCATTTATCAGGTGGATTTGAGTAAGTAAAAACTTCTTTTTCGCGCTATAAATCATCAACCGCATCCGGCATTT  
ATTGGCGATGCGATGCTGGCGCATTTATCCGCCCTACAAGCATGCACCGTAGACCAGATAAGCTCAGCGCATCCGGCAGTTATGCCGACG  
TTCATCCGCAACCGCAACCTCGGGCAAAATTCAACTCAATCCAGTCCGCGCAGTAAACCTTTTTCGCTCACCCTGCTCGCCGACCGCGTG  
AGGCTATATTTCCACATGCGGGCGCACCCCGGATACCGGATACGGTTAAGAAAACCTCTGTTCTAACGCTGTAAACGACTGCGCAAGCATCT  
TTTCGCTCACCCCGCGATTTTGGCGCGAGGTGCTAAAGCATGAGTACCTTCGCGTAGCGCCACCAGAAATCAACACCCCAACCGCTGAT  
GAGCTGTTTCAACACTCGCGCAGGGCACTGTTCCGCAAGAGATTACCCTTTTCACTGTTGGCAGCGGTAACCTGACTCATTTCATCT  
TTACTTTTGTGCGTACTCATGATAAAGTAAAGTTAGTGTGTTAGCGTATTTAAACAGACAAACAGTGGAGACTTCCCATCGTATCTGATTA  
CTGGTGCCTGCGCAACTTGGTCACTATGTTATGAACTCTGAAAGAAAACGGTTCCTGTGACGCAAAATAGTGGCTATCGTTGTAATCCGGC  
AAAAGCCAGGCTCTGGCAGCAAGGATTAACCGTGGCTCAGCTGACTACGGCGATGAAGCCGCACTGACATCTGCACTTACGGGAGTGGAA  
AACTACTGCTGATCTTCCAGCGAAGTGGTCAACCTGCCCGCAGCATCTGAATGTTAATGCGCAAGGGCGGTGGTGTGAATTTA  
TCGCTTATACCAGCTGCTACATGACAGATACCTCCCGCTCGCCACTCGCAATGAGCACATCGAGACGAGAAAAATGTTGGCTGATTTGCGCAT  
CGTTTACACCCTGCTGCGCAACGGCTGGTACAGCGAAAACTACCTCGCCAGCGCCCGGCGAGCACTGGAACACGGCGTATTTATCGGTGCGCG  
GGCGATGGCAAAATCGCCTCAGCAACCGGGCAGATTATGCGGAGCTGCGGCAACGCGTATTAGCGAAGCCGCTCAGCAAGGCAAGGTTTACG  
AACTGCGGGCGGATAGTGCCTGGACATTGACACAGTTAGCGGCAGAGCTGCACAAACAGAGCGGCAAAACAGGTTACCTAATAAATCTGAGCGA  
AGCGGATTTCCGCGCGGCACTGAAAAGCGTCCGACTGCCGACGAGCTGCGGATATGCTGGCGGATTTCTGAGTGGTGGCGGATTTGAGTGGT  
CTGTTTGATAACAGCAAAACGCTTAGCAAAATGATTGGCCGCCAACGACAACGTTAGCCGAAAGCGTAAAGCCATCTTTTAAATGTTAATAACT

AGTTAATTAAAGTGGCATCTCCCGCATCTCTCTAATAAATGACGGGATGCGGGGAGCAATCATGTCTGCTTCTGTAACCTTCTTCTGACAGAC  
CAATGGATGCCAGTAATAATAGCGCGTGTCTGACCCCTGTTAGCAGGGTGTAGTGGGGGCTTATTTTTGTGCGGGCGTGTATCGTACCCG  
AATACCCGCGATGTGGCAGTGCATGGGGCGTTATCTGGCGTTAGGGTAAATTCGCTGCCATGCGCTGGCTGCCAGCGCTGCTGCTGCTCA  
GTTGGCGCGTGGGACTGGCTTACCGCCTTGATGCTCACTATGATGGGCAACCTCATTATTTACTTCTGCCTTGCCAGTGCCATTCAACGTACT  
GGCGCGCTGTTTCCACGATGATATCGGCACCTGCGCGTGGTTCATTCTGTCTTTGCCAATCTGCTTTATAGCCAGCGCGACGGCAAACCTCG  
CGTGGGAAAAACTCGCCCCGGCACTGATTTGTATTGGCATCGGCTGGCGAGTGTGAATATTGCTGAGTTAAACCAGGGACTCCCCGATTTTGA  
CTGGGCACGTTTACTCTGGCATCGTCTAGCGTTAGTTTCCGTTGCTGCTGGGCATGGTATGCCCTGCGCAACGCCCGTGGCTGGCGGAA  
AATCCCGACAAACATCCGATGATGTGGGCGACGGCGCAGGCGTGGTCAAGCTGCCGTTTCTCTCATCGGCTATCTCGTCCGCTGTTACTGGC  
TGAATATACAAACGCCGACTTCTCCTTACCTTTTGGCCCCGTCGCTGGTGTATTATAGTCTGATGTTGCGATAGCCGCTGCTTGTCTCATG  
GGTTGGCGCACTCTGCTGGAACGTCGCCAGCCAGCGATTACCGACAGTGTCTCGGGCCGCTGATTGTTTTCGAAACGCTGGCAGGTTTGTCTG  
TACACCTTTTTACTCCGCGAGCAAAATGCCGCCCTAATGACGCTGAGCGGTATCGCGCTGTAGTATTGGCGTGGTCAATTGCGGTCAGAGCAA  
AACCAGAAAAACCTTAACTGAACTGTCTCAGAAAGTTGACACGCTGGCAGTGTAAATAAGCCTCTGCTACGTAAGGGTTATAGCTTTTG  
CCTTAAAGATGCATTTAAATAACATCTTATCTTATTAAGAATGAGGTATCAGCTATGGCTTATCGCGACCAACCTTTAGTGAACTGGCGCTCT  
CTATCCCTCGCGCTTGGCTCTGTTTCGTAAATATGATATGGATTACTGCTGTGGCGGTAAGCAGACGCTGGCGCGCGGGCGGCAGTAAAGA  
ACTGGATTGTGAGGCTATTGAAGCTGAACTGGCAAAGCTTGTCTGAACAACCGATTGAGAAAGACTGGCGTAGCCCGCCGCTGGCAGAAATCATT  
GACCATATCATCTGCGCTAC

### PRÍKLAD 3

TTAAGAAATCCGGTTATCAACTTCGAGGCCCTTTCAGGTACCCTGGCGTATTTTCGGGGCGATAGATTTGTATTTGTAGCCGCAACTG  
CGATCGCCCCGGCGAAAAACGCAGATCCGGCGTACTTTGTGCAATCACCAGTTTCAATCCAGCCAGCGTGGCGGTGTAACGAAACGAT  
GGCGATCGTACAGTGTGGGATCCAGACGTTCTGCGCCTGGAAACCGGAGAAATCGACATCTGAGCCCCAGGCGGTGTTGGAGTGTGATGATA  
GTAGCGTTGGCAGTCCGCCCTGGTTGGCATCAGGCCACCAGAAACCGTGGCGTAAATCATACCCAGGATAAAACAGCATCGTGGTGTG  
GTAATTTACCCTGAGTAAAGTGGTCGAGACTCCAGCGCAGTTAATGACGCTGAGCGGTGCCAGCGTGGAGAGATCCAGCAGTAAGCCACCA  
GCGTAGTGGAGTCAAGTTCGGTATCGTTCAGGTGAGTGGCTTAAACCGCCCTGCACCAATAATATTGTTCCAGTGGATTCTTTCAGCAGCGG  
CATTTTTATAGCTGAAGTGGAGGCTGTGTTCCGGCGCGGAAATACTGGTACTGGTGGTCAAGTGTGACCATAAAGAGTTTCAACAGCGCTTTTC  
CACGTAGCTTTCACGCGCGTCCACGTCGCTAGAGTAAACCGCCCGTTCGATGGTGTGTTTTCAGTTCGCGCGGAAACACCGCCGTCAGG  
GTAATACTTTCCGTTTCGCGCTTATCAAATTTGGAGCCACTACGACCGAGTTAAACCGCCGTTAGCAGAAAGTTCGACGTTAAGTTCCGC  
CAGATCTTTGATTCGTACTCATCGCCCTTTTAAACGGCACCAGATTTGACAGGTATTCATCGCGGATTTGTGATCTTCAAAGGTCACATGC  
CCAAAGCGGTAACGTTCCGCACTGTTATAATCAATATCCAGAAAGGCTTATGACAGCCGAGCGCAATGCCAGCTGCGCTTTGGTAAATTCGC  
TATCGAAATAACCTTACGCAACGCAATGCTGGTTAAGGACTTTTGAATTTTCATAATCGCCCTGGTTCAGTACCCTGCCAATAGCCGGGCG  
AGTATCGAGCAATTTCAAATAGTCTTTTCCGGTCCGCGCGCGCAATACCACATCGGTGCCCAATTAACACCGGCTGCAAGTTCAGCGT  
ACTTTGGCGATCAATACCTGCCGCCCTTTCTTTGGCGGTGGACGGAGATCAAATCAATGGTCCGCTGGTAAATAACCCAGCGCTTTCAGACCTT  
CGCGGATGGCATCATCGACGCTGCGCGAAAGCGACGCTGCTGGGCTCACTTCACTCACTTTCAATCGTAGAAAGCTGCGCACGAACGTTCTTTTC  
CAGCTGTCCCGATAACCCCTCGACTGTAGACGGAGCTTCCGCGGCGACGGCAGATCCGCTTAAAGCAGATAAGCTTACACAGCATAACTGTGCG  
ATATAGCCGATATTTCTCGCTGAAATCTTTTCCGTTCCGCGCGCGCAATACCACATCGGTGCCCAATTAACACCGGCTGCAAGTTCAGCGT  
CAGACTAAACCCCAAATATTTCTTATGTTTTACTTTAGACCTATTCACGGTGGTATTGTTGTGCAAAATACGCCTCTTGTGTACAACCTTAAACCC  
AATGACCGATTTTCGGGAGAGCGACACCATGATTTTATTTGATAAAAAGCATCTGGTTCGCCCGCCGATGCCCTGCCTGGACGTTAACACCCCG  
ATGCCCGTAGCCAGCTGCATGCGGTCAACGGTCACTCAATGACCAATGTACCTGACGGAAATGGAGATTGCCATTTTGGCGATGGGTTGTTTTCT  
GGGTTGGAGCGCTGTTTGGGATTTACCCGCGTTCACGACCGGATGAGTACCGGCTATACCAGCGGCTATACCAGCGGCTATACCAGCGGCTGCG  
AGTGTGCTCCGGTGATACGGGTGATGCCGAAGCGGTACGCATTTGTTACGATCTTCTGTATCAGCTATGAGCAGTGTGTACAGGTATTTTGG  
GAGAATACGATCCCGCCAGGGAATGCGTCAAGGCAATGACACGGCAGCAGTATCGTTCAGCGATTATCCGCTGACCCAGAACAGGATG  
CCGACGCTCGCGCCAGTCTGGAACGTTTTCAGGCGCGGATGCTTCCGCGGATGACGATCGTCACTACACCGGAAATCGTAAACCGCACCC  
GTTTTATATGCGGAAATGACCAACGCAATCTGCATAAAAAGCGTATGTTTAAACAGTATTTAGTGGAAATGGCGAATTCGCTGCTGCTGCTG  
GAAGCATAGCGTTACGGGTACAAATGTAGATTGTTGATAAAGTGGTTTTGTATTGCGCAATGACGCGTGAATGCCCTTACCAGGCTACAAAA  
TCGTCCAAATTCATATATGTCAGGACTGCGTAGGCTGATCGGCATAGCGCATAGCGCAGTATGCAATTTATCAGCAGTCTGAATTTTAAAC  
CTCTGGGCACTTACAGTACTTACGCTATACCTAGCCACTGAAAATCGCGGATCACTTTCTTTCGAATCGGCTTCAATGTGATTTTACACAAA  
ACTGCACTTCCCTCCGATCTGGCTGAAAGTTCGATAAAGTATGTTTAAACAGTATTTAGTGGAAATTCGCTGCTGCTGCTGCTGCTGCTGCTG  
GCGTCTTCTCGATGTCTGAGATCTCACTTGCAGCCTCACGCAAAATCAAACCTTAAACTGCTGGCTGATGAAAGCAATATAAATGCCAACGCG  
TCTGAAATATGACGAAATCCCGCATGTTCTTACGGTGGTCCAAATCGGTGCTGAACGCGATGCGATTTCTGCGCGGTATCGTGGTGTGTC  
GGCATTTTCTCCAGCTTTTACAGCCTGTTCTCCCGCTATATGTCGGCAGAATCTCTGAGCAACTGAGCTTTATTTCTCTCTTCTGCTGTTAGT  
ACTGGCATGTTTACTCTGTTTGGGATTTAACCCCGAAACGCATCGGATGATGTTGCGCCAGAAGCGGTGGCTTTGGCTGCTGCTGCTGCTGCTG  
GCTTCTGCTGTACGTTTGCACCCCGTGGTGTGGTCTTCAACCGCCTGGCGAATGATCTTCCGATTTTCAAACCTGCCAATGGTAGCTGAA  
AGATGACATCACTTCTGATGACATCTACGCGGTAGTGAAGCCGGTGGCTGGCGGGCGTGTGATGATAACAGGAACACGAGCTGATTGAAAC  
GTCTTTGAGTGGAAATCCCGTACCGTTCGCTTCAATGACACCGCGTGAACCGTGAATTTGGTTGATCTCCAGGAAGTGAAGCAAGCCTGA  
AGAATAAGGTGGCGAATGACCACTCAAGTTCCTGCTGTAAGTGAAGATATTGACCATCATCGGTTATGTCGATTTAAAGACCTGCTG  
GAACCGGCTGCTGGCTAACCAAGCCTGGCACTGAACAGCGCGGTACAAATTCGCAACACGCTGATTGTGCCGGATACGTTAACCTTTACAG  
GGTTGAAAGTTTTAAACCGAGGTGAAGACTTCGCGGTGATCATGAACGAGTACGCGCTGTTGGTGGGGATCATCCCTCAACGACGCTGA  
TGACCAGCTGATGGCGATCTGGTCCGTCAGGGGCTGGAAGAGCAGATGTCGCCCGTGTGATGAACTCATGGCTGATGACGCGCGCACCC  
AATTGACGACTGATCGCGGTGTTGATTTGACGAGTTCCCGAGTCCGGCACTACGAAACCATCGGAGGCTTATGATGTTTATGCTGCTGCT  
AAGATCCGAAACGCACGATTCGGTGAATTCGCGCGCTACAAATTTGAAGTGGTGGATATCGATAAATACCAGCATCGACCGTCTGCTGGTGA  
CCCGATCGACAGCAAGGCCACCGCCCTTTCGCAAAACTGCCTGACGCTAAAGATAAAGAAAGCGTCCGCTAACCCCGAAGCATCAAC  
GGCTCCTGAATCAGGAGCGCTTATTACTGCTAGCACTTTGGTTAAGCCATCTCTGTTTGCAGACGCATAACCTGACGTTAACTTCGGACA  
TCACTGACAAATGACGCTTACTCTTCACTTTGGGATAAAGATCTTACCCTTATCAAATTCAAAAGCGCAACGCTTCTGATGATAACCGTCC  
ACGGAACAGGATCTTACTGATCTTCGCACTTGGCACTGAAGTGGGTTGACCTTGGAAATTTTCAATTTGTTATCTCCTGCTGATTAACGCTGTG  
GGTGCCCAATCGGCCAACTATTATGAGCGCAATTTTAAATGCCTAGTACTATAGACTATCCGGCAATGTTTCCACCGCGTATAACTTTT  
TTTACCTAAAGTTTACAATATTCAGAAATATCTTTTACCAGCGCGGCTTTCAGTATAAGCATTCATTTTTATATGATTTGCTGCGCTTGA  
CCGAAACTGGCATCACTTCCGGGAAATTCGATAAATGACCATGATGATAAATCAACAGCAAGTGGTGGTACCTGATGATGATGATGATGATAA  
AAGAAACACCATGACCTACGCAAGATTCGGCACTACCTGCTGTTGGCCGATGATGGCTTCCGCAATCAGTTCGAAACCGGTCAGCG  
AGTGGCGGATTTGGCATACCGATCGTGGCGAGTGGTGGTGTGATAAAGATCAGTTTAGCTACAAAACCTGGAACAGCGCGCAGTTAGTGGGA  
AAAGTGGGAGTACTGCAACATATGCTGGTCCGACCTTCGAAAGAGAAACCGCGACGCTGATTGAAGCGATTAAATCAGCGAAGTTACCG  
ACGATCTTACCAGACCCAACTTGTAAACACCGACGCAATTCGGGTTCCGGGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  
GCTTTACTCTGCTGCGAGTTATTGTTGATAGCAAGTGGCTCCGACGCGTGGCTGGCAGCTGGATGAAAAAAGTTCGCTGATGATGATGCTG  
GATAAAGACGCTGCGTGAATGGCCAAAGACAGGGCGCTTACTCAGGAAGAGGTGCAGCAAGTGTGACCTGCTGCTGCTGCTGCTGCTGCTG  
AATAGACTCTGAAACCGGATTCAGGAACGACTCGCGCGGAGTGAATCCAGCGGTTTACCCTGCCAGTCTGTAACGCTGCGCTCCGGCAGT

GCGGCCACAGCATGTCCAGCGCGGTGTCCAAATATTCGTTGGTCCGAAGCGCGGGTACAGCTGCGCCTGTCTCCGCCACCAGGCAGAATT  
TCAGCGAAGAGCCGATGGACGTGGTCTGATGTTTCGCCAAGCTGTTGCAGATACTCTTTACAGCTCCGCATCTGCATGGGAACGGCTGATAACCAC  
CAGCGGGCGGCGCATCGCGGACCTGAATCTGCTTGGCCACACCCGACTCTTCTTTCCAGGCCTTGCCTTCTGCCCGCGTGTACATTAACGTTT  
ATTACCGGCGCATACACCACACCTAAAAATCGGTTTGCCATGGTCAATGAGCGCAATGTTAACGGTGAATTCGCCATTGCGTTTAAATAACTCTT  
TAGTACCATCCAGCGGGTCTACAGCCAGTAACGCTGCCAGTGTGACGGACTTCCCAACCGGGAGGATCTTCTTCAGAAAGGACCGGAATATC  
CGGTGTACCGTACGTAACCCGTCATGATAACGGTGTGAGCGGCAATATCCGCTGCCGTTACCGGAGAATTTGTCGCTTTGTGACGACGCTCC  
ATCGGTTTCGTCGCCGTCGTAGACCTGCATAAATGGCATCGCTGCATTCGCTGCAAGCTGGCATACTTGATCTAACATTTCTCCACCTCGTCTCT  
GTGAGCGGTGTTAACTTATTGTTTTACTTATACCCATCGTTAATGAATGCGCAACTGTGATAGTGCATCATTTTTCAAAGCGTAAAAATGTG  
GCATTCCTCACTGTTCTATAAGTAAGACGTTTATCTTCTCTTTCTTTCTGATTCGCCATGATAAAAAGGATGTCCTGATGATTAAGTTTAGCG  
CAACGCTCCTGGCCACGCTGATTGCCGCCAGTGTGAATGCAGCGACGGTGCATCTGCGTATCATGGAAACCCTGATCTGCATAGCAACATGAT  
GGATTTTCGATTAATACAAAAGACACCGCCACGGAAAAATTCGGACTGGTACGTACGGCAAGCCTGATTAACGATGCCCGCAATGAAGTAAAAAC  
AGCGTACTGGTCGATAACGGCGATTGATTACGGGGAGTCCGCTGGCCGATTACATGTCCGGCAAAGGATTAAGAGCAGGTGATGTTTATCCGG  
TTTATAAGGCGCTGAATACGCTGGATTATACGCTCGGTACACTCGGCAATCATGAATTAACCTACGGTCTGGATTACCTGAAAAATGCGCTGGC  
GGGAGCGAAAAATCCCTTATGTAATGCCAACGTCATTGACGCCAGAACCAACAGCCTAATGTTTACACCGTATTTAATTAAGATACCGAAGTG  
GTCGATAAAGACGAAAAAACAGCTAGCTGAAGATTGGCTATATTTGGCTGTCGCGCCGCAAAATCATGGCTTCCGGATAAAGCTAATTTATCCG  
GAAAAAGTACGGTGAATGATATTACCGAAACCGTGGCCAAATACGTGCCTGAAATGCGCGAGAAAGGTGCCGATGTGCTTGTGCTTCTGGCGCA  
TTCCGGGCTGTCTGCCGATCCGTATAAAGTAATGGCGAAAACTCAGTPTAATACCTCAGTGAATTCGGGCGGTAACGCCATTATGTTTGGT  
CATGCTCAGCCGTTTTCCCGGGTAAAGATTTTGTGTATCGAAGGGGCTGATATACCAAAGGCACGCTGAATGGTGTTCGGCGGTAATGC  
CGGACTGTGGGATGATCTGGGCTGACTTCAACTTACACTGACAGCGGTAATGGCAGGTGAAATGGCAGGTGACAGCGGTAATGGCAGGTGACAG  
GCCGATTTACGACATCGCAATAAAAAATCCCTCGCGGCGGAAGACAGCAAGCTGGTAGAAAACACTCAAAGCCGATCACGATGCCACACGCCAG  
TTCGTCAGCAAGCAATCGGTAATCTGCCGACAATATGTATAGCTATCTGGCGCTGGTGCAGGACGATCCGACCGTGCAGTAGTGAACAACG  
CGCAAAAAGCGTATGTCGAACATTACATTCAGGGCGATCCGGATCTGGCAAACTGCCGGTGTCTTACGCTGCCCGCAGCTTTAAAGTTGGTGG  
TCGCAAAAATGACCGCAAGCTATGTGGAGGTGAAAAAGCCGATTTCCGTAATGCGCGCACTTTTATCTCTACCCCAATACCGCTG  
ATTGTGGTGAAGCCAGCGGTAAAGAGGTGAAAGAGTGGTGGAGTGTCCGCGGACAGTTTAACAGATTGATCTAACAGCACGAAACCCG  
AATCACTCATCAACTGGGATGGTTTCCGCACCTATAACTTTGATGTGATTGATGGTGAATTAATCAGATTGATGTACCCAACCTGCCCGTTA  
TGACGGCAGTGGCAGATGATTAATGCCAATGCGGAAAGGATTAAGAACCTGACCTTTAATGGCAAGCCGATTGATCCGAACGCCATGTTCTCT  
GTTGGCAAAATAACTGAAACAGCTAGCTGAAATTTGCCGCGCAAAATTTGCCGCGGCAATATCGCTTTTGTCTACCGGATGACAGCCGCTCG  
TGCTGGCAGCGTGGATTGCTGATGAGTCGAAACGTCGCGGGGAAATTCACCGCGCGCAGATAACAACCTGGCGTTTAGCACCGATAGCTGGCGA  
TAAGAAACTGGATATCCGTTTCCGAAACTTCTCCGTCAGATAAAGCCGCGAGCTTTATTAAGAGAAAGGGCAATATCCGATGAATAAAGTCGCG  
ACCGATGATATCGGGTTTGGCATTATCAGGTGGATTTGAGTAAGTAAAAACACTTCTTTTTCCGCCATATAAATCATCAACCGCATCCGGCATT  
ATTGGCGGATGGGATGATCTGGCGACTTATCCGCCCTACAAGCCATGCACCGTAGACCAGATAAGCTCAGCGCATCCGGGAGTTATGCCGACG  
TTCATCCCGCACCGCAACACCTCGGGCAAAATCAACTCAATCCAGTCCGCCAGTGCAGCAACCTTTTCGCTCACCTGCTCGCCAGCGCGTG  
AGGCTATATTCCACATGGCGCGCACACCACCGGATACGGTAAAGAAAACCATCTGTTCTAACGCTGTAAACGACTGCGCAAGCATCT  
TTTCGCTCACCCCGCGATTTCGCGCCGAGGTGCTAAAGCGATGAGTACCTTCGCGTAGCGCCACCAGAATCAACACCCCAACGGCTGTT  
GACGTGTTTCAACACTCGCGGACGGGCACTTTCGCGAAAGAGATTACCTCTTTTCAGTTGTTGCGACAGGTAACCTACCGCTGCGCTATAC  
TTACCTTTTTGTGCGTACTTACTAAAAGTAAGTTAGGTGTTAGCGTATTTAAACACAAGACAAAACGATGGAGACTTCCCATGATCGCTATTA  
CTGGTGCCACTGGCAACTTGGTCACTATGTTATGAAATCCTTGAAGAAAACGGTTTCCGTGTCAGCCAAATAGTGGCTATCGTTGTAATCCGGC  
AAAAGCCAGGCTCTGGCAGCAACAGGCAATACCGTGGTTCAGCTGACTACGGCGATGAAGCCGCACTGACATCTGCATTCAGGGAGTGGAA  
AAACTACTGCTGATCTCTTCCAGCGAAGTGGGTCAACGTGCCCGCAGACTCGTAAATGTTAATGCGCGAAGCGGCTGGTGTGAATTTA  
TCGCTTATACAGCCTGTACATGCAGATACCTCCCGCTCGGCTCGCAATGAGCACATCGAGACGGAGAAAAATGTTGGCTGATTTCTGCAT  
CGTTTACACCCTGCTGCGCAACGGCTGTTACAGCGAAAACTACCTCGCCAGCGCCCGGCGAGCACTGGAACACGGCGTATTTATCGGTGCGCG  
GGCGATGGCAAAATCGCTCAGCAACCGCGGCGAGATTATGCGGCACTGCGGCAAGCGTATTAGCGAAGCCGGTACAGAAAGCAAGGTTTACG  
AACTGGCGGCGATAGTGCCTGGACATTGACACAGTTAGCCGAGAGTGCACAAACAGAGCGGCAACAGGTTACCTATCAAAATCTGAGCGA  
AGCCGATTTTCGCGCGGCAAGTAAAAAGCGTCCGACTGCGGCAAGCCGACTGGCGGATATGCTGGCGATTCTGACGTTGGCGCATGAAAGTGGT  
CTGTTTGATAACAGCAAAACGCTTAGCAAAATGATTGGCGCCCAACGACAACGTTAGCCGAAAGCGTAAAGCCATCTTTTAAATGTTAATACT  
AGTTAATTAAGTGGCATCTCCCGCATCTCTCTAATAATGACGGGATGCGGGGAGCAATCATGCTGCTTCTGAACTTTCTTCTGACAGAC  
CAATGGATGCCAGTAATGATTAGCGCGTGTGTACGCGCTGTTAGCAGGTTGATGTGGGGCTATTTTTGTGCGGCGGTTGATCGTACCGG  
AATACCCGGCGATGTTGCGATCGATGGGCGTTATCTGGCGTTAGGTTAATTTCCGCTGCTGCGGATGCCCATTGCTGGCTGGGACCGCTGCTGCGTCA  
GTTGGCGCGTGGGACTGGCTTACCCTTGTGCTCACTATGATGGCAACCTCATTATTAATCTTCTGCTTCCAGTGCATTAACGTA  
GGCGCGCTGTTTCCACGATGATATCGGCACCTGCGGTTGTCATTCCTGTCTTTGCCAATGCTTTATAGCCAGCGGACGGCAAACTCG  
CGTGGGAAAAACTCGCCCGGCACTGATTTGTATGGCATCGGCTGGCGAGTGAATATTGCTGAGTTAAACCACGGACTCCCGGATTTTGA  
CTGGGACGTTATACCTCTGGCATCGTCTAGCGTTAGTTTCCGTTGGTCTGCTGGGCATGGTATGCCCTGCGCAACCGCCGCTGGCTGCGGGAA  
AATCCCGACAAACATCCGATGATGTGGGCGACGGCGAGGCGTGGTACCGCTGCCGTTTCTCTCATCGGCTATCTCGTCCCTGTTACTGGC  
TGAATATACAAACCGCGGACTTCTCCTTACCTTTTGGCCCGCTCCGCTGGTGTATTAGTCTGATGGTTGCGATAGCCGCTGCTTGTGCTATG  
GGTTGGCGCACTCTGCTGGAACGTCGCCAGCCAGGATACCAGACAGTATTCGCGGCGCTGATTGTTTTCGAAACGCTGGCAGGTTGTGCT  
TACACCTTTTTACTCCGCCAGCAAAATGCCGCCGCTAATGACGCTGAGCGGATCGCGCTGTTAGTATTGGCGTGGTCAATGCGGTCAGAGCAA  
AACCAGAAAAACCTTAACTGAATCTGTCTCAGAAAGTTGACACGCTGGCAGTGAAGTAAATAAGCCTCTGCTACGTAAGGGTTATAGCTTTT  
CCTTAAAGATGCATTTAAATACATCTTATCTTATTAAGAATGAGGTATCAGCTATGGCTATCGCGACCAACCTTTAGTGAACTGGCGCTCT  
CTATCCCTCGCGCTTCCGCTCTGTTTCCGTAATAATGATATGGATTACTGCTGTGGCGTAAAGCAGACGCTGGCGCGCGGGCGGACGTAAGA  
ACTGGATGTTGAGGTCAATGAAGCTGAAGTGGCAAGCTTGTGAAACACCGATTGAGAAAGACTGGCGTAGCGCCCGGCTGGCAGAAATCATT  
GACCATATCATCGTGGCTAC

## PREDIKCE GENŮ U EUKARYOT

Genomy jednobuněčných eukaryot se navzájem výrazně liší (frekvence intronů, jak velká část genomu je tvořená geny kódujícími proteiny). Pro některá jednobuněčná eukaryota (s nízkou frekvencí intronů) je možné použít stejné postupy jako pro prokaryota.

### ÚKOL 2

Charakterizujte část genomu *Candida albicans* (kvasinka) pomocí programu **GeneMark** (<http://exon.gatech.edu/GeneMark>), verze pro prokaryota. Vyzkoušejte modely pro různé organismy a rovněž heuristický model. Věrohodnost predikovaných genů ověřte pomocí aplikace **BLAST** (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

```
TCAATGTCCTTTGTGATTAATCTTCGGACATATGCAATTTTGAATATTAATATGTTTCATATCAACATCACGAGCTTGGAACTGAATATTGAGT
AATCAGCATATTGCACCGACTCAACCTTGATTAATCTGCTAGTGTACTAATGTGACCTTTCTATTAGATCATTATCTTTCTTAAATGTGT
TGAAATCACGAGAAATTTGAAATCTACTAAGCCTGTATATTCAACACAGAAATACATCATAGATGTTTTTTTTTCTTATATTGTGTGTCAT
TAACCTGTCTCCCTCTCCTTTGAACAATCAATTAGTACTTTACAACCTCACAATTTGTAACCCCTGTCCGGATACACATTTTATTTCAAATTACC
GTGAGTGCATGTAATAATCAACTGGTTCGTCGTTTCATTTCAAGGTCACGCATAGTCTAATAAGTATGTCGTTTCGAGAACCTACTGTCA
TCTCTCGTGTAAAGAAGTATTTTTTTTTTGGTCCAGTTTGTCAACAAATATACTTTAAATCTCATAGAATGTTTATTTTCATTTACTACAGA
CCTTACAAAATCCTAACAGACTAGATGCTATCTATGGAGCATCAGTACATCCCAATCTGTTGGTTCATAAGCATCAGGTTGATCAATAGCC
TTAATTTCTTCTTAGTCAATTCAAAATCATCAACAGAAAGATTTCTTTAATCGAGATGGAGTTTTGTGTTTTGGTAAATGGTATATAACCTT
TTTGTAAATGACATTTAAATCAATTTTGGAGCAGTGAATTTATTTACTTTTGGCAATTTCTTGAATTCAGTATTGTTGACTTAATTTGTT
ACCATGGGTTAATGGTGCATATGCTCAACATTGATACCTTTACTTAAACACCAAGTAGCTAAATCCTGTCTCATACACCAAGGACTAATTTCA
ATTTGATGACAGCTGGAGGGATCGTTGCATTGGTCAACAATTTCTCAATATGATGTTTACCATAAATAGAAACCCCAATGTTTTTAAATCCAT
CTTTTTCCACAGCATCTGCAAACTTTCCAGCTTTCTAAACGTTTGGTCTTACCTGGTAATGGAGAATGAATCAATAAATAAATCAATATATTC
TAATTTATCACCACCTTGGCCATCATTGTTGAAATGGCTTGTAGTCTTGAAGTACCCCAATGATTATCCAAAGCTTTGTGGTGTAAAAA
AACTCAGATCGTGGTATATTGGGGTCTCTCGTAAGAATTTACTTATACTTCAATGACTTCTTCTCGTTCATATAACACTGCAGTATCAA
AATGACGATATCCAACCTTACAAGCTTCATAAACTACCGAAACCGTTTTATTTCTTGGGATATCATAACATCCTAATCCAATTTGATGGAATGGT
ATGACCCGAATTGAGTTGATTAATCGATATGACATTTTCTTAAAGCAATTTAGCAATTTGAGTGAATATAAAGAGATTTGAAATCTAGAG
AATTAATAAAGAAAGAGTGGTTTGTGTAAGAACAATTTGACCAATTTGGGTGGAACCTTTTTTTTATGGTTTCCGAACCTGGGGATGGCAAGC
AAAGCAATCTTTGTACAGTCTCAAACTTCTTCTATATCGTTCCTTCTGTGTAATACTGCACATTTGCATTTGCCAATTTTGTATTATATA
TCATTTCTGTTACAATGAATTTGACAGACGAGTATTTGGTTATAGTATTTGGACTATTTGGTCTTGTAGATCACCATTTACACTTATAATCAAAA
TTCTCTGGAATTAAGATCTAATTTGAAACAATACAGACGATTTCCAGCCTGTTATTTCCCGGATTTGGATCTTTTGGTTGCGTTAGTCCGAGGA
TATATGACAGTGAATTAATGGTGAAGCTATAAGACCACTGTTCTGTAATGGATGAGCAAAAATGGAATAGAGCAAGAAAGAGTGTGCTAGGGT
TGAGAAAAATAATAGATACTGTAGTGGTAGTAGTGGTATGATGCAAAATTTATGTTTGTATTTTTTTTTTTTGTGCTCCTCGTCGACAACG
ACTGCTGATTCACACCACCACCCTTTGTAGTCAAAATTAATACTGCTGGTGGCAAAAAAACACAGACACACATACAGACAATTTGCCGATA
TTTTATTTTAGGTCATTTGGAACGTCACCATACAGTATTCAGTATGATTAATTTGACTTGCAAAATACAAATCGCGTGGTTCATTTCCGAAA
AATTTGATAGTATCCATGAAAAATACAGGAGAGCCGTGAATTTATAACAAATTTACTCACATTTTGGCTGGCAATTTCTTAAACCCCTTAAT
TTTTTTCATGCAGTTTATTTTACGTAGAAGCTTATTTCTTAAACACTCGGTTATCCGAATATAATTTGCTTTCAGAGCAATTTGACTGATAT
AATACCATCAATTTGAATAAATCAGTATCATCAATGCTAGTTTGGTTCATATAAAGGTATAAATCGGTTTAACTGAAAGGTATTTGCTAATTA
ATTAGTTTGGGTTTTTGGATCAGCTGTTTCAAATCCAGGCTCGTGGTGCACATTTTTTATTTTTTGTGATGATTTTTTTCTTTCTTTGCTT
AATGTTGTTGCTCCATGAAAAATACAGCAGCAAAATGCTTCTGTAATGGATGAGCAAAAATGGAATAGCCCGGATGCGCCTGCAAAATGAT
GCAAAAATAGCCAAAATAGCAAAAATAGGTGGAATTTGTTCAACCAACCCACTCAACCACTAAAATTTGCTTGTCAAAGCCAATTTCAACG
AAAATTTGATTTTTTTTACTTAAATATATATAACCCCTGAAATTTCCCTTCAATTTGAAAAATTTTTTAAATTTATTTATTTCTTCTTTCTTT
TTCTTTTTCTTTTTCTTTTATTTCATCACAATTTGAATTCAAATCAATTAACATCAACAACTTTACAATCAAATTAATAACAAATGGCTAT
TAAATTTGGTATTAACGTTTCCGTAGAATCGGTAGATTTAGTCTTAAAGAGTGTCTTTGGCGAGAAAAGACATTTGAGTTTCTGCCCTCAACGAT
CCATTCATTTGCTCCAGACTATGCTGCTTACATGTTCAAAATACGATTTCTACTCACGGTAGATACAAGGGTGAAGTCACTGCTTCTGGTGCAGGACT
TGGTCATTTGATGGTCAAGATTTAAAGTTTCCAAAGAAAGAGACCCAGCTAATTTCCATGGGGTAAATCTGGTGTGACTACGTTATTGAATC
CACCGGTGTTTTTCAACCAACTCGAAGGTGCTCAAAAACACATTTGATGCTGGTGCACAAAAGTTATCATCACTGCTCCATCTGCTGATGCCCCA
ATGTTTTGTTGCTCGGTTTAAACGAGCAAAATACACTCAGACTTGAAGATTTATCTCAATGCTTCTTTGACCACCAATTTGATGCTCCATTTAG
CTAAAGTCTGCAACGATACTTTCCGTTATGAAGAAGTTTGTATGACCACTGTCCACTCCATCACTGCTACCCAAAAGACCGTTGACGGTCCATC
CCACAAGGACTGGAGAGTGGTGAAGTCTTCTGGTAAACATTTCCATCTTCCACTGGTGTGCTAAAGCCGTTGGTAAAGTTATTCCAGAA
TTGAACCGTAAATTTGACTGGTATGCTTTGAGAGTCCCAACCCAGATGTTTCCGTTGTTGACTTGTGCTGATGAGAAAGCTGCTTCTT
ACGAAGAATTTGCTCAAGCTATCAAGAAAGCTTCTGAAGTCCATTTGAAGGTTGTTTGGGCTACACTGAAGATGCTGTTGCTCCACCGATTT
CTTGGGTTCAAGTACTCATCTATCTTTGATGAAAAAGCCGGTATCTTGTGTTCCCAACTTTCTGCAAAATTTGATTTCTTGGTACGATAACGAA
TAGGGTTACTCCACAGAGTTGTTGACTTGTGGAACAGCTTGTCTAAAGCTTCTGCTTGAAGTAAAACAGACTTTGATTTGATTTTTTATAGTT
AATGAAAGTTATATAATGAGTTTTTGAATAAATATCTTACATTTGAAGAGAATATGATGTTGTTTTTTTTTGTTTTTATATAGGTGTTTGCAG
TTATAAGTAAAGCAATATCATACTGGAATATATAAATTAACAAAAAATTAATAATTAACAAGCGCGTATAAAGTAGTTTATGATTTGAAAC
TGATTTTTTTTTTTCAAATATAAAACAAACCACATCTTTCTTTTACTTCCATTTACCAATGCTCCAGGATAGAGATATATCGTCTGTTGATAATA
AATCAAACATTTAAGATCTATAAATCCAATTCATCCTATCGTCTATATAAACTCACTAAAAAACAATTTGGAGAAATTTGCAATAACCTAT
GTAAATACTTGTACCAAGATATCAACAACTTTCAAAATAGAAGATGATTTACTAGTTCAATATGAATATATCGTGTATTTGACTTACCTGAAC
CACTATTTTATAACAATTTCAATTTGAAGACATTTCAAGTATAAATCCAATCTAATAATCAAAATGTTGCAAAATGTTGCAATTTGATTTGATTT
GTGTTTCATATTTGTTAGTAAATGATGATGATTTCACTTTGGTATTTGAATAGAATTTACCAATCACCCCTAGCATAAATGAATATATTTTGAATTA
TAGAAGATATAATGATTTATCCATTGATTTATAAAACCACTTTGATTTACTGATTTATCTAATACCAACTTTGGTGGATCAATTTGGTAAATAAGTT
GGATTTCTTTGGGTTTACAATTTAGTATATTCATCTTCTACATCTAATAGTTGAGAAATTTTATAATTTGATTTCCGAAAACCGATTTTACCA
AAATTTGCTAGTGAAGTAAATTAATCAAGATATTTGAAATCTAGTATAAATTAATACCAAAATTAATAATTAATAATTAATAATTAATAATTAATA
TTATTAATAATTTAATTAATTTGCTAGTGTGGGAAATTTAAATTAATAATTAATGATAATCAATTAATTTGGTTTTGATTTGAATCTATAAAA
ATCATCTTTACAATAAGTAGTAAAGTGTACGTGGTCAATAAAGTAAATGATATATGTCGCTGACTGAGTGTGATTTTTCTATTTTTATCTCGCGC
GTTGTTATTTTTGTTGTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
ATTTTTTTTTTCTCTCTCCACTTTCTACTCTTTGAAACAATTTGAGTGTGAAAACAAAAAATAAAGAAATAGAACAGACTTTTTTTTATTTCTCATAT
TTCAAAACCGACTTGTTTCTTTTTTTCTTTTACAACCAATTTAATCAAAAAGAAAGCAACAATCATAAATGCTACTCTGTTGAAACCAATGAA
ACAGAAGCTTTGTTGAGAAAGCAGAATGATCTTTCCACAACCTGCTCAATTTGAAGAAAAATATCTCACCACAAGGAGGCTGCAGAAAGCAG
ATGACGACACTCTTAAAGAACCAATATGATGAAGCTAAAGAACCGCTGAATCTTTAAACAAGTTGAATCGATATTAGCACCTATTGTTTT
```

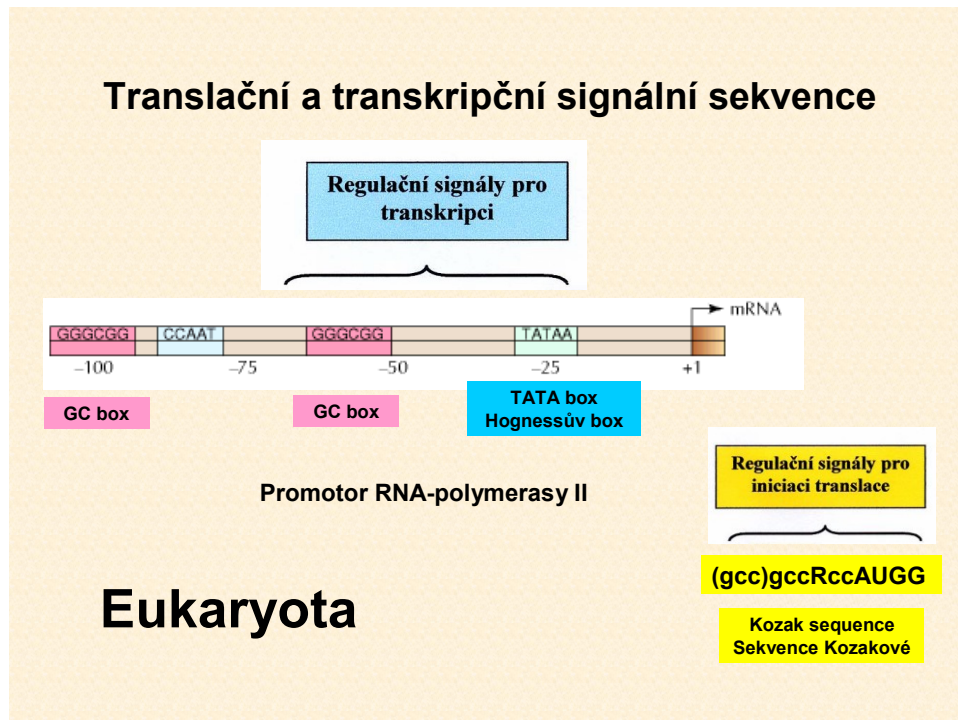


CACTGCATTGTCATTTTTTCGTGAGATTTTATCGTATTTTCAGTGAATGACCATGTTGTTTGGGATGAAGCTCATTTTGGTAAATTTGGATCCTAT  
TATTTACGACACGAATTTTATCATGATGTTTCATCCTCCATTGGGTAATAATGTTAGTTGGTTTATCTGGTTATTTGGCTGGATACAATGGATCTT  
GGGATTTCCCAAGTGGTGAAAAATACCCTGATTATATTGATTACTAAAATGAGATTGTTTAAATGCCACTTTCTCTGCCTTGTGTGTACCATT  
GGCTTATTTACTGGGAAAGAAGTTGGATTTTCCATGTTTACTACTTGGTTATTTACTTTGATGGTGGCTCTTGAATCAAGTTATGTCACTTTA  
GGTAAATTCATTTTGTGGATTCAATGTGTTATTTCTTCACCGTTGCTACTGTTTCTGTTTTTCACGTTTCAACAATTTAACAATAAATCAC  
AAGAATTTCTAGAAAATGGTGGAAATGGATTCTTTAACTGGTGTTCATTGGTTGTTACTTGTCTGTAAAAATGGTTGGATTGTTTGTAC  
CACATTGGTCCGGATTTACACTGTTGTTGACCTTTGGAATAAATGAGTGATAAATCTATTTTCATGGACAAAATACATTCAACATGGTTGCT  
AGAATTTGTGCTTTGATTCTTGTCCCAATTTTCATTTTCATGCTTTTCAATTTAAAGTTTCAATTTGATTTGTTGTATAAAATCGGGTACTGGTGATG  
CCAATATGTCATCACTTTTCCAAGCTAATTTGGCTGGTTCGGATGTTGGTGGTGGCCACGTTGAAGTATCCATGTTCCACTCGGTTATCACTTT  
AAAGAATCAAGGTTTAAAGTGGTGGCCTTTTACACTCCCACGTTCAAACATTTCCAGAAGTTCAAACAACAACAAGTTACTACTTATGGTCAC  
AAAGATTCAAAACAACAATTTGGATTTTCCAAAGAGCTAGAGGACAACCTTATTATGATACTTCTGTTAACACCCTGCATTTGAATATATTTTGG  
ACGGTATGCATGTAAGATTGATGCATCCACAACTGGTAGAACTTACATACTCATGATTTCCAGCTCCAGTGTCTAAATCTGAATATGAAGT  
TGCATGTTATGGTAAATTTGACTATTTGGTATCTTAAAGATAATTTGGACTGTTGAAATTTATGGAACAAGCAAGTATGAAGATAAATGAGATTA  
CATCTTTGACTTCGTCAATTTAGATTGAAGAATGAAGTATGTAATTTGTTATTTGGGGTCACTGGTACTACATTACCTCAATGGGGTTCAGAC  
AAGGTGAAGTTGTTTGTACAAAGAACCCATTTAAAAAGACAAGAGAAGCTT

### ÚKOL 3

Charakterizujte předcházející sekvenci části genomu *Candida albicans* (kvasinka) také pomocí programu **GeneMark** (<http://exon.gatech.edu/GeneMark>), verze pro nižší eukaryota. Výsledky porovnejte. U identifikovaných genů určete funkci.

### PREDIKCE GENŮ U MNOHOBUNĚČNÝCH EUKARYOT



Mnohobuněčná eukaryota se vyznačují komplexní organizací genomu, geny jsou separovány dlouhými intergenovými úseky, geny obsahují mnoho intronů, i velmi dlouhých. Exony/introny jsou identifikovány pomocí míst setřihu (GT na 5'konci intronu, AG na 3'konci). Vzniká velké množství chyb! Dlouhé introny jsou určeny jako intergenové úseky, krátké intergenové úseky jako introny, krátké exony nemusí být identifikovány.





GTACACGGCGTGGCGGGCCCGGCCCGCTGCCCGCCACTCACCTGACGAGGGAGATCTGCTCTTTCTGCTCGTGGTTCGAGTACTTCTCGATGTGAAAGAAGG  
TGTCCAAAGAACCTTAGCCAGCTTGACGCCGCTTCAGGTCCTGCAGCGTACTTCCCTCGGGGGAGGGAGTCCGCTCAAGGCCGCTAGCCCGGCTCACCTTCG  
GGGCTGGTGTGGGGTGGCCGCTCCTGTACACAGTGTCTAAGGACCGCCGACCAAGCTCCGGGCTCACGCCACGTGTGACGTGAAAGGGGACACCGGAGT  
GTCCGTGTGTGGTGGGCTGAGCCGCTGAGGCGCTGGGAGCCGCTGATGCGGAGGCGCTAGGCGGACCTCCGGTCCCGCCAGGATCGGCTGGGACTAGCCAGGGAGGCC  
GGCCAGACCTGGGCTGCTGATCTGAAATTACAGAGAACGGAGTCCCGCTGACGGACAGGGGTTCCCGGGGAGCTGACGACGAGGGGAGGGGAGGGG  
TGGCGGGGGGGGAGGGCTGCAGGGGCTGCCCGGCCTCAGAGCCATTTCCAAGTGAAAGTCCAGAGCTGGATTCTGAGATGCAGAAAGAGCCAGGGGGCGT  
GGACGGTGGGGTCCCTTTGCTTAGAGGGACCGGGGTTGGTGGGAGGACTGGGAGATGTTGGGGGAAACAGGGGTGAGGGGAGGAGCCCGGGGGGGGTTG  
GAGGGAGGAGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCC  
CCGGGGGGGATGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGG  
TGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGG  
AGGAGCCCGAGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCC  
GGGGGGTGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGATGGGGGGTG  
GGAGGAGATGGGGGGTGTGAGAGGGAGGAGATGGGGGGGTGAGAGGGAGGAGATGGGGGGTGTGAGAGGGAGGAGATGGGGGGTGTGAGAGGGAGGAGATGG  
GGGGGAGAGGGGGGGTGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGGAG  
TGTGTGGGATGACAGCAACGTGGCACTGTCCCTCAGCCAGTGCACAGTCCCGGTGAGCTCCGGCTTAGTTCCTAGGTACCCTTCGGGGGTGTCGA  
CGCCCGAAATAGGAGTTCTGCACTGTGAGACTAGGCCACCTGTGCTGGCCCTCCACAACTCAGTATGACAGCCGTCCAGAGGACGTCCCTCACCCTGTG  
TCCGTCCCTCCCTGTGTCCTTCCCTTGCCCTGCCCCCTCAGCCCTGGGCCCTCCCTTCAGCCCTGGGCGTCCCTCCCTCCCTCCCTCCCTCCCTCGG  
CCCTCCCTCCCTTCCCTCAGCCCTGGCCGTCTCCCTCCCTCCTCCCTCCTCCCTGGCCGTCTCCCTCCCTCCTCCCTCCTCCCTGGGCTGTCAATCC  
GTCCCTCCTCCCTGGCCGTCTCCTCCCGCTCCCTCCCTCAGCCCTGGCCGTCTCCTCCCGCTCCCTCCCTCAGCCCTGGGCTATGCCCTCAGGGGATCCAC  
TTCAGTCTCGGCTTCGCTTCAGGATCCAGCATCTGGCAGAGCCAGTCTGGAAGGGACGGCCCTCGATGGCCATGCTTCCAGCCCTCCGGACTCTCGTAGAAGT  
ACTCGAGCTCGAACATGAGACAGGGGGCTCCCGTCCAGGTCCATGCAGCGAAACGATCTCGATGTGCGGCACGGGGAGCTGTGTCAGCCTTCCTGGCC  
TCCCAGCCCGTGCCTGCAGCCCTGAGGAGCCCGCCCGGGGGTGCAGCGTCCCGCTGTGCTTGCAGCCCCACCGGGGCTGCAGGCATCCCTGCCCCCT  
GGCCCGCCACCGGGCGCCGAGCTCCCTGTGCCGTGACGCCGTTGAGAGGGAGGAGCCCGGGGGTGAGAGGGAGGAGCCCGGGGGGGTGAGAGGGAGG  
CCTGTGCCGTGCAGCCCCACAGCGTGCAGCATCCCGCTGGGACACATGTCACATGGGGGGTCCCGGCCCTCCACTGGGACAAACGATGCCGAGCAGGA  
ACCCACTGGTGGTGTGTTTTGTCTCTCAGAGATCAAAAACAGCAAAAGTCCGATAGCTGATCTTCCCTCCTTCTGCACTTTTTGTCCCTATAGATGAA  
GCCAGATGGAGACAGTGCCTCCCTCAGAGGGAGGAGTGTGAGACTTACAGGAGAGCTTCCAGGAGGAGCCCTGTCACCGGCTCGGTGAGGGAGAGCCCG  
TCCCTGCAGGGCAATGCTGCCGTGAGCCGTGAGCCGTGACCTGACCTGTGACTGCTCAGTGAAGATCTGTCTATCATCTTGTGAGAAAGGGTGGAAAGGAA  
TGCGGTGTGAGGGACGCCGTGCCCTCGGCCGAGCCTACCAGCCCGCCCGGGAGCCGAGTGGATGGGTTGGGACCGAAAAGCTGAGAGGACCGCTGGGT  
GGAGAGCCCGGGGGCGAGTCCCTCAGCCCTGACCGGCCATCTTAGGGCCCTCACAGCTTTCCGCTCGAGCAGAGAAATCCAGCATGAAATCCACAGAGC  
CTGACCGGGCTCCAAAGCCGACTTTGAAAGAGCAGAAATATTAGGCTCAAAGGTACAGCTTTCAGAGCAGAGAGGAGACTTCAGGTGTGATCACGGA  
CGTTTCAACAAGAGTCCAAACGGGAGAGCGGAGTGAAGCTCAGCAACGGGAGGGGAGTGAAGCTCAGCAACGGGAGGGGGAGTGAAGCTCAGCAACGG  
GAGGGGAGTGAAGCTCAGCAACGGGAGGGAGGGAGTGAAGCTCAGCAACGGGAGGGAGGAGACTCAGCAACGGGAGGGAGGAGACTCAGCAACGGG  
GGAGGGAGGAGGAGACTTCGCCAACGGGAGGGAGGAGACTTCGCCAACGGGAGGGAGGAGGAGACTTCGCCAACGGGAGGGAGGAGGAGACTTCGCC  
GGGAGGGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAGCTCG  
CAGCCAGGGAGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAG  
ACGGAGGGGGAGGAGGAGACTTCAGCAACGGGAGGGAGGGAGGAGACTTCAGCAACGGGAGGGAGGGAGGAGACTTCAGCAACGGGAGGGAGGGAGG  
GCCAACGGGAGGCGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGTGAAGCTCGCAACGGGAGGGAGGGAGT  
CGCAACGGGAGGGAGGAGGAGACTTCGCCAACGGGAGGGAGGGAGGAGACTTCGCCAACGGGAGGGAGGGAGGAGGAGACTTCGCCAACGGGAGGGAG  
CGCCAACGGGAGGGAGGAGGAGACTTCGCCAACGGGAGGGAGGGAGGAGACTTCGCCAACGGGAGGGAGGGAGGAGGAGACTTCGCCAACGGGAGGG  
CTCGCAACGGGAGGGAGGAGGAGACTTCGCCAACGGGAGGGAGGGAGGAGACTTCGCCAACGGGAGGGAGGGAGGAGGAGACTTCGCCAACGGGAGGG  
AGTGAAGCTCAGCAACGGGAGGGAGGAGACTTCAGCAACGGGAGGGAGGGAGGAGGAGACTTCAGCAACGGGAGGGAGGGAGGAGGAGACTTCAG  
GAGACTCAGCAACGGGAGGGAG  
GAGACTCAGCAACGGGAGGGAGG  
GGGAG  
GAGG  
GTGAGACTCAGCAACGGGAGGGAGGGAGG  
AGTGAAGCTCAGCAACGGGAGGGAG  
CGCCGCTTCCTGCGGTACAGGGTTCCTGGGAG  
GAGTTCGCTGCCAGGAGTGGCTCCAGAGGAACGGGAGG  
CCAGCTGTGAGGTGCTGGCCAGGCCACAGGGCCGCGCTGTGGAGTGTGCTCTCCTCGTGGCTAGTTAGGGGCCAGGGCCAGGTGGCTCCTGAACTGAGC  
GCAGGACAACTCAGGAAATGCTAGCTGCTGTGCTAGTGCCTCCGAGCCAGGGAGTACTGCGGCCACGCGACCCAGGACAGCCCGGGAGACGCGAGCGGTG  
AGCTACAGTTCGGGAACCAAGGCTATTAGTGTGACACACACAGATGCCAGCTTAATTCGGAACAGCTTTGGTATATGATGTTTTGGGTAAGAAGGAC  
ACTTAAAGGAATCCTTAAAGAACTCAAGCCACAGGCCACAGGCCGCTTTGAGTTCCTGCAAGTGTGGATGACTTTAGAGGGCCGGGAGGCAATGGCTGTG  
GAGGACACTTCGGGCGCCGCTCCCGGGAGGATATGCAAGGAGGATTAAGCTGGGCTGGGCTCCACTAGCCGCGGATGGAATGGAAGTCTTCAGGCGCTGGGG  
TGTAACTGAGCTTCAGGGAAGGCTGTGGCTGCCCCCTGTCTGGGTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
ACCAGCGCCACACTGACCTGTAGACGCCAGGAGGCTCGGGCAGAGGCGCAGCGGGACCCAGACAGCGGGCAGCGACTGGGAAGGAGAGGAGGAGGAG  
GCAGACAGAGGGGCGAGTGGGAGG  
TCGGTGTCCGGCGGCGCCCTGTGCCCTCAGCCACAGCCCTGATGTCAGTGTGCGCCAGGAGTGTGCGCCGCGCAGGTCGTCGCGCTCGATGAGAGCTG  
CAGAATTCAGTATGAGTAGAATGCTCGTAGAGGAATTCGTCAGTGGTGTGATGCTCGCCTCCCTCCAGCAGCCAGTTCCTGCCAAGGAGGAGGAGG  
GGGGCGTGCAGCGGGGGCTTCGGGGATGCCCAAGTCCCGTGGCTGGGGCGGGGAGGAAAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG  
GGGAG  
TGGAGTGGAGGGGTAGGGACAAGGAGGGGAGGGAGG  
CAGGACCGGTTACCGCTGAGAATCCGCTGGATGACTGCGGGGCGCTGTGTCAGTGGGTGGTGCAGAGACCCCGAGGAGCCTCGCCCGCAGGAGACCG  
GAGCCGGGAGAGGGGCGCCCTTGGTCTACGCGCACGGGCGCCAGGAGCAGGCGGGGAGAGGAAAGGCGCTCGGGAGGAGGAGGAGGAGGAGGAGGAG  
AGCGGGAGGGTTCGCGGGGAGG  
CCGGCGGGTGGAGACTTCAGAGGCGGCGGCTCCTCCGAAAGTTCAGGAAACCGGGCCCGCCCGCCCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAG  
AACCCCGCGCCCGGCTGGAACCGAGGCGCCCTCCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
GTGCTGTTTACCAAGTGGGTTGGAGACAGAGGCGGCTGTCACCTAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG  
GTGAGCGATGAGGTTGCGGTTAGACGCGGCGCTCCCGTAGAGGATAGGGCTATGTTGATAGACGCGGCGCTCCCGTGGAGGATGAGGAGTGTGGTGTAGACG  
GGCCCTCCCTGAGGTTGCTGAGTGTAGACCGGAG  
CGCGGGCTCCCTGAGGATGAGGCTGTGAG  
GGCCCTCCCGTGAAGGATGAGGCTGTGTTGTGAGAGG  
GATGGAGAACTGGCCGCTGCTGCGGAGTTCGAGGAGTTCGGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
AGGGTGGAAACAGTGTGGTGTTCACAGTGGTGGTGGAG  
AGGGGACCGGCTTCGCGCACCGGGGAGG  
TGAGCAGACCGGAGGCGCTGCCCTCCCGCCGCAATAGAAAGCAGCGCCCTTCCAGTAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
GTCAGAGGGCTGGAGCGCGGCTCCTGTCTGCTTACCTCGCCCTGACCGAGCAGCGCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
CTGGGCATTCGAACATCAAGGAG  
GCTGGCACGTTGAACAGAGGCGGCTGCAATCCCGACGCTTAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
CGCGGTGAAATCGGAGGCGACAGAACTTCATGAAATGTGTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG



## **ÚKOL 5**

Analyzujte část genomu mnohobuněčného eukaryotického organismu také pomocí programu **GENSCAN** (<http://genes.mit.edu/GENSCAN.html>). Porovnejte s výsledky z úkolu 4.