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 EMBOSS Needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html).

SEKVENCE *E. COLI*

```
TTAAGAAATTCCGGTTATCAACTTCGAGGCCCCTTTCAGGTCACCGTTGGCGTATTTCGGGGCGATAGATTTGTATTTGTAGCCGCGAATACTG
CGATCGCCCCCGGCGAAAAAACGCAGATCCGGCGGTACTTTGTCGAAATCACCGGTTTCAATCCAGCCCAGCGTGCCGCGTGTAACGAAACGAT
GGCGATCGTACAGTGTGCGGATCCAGACGTTCTGCGCCTGGAAAACGGAGAAATCGACATCTGAGCCCCAGGCCGTGTTGGAGTAGTCGATAGA
GTAGCGTTGCGAGTCGCCCCTGGTTGGCATCAGGCCACCACGAGAACGCGTGCGGCTAATCATCACCCCAGGATAAAACAGCATCGTGGTGTTG
GTAATTTCACCCTGAGTAAAGTGGTCGAGACTCCAGCGCAGGTTAATGGCACGCTGCCAGCCGCTGGAGAGATCCCAGTAGCGAGAAGCCACCA
GCGTAGTGGAGTCAGATTCGGTATCGTTCAGGTCAGTGCGCTTAAAACCGCCCTGCACCAAATAATATTGTTCCAGTGGATTCTTCAGCAGCGG
CATTTTATAGCTGAAGTCGAGGGTCTGTTCCGGCGCGGAAATACTGGTACTGGTGGTCAGACTGTGACCATAAGAGTTCATCCACGGCTTTTTC
CACGTAGCTTTCACGCGCGGTCCCACGTCCGTAGAGTAACCGACCCCGGTTTCGATGGTGTTTTCAGTTCGCGGCGAAACCACGCCCGTCAAGG
GTAATACTTTCGTTTCGCGCGCTTTATCAAATTGTGGAGCCACTACGACCGAGTTAAACCAGCCGGTAGCAGAAAGTCGACGGTTAAGTTCCGC
CAGATCTTTCGATTCGTACTCATCGCCCTCTTTAAACGGCACCAGATTTTGCAGGTATTCATCGCGGATTTGTGATCCTTCAAAGGTCACATGC
CCAAAGCGGTAACGTTCGCCACTGTTATAATCAATATCCCAGAAGGCTTTATGCAGGCCGAGCGCAATGCCCAGCTGCGCTTTGGTAAATTCGC
TATCGAAATAACCTTTACGCAACGCAATGCTGGTTAAGGACTTTTTGAAATTTTCATAATCGCCCTGGTTCAGTACCGTGCCAATAGCCGGGCG
AGTATCGAGCAATTTCAAATAGTCTTTATCGGTCCGCGCGCCGCCGCGCAATACCACATCGGTGCCGCCAATTAACACCGGCACGCCTGGCGTG
ACTTTGGCGATCAATACCTGCCGCCCTTTCTTTGGCGGTGGACGGAGATCAAATTCAATGGTCGGCTGGTAATAACCCAGCGCTTTCAGACCTT
CGCGGATGGCATCATCGACGCGTGCGCGAAAGCGACGGTCTGGCGTCACTTCATCACTTTCAATCGTAGAAAGCTGCGCACGAACGTTCTTTTC
CAGCTGTCCCGATAACCCCTCGACCTGTAGACGGACGTTCGCGGCGACGGCAGATCCGCTTAAGCAGAGTAAGCTTACACAGCATAACTGTCGG
ATATAGCGCATATTTTCTCCTGAATATCCTTTTTTCCTGCCCCTGGAACGCCGTTAAACGGCTTAACAAAAATCCAGTAATATGGATTAAAAAG
CAGACTAAACCCCAAATATTTCTTATGTTTTACTTTAGACCTATTCACGGTGGTTATTGTGTGCAAATACGCCTCTTGTTACAACCTTAACCCC
AATGACCGATTTTCGGGAGAGCGACACCATGAGTTTATTTGATAAAAAGCATCTGGTTTCCCCCGCCGATGCCCTGCCTGGACGTAACACCCCG
ATGCCCGTAGCCACGCTGCATGCGGTCAACGGTCACTCAATGACCAATGTACCTGACGGAATGGAGATTGCCATTTTTGCGATGGGTTGTTTCT
GGGGTGTGGAGCGTCTGTTCTGGCAGTTACCCGGCGTTTACAGCACCGCCGCAGGCTATACCGGCGGCTATACGCCAAATCCGACTTATCGGGA
AGTGTGCTCCGGTGATACGGGTCATGCCGAAGCGGTACGCATTGTTTACGATCCTTCTGTCATCAGCTATGAGCAGTTGCTACAGGTATTTTGG
GAGAATCACGATCCCGCCCAGGGAATGCGTCAGGGCAATGACCACGGCACGCAGTATCGTTCAGCGATTTATCCGCTGACCCCAGAACAGGATG
CCGCAGCTCGCGCCAGTCTGGAACGTTTTCAGGCGGCGATGCTTGCCGCCGATGACGATCGTCACATCACCACGGAAATCGCTAACGCCACACC
GTTTTATTATGCCGAAGATGACCACCAGCAATATCTGCATAAAAACCCGTATGGTTACTGTGGAATTGGCGGAATTGGCGTCTGTCTGCCACCG
GAAGCATAGCGTTACGGGTACAAATGTAGATTGTTGATAAAGTGCGTTTTGTTTATGCCGAATGCAGCGTGAATGCCTTACCAGGCCTACAAAA
TCGTCCAAATTCAATATATTGCAGGGACTGCGTAGGCCTGATCGGCATAGCGCATCAGGCAGTATTGCATTTATCAGCAGTCTGAATTTTAACC
CTCTGGCGACTTTACAGTACCTTACGCTATACTAGCCACTGAAAATGCCGGATCACTTTCTTCGAATCGGCTTTCAATGTGTATTTCACACAAA
TTAATCAACTTCCCCTTCCGAGGATCTGGCCTGAAAGGTCGGATAAGATATGTTAAACAGTATTTTAGTCATACTCTGCTTGATCGCTGTAAGT
GCGTTCTTCTCGATGTCTGAGATCTCACTTGCCGCCTCACGCAAAATCAAACTTAAACTGCTGGCTGATGAAGGCAATATAAATGCCCAACGCG
TTCTGAATATGCAGGAAAATCCCGGCATGTTCTTTACGGTGGTCCAAATCGGTCTGAACGCAGTTGCGATTCTCGGCGGTATCGTCGGTGATGC
GGCATTTTCTCCAGCTTTTCACAGCCTGTTCTCCCGCTATATGTCGGCAGAACTCTCTGAGCAACTGAGCTTTATTCTCTCTTTCTCGTTAGTG
ACTGGCATGTTTATCCTGTTTGCGGATTTAACCCCGAAACGCATCGGTATGATTGCGCCAGAAGCGGTGGCTTTGCGTATCATCAACCCGATGC
GCTTCTGCCTGTACGTTTGCACCCCGCTGGTGTGGTTCTTCAACGGCCTGGCGAACATGATCTTCCGTATTTTCAAACTGCCAATGGTACGTAA
AGATGACATCACTTCTGATGACATCTACGCGGTAGTGGAAGCCGGTGCGCTGGCGGGCGTGTTACGTAAACAGGAACACGAGCTGATTGAAAAC
GTCTTTGAGCTGGAATCCCGTACCGTTCCGTCTTCAATGACACCGCGTGAAAACGTGATTTGGTTTGATCTCCACGAAGATGAGCAAAGCCTGA
AGAATAAGGTGGCGGAACATCCGCACTCTAAGTTCCTCGTCTGTAATGAAGATATTGACCACATCATCGGTTATGTCGATTCTAAAGACCTGCT
```

GAACCGCGTGCTGGCTAACCAAAGCCTGGCACTGAACAGCGGCGTACAAATTCGCAACACGCTGATTGTGCCGGATACGTTAACCCTTTCAGAG
GCGTTGGAAAGTTTTAAAACCGCAGGTGAAGACTTCGCGGTGATCATGAACGAGTACGCGCTGGTGGTGGGGATCATCACCCTCAACGACGTGA
TGACCACGCTGATGGGCGATCTGGTCGGTCAGGGGCTGGAAGAGCAGATTGTCGCCCGTGATGAGAACTCATGGCTGATTGACGGCGGCACCCC
AATTGACGACGTCATGCGCGTGCTGGATATTGACGAGTTCCCGCAGTCGGGCAACTACGAAACCATCGGAGGCTTTATGATGTTTATGCTGCGT
AAGATCCCGAAACGCACCGATTCGGTGAAATTCGCCGGCTACAAATTTGAAGTGGTGGATATCGATAACTACCGCATCGACCAGCTGCTGGTGA
CCCGGATCGACAGCAAGGCCACCGCCCTTTCGCCAAAACTGCCTGACGCTAAAGATAAAGAAGAAAGCGTCGCGTAACCCCCAGAAACATCAAC
GGCTCCTGAATCAGGAGCCGTCTTATTACTGCATAGCACTTTGGTTAAGCCATCTCTGTTTGCAGACGCATAACCTGACGGTTAACTTCGGACA
TCACTGACAAATGCAGCTTATCCTTCACTTTTGGGATAAGAATCTTACCCTTATCAAATTCAAAAGCGCCAACGTCCTTGATGTATAACCGTCC
ACGGAACAGGATCTTTACGTACTTCGCCACCTGAAGTGGGTTGTACCGTTGGAAAATTTTCATTGTTATCTCCTGCTGAGTATTACGCCTGTGC
GGTGCCACAATCGGCCCAACTATTATGAGGCGCAAATTTTAATGCCTAGTGACTATAGACTATCCGGGCAATGTTTCCACCGCGTATAACTTTT
TTTACCTAAAGGTTACAATTATTCAGAATTATCTTTTTACCGAAGCGCGGTCTTCAGTATAAGCATTCATTTTTCATATGATTTGTGCGCTTGA
CCGCAAACTGGCATCACACTTGCGGGAAATTCGATAAATAGCACATATGATTAAAACTCAGACCCAAGTGGTCGGATCACCTGCATATCATAAG
AAGGAAACACCATGACCCTACGCAAGATTCTGGCACTCACCTGCCTGCTGTTGCCGATGATGGCTTCCGCACATCAGTTCGAAACCGGTCAGCG
AGTGCCGCCGATTGGCATCACCGATCGTGGCGAGTTGGTGCTTGATAAAGATCAGTTTAGCTACAAAACCTGGAACAGCGCGCAGTTAGTGGGA
AAAGTGCGAGTACTGCAACATATTGCTGGTCGCACCTCTGCAAAAGAGAAAAACGCGACGCTGATTGAAGCGATTAAATCAGCGAAGTTACCGC
ACGATCGTTACCAGACCACCACCATTGTTAACACCGACGACGCAATTCCGGGTTCCGGCATGTTTGTGCGCAGCAGTCTGGAGAGTAATAAAAA
GCTTTATCCCTGGTCGCAGTTTATTGTTGATAGCAATGGCGTCGCACGCGGTGCCTGGCAGCTGGATGAAAAAAGTTCCGCTGTAGTGGTGCTG
GATAAAGACGGTCGCGTGCAATGGGCCAAAGACAGGGCGCTTACTCAGGAAGAGGTGCAGCAAGTGATGGACCTGCTGCATAAATTAATTAATA
AATAGACACTCTGAACCCAGGATTCAGGAACGACTCGCGCGGAGTGTAATCCAGCGGTTTACCCTGCCAGTCGTGAACGTGCGCTCCGGCAGCT
GCGGCCACAGCATGTCCAGCGGCGGTGTCCCAAATATTCGTTGGTCCGAAGCGCGGGTACAGCTGCGCCTGTCCTTCCGCCACCAGGCAGAATT
TCAGCGAAGAGCCGATGGACGTGGTCTGATGTTCGCCAAGCTGTTGCAGATACTCTTTCAGCTCCGCATCTGCATGGGAACGGCTGATAACCAC
CAGCGGCGGGCGCGCATCGCGGACCTGAATCTGCTTGCGCACACCGCACTCTTCTTTCCAGGCCTTGCCTTCTGCCGCGCTGTACATTACGTTC
ATTACCGGCGCATACACCACACCTAAAATCGGTTTGCCATGGTCAATGAGCGCAATGTTAACGGTGAATTCGCCATTGCGTTTAATAAACTCTT
TAGTACCATCCAGCGGGTCTACCAGCCAGTAACGCTGCCAGTGCTGACGGACTTCCCAACCGGGAGGATCTTCTTCAGAAAGGACCGGAATATC
CGGTGTCAGCGTACGTAAACCGTCCATGATAACGGTGTGAGCGGCAATATCCGCTGCCGTTACCGGAGAATTGTCCGCTTTGCTGACGACGTCC
ATCGGTTTCGTCCCGTCGTAGACCTGCATAATGGCATCGCCTGCATTCCGTGCAAGCTGGCATACTTGATCTAACATTTCTCCACCTCGTCTCT
GTGAGCGGTGTTAACTTATTGTTTTACTTATACCCTATCGTTAATGAATGCGCCAACTGTGATAGTGTCATCATTTTCAAAGCGTAAAATTGTG
GCATTCTTCACTGTTCTATAAGTAAGACGTTTATTCTTCCTTTTCTTTCGTATTCCCGATGATAAAAGGATGTCCCTGATGATTAAGTTTAGCG
CAACGCTCCTGGCCACGCTGATTGCCGCCAGTGTGAATGCAGCGACGGTCGATCTGCGTATCATGGAAACCACTGATCTGCATAGCAACATGAT
GGATTTCGATTATTACAAAGACACCGCCACGGAAAAATTCGGACTGGTACGTACGGCAAGCCTGATTAACGATGCCCGCAATGAAGTGAAAAAC
AGCGTACTGGTCGATAACGGCGATTTGATTCAGGGGAGTCCGCTGGCCGATTACATGTCGGCGAAAGGATTAAAAGCAGGTGATGTTCATCCGG
TTTATAAGGCGCTGAATACGCTGGATTATACGGTCGGTACACTCGGCAATCATGAATTTAACTACGGTCTGGATTACCTGAAAAATGCGCTGGC
GGGAGCGAAATTCCCTTATGTAAATGCCAACGTCATTGACGCCAGAACCAAACAGCCAATGTTTACACCGTATTTAATTAAAGATACCGAAGTG
GTCGATAAAGACGGAAAAAAACAGACGCTGAAGATTGGCTATATTGGCGTCGTGCCGCCGCAAATCATGGGCTGGGATAAAGCTAATTTATCCG
GAAAAGTGACGGTGAATGATATTACCGAAACCGTGCGCAAATACGTGCCTGAAATGCGCGAGAAAGGTGCCGATGTCGTTGTCGTTCTGGCGCA
TTCCGGGCTGTCTGCCGATCCGTATAAAGTAATGGCGGAAAACTCAGTTTATTACCTCAGTGAAATTCCGGGCGTTAACGCCATTATGTTTGGT
CATGCTCACGCCGTTTTCCCGGGTAAAGATTTTGCTGATATCGAAGGGGCTGATATCACCAAAGGCACGCTGAATGGTGTTCCGGCGGTAATGC
CGGGCATGTGGGGCGATCATCTTGGGGTGGTCGACTTACAACTCAGTAATGACAGCGGTAAATGGCAGGTGACGCAGGCGAAAGCGGAAGCTCG
GCCGATTTACGACATCGCCAATAAAAAATCCCTCGCGGCGGAAGACAGCAAGCTGGTAGAAACACTCAAAGCCGATCACGATGCCACACGCCAG
TTCGTCAGCAAGCCAATCGGTAAATCTGCCGACAATATGTATAGCTATCTGGCGCTGGTGCAGGACGATCCGACCGTGCAAGTAGTGAACAACG
CGCAAAAAGCGTATGTCGAACATTACATTCAGGGCGATCCGGATCTGGCAAAACTGCCGGTGCTTTCAGCTGCCGCACCGTTTAAAGTTGGTGG
TCGCAAAAATGACCCGGCAAGCTATGTGGAGGTGGAAAAAGGCCAGTTGACCTTCCGTAATGCCGCCGATCTTTATCTCTACCCCAATACGCTG
ATTGTGGTGAAAGCCAGCGGTAAAGAGGTGAAAGAGTGGCTGGAGTGCTCCGCCGGACAGTTTAACCAGATTGATCCTAACAGCACGAAACCGC
AATCACTCATCAACTGGGATGGTTTCCGCACCTATAACTTTGATGTGATTGATGGTGTGAATTATCAGATTGATGTTACCCAACCTGCCCGTTA
TGACGGCGAGTGCCAGATGATTAATGCCAATGCGGAAAGGATTAAGAACCTGACCTTTAATGGCAAGCCGATTGATCCGAACGCCATGTTCCTC
GTTGCCACCAATAACTATCGCGCTTACGGCGGCAAATTTGCCGGTACGGGCGACAGCCATATCGCTTTTGCTTCACCGGATGAGAACCGCTCGG
TGCTGGCAGCGTGGATTGCTGATGAGTCGAAACGTGCGGGGGAAATTCACCCGGCGGCAGATAACAACTGGCGTTTAGCACCGATAGCTGGCGA
TAAGAAACTGGATATCCGTTTCGAAACTTCTCCGTCAGATAAAGCCGCAGCGTTTATTAAAGAGAAAGGGCAATATCCGATGAATAAAGTCGCG
ACCGATGATATCGGGTTTGCGATTTATCAGGTGGATTTGAGTAAGTAAAACACTTCTTTTTCGGCCTATAAATCATCAACCGCATCCGGCATTT
ATTGGCGGATGCGATGCTGGCGCATCTTATCCGCCCTACAAGCCATGCACCGTAGACCAGATAAGCTCAGCGCATCCGGCAGTTATGCCGCACG
TTCATCCCGCACCGCCAACACCTCGGGCAAATTCAACTCAATCCAGTCCGCCAGTGCAGCAACCTTTTCGCTCACCTGCTCGCCCAGCGGCGTG
AGGCTATATTCCACATGCGGCGGCACCACCGGATACGCGATACGGTTAAGAAAACCATCCTGTTCTAACGCCTGTAACGACTGCGCAAGCATCT
TTTCGCTCACCCCGCCGATTTTGCGCCGCAGGTCGCTAAAGCGATGAGTACCTTCGCGTAGCGCCACCAGAATCAACACCCCCCAACGGCTGGT
GACGTGTTTCAACACCTCGCGCGACGGGCACTGTTCCGCAAAGAGATTACCCTCTTTCAGTTGTTGCGACAGGCTAACCTGACTCATTTCATAC
TTACCTTTTTGTGCGTACTTACTAAAAGTAAGTTTAGGTGTTAGCGTATTTAAACACAAGACAAAACGATGGAGACTTCCCATGATCGCTATTA
CTGGTGCCACTGGCCAACTTGGTCACTATGTTATTGAATCCTTGAAGAAAACGGTTCCTGTCAGCCAAATAGTGGCTATCGTTCGTAATCCGGC
AAAAGCCCAGGCTCTGGCAGCACAAGGCATTACCGTGCGTCAGGCTGACTACGGCGATGAAGCCGCACTGACATCTGCACTTCAGGGAGTGGAA
AAACTACTGCTGATCTCTTCCAGCGAAGTGGGTCAACGTGCCCCGCAGCATCGTAATGTTATTAATGCCGCAAAGGCGGCTGGTGTGAAATTTA
TCGCTTATACCAGCCTGCTACATGCAGATACCTCCCCGCTCGGCCTCGCCAATGAGCACATCGAGACGGAGAAAATGTTGGCTGATTCTGGCAT
CGTTTACACCCTGCTGCGCAACGGCTGGTACAGCGAAAACTACCTCGCCAGCGCCCCGGCAGCACTGGAACACGGCGTATTTATCGGTGCGGCG
GGCGATGGCAAAATCGCCTCAGCAACGCGGGCAGATTATGCGGCAGCTGCGGCACGCGTGATTAGCGAAGCCGGTCACGAAGGCAAGGTTTACG
AACTGGCGGGCGATAGTGCCTGGACATTGACACAGTTAGCGGCAGAGCTGACCAAACAGAGCGGCAAACAGGTTACCTATCAAAATCTGAGCGA
AGCCGATTTCGCCGCGGCACTGAAAAGCGTCGGACTGCCCGACGGACTGGCGGATATGCTGGCGGATTCTGACGTTGGCGCATCGAAAGGTGGT
CTGTTTGATAACAGCAAAACGCTTAGCAAATTGATTGGCCGCCCAACGACAACGTTAGCCGAAAGCGTAAGCCATCTTTTTAATGTTAATAACT
AGTTAATTAAAGTGGCATCCTCCCGCATCCTCTCTAATAATGACGGGATGCCGGGAGCAATCATGTCTGCTTCCTGAACTTTCTTCTGACAGAC
CAATGGATGCCAGTAATGATTAGCGGCGTGCTGTACGCCCTGTTAGCAGGGTTGATGTGGGGGCTTATTTTTGTCGGGCCGTTGATCGTACCGG
AATACCCGGCGATGTTGCAGTCGATGGGGCGTTATCTGGCGTTAGGGTTAATTGCGCTGCCCATTGCCTGGCTGGGACGCGTGCGTCTGCGTCA
GTTGGCGCGTCGGGACTGGCTTACCGCCTTGATGCTCACTATGATGGGCAACCTCATTTATTACTTCTGCCTTGCCAGTGCCATTCAACGTACT
GGCGCGCCTGTTTCCACGATGATTATCGGCACCCTGCCGGTGGTCATTCCTGTCTTTGCCAATCTGCTTTATAGCCAGCGCGACGGCAAACTCG
CGTGGGGAAAACTCGCCCCGGCACTGATTTGTATTGGCATCGGCCTGGCGAGTGTGAATATTGCTGAGTTAAACCACGGACTCCCCGATTTTGA
CTGGGCACGTTATACCTCTGGCATCGTGCTAGCGTTAGTTTCCGTGGTCTGCTGGGCATGGTATGCCCTGCGCAACGCCCGCTGGCTGCGGGAA
AATCCCGACAAACATCCGATGATGTGGGCGACGGCGCAGGCGCTGGTCACGCTGCCGGTTTCTCTCATCGGCTATCTCGTCGCCTGTTACTGGC
TGAATATACAAACGCCGGACTTCTCCTTACCTTTTGGCCCCCGTCCGCTGGTGTTTATTAGTCTGATGGTTGCGATAGCCGTGCTTTGCTCATG
GGTTGGCGCACTCTGCTGGAACGTCGCCAGCCAGCGATTACCGACAGTGATTCTCGGGCCGCTGATTGTTTTCGAAACGCTGGCAGGTTGCTG
TACACCTTTTTACTCCGCCAGCAAATGCCGCCGCTAATGACGCTGAGCGGTATCGCGCTGTTAGTGATTGGCGTGGTCATTGCGGTCAGAGCAA
AACCAGAAAAACCTTTAACTGAATCTGTCTCAGAAAGTTGACACGCTGGCAGTGAGTTAAATAAGCCTCTGCTACGTAAGGGTTATAGCTTTTG
CCTTAAAGATGCATTTAAAATACATCTTATCTTATTAAGAATGAGGTATCAGCTATGGCTTATCGCGACCAACCTTTAGGTGAACTGGCGCTCT

CTATCCCTCGCGCTTCGGCTCTGTTTCGTAAATATGATATGGATTACTGCTGTGGCGGTAAGCAGACGCTGGCGCGCGCGGCGGCACGTAAAGA
ACTGGATGTTGAGGTCATTGAAGCTGAACTGGCAAAGCTTGCTGAACAACCGATTGAGAAAGACTGGCGTAGCGCCCCGCTGGCAGAAATCATT
GACCATATCATCGTGCGCTAC

PŘÍKLAD 1

TTAAGAAATTCCGGTTATCAACTTCGAGGCCCCTTTCAGGTCACCGTTGGCGTATTTCGGGGCGATAGATTTGTATTTGTAGCCGCGAATACTG
CGATCGCCCCCGGCGAAAAAACGCAGATCCGGCGGTACTTTGTCGAAATCACCGGTTTCAATCCAGCCCAGCGTGCCGCGTGTAACGAAACGAT
GGCGATCGTACAGTGTGCGGATCCAGACGTTCTGCGCCTGGAAAACGGAGAAATCGACATCTGAGCCCCAGGCCGTGTTGGAGTAGTCGATAGA
GTAGCGTTGCGAGTCGCCCCTGGTTGGCATCAGGCCACCACGAGAACGCGTGCGGCTAATCATCACCCCAGGATAAAACAGCATCGTGGTGTTG
GTAATTTCACCCTGAGTAAAGTGGTCGAGACTCCAGCGCAGGTTAATGGCACGCTGCCAGCCGCTGGAGAGATCCCAGTAGCGAGAAGCCACCA
GCGTAGTGGAGTCAGATTCGGTATCGTTCAGGTCAGTGCGCTTAAAACCGCCCTGCACCAAATAATATTGTTCCAGTGGATTCTTCAGCAGCGG
CATTTTATAGCTGAAGTCGAGGGTCTGTTCCGGCGCGGAAATACTGGTACTGGTGGTCAGACTGTGACCATAAGAGTTCATCCACGGCTTTTTC
CACGTAGCTTTCACGCGCGGTCCCACGTCCGTAGAGTAACCGACCCCGGTTTCGATGGTGTTTTCAGTTCGCGGCGAAACCACGCCCGTCAAGG
GTAATACTTTCGTTTCGCGCGCTTTATCAAATTGTGGAGCCACTACGACCGAGTTAAACCAGCCGGTAGCAGAAAGTCGACGGTTAAGTTCCGC
CAGATCTTTCGATTCGTACTCATCGCCCTCTTTAAACGGCACCAGATTTTGCAGGTATTCATCGCGGATTTGTGATCCTTCAAAGGTCACATGC
CCAAAGCGGTAACGTTCGCCACTGTTATAATCAATATCCCAGAAGGCTTTATGCAGGCCGAGCGCAATGCCCAGCTGCGCTTTGGTAAATTCGC
TATCGAAATAACCTTTACGCAACGCAATGCTGGTTAAGGACTTTTTGAAATTTTCATAATCGCCCTGGTTCAGTACCGTGCCAATAGCCGGGCG
AGTATCGAGCAATTTCAAATAGTCTTTATCGGTCCGCGCGCCGCCGCGCAATACCACATCGGTGCCGCCAATTAACACCGGCACGCCTGGCGTG
ACTTTGGCGATCAATACCTGCCGCCCTTTCTTTGGCGGTGGACGGAGATCAAATTCAATGGTCGGCTGGTAATAACCCAGCGCTTTCAGACCTT
CGCGGATGGCATCATCGACGCGTGCGCGAAAGCGACGGTCTGGCGTCACTTCATCACTTTCAATCGTAGAAAGCTGCGCACGAACGTTCTTTTC
CAGCTGTCCCGATAACCCCTCGACCTGTAGACGGACGTTCGCGGCGACGGCAGATCCGCTTAAGCAGAGTAAGCTTACACAGCATAACTGTCGG
ATATAGCGCATATTTTCTCCTGAATATCCTTTTTTCCTGCCCCTGGAACGCCGTTAAACGGCTTAACAAAAATCCAGTAATATGGATTAAAAAG
CAGACTAAACCCCAAATATTTCTTATGTTTTACTTTAGACCTATTCACGGTGGTTATTGTGTGCAAATACGCCTCTTGTTACAACCTTAACCCC
AATGACCGATTTTCGGGAGAGCGACACCATGAGTTTATTTGATAAAAAGCATCTGGTTTCCCCCGCCGATGCCCTGCCTGGACGTAACACCCCG
ATGCCCGTAGCCACGCTGCATGCGGTCAACGGTCACTCAATGACCAATGTACCTGACGGAATGGAGATTGCCATTTTTGCGATGGGTTGTTTCT
GGGGTGTGGAGCGTCTGTTCTGGCAGTTACCCGGCGTTTACAGCACCGCCGCAGGCTATACCGGCGGCTATACGCCAAATCCGACTTATCGGGA
AGTGTGCTCCGGTGATACGGGTCATGCCGAAGCGGTACGCATTGTTTACGATCCTTCTGTCATCAGCTATGAGCAGTTGCTACAGGTATTTTGG
GAGAATCACGATCCCGCCCAGGGAATGCGTCAGGGCAATGACCACGGCACGCAGTATCGTTCAGCGATTTATCCGCTGACCCCAGAACAGGATG
CCGCAGCTCGCGCCAGTCTGGAACGTTTTCAGGCGGCGATGCTTGCCGCCGATGACGATCGTCACATCACCACGGAAATCGCTAACGCCACACC
GTTTTATTATGCCGAAGATGACCACCAGCAATATCTGCATAAAAACCCGTATGGTTACTGTGGAATTGGCGGAATTGGCGTCTGTCTGCCACCG
GAAGCATACCGTTACGGGTACAAATGTAGATTGTTGATAAAGTGCGTTTTGTTTATGCCGAATGCAGCGTGAATGCCTTACCAGGCCTACAAAA
TCGTCCAAATTCAATATATTGCAGGGACTGCGTAGGCCTGATCGGCATAGCGCATCAGGCAGTATTGCATTTATCAGCAGTCTGAATTTTAACC
CTCTGGCGACTTTACAGTACCTTACGCTATACTAGCCACTGAAAATGCCGGATCACTTTCTTCGAATCGGCTTTCAATGTGTATTTCACACAAA
TTAATCAACTTCCCCTTCCGAGGATCTGGCCTGAAAGGTCGGATAAGATATGTTAAACAGTATTTTAGTCATACTCTGCTTGATCGCTGTAAGT
GCGTTCTTCTCGATGTCTGAGATCTCACTTGCCGCCTCACGCAAAATCAAACTTAAACTGCTGGCTGATGAAGGCAATATAAATGCCCAACGCG
TTCTGAATATGCAGGAAAATCCCGGCATGTTCTTTACGGTGGTCCAAATCGGTCTGAACGCAGTTGCGATTCTCGGCGGTATCGTCGGTGATGC
GGCATTTTCTCCAGCTTTTCACAGCCTGTTCTCCCGCTATATGTCGGCAGAACTCTCTGAGCAACTGAGCTTTATTCTCTCTTTCTCGTTAGTG
ACTGGCATGTTTATCCTGTTTGCGGATTTAACCCCGAAACGCATCGGTATGATTGCGCCAGAAGCGGTGGCTTTGCGTATCATCAACCCGATGC
GCTTCTGCCTGTACGTTTGCACCCCGCTGGTGTGGTTCTTCAACGGCCTGGCGAACATGATCTTCCGTATTTTCAAACTGCCAATGGTACGTAA
AGATGACATCACTTCTGATGACATCTACGCGGTAGTGGAAGCCGGTGCGCTGGCGGGCGTGTTACGTAAACAGGAACACGAGCTGATTGAAAAC
GTCTTTGAGCTGGAATCCCGTACCGTTCCGTCTTCAATGACACCGCGTGAAAACGTGATTTGGTTTGATCTCCACGAAGATGAGCAAAGCCTGA
AGAATAAGGTGGCGGAACATCCGCACTCTAAGTTCCTCGTCTGTAATGAAGATATTGACCACATCATCGGTTATGTCGATTCTAAAGACCTGCT
GAACCGCGTGCTGGCTAACCAAAGCCTGGCACTGAACAGCGGCGTACAAATTCGCAACACGCTGATTGTGCCGGATACGTTAACCCTTTCAGAG
GCGTTGGAAAGTTTTAAAACCGCAGGTGAAGACTTCGCGGTGATCATGAACGAGTACGCGCTGGTGGTGGGGATCATCACCCTCAACGACGTGA
TGACCACGCTGATGGGCGATCTGGTCGGTCAGGGGCTGGAAGAGCAGATTGTCGCCCGTGATGAGAACTCATGGCTGATTGACGGCGGCACCCC
AATTGACGACGTCATGCGCGTGCTGGATATTGACGAGTTCCCGCAGTCGGGCAACTACGAAACCATCGGAGGCTTTATGATGTTTATGCTGCGT
AAGATCCCGAAACGCACCGATTCGGTGAAATTCGCCGGCTACAAATTTGAAGTGGTGGATATCGATAACTACCGCATCGACCAGCTGCTGGTGA
CCCGGATCGACAGCAAGGCCACCGCCCTTTCGCCAAAACTGCCTGACGCTAAAGATAAAGAAGAAAGCGTCGCGTAACCCCCAGAAACATCAAC
GGCTCCTGAATCAGGAGCCGTCTTATTACTGCATAGCACTTTGGTTAAGCCATCTCTGTTTGCAGACGCATAACCTGACGGTTAACTTCGGACA
TCACTGACAAATGCAGCTTATCCTTCACTTTTGGGATAAGAATCTTACCCTTATCAAATTCAAAAGCGCCAACGTCCTTGATGTATAACCGTCC
ACGGAACAGGATCTTTACGTACTTCGCCACCTGAAGTGGGTTGTACCGTTGGAAAATTTTCATTGTTATCTCCTGCTGAGTATTACGCCTGTGC
GGTGCCACAATCGGCCCAACTATTATGAGGCGCAAATTTTAATGCCTAGTGACTATAGACTATCCGGGCAATGTTTCCACCGCGTATAACTTTT
TTTACCTAAAGGTTACAATTATTCAGAATTATCTTTTTACCGAAGCGCGGTCTTCAGTATAAGCATTCATTTTTCATATGATTTGTGCGCTTGA
CCGCAAACTGGCATCACACTTGCGGGAAATTCGATAAATAGCACATATGATTAAAACTCAGACCCAAGTGGTCGGATCACCTGCATATCATAAG
AAGGAAACACCATGACCCTACGCAAGATTCTGGCACTCACCTGCCTGCTGTTGCCGATGATGGCTTCCGCACATCAGTTCGAAACCGGTCAGCG
AGTGCCGCCGATTGGCATCACCGATCGTGGCGAGTTGGTGCTTGATAAAGATCAGTTTAGCTACAAAACCTGGAACAGCGCGCAGTTAGTGGGA
AAAGTGCGAGTACTGCAACATATTGCTGGTCGCACCTCTGCAAAAGAGAAAAACGCGACGCTGATTGAAGCGATTAAATCAGCGAAGTTACCGC
ACGATCGTTACCAGACCACCACCATTGTTAACACCGACGACGCAATTCCGGGTTCCGGCATGTTTGTGCGCAGCAGTCTGGAGAGTAATAAAAA
GCTTTATCCCTGGTCGCAGTTTATTGTTGATAGCAATGGCGTCGCACGCGGTGCCTGGCAGCTGGATGAAAAAAGTTCCGCTGTAGTGGTGCTG
GATAAAGACGGTCGCGTGCAATGGGCCAAAGACAGGGCGCTTACTCAGGAAGAGGTGCAGCAAGTGATGGACCTGCTGCATAAATTAATTAATA
AATAGACACTCTGAACCCAGGATTCAGGAACGACTCGCGCGGAGTGTAATCCAGCGGTTTACCCTGCCAGTCGTGAACGTGCGCTCCGGCAGCT
GCGGCCACAGCATGTCCAGCGGCGGTGTCCCAAATATTCGTTGGTCCGAAGCGCGGGTACAGCTGCGCCTGTCCTTCCGCCACCAGGCAGAATT
TCAGCGAAGAGCCGATGGACGTGGTCTGATGTTCGCCAAGCTGTTGCAGATACTCTTTCAGCTCCGCATCTGCATGGGAACGGCTGATAACCAC
CAGCGGCGGGCGCGCATCGCGGACCTGAATCTGCTTGCGCACACCGCACTCTTCTTTCCAGGCCTTGCCTTCTGCCGCGCTGTACATTACGTTC
ATTACCGGCGCATACACCACACCTAAAATCGGTTTGCCATGGTCAATGAGCGCAATGTTAACGGTGAATTCGCCATTGCGTTTAATAAACTCTT
TAGTACCATCCAGCGGGTCTACCAGCCAGTAACGCTGCCAGTGCTGACGGACTTCCCAACCGGGAGGATCTTCTTCAGAAAGGACCGGAATATC
CGGTGTCAGCGTACGTAAACCGTCCATGATAACGGTGTGAGCGGCAATATCCGCTGCCGTTACCGGAGAATTGTCCGCTTTGCTGACGACGTCC
ATCGGTTTCGTCCCGTCGTAGACCTGCATAATGGCATCGCCTGCATTCCGTGCAAGCTGGCATACTTGATCTAACATTTCTCCACCTCGTCTCT
GTGAGCGGTGTTAACTTATTGTTTTACTTATACCCTATCGTTAATGAATGCGCCAACTGTGATAGTGTCATCATTTTCAAAGCGTAAAATTGTG
GCATTCTTCACTGTTCTATAAGTAAGACGTTTATTCTTCCTTTTCTTTCGTATTCCCGATGATAAAAGGATGTCCCTGATGATTAAGTTTAGCG
CAACGCTCCTGGCCACGCTGATTGCCGCCAGTGTGAATGCAGCGACGGTCGATCTGCGTATCATGGAAACCACTGATCTGCATAGCAACATGAT
GGATTTCGATTATTACAAAGACACCGCCACGGAAAAATTCGGACTGGTACGTACGGCAAGCCTGATTAACGATGCCCGCAATGAAGTGAAAAAC
AGCGTACTGGTCGATAACGGCGATTTGATTCAGGGGAGTCCGCTGGCCGATTACATGTCGGCGAAAGGATTAAAAGCAGGTGATGTTCATCGG
TTTATAAGGCGCTGAATACGCTGGATTATACGGTCGGTACACTCGGCAATCATGAATTTAACTACGGTCTGGATTACCTGAAAAATGCGCTGGC
GGGAGCGAAATTCCCTTATGTAAATGCCAACGTCATTGACGCCAGAACCAAACAGCCAATGTTTACACCGTATTTAATTAAAGATACCGAAGTG
GTCGATAAAGACGGAAAAAAACAGACGCTGAAGATTGGCTATATTGGCGTCGTGCCGCCGCAAATCATGGGCTGGGATAAAGCTAATTTATCCG

GAAAAGTGACGGTGAATGATATTACCGAAACCGTGCGCAAATACGTGCCTGAAATGCGCGAGAAAGGTGCCGATGTCGTTGTCGTTCTGGCGCA
TTCCGGGCTGTCTGCCGATCCGTATAAAGTAATGGCGGAAAACTCAGTTTATTACCTCAGTGAAATTCCGGGCGTTAACGCCATTATGTTTGGT
CATGCTCACGCCGTTTTCCCGGGTAAAGATTTTGCTGATATCGAAGGGGCTGATATCACCAAAGGCACGCTGAATGGTGTTCCGGCGGTAATGC
CGGGCATGTGGGGCGATCATCTTGGGGTGGTCGACTTACAACTCAGTAATGACAGCGGTAAATGGCAGGTGACGCAGGCGAAAGCGGAAGCTCG
GCCGATTTACGACATCGCCAATAAAAAATCCCTCGCGGCGGAAGACAGCAAGCTGGTAGAAACACTCAAAGCCGATCACGATGCCACACGCCAG
TTCGTCAGCAAGCCAATCGGTAAATCTGCCGACAATATGTATAGCTATCTGGCGCTGGTGCAGGACGATCCGACCGTGCAAGTAGTGAACAACG
CGCAAAAAGCGTATGTCGAACATTACATTCAGGGCGATCCGGATCTGGCAAAACTGCCGGTGCTTTCAGCTGCCGCACCGTTTAAAGTTGGTGG
TCGCAAAAATGACCCGGCAAGCTATGTGGAGGTGGAAAAAGGCCAGTTGACCTTCCGTAATGCCGCCGATCTTTATCTCTACCCCAATACGCTG
ATTGTGGTGAAAGCCAGCGGTAAAGAGGTGAAAGAGTGGCTGGAGTGCTCCGCCGGACAGTTTAACCAGATTGATCCTAACAGCACGAAACCGC
AATCACTCATCAACTGGGATGGTTTCCGCACCTATAACTTTGATGTGATTGATGGTGTGAATTATCAGATTGATGTTACCCAACCTGCCCGTTA
TGACGGCGAGTGCCAGATGATTAATGCCAATGCGGAAAGGATTAAGAACCTGACCTTTAATGGCAAGCCGATTGATCCGAACGCCATGTTCCTC
GTTGCCACCAATAACTATCGCGCTTACGGCGGCAAATTTGCCGGTACGGGCGACAGCCATATCGCTTTTGCTTCACCGGATGAGAACCGCTCGG
TGCTGGCAGCGTGGATTGCTGATGAGTCGAAACGTGCGGGGGAAATTCACCCGGCGGCAGATAACAACTGGCGTTTAGCACCGATAGCTGGCGA
TAAGAAACTGGATATCCGTTTCGAAACTTCTCCGTCAGATAAAGCCGCAGCGTTTATTAAAGAGAAAGGGCAATATCCGATGAATAAAGTCGCG
ACCGATGATATCGGGTTTGCGATTTATCAGGTGGATTTGAGTAAGTAAAACACTTCTTTTTCGCGCCTATAAATCATCAACCGCATCCGGCATTT
ATTGGCGGATGCGATGCTGGCGCATCTTATCCGCCCTACAAGCCATGCACCGTAGACCAGATAAGCTCAGCGCATCCGGCAGTTATGCCGCACG
TTCATCCCGCACCGCCAACACCTCGGGCAAATTCAACTCAATCCAGTCCGCCAGTGCAGCAACCTTTTCGCTCACCTGCTCGCCCAGCGGCGTG
AGGCTATATTCCACATGCGGCGGCACCACCGGATACGCGATACGGTTAAGAAAACCATCCTGTTCTAACGCCTGTAACGACTGCGCAAGCATCT
TTTCGCTCACCCCGCCGATTTTGCGCCGCAGGTCGCTAAAGCGATGAGTACCTTCGCGTAGCGCCACCAGAATCAACACCCCCCAACGGCTGGT
GACGTGTTTCAACACCTCGCGCGACGGGCACTGTTCCGCAAAGAGATTACCCTCTTTCAGTTGTTGCGACAGGCTAACCTGACTCATTTCATAC
TTACCTTTTTGTGCGTACTTACTAAAAGTAAGTTTAGGTGTTAGCGTATTTAAACACAAGACAAAACGATGGAGACTTCCCATGATCGCTATTA
CTGGTGCCACTGGCCAACTTGGTCACTATGTTATTGAATCCTTGAAGAAAACGGTTCCTGTCAGCCAAATAGTGGCTATCGTTCGTAATCCGGC
AAAAGCCCAGGCTCTGGCAGCACAAGGCATTACCGTGCGTCAGGCTGACTACGGCGATGAAGCCGCACTGACATCTGCACTTCAGGGAGTGGAA
AAACTACTGCTGATCTCTTCCAGCGAAGTGGGTCAACGTGCCCCGCAGCATCGTAATGTTATTAATGCCGCAAAGGCGGCTGGTGTGAAATTTA
TCGCTTATACCAGCCTGCTACATGCAGATACCTCCCCGCTCGGCCTCGCCAATGAGCACATCGAGACGGAGAAAATGTTGGCTGATTCTGGCAT
CGTTTACACCCTGCTGCGCAACGGCTGGTACAGCGAAAACTACCTCGCCAGCGCCCCGGCAGCACTGGAACACGGCGTATTTATCGGTGCGGCG
GGCGATGGCAAAATCGCCTCAGCAACGCGGGCAGATTATGCGGCAGCTGCGGCACGCGTGATTAGCGAAGCCGGTCACGAAGGCAAGGTTTACG
AACTGGCGGGCGATAGTGCCTGGACATTGACACAGTTAGCGGCAGAGCTGACCAAACAGAGCGGCAAACAGGTTACCTATCAAAATCTGAGCGA
AGCCGATTTCGCCGCGGCACTGAAAAGCGTCGGACTGCCCGACGGACTGGCGGATATGCTGGCGGATTCTGACGTTGGCGCATCGAAAGGTGGT
CTGTTTGATAACAGCAAAACGCTTAGCAAATTGATTGGCCGCCCAACGACAACGTTAGCCGAAAGCGTAAGCCATCTTTTTAATGTTAATAACT
AGTTAATTAAAGTGGCATCCTCCCGCATCCTCTCTAATAATGACGGGATGCCGGGAGCAATCATGTCTGCTTCCTGAACTTTCTTCTGACAGAC
CAATGGATGCCAGTAATGATTAGCGGCGTGCTGTACGCCCTGTTAGCAGGGTTGATGTGGGGGCTTATTTTTGTCGGGCCGTTGATCGTACCGG
AATACCCGGCGATGTTGCAGTCGATGGGGCGTTATCTGGCGTTAGGGTTAATTGCGCTGCCCATTGCCTGGCTGGGACGCGTGCGTCTGCGTCA
GTTGGCGCGTCGGGACTGGCTTACCGCCTTGATGCTCACTATGATGGGCAACCTCATTTATTACTTCTGCCTTGCCAGTGCCATTCAACGTACT
GGCGCGCCTGTTTCCACGATGATTATCGGCACCCTGCCGGTGGTCATTCCTGTCTTTGCCAATCTGCTTTATAGCCAGCGCGACGGCAAACTCG
CGTGGGGAAAACTCGCCCCGGCACTGATTTGTATTGGCATCGGCCTGGCGAGTGTGAATATTGCTGAGTTAAACCACGGACTCCCCGATTTTGA
CTGGGCACGTTATACCTCTGGCATCGTGCTAGCGTTAGTTTCCGTGGTCTGCTGGGCATGGTATGCCCTGCGCAACGCCCGCTGGCTGCGGGAA
AATCCCGACAAACATCCGATGATGTGGGCGACGGCGCAGGCGCTGGTCACGCTGCCGGTTTCTCTCATCGGCTATCTCGTCGCCTGTTACTGGC
TGAATATACAAACGCCGGACTTCTCCTTACCTTTTGGCCCCCGTCCGCTGGTGTTTATTAGTCTGATGGTTGCGATAGCCGTGCTTTGCTCATG
GGTTGGCGCACTCTGCTGGAACGTCGCCAGCCAGCGATTACCGACAGTGATTCTCGGGCCGCTGATTGTTTTCGAAACGCTGGCAGGTTTGCTG
TACACCTTTTTACTCCGCCAGCAAATGCCGCCGCTAATGACGCTGAGCGGTATCGCGCTGTTAGTGATTGGCGTGGTCATTGCGGTCAGAGCAA
AACCAGAAAAACCTTTAACTGAATCTGTCTCAGAAAGTTGACACGCTGGCAGTGAGTTAAATAAGCCTCTGCTACGTAAGGGTTATAGCTTTTG
CCTTAAAGATGCATTTAAAATACATCTTATCTTATTAAGAATGAGGTATCAGCTATGGCTTATCGCGACCAACCTTTAGGTGAACTGGCGCTCT
CTATCCCTCGCGCTTCGGCTCTGTTTCGTAAATATGATATGGATTACTGCTGTGGCGGTAAGCAGACGCTGGCGCGCGCGGCGGCACGTAAAGA
ACTGGATGTTGAGGTCATTGAAGCTGAACTGGCAAAGCTTGCTGAACAACCGATTGAGAAAGACTGGCGTAGCGCCCCGCTGGCAGAAATCATT
GACCATATCATCGTGCGCTAC

PŘÍKLAD 2

TTAAGAAATTCCGGTTATCAACTTCGAGGCCCCTTTCAGGTCACCGTTGGCGTATTTCGGGGCGATAGATTTGTATTTGTAGCCGCGAATACTG
CGATCGCCCCCGGCGAAAAAACGCAGATCCGGCGGTACTTTGTCGAAATCACCGGTTTCAATCCAGCCCAGCGTGCCGCGTGTAACGAAACGAT
GGCGATCGTACAGTGTGCGGATCCAGACGTTCTGCGCCTGGAAAACGGAGAAATCGACATCTGAGCCCCAGGCCGTGTTGGAGTAGTCGATAGA
GTAGCGTTGCGAGTCGCCCCTGGTTGGCATCAGGCCACCACGAGAACGCGTGCGGCTAATCATCACCCCAGGATAAAACAGCATCGTGGTGTTG
GTAATTTCACCCTGAGTAAAGTGGTCGAGACTCCAGCGCAGGTTAATGGCACGCTGCCAGCCGCTGGAGAGATCCCAGTAGCGAGAAGCCACCA
GCGTAGTGGAGTCAGATTCGGTATCGTTCAGGTCAGTGCGCTTAAAACCGCCCTGCACCAAATAATATTGTTCCAGTGGATTCTTCAGCAGCGG
CATTTTATAGCTGAAGTCGAGGGTCTGTTCCGGCGCGGAAATACTGGTACTGGTGGTCAGACTGTGACCATAAGAGTTCATCCACGGCTTTTTC
CACGTAGCTTTCACGCGCGGTCCCACGTCCGTAGAGTAACCGACCCCGGTTTCGATGGTGTTTTCAGTTCGCGGCGAAACCACGCCCGTCAAGG
GTAATACTTTCGTTTCGCGCGCTTTATCAAATTGTGGAGCCACTACGACCGAGTTAAACCAGCCGGTAGCAGAAAGTCGACGGTTAAGTTCCGC
CAGATCTTTCGATTCGTACTCATCGCCCTCTTTAAACGGCACCAGATTTTGCAGGTATTCATCGCGGATTTGTGATCCTTCAAAGGTCACATGC
CCAAAGCGGTAACGTTCGCCACTGTTATAATCAATATCCCAGAAGGCTTTATGCAGGCCGAGCGCAATGCCCAGCTGCGCTTTGGTAAATTCGC
TATCGAAATAACCTTTACGCAACGCAATGCTGGTTAAGGACTTTTTGAAATTTTCATAATCGCCCTGGTTCAGTACCGTGCCAATAGCCGGGCG
AGTATCGAGCAATTTCAAATAGTCTTTATCGGTCCGCGCGCCGCCGCGCAATACCACATCGGTGCCGCCAATTAACACCGGCACGCCTGGCGTG
ACTTTGGCGATCAATACCTGCCGCCCTTTCTTTGGCGGTGGACGGAGATCAAATTCAATGGTCGGCTGGTAATAACCCAGCGCTTTCAGACCTT
CGCGGATGGCATCATCGACGCGTGCGCGAAAGCGACGGTCTGGCGTCACTTCATCACTTTCAATCGTAGAAAGCTGCGCACGAACGTTCTTTTC
CAGCTGTCCCGATAACCCCTCGACCTGTAGACGGACGTTCGCGGCGACGGCAGATCCGCTTAAGCAGAGTAAGCTTACACAGCATAACTGTCGG
ATATAGCGCATATTTTCTCCTGAATATCCTTTTTTCCTGCCCCTGGAACGCCGTTAAACGGCTTAACAAAAATCCAGTAATATGGATTAAAAAG
CAGACTAAACCCCAAATATTTCTTATGTTTTACTTTAGACCTATTCACGGTGGTTATTGTGTGCAAATACGCCTCTTGTTACAACCTTAACCCC
AATGACCGATTTTCGGGAGAGCGACACCATGAGTTTATTTGATAAAAAGCATCTGGTTTCCCCCGCCGATGCCCTGCCTGGACGTAACACCCCG
ATGCCCGTAGCCACGCTGCATGCGGTCAACGGTCACTCAATGACCAATGTACCTGACGGAATGGAGATTGCCATTTTTGCGATGGGTTGTTTCT
GGGGTGTGGAGCGTCTGTTCTGGCAGTTACCCGGCGTTTACAGCACCGCCGCAGGCTATACCGGCGGCTATACGCCAAATCCGACTTATCGGGA
AGTGTGCTCCGGTGATACGGGTCATGCCGAAGCGGTACGCATTGTTTACGATCCTTCTGTCATCAGCTATGAGCAGTTGCTACAGGTATTTTGG
GAGAATCACGATCCCGCCCAGGGAATGCGTCAGGGCAATGACCACGGCACGCAGTATCGTTCAGCGATTTATCCGCTGACCCCAGAACAGGATG
CCGCAGCTCGCGCCAGTCTGGAACGTTTTCAGGCGGCGATGCTTGCCGCCGATGACGATCGTCACATCACCACGGAAATCGCTAACGCCACACC
GTTTTATTATGCCGAAGATGACCACCAGCAATATCTGCATAAAAACCCGTATGGTTACTGTGGAATTGGCGGAATTGGCGTCTGTCTGCCACCG
GAAGCATAGCGTTACGGGTACAAATGTAGATTGTTGATAAAGTGCGTTTTGTTTATGCCGAATGCAGCGTGAATGCCTTACCAGGCCTACAAAA
TCGTCCAAATTCAATATATTGCAGGGACTGCGTAGGCCTGATCGGCATAGCGCATCAGGCAGTATTGCATTTATCAGCAGTCTGAATTTTAACC
CTCTGGCGACTTTACAGTACCTTACGCTATACTAGCCACTGAAAATGCCGGATCACTTTCTTCGAATCGGCTTTCAATGTGTATTTCACACAAA

TTAATCAACTTCCCCTTCCGAGGATCTGGCCTGAAAGGTCGGATAAGATATGTTAAACAGTATTTTAGTCATACTCTGCTTGATCGCTGTAAGT
GCGTTCTTCTCGATGTCTGAGATCTCACTTGCCGCCTCACGCAAAATCAAACTTAAACTGCTGGCTGATGAAGGCAATATAAATGCCCAACGCG
TTCTGAATATGCAGGAAAATCCCGGCATGTTCTTTACGGTGGTCCAAATCGGTCTGAACGCAGTTGCGATTCTCGGCGGTATCGTCGGTGATGC
GGCATTTTCTCCAGCTTTTCACAGCCTGTTCTCCCGCTATATGTCGGCAGAACTCTCTGAGCAACTGAGCTTTATTCTCTCTTTCTCGTTAGTG
ACTGGCATGTTTATCCTGTTTGCGGATTTAACCCCGAAACGCATCGGTATGATTGCGCCAGAAGCGGTGGCTTTGCGTATCATCAACCCGATGC
GCTTCTGCCTGTACGTTTGCACCCCGCTGGTGTGGTTCTTCAACGGCCTGGCGAACATGATCTTCCGTATTTTCAAACTGCCAATGGTACGTAA
AGATGACATCACTTCTGATGACATCTACGCGGTAGTGGAAGCCGGTGCGCTGGCGGGCGTGTTACGTAAACAGGAACACGAGCTGATTGAAAAC
GTCTTTGAGCTGGAATCCCGTACCGTTCCGTCTTCAATGACACCGCGTGAAAACGTGATTTGGTTTGATCTCCACGAAGATGAGCAAAGCCTGA
AGAATAAGGTGGCGGAACATCCGCACTCTAAGTTCCTCGTCTGTAATGAAGATATTGACCACATCATCGGTTATGTCGATTCTAAAGACCTGCT
GAACCGCGTGCTGGCTAACCAAAGCCTGGCACTGAACAGCGGCGTACAAATTCGCAACACGCTGATTGTGCCGGATACGTTAACCCTTTCAGAG
GCGTTGGAAAGTTTTAAAACCGCAGGTGAAGACTTCGCGGTGATCATGAACGAGTACGCGCTGGTGGTGGGGATCATCACCCTCAACGACGTGA
TGACCACGCTGATGGGCGATCTGGTCGGTCAGGGGCTGGAAGAGCAGATTGTCGCCCGTGATGAGAACTCATGGCTGATTGACGGCGGCACCCC
AATTGACGACGTCATGCGCGTGCTGGATATTGACGAGTTCCCGCAGTCGGGCAACTACGAAACCATCGGAGGCTTTATGATGTTTATGCTGCGT
AAGATCCCGAAACGCACCGATTCGGTGAAATTCGCCGGCTACAAATTTGAAGTGGTGGATATCGATAACTACCGCATCGACCAGCTGCTGGTGA
CCCGGATCGACAGCAAGGCCACCGCCCTTTCGCCAAAACTGCCTGACGCTAAAGATAAAGAAGAAAGCGTCGCGTAACCCCCAGAAACATCAAC
GGCTCCTGAATCAGGAGCCGTCTTATTACTGCATAGCACTTTGGTTAAGCCATCTCTGTTTGCAGACGCATAACCTGACGGTTAACTTCGGACA
TCACTGACAAATGCAGCTTATCCTTCACTTTTGGGATAAGAATCTTACCCTTATCAAATTCAAAAGCGCCAACGTCCTTGATGTATAACCGTCC
ACGGAACAGGATCTTTACGTACTTCGCCACCTGAAGTGGGTTGTACCGTTGGAAAATTTTCATTGTTATCTCCTGCTGAGTATTACGCCTGTGC
GGTGCCACAATCGGCCCAACTATTATGAGGCGCAAATTTTAATGCCTAGTGACTATAGACTATCCGGGCAATGTTTCCACCGCGTATAACTTTT
TTTACCTAAAGGTTACAATTATTCAGAATTATCTTTTTACCGAAGCGCGGTCTTCAGTATAAGCATTCATTTTTCATATGATTTGTGCGCTTGA
CCGCAAACTGGCATCACACTTGCGGGAAATTCGATAAATAGCACATATGATTAAAACTCAGACCCAAGTGGTCGGATCACCTGCATATCATAAG
AAGGAAACACCATGACCCTACGCAAGATTCTGGCACTCACCTGCCTGCTGTTGCCGATGATGGCTTCCGCACATCAGTTCGAAACCGGTCAGCG
AGTGCCGCCGATTGGCATCACCGATCGTGGCGAGTTGGTGCTTGATAAAGATCAGTTTAGCTACAAAACCTGGAACAGCGCGCAGTTAGTGGGA
AAAGTGCGAGTACTGCAACATATTGCTGGTCGCACCTCTGCAAAAGAGAAAAACGCGACGCTGATTGAAGCGATTAAATCAGCGAAGTTACCGC
ACGATCGTTACCAGACCACCACCATTGTTAACACCGACGACGCAATTCCGGGTTCCGGCATGTTTGTGCGCAGCAGTCTGGAGAGTAATAAAAA
GCTTTATCCCTGGTCGCAGTTTATTGTTGATAGCAATGGCGTCGCACGCGGTGCCTGGCAGCTGGATGAAAAAAGTTCCGCTGTAGTGGTGCTG
GATAAAGACGGTCGCGTGCAATGGGCCAAAGACAGGGCGCTTACTCAGGAAGAGGTGCAGCAAGTGATGGACCTGCTGCATAAATTAATTAATA
AATAGACACTCTGAACCCAGGATTCAGGAACGACTCGCGCGGAGTGTAATCCAGCGGTTTACCCTGCCAGTCGTGAACGTGCGCTCCGGCAGCT
GCGGCCACAGCATGTCCAGCGGCGGTGTCCCAAATATTCGTTGGTCCGAAGCGCGGGTACAGCTGCGCCTGTCCTTCCGCCACCAGGCAGAATT
TCAGCGAAGAGCCGATGGACGTGGTCTGATGTTCGCCAAGCTGTTGCAGATACTCTTTCAGCTCCGCATCTGCATGGGAACGGCTGATAACCAC
CAGCGGCGGGCGCGCATCGCGGACCTGAATCTGCTTGCGCACACCGCACTCTTCTTTCCAGGCCTTGCCTTCTGCCGCGCTGTACATTACGTTC
ATTACCGGCGCATACACCACACCTAAAATCGGTTTGCCATGGTCAATGAGCGCAATGTTAACGGTGAATTCGCCATTGCGTTTAATAAACTCTT
TAGTACCATCCAGCGGGTCTACCAGCCAGTAACGCTGCCAGTGCTGACGGACTTCCCAACCGGGAGGATCTTCTTCAGAAAGGACCGGAATATC
CGGTGTCAGCGTACGTAAACCGTCCATGATAACGGTGTGAGCGGCAATATCCGCTGCCGTTACCGGAGAATTGTCCGCTTTGCTGACGACGTCC
ATCGGTTTCGTCCCGTCGTAGACCTGCATAATGGCATCGCCTGCATTCCGTGCAAGCTGGCATACTTGATCTAACATTTCTCCACCTCGTCTCT
GTGAGCGGTGTTAACTTATTGTTTTACTTATACCCTATCGTTAATGAATGCGCCAACTGTGATAGTGTCATCATTTTCAAAGCGTAAAATTGTG
GCATTCTTCACTGTTCTATAAGTAAGACGTTTATTCTTCCTTTTCTTTCGTATTCCCGATGATAAAAGGATGTCCCTGATGATTAAGTTTAGCG
CAACGCTCCTGGCCACGCTGATTGCCGCCAGTGTGAATGCAGCGACGGTCGATCTGCGTATCATGGAAACCACTGATCTGCATAGCAACATGAT
GGATTTCGATTATTACAAAGACACCGCCACGGAAAAATTCGGACTGGTACGTACGGCAAGCCTGATTAACGATGCCCGCAATGAAGTGAAAAAC
AGCGTACTGGTCGATAACGGCGATTTGATTCAGGGGAGTCCGCTGGCCGATTACATGTCGGCGAAAGGATTAAAAGCAGGTGATGTTCATCCGG
TTTATAAGGCGCTGAATACGCTGGATTATACGGTCGGTACACTCGGCAATCATGAATTTAACTACGGTCTGGATTACCTGAAAAATGCGCTGGC
GGGAGCGAAATTCCCTTATGTAAATGCCAACGTCATTGACGCCAGAACCAAACAGCCAATGTTTACACCGTATTTAATTAAAGATACCGAAGTG
GTCGATAAAGACGGAAAAAAACAGACGCTGAAGATTGGCTATATTGGCGTCGTGCCGCCGCAAATCATGGGCTGGGATAAAGCTAATTTATCCG
GAAAAGTGACGGTGAATGATATTACCGAAACCGTGCGCAAATACGTGCCTGAAATGCGCGAGAAAGGTGCCGATGTCGTTGTCGTTCTGGCGCA
TTCCGGGCTGTCTGCCGATCCGTATAAAGTAATGGCGGAAAACTCAGTTTATTACCTCAGTGAAATTCCGGGCGTTAACGCCATTATGTTTGGT
CATGCTCACGCCGTTTTCCCGGGTAAAGATTTTGCTGATATCGAAGGGGCTGATATCACCAAAGGCACGCTGAATGGTGTTCCGGCGGTAATGC
CGGGCATGTGGGGCGATCATCTTGGGGTGGTCGACTTACAACTCAGTAATGACAGCGGTAAATGGCAGGTGACGCAGGCGAAAGCGGAAGCTCG
GCCGATTTACGACATCGCCAATAAAAAATCCCTCGCGGCGGAAGACAGCAAGCTGGTAGAAACACTCAAAGCCGATCACGATGCCACACGCCAG
TTCGTCAGCAAGCCAATCGGTAAATCTGCCGACAATATGTATAGCTATCTGGCGCTGGTGCAGGACGATCCGACCGTGCAAGTAGTGAACAACG
CGCAAAAAGCGTATGTCGAACATTACATTCAGGGCGATCCGGATCTGGCAAAACTGCCGGTGCTTTCAGCTGCCGCACCGTTTAAAGTTGGTGG
TCGCAAAAATGACCCGGCAAGCTATGTGGAGGTGGAAAAAGGCCAGTTGACCTTCCGTAATGCCGCCGATCTTTATCTCTACCCCAATACGCTG
ATTGTGGTGAAAGCCAGCGGTAAAGAGGTGAAAGAGTGGCTGGAGTGCTCCGCCGGACAGTTTAACCAGATTGATCCTAACAGCACGAAACCGC
AATCACTCATCAACTGGGATGGTTTCCGCACCTATAACTTTGATGTGATTGATGGTGTGAATTATCAGATTGATGTTACCCAACCTGCCCGTTA
TGACGGCGAGTGCCAGATGATTAATGCCAATGCGGAAAGGATTAAGAACCTGACCTTTAATGGCAAGCCGATTGATCCGAACGCCATGTTCCTC
GTTGCCACCAATAACTATCGCGCTTACGGCGGCAAATTTGCCGGTACGGGCGACAGCCATATCGCTTTTGCTTCACCGGATGAGAACCGCTCGG
TGCTGGCAGCGTGGATTGCTGATGAGTCGAAACGTGCGGGGGAAATTCACCCGGCGGCAGATAACAACTGGCGTTTAGCACCGATAGCTGGCGA
TAAGAAACTGGATATCCGTTTCGAAACTTCTCCGTCAGATAAAGCCGCAGCGTTTATTAAAGAGAAAGGGCAATATCCGATGAATAAAGTCGCG
ACCGATGATATCGGGTTTGCGATTTATCAGGTGGATTTGAGTAAGTAAAACACTTCTTTTTCGGCCTATAAATCATCAACCGCATCCGGCATTT
ATTGGCGGATGCGATGCTGGCGCATCTTATCCGCCCTACAAGCCATGCACCGTAGACCAGATAAGCTCAGCGCATCCGGCAGTTATGCCGCACG
TTCATCCCGCACCGCCAACACCTCGGGCAAATTCAACTCAATCCAGTCCGCCAGTGCAGCAACCTTTTCGCTCACCTGCTCGCCCAGCGGCGTG
AGGCTATATTCCACATGCGGCGGCACCACCGGATACGCGATACGGTTAAGAAAACCATCCTGTTCTAACGCCTGTAACGACTGCGCAAGCATCT
TTTCGCTCACCCCGCCGATTTTGCGCCGCAGGTCGCTAAAGCGATGAGTACCTTCGCGTAGCGCCACCAGAATCAACACCCCCCAACGGCTGGT
GACGTGTTTCAACACCTCGCGCGACGGGCACTGTTCCGCAAAGAGATTACCCTCTTTCAGTTGTTGCGACAGGCTAACCTGACTCATTTCATAC
TTACCTTTTTGTGCGTACTTACTAAAAGTAAGTTTAGGTGTTAGCGTATTTAAACACAAGACAAAACGATGGAGACTTCCCATGATCGCTATTA
CTGGTGCCACTGGCCAACTTGGTCACTATGTTATTGAATCCTTGAAGAAAACGGTTCCTGTCAGCCAAATAGTGGCTATCGTTCGTAATCCGGC
AAAAGCCCAGGCTCTGGCAGCACAAGGCATTACCGTGCGTCAGGCTGACTACGGCGATGAAGCCGCACTGACATCTGCACTTCAGGGAGTGGAA
AAACTACTGCTGATCTCTTCCAGCGAAGTGGGTCAACGTGCCCCGCAGCATCGTAATGTTATTAATGCCGCAAAGGCGGCTGGTGTGAAATTTA
TCGCTTATACCAGCCTGCTACATGCAGATACCTCCCCGCTCGGCCTCGCCAATGAGCACATCGAGACGGAGAAAATGTTGGCTGATTCTGGCAT
CGTTTACACCCTGCTGCGCAACGGCTGGTACAGCGAAAACTACCTCGCCAGCGCCCCGGCAGCACTGGAACACGGCGTATTTATCGGTGCGGCG
GGCGATGGCAAAATCGCCTCAGCAACGCGGGCAGATTATGCGGCAGCTGCGGCACGCGTGATTAGCGAAGCCGGTCACGAAGGCAAGGTTTACG
AACTGGCGGGCGATAGTGCCTGGACATTGACACAGTTAGCGGCAGAGCTGACCAAACAGAGCGGCAAACAGGTTACCTATCAAAATCTGAGCGA
AGCCGATTTCGCCGCGGCACTGAAAAGCGTCGGACTGCCCGACGGACTGGCGGATATGCTGGCGGATTCTGACGTTGGCGCATCGAAAGGTGGT
CTGTTTGATAACAGCAAAACGCTTAGCAAATTGATTGGCCGCCCAACGACAACGTTAGCCGAAAGCGTAAGCCATCTTTTTAATGTTAATAACT
AGTTAATTAAAGTGGCATCCTCCCGCATCCTCTCTAATAATGACGGGATGCCGGGAGCAATCATGTCTGCTTCCTGGACTTCTTCTGACAGAC
CAATGGATGCCAGTAATAATTAGCGGCGTGCTGTACGCCCTGTTAGCAGGGTTGATGTGGGGGCTTATTTTTGTCGGGCCGTTGATCGTACCGG
AATACCCGGCGATGTTGCAGTCGATGGGGCGTTATCTGGCGTTAGGGTTAATTGCGCTGCCCATTGCCTGGCTGGGACGCGTGCGTCTGCGTCA
GTTGGCGCGTCGGGACTGGCTTACCGCCTTGATGCTCACTATGATGGGCAACCTCATTTATTACTTCTGCCTTGCCAGTGCCATTCAACGTACT

GGCGCGCCTGTTTCCACGATGATTATCGGCACCCTGCCGGTGGTCATTCCTGTCTTTGCCAATCTGCTTTATAGCCAGCGCGACGGCAAACTCG
CGTGGGGAAAACTCGCCCCGGCACTGATTTGTATTGGCATCGGCCTGGCGAGTGTGAATATTGCTGAGTTAAACCACGGACTCCCCGATTTTGA
CTGGGCACGTTATACCTCTGGCATCGTGCTAGCGTTAGTTTCCGTGGTCTGCTGGGCATGGTATGCCCTGCGCAACGCCCGCTGGCTGCGGGAA
AATCCCGACAAACATCCGATGATGTGGGCGACGGCGCAGGCGCTGGTCACGCTGCCGGTTTCTCTCATCGGCTATCTCGTCGCCTGTTACTGGC
TGAATATACAAACGCCGGACTTCTCCTTACCTTTTGGCCCCCGTCCGCTGGTGTTTATTAGTCTGATGGTTGCGATAGCCGTGCTTTGCTCATG
GGTTGGCGCACTCTGCTGGAACGTCGCCAGCCAGCGATTACCGACAGTGATTCTCGGGCCGCTGATTGTTTTCGAAACGCTGGCAGGTTTGCTG
TACACCTTTTTACTCCGCCAGCAAATGCCGCCGCTAATGACGCTGAGCGGTATCGCGCTGTTAGTGATTGGCGTGGTCATTGCGGTCAGAGCAA
AACCAGAAAAACCTTTAACTGAATCTGTCTCAGAAAGTTGACACGCTGGCAGTGAGTTAAATAAGCCTCTGCTACGTAAGGGTTATAGCTTTTG
CCTTAAAGATGCATTTAAAATACATCTTATCTTATTAAGAATGAGGTATCAGCTATGGCTTATCGCGACCAACCTTTAGGTGAACTGGCGCTCT
CTATCCCTCGCGCTTCGGCTCTGTTTCGTAAATATGATATGGATTACTGCTGTGGCGGTAAGCAGACGCTGGCGCGCGCGGCGGCACGTAAAGA
ACTGGATGTTGAGGTCATTGAAGCTGAACTGGCAAAGCTTGCTGAACAACCGATTGAGAAAGACTGGCGTAGCGCCCCGCTGGCAGAAATCATT
GACCATATCATCGTGCGCTAC

PREDIKCE GENŮ U EUKARYOT

Genomy j**ednobuněčných** 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 nižší eukaryota. U identifikovaných genů určete možnou funkci.

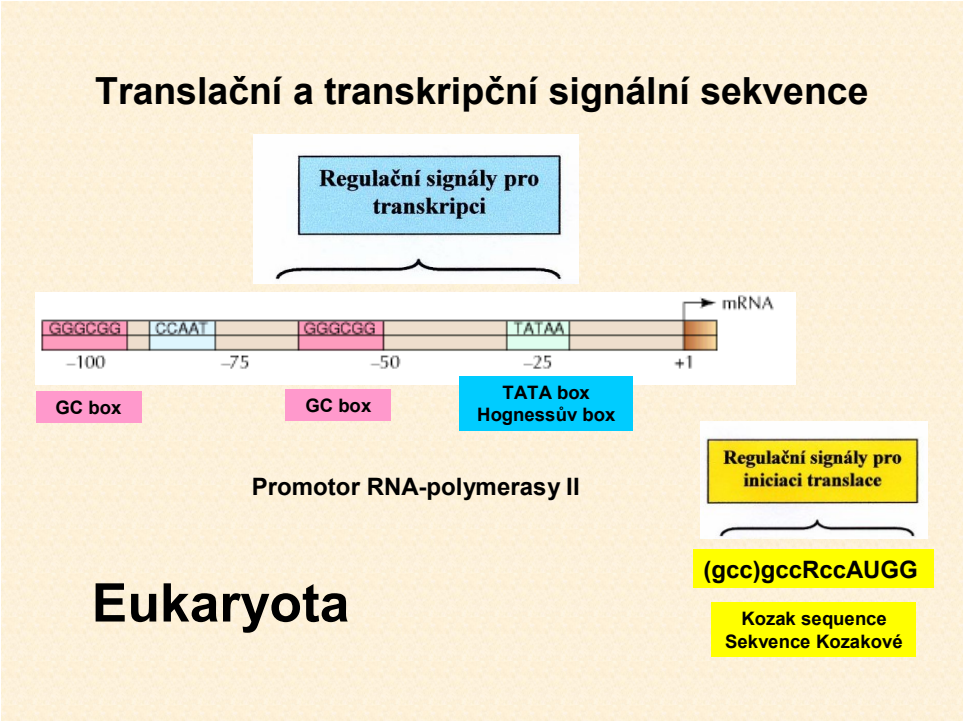
TCAATGTCTTTGTGATTAATCTTCGGACATATGCAATTTTGAAATATTAATTATGTTCATATCAACATCACGAGCTTGGAACTGAATATTGAGT
AATCACGATATTGCACCGACTCAACCTTGATTAAATCTGCTAGTGTACACTAATTGTGACCTTTCTATTAGATCATTATCTTTCTTAAATGTGT
TGAAATCACGAGAAATTGTAAATACTACTAAGCCTGTATATTCAACACAGAAATACATCATAGATGTTTTTTTTTCTTATTATTGTGTGTCAAT
TAACCTGTCTCCCTCTCCTTTGAACAATCAATTAGTACTTTACAACTCACAATTTGTAAACCCTGTCCGGATACACATTTTATTTCAAATTACC
GTGAGTGCATGTACTAAATTCAACTGGTCGTCGTTTCATTTCAAGGTCAACGCATAGTCTAATAAGTATGTCGTTTCGAGAACTCTACTGTCAA
TCTCTCGTGTAAAGAAGTATTTTTTTTTTGGTCCAGTTTTGCAACAAATTATACTTTAAATTCTCATAGAATGTTTATTTTCATTTACTACAGA
CCTTACAAAATCCTAACCAGACTAGATGTCTATCTATGGAGCATCAGTACATTCCCAATCTGTTGGTTCATAAGCATCAGGTTGATCAATAGCC
TTAATTTCTTCATTAGTCAATTCAAAATCATCAACAGAAAGATTTTCCTTTAATCGAGATGGAGTTTTTGTTTTTGGTAATGGTATATAACCTT
TTTGTAATGACCATTTAATCAATATTTGAGCAGCTGATTTATTATACTTTTGCATAATTTCTTGAAATTCAGTATTGTTGACTTGTAATTTGTT
ACCATGGGTTAATGGTGCATATGCCTCAACATTGATACCTTTACTTAAACACCAAGTAGCTAAATCCTGTCTCATACACCAAGGACTAATTTCA
ATTTGATTGACAGCTGGAGGGATCGTTGCATTGGTCAACAATTCTTCAATATGATGTTTACCATAATTAGAAACCCCAATGTTTTTAATCCATC
CTTTTTCCACAGCATCCTGCAAAACTTTCCAGCTTTCTAAACGTTTGGTCTTACCTGGTAATGGAGAATGAATCAATAATAAATCAATATATTC
TAATTTATCACCAACTTGAGCCATCATTGTTGAAATGGCTTGTTTAGTGCTTGAAGTACCCAATTGATTATTCCAAAGCTTTGTGGTGTAAAAA
AACTCAGATCGTGGTATATTGGGGTTCTCTCGTAAGAATTTACTTATACCTTCAATGACTTCTTCTTCGTTTCCATATAACACTGCAGTATCAA
AATGACGATATCCAACTTTACAAGCTTCATAAACTACCGAAACCGTTTTATTTCTTGGGATATCATAACATCCTAATCCAATTGATGGAATGGT
ATGACCGGAATTGAGTTTGATTAATCGATATGACATTTTTTCTTAAGCAAATTGAGCAATTGAGTGAAATATAAAGAGTATTTGAAATCTAGAG
AATTAAAAAAAAAGAAGTGGTTGTTGTAAGAACAATTGACCAATTTGGGTGGAACTCCTTTTTTATTGGTTTCGAACTGTGGGGATGGCAAAGC
AAAGCAATCTTTTGTACAGTCTCAAAACTTCTTTCTATTATCGTTCCTTCGTGTAAAATCTGCACATTTGCATTGCCCAATTTTGTTTATTATA
TCATTTCGTTACAATGAATTGACAGACGAGTATTTGGTTATAGTATTTGGACTATTTGGTCTTGTAGATCACCATTTACACTTATAATCAAAAT
TTCTCTGGAATTAAGATCTAACTTTGAAAACAATACAGACGATTCCAGCCTGTTATTCCCGGGATTGGATCTTTTGGTTGCGTTAGTCCGAGGA
TATATGCACGTGAATTGATGGTGAAGCTATAAGACCAGTGGTAGTGTACTGAGACGAAAAATGGAATAGAGCAAGAAAGAAGTGTGCTGAGGGT
TGAGAAAAATAATAGATACTGTAGTGGTAGTAGTAGTGGTATGATGCAAATTTATTGTTTGTATTTTTTTTTTTTGTTGCTCCTCGTCGACAACG
ACTGCTGATTCACACCACCACCACCTTTGTAGTCAAATTAAAATCGTCGGTGGCAAAAAAAAACACAGACACACACATACAGACAATTGCGATA
TTTTATTTTAGGGTCAATTTGGAACGTCACCATACAGTATTCAGTATGATTAATTTGACTTGCAAAATACAAATCGCGTGGTTCAATTCCGAAA
AATTGATAGTGTTCCATGAAAAATTACAAGGAGAGCCGTGAATTATAACAAATTTACTCACATTTTGGCTGGCAATTTCTTTAACCCCCTTAAT
TTTTTTCATGCAGTTTATTATTTTACGTAGAAGCTTATTCTCTAAACACTCGGTTATCCGAATATATTTTGCTTTCAGAGCAATTGACTGATAT
AATACCATCAATTGAATAAATCAGTATCATCAAATGCTAGTTTGAGTTCAATAAAAGGTATAATCGGTGTTAACTGAAAGGTATTGTCTAATTA
ATTAGGTTGCGGTTTTTGGATCAGCTGTTTCAAATCCAGGCTCGTGGTGCACATTTTTTATTTTTGTTGTGATGATTTTTTCTTTTCTTTGCTT
AATAGTTAGTTAGTAAAATGGATAATGACGCAAAATTGCTTCTGAATGGATGATAATGCCCCGGGTATGGCGTGTCCCGTACAAATCATATTAT
GCAAAAATAGCCAAAATAGACAAAAATAGGTGGAAATTGTTTCAACCAACCCACTCAACCACTAAAATTTGTCTTGTCAAAGCCAATTCCAACG
AAAATTGGATTTTTTTTGACTTAATATATATTAACCCTTGAAATTCCCTTCAATTGGAAAATTTTTTTAATTTATTTATTTCTTCTTTCTTTTT
TTCTTTTTCTTTTTCTTTCTTTTATTCATCACAATTGAATTCAAATCAATTAACATCAACAAACTTTACAATCAAATTAATTAACAATGGCTAT
TAAAATTGGTATTAACGGTTTCGGTAGAATCGGTAGATTAGTCTTAAGAGTTGCTTTGGGCAGAAAAGACATTGAAGTTGTCGCCGTCAACGAT
CCATTCATTGCTCCAGACTATGCTGCTTACATGTTCAAATACGATTCTACTCACGGTAGATACAAGGGTGAAGTCACTGCTTCTGGTGACGACT
TGGTCATTGATGGTCACAAGATTAAAGTTTTCCAAGAAAGAGACCCAGCTAACATTCCATGGGGTAAATCTGGTGTTGACTACGTTATTGAATC
CACCGGTGTTTTCACCAAACTCGAAGGTGCTCAAAAACACATTGATGCTGGTGCCAAAAAAGTTATCATCACTGCTCCATCTGCTGATGCCCCA
ATGTTTGTTGTCGGTGTTAACGAAGACAAATACACTCCAGACTTGAAGATTATCTCCAATGCTTCTTGTACCACCAACTGTTTGGCTCCATTAG
CTAAAGTCGTCAACGATACTTTCGGTATTGAAGAAGGTTTGATGACCACTGTCCACTCCATCACTGCTACCCAAAAGACCGTTGACGGTCCATC
CCACAAGGACTGGAGAGGTGGTAGAACTGCTTCTGGTAACATTATCCCATCTTCCACTGGTGCTGCTAAAGCCGTTGGTAAGGTTATTCCAGAA
TTGAACGGTAAATTGACTGGTATGTCTTTGAGAGTCCCAACCACCGATGTTTCCGTTGTTGACTTGACTGTCAGATTGAAGAAAGCTGCTTCTT
ACGAAGAAATTGCTCAAGCTATCAAGAAAGCTTCTGAAGGTCCATTGAAAGGTGTTTTGGGCTACACTGAAGATGCTGTTGTCTCCACCGATTT
CTTGGGTTCAAGCTACTCATCTATCTTTGATGAAAAAGCCGGTATCTTGTTGTCCCCAACTTTCGTCAAATTGATTTCCTGGTACGATAACGAA
TACGGTTACTCCACCAGAGTTGTTGACTTGTTGGAACACGTTGCTAAAGCTTCTGCTTGAAGTAAAACCAGACTTTGATTTGATTTTTTTAGTT
AATTGAAGTTATATAATGAGTTTTTGAAATAAATATCTTACATTGAAGAGAATATTGATGTGTTTTTTTTTTGTTTTTATATAGGTGTTTGCAG
TTATAAGTAAGACCAATATCATATCTGGAATATATATAATTACAAAAAAAATTAAATATTACAAGCGCGTATAAAGTAGTTTAGATTTTGAAAC
TGATTTTTTTTTTTCAAATATAAACAAACCACATCTTTCTTTTGACTTCCTATTACCAATGTCCCAGGATAGAGATATATCGTCGTTGATAATA
AATCAAACATTTAAGATCTATAAATCCAATTCATCACTATCGTCATCTATAAAACTCACTAAAAAACAATTGGAGAAATTTCGAATAACCCTAT
GTAAATACTTGTACCAAAGATATCAACAAACTTTACAAATAGAAGATGATTTACTAGTTCAATATGAATTATATCGTGATATTGTACCTGAACT

CACTATTTATAACAATTCAATTGAAGAGCATTCACAGTATAAAATCCAATATCTAATAATCAAATGTGCCAATAATCGATTGATATCTCGTCAA
GTGTTCATATTTGTTAGTAATGATGATGATTTCACTTTGGTATTGAATAGAATTACCAATCACCCTAGCATAAATGAATATATTTTGGAATTAT
TAGAAGATATTAATGATTATCCATTGATTATAAAACCACTTGATTTACCTGATTATCTAATACCCAACTTGGTGGATCAATTGGGTAATAAGTT
GGATTCTTTGGGTGATTTACAATTAGTATATTCATCTTCTACATCTAATAGGTTGAGAAATTTTATAATTGATATTCCGAAAAACGATTTACCA
AAATTGTCTAGTGATGGGAAATTATATCAAGATATTGTTAAATTCATGTATAATAATACCAAAATAAAATTTGAAAAATTAAGAATAGAAAAAT
TTATTAATAATTTAATTAATATTGCTAGTGATGGGAAATTTAAATTAATAATTAATGATAATCAATTAATTTGGTTTTTGATTGAATCTATAAA
ATCATCTTTACAATAAGTAGTAAGTGTACGTGGTCAATAAAGTAATGTATATATGTGCGTGTACTGAGTGATGATTTTCTATTTTATCTCGCGC
GTTGTTATTTTTGTTGTTATTGTTGTTGTTGTCTGAGACTTAATTAATAACTTACGTGTTCACTTACATTTATTTGCTTATGAGCCTAGGCGTT
ATTTTTTTTTTCTCTCTCTCTCCCACTTTCTACTCTTTGAACAATGTGGTGAAAAAAAAAAAAAAATAGAACCAGATCTTTTTTATTCCATATAT
TTCAAAACCGACTTGTTTCTTTTTTTCTTTTACAACCAATTAATACAAAAGAAAGACAACAATCATAATGTCTACTTCTGTTGAACCCAATGAA
ACAGAAGCTTTGTTGAGAAAGCAGAATGATCTTTCCACAACTGCCTCAATTGAAGAAAAATATCCTCACCAACAAGGAGAGGCTGCAGAAGACG
ATGACGACACTCTTAAAAGAACCCAATATGATGAAGCTAAAGAAACCGCTGAATCTTTAAAACAAGTTGAATCGATATTAGCACCTATTGTTTT
CACTGCATTGTCATTTTTCGTGAGATTTTATCGTATTTCAGTGAATGACCATGTTGTTTGGGATGAAGCTCATTTTGGTAAATTTGGATCCTAT
TATTTACGACACGAATTTTATCATGATGTTCATCCTCCATTGGGTAAAATGTTAGTTGGTTTATCTGGTTATTTGGCTGGATACAATGGATCTT
GGGATTTCCCAAGTGGTGAAAAATACCCTGATTATATTGATTATACTAAAATGAGATTGTTTAATGCCACTTTCTCTGCCTTGTGTGTACCATT
GGCTTATTTCACTGGGAAAGAAGTTGGATTTTCCATGTTTACTACTTGGTTATTTACTTTGATGGTGGCTCTTGAATCAAGTTATGTCACTTTA
GGTAAATTCATTTTGTTGGATTCAATGTTGTTATTCTTCACCGTTGCTACTGTTTTCTGTTTTTCACGTTTCAACAATTTTAACAATAAATCAC
AAGAATTTTCTAGAAAATGGTGGAAATGGATTCTTTTAACTGGTGTTTCCATTGGTTGTACTTGTTCTGTAAAAATGGTTGGATTGTTTGTTAC
CACATTGGTCGGGATTTACACTGTTGTTGACCTTTGGAATAAATTGAGTGATAAATCTATTTCATGGACAAAATACATTCAACATTGGTTTGCT
AGAATTGTTGCTTTGATTCTTGTCCCAATTTTCATTTTCATGCTTTCATTTAAAGTTCATTTTGATTTGTTGTATAAATCGGGTACTGGTGATG
CCAATATGTCATCACTTTTCCAAGCTAATTTGGCTGGTTCCGATGTTGGTGGTGGCCCACGTGAAGTATCCATGTTCCACTCGGTTATCACTTT
AAAGAATCAAGGTTTAAGTGGTGGCCTTTTACACTCCCACGTTCAAACATTCCCAGAAGGTTCAAAACAACAACAAGTTACTACTTATGGTCAC
AAAGATTCAAACAACAATTGGATTTTCCAAAGAGCTAGAGGACAACCTTATTATGATACTTCTGGTAACACCACTGACATTGAATATATTTTTG
ACGGTATGCATGTAAGATTGATGCATCCACAAACTGGTAGAAACTTACATACTCATGATATTCCAGCTCCAGTGTCTAAATCTGAATATGAAGT
TGCATGTTATGGTAATTTGACTATTGGTGATCCTAAAGATAATTGGACTGTTGAAATTATGGAACAAGCAAGTGATGAAGATAAAATGAGATTA
CATCCTTTGACTTCGTCATTTAGATTGAAGAATGAAGTGATGAATTGTTATTTGGGGGTCACTGGTACTACATTACCTCAATGGGGGTTCAGAC
AAGGTGAAGTTGTTTGTTACAAGAACCCATTTAAAAAAGACAAGAGAACTTG

ÚKOL 3

Charakterizujte část genomu *Candida albicans* (kvasinka) rovněž pomocí heuristického modelu.

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.

ÚKOL 4

Analyzujte část genomu mnohobuněčného eukaryotického organismu pomocí programu **GeneMark** (http://exon.gatech.edu/GeneMark), verze pro eukaryota.

```
GATCGGGGTATCCCAGCTGCTAATTCCCCTCCTAACAGCCCTGCAGCGCCTGGATTGGAGGTGTGACAGGTGGCATTTTTGGCCCGAAGTCCCTAGGCAGGACGGGT
GGCTGTAGCCTGCGCGTTTGTGTAAATTAACCTCCCTTTTTAGTTCTGTGAAGGGAGAACGTGATTAGTTTAAAATCACCTGCGTGAAAGAGGCTCTTGATTAGTTT
AAAATCACCCACGTGAAAGGGGCTCTTGATTAGCTTAAAATCACCTGGGTGTTTCGTTTTATCTCAGTTGGCCTCTGAATGTTTATCCTACAGCTTCCTGCAAGCTC
CAGTCTACCTTCTAAGGGAATAACGGGGCTCGTCTTTCTACGTTACTTTATGTGGGAGTAGTTGTAAAACAGAATGAGGGAGACCGCGGGGATTGTGCGGAAATGTA
GCCACCGTTTACATGTCGGGAGGCTGAGGTGGGAGGATCGCTTGAGCCCAGGAGTTTGACACCAGCCTGGGCAACATGGTAAGACCCCTTCTCTACCCTCCCCCCCA
AAAAAAAAATTTTAGTGTGGCATGGTGGTGGGTGCCTGTGGTCCCAGCTACTCGGGACGCTGAGGTGGGAGGATTGCTTGAGCCCAGGAGTTTGACACCAGCATGGG
CAATATGGTAAGACCCCATTATCTACAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAAATTAGCCGGGCGTGGTGGTGGGTGCCTGTGGTCCCAGCTGCTCGGGACGCTG
AGGTGGGAGGATTGCTTGAGCCCAGGAGTTTGACACCAGCATGGGCAATATGGTAAGACCCCATTATCTACAAAAAAAAAAAAAAAAAAAAGAAAAGCCGGGCGTGGTG
GTGGGTGCCTGTGGTCCCAGCTGCTCGGGACGCTGAGGTGGGAGGATTGCTTGAGCCCAGGAGTTTGACACCAGCATGGGCAATATGGTAAGACCCCATTATCTACA
AAAAAAAAAAAAAAAAGAAAAGCCGGGCGTGGTGGTGGGTGCCTGTGGTCCCAGCTGCTCGGGACGCTGAGGTGGGAGGATTGCTTGAGCCCAGGAGTTTGACACCAG
CATGGGCAATATGGTAAGACCCCATTATCTACAAAAAAAAAAAAAAAAAAGAAAAGCCGGGCGTGGTGGTGGGTGCCTGTGGTCCCAGCTGCTCGGGACGCTGAGGTGG
GAGGATTGCTTGAGCCCAGGAGTTTGACACCAGCATGGGCAATATGGTAAGACCCTATTATCTACAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAAATTAGCCTGGTG
TGGTGGTGGGTGCCTGTGGTCCCAGCTGCTCGGGACGCTGAGGTGGGAGGATTGCTTGAGCCCAGGAGTTTGACGCCAGCATGGGCAATATGGTAAGACCCTATTAT
CTACAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAAATTAGCCTGGTGTGGTGGTGGGTGCCTGTGGTCCCAGCTGCTCGGGACGCTGAGGTGGGAGGATTGCTTGAG
CCCAGGAGTTTGACACCAGCATGGGCAATATGGTAAGACCCCATTATCTACAAAAAAAAAAAAAAAAAAAAGAAAAGCCGGGCGTGGTGGTGGGTGCCTGTGGTCCCAG
CTGCTCGGGACGCTGAGGTGGGAGGATTGCTTGAGCCCAGGAGTTTGACGCCAGCATGGGCAATATGGTAAGACCCTATTATCTACAAAAAAAAAAAAAAAAAAAAAA
AAAAGAAAGAAAAATTAGCCTGGTGTGGTGGTGGGTGCCTGTGGTCCCAGCTGCTCGGGACGCTGAGGTGGGAGGATTGCTTGAGCCCAGGAGTTTGACGCCAGCAT
GGGCAATATGGTAAGACCCTATTATCTACAAAAAAAAAAAAAAAAGAAAGAAAGAAAAATTAGCCTGGTGTGGTGGTGGGTGCCTGTGGTCCCAGCTGCTCGGG
ACGCTGAGGTGGGAGGATCACCTGAGCCCAGGAGGTGGAGGCTGCAGTGAGCTTAGATTGCAGCACTGTGCTCCAGCCTGGGTGACAGAGTGAGACCCCATCTCAAA
CAAGAAGCACCAAAAGGTGTCTGAGATCAGGCACGGTGGCTCACGCCTGTCATTCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGTCAGGCATTTGAG
ACCAGCCTGGCCAACATGGAGAAACACTGACTCTACTAAATGTGCAAAAATTAGCCGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTCCAGAGGCTGAGGTAG
GAGAATTGCTTGAACCTGGGAGGTGGAGGTTGCAGTGAGCTGAGATTGTACCACTCCAGCCTGGGAGCTGAGATTGTACCACTGCACTCCAGCCTGGGCCCT
AAAGCGAGAGCTTGACGGTGGGGTGGGGGTGTGGGATGGACCTTCTGTGGTCCCAGCTACTCGGGAGGCTGAGGTGGGAGGATCGCTTGAGCCCAGGAGTTTGACAC
CAGCATGGGCAACATGGTAAGAACCCATATCTACAAAAAAAAAAATTTAAAAATTAGCCGGGTGTGGTGTGTGGGTGCCTGTGGTCCCAGCTACACCACGAGTGGAAT
TGTGACGTGCACAGCCTGGGTAAATGGGGTTAGCTATATTCACCCTGAAGGGGGCACATACCCTATTATTTCTAACACCACTGAACTCAGAATCGGGACCCAGATTC
GCCGGGGCAGCGCGGGACGAGGCCTCTCCCGGGCTCCGTCTTCCCTGTTAATAGACTGAGAGACGGAATGAGAGCTTCGAAGGTGCTGGGGGAGGAGCATTGGGTGT
CCCCAGGTGGTGAGCTTGGGTGTGTGTGTAGCCGACAGTCTCCCAGGCTCAGTATTCCCTGTTAATAGACTGAGAGACGGAATGAGGGCGTCCCCAGGTGGTGAG
CTCAGGGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAGCCGTCAATCTCCTGGGCTCCTGGGTTCCGTCTGCCCTGTTCATAGAGACAGA
ATGAGGGCGTCCCCAGGTGGTGAGCTCAGGGACGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAGCCGTCAATCTCCTGGGCTCCTGGGTTC
CGTCTGCCCTGTTCATAGAGACAGAATGAGGGCGTCCCCAGGTGGTGAGCTCAGGGACGTGTGTGTGTGTGTGTGTGTAACCATCAATCTCCTGGGCTCCTGG
GTTCCGTCTGCCCTGTTCATAGGGATGGAATGAGGGCGTCCCCAGGTGGTGAGCTCAGGGACGTTGTGTGTGTGTGTGTGTGTGTGTAACCATCAATCTCCTGGG
CTCCTGGGTTCCGTCTGCCCTGTTCATAGAGACGGAATGAGGGCGTCCTCAGGTGGTGAGCTCAGGGACGTGTGTGTAGCCATGATTATCCTTTAGTACCTGGTCCC
CGTGCCCTCCAAGCCATCCAAGGGCTTCAGGAGGCCGCGCAGAAGGAGGGTCGGTAAGCGAGTCCACGGACCACCACGCAGGCCCTGACCCCGCCAGGGGCCCTTTG
GAGGGTGAGCCCTGCGTGTTGGGGGGATGATGGCTCCGGAGGGTGAGCTCTGGGATGAGAGGGGACGTGCAGGTGGCCGCATTGCAGACACAGCTGGCTCGGGACCT
GGGCGCCATCTGCCGTCGACTCAGCTCGCTCCTGAGCTGTGCCTCTTCCTTGCCTCACGTTGGAGATAAAAAGCCATAGATGGCCGGGGGCGGTGGCTCAGGCCTGT
CATCCCAGCACTGTGGGATACCAGCGCAGGAGGGTTGCCTCAGCTCAGGGGTTCGAGACCAGCCTGGCCAGCATAGCGAGACCCTCTTCTCTGTAACAATTTTTTTT
TCTTTTCTTTTTTTTTTTTTTGAAACAGAGTCTCGCTCTGTCACCCAGGCTGGACTGTAGTGGTATGATCTCAGCTCACGGAAACCTCTGCCTCCCGGGTTCCCGGG
TTCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCAGCTGCCACCATGCCTGGCTAAATTCTGTATTTTTAGTAGAGACGGGGTTTCTCCATG
TTGGTCAGGCTGGTCTTGAACTCCCGACCTCAGGTGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATGACAGGCGTGAGCCACCGCGCCCGGCCTCTTTCAGGT
TTTTATATGTAACTTTGCCACTCAGTCGGTACTGAAGCCGTTATGGAGGCCTGCCTGGTGAGATGCGGCTGCCACAGTGTGGCTGGTATTCGAAACACACAGGAGGC
CGCAGCCAGCTTAGCAGAAGCACCAGGGCTTTCGGAACTTGGTTTTGAGAGTTTCCTGTTTTCTTGACTTTCATATACCATGGATGGTAATAACTGTTGGGGGGAAA
AAGCTATTAAAGATTCCTGTGGATGGTGAGATGACCTTGGTTTCAGAAACACCCAGACACAGAGAATTGTGCATTTTAGAATCCAATAAATCTACGCTGGGCACGGT
GGCTCATGCCTGTCATCCCAGCACTTTGGGAGGCCGAGGTAGGTGGATCACCTGAGGTCAGGAGTTCGAGACCAACCTGGCCAACATAGTGAGAACCTGTCTCTACT
AAAAATACAAAAAATTAGCGGGGCGTTGTGGCAGGTGCCTGTAATCCCAGCTGCTCTGGAGGCTGAGGTGGGAGAATTGCTTGAACCCGGGAGGCAGAGGTTGCAGT
GAGCTGAGACCATGCTGCTGTACTCCAGCCTGGGTGACAGGGCGAGGCTCCATCTTAAAAAAAAAAAAAATCTGATCATTTACCTGTGCACCTTCACAAGTGTTTAAA
AAAAAAAAAAAACCGGGCACGGTGGCTCACGCCTGTCATCCCAGCACTTTGGGAGGCTGAGGTGGGTGGATCACGAGGTCAGGAGATCGAGACCATCCTGGCTAATT
CGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCTGGGTGTGGTGGAAGGTGCCTGTAGTCCCAGCTGCTCAGGAGGCTGAGGCGGGAGAATGGCGTGAAC
CCAGGAGGCGGAGCTTGCAGTGAGCTGAGATTGCAGCACTGCACTCCAGCCTGGGTGACAGAGCCAGACTCTGTCTCAAAAAAAAAAAAAAAAAGATTGGGTTGATAT
TTTACCTACTATAATCTTATATGTTTGCAATTTCCTGTCATTGAGTTTTATAACTTTCTGCCCTTTGCTGACAAAAACATTTAAAATGTCTCATGGCAGCTTGTGTT
CTTAAAATTCAAACTGTTTATCCATAGCCCTTGAGAGAAGGGTTTAATGTGAGTTGGTTGGGAGTGACCGTTGTCTGTTTTTTTGAGACAGAGTGTCGCTCGCTCGC
CCAGGCTGGAGTGTAGTGGCGCGATCTCGGCTCACTGCAAACTCTGCCTTCCAGGTTCAAGTGATCCTTCTGCCCCAGCCCTCCAGCTAGCTTGGATTACAAACATG
CACCACCACACCTGGCTAACTTTTGTATTTTTGGTAGAGACAGCGTTTTACCACGTTGGCCAGGCTGGTCTCAAAATCCTGAGCTCAAGGGATCCGCCTGCCTCAGC
CTCCCAAAGTGCTGGGATTACAGGCGTGAGGCACGACGTCAGGCCAAGGGGGTGGCTGTTTTATAACAGGTTTGCAAGTTTGTTTTTGTTTTTGTTTTTCTTTTTTT
TTGAGGTGGAGTTTTGCTCTTGTTCCCCAGGCTGGAGTGCAATGGCAAGATCTCGGCTCACTGCAGCCTCTGCCTCCGGGGTTCAAGGGATTCTCCTGCCTCAGCCT
CTCGAGTAGCTGGGATTATAGGCACCTGCCACCATGCCTGGTTAATTTTGTATTTTTAGTAGAGATAGGGGTTTTTCCGTGCTGGCCAGGCTGGTCTCAAACTCCCG
AACTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCCCCGTGCCCAGCCAGGGGGTTGCTGTTTTATAATAGGTTTGCAAGTTTG
TTTTTTGTTTTGTTTTGTTTTGTTTTTTTGAGACGGAGTTTTGCTCTTGTTGCCCAGGCCGGAGTGCAATGGCACAACCTTGGCTCACTGCAACCTCCGCCTCCCGG
GTTCAAGCTATTCTCCTGCCTCAGTCTCCTGAGTAGCTGGGATTACAGGCGTCCGCCACCACACCCGGCCAAGGGGGTGGCTGTTTTATAACCAGTTTGCAAGTTTG
TTTTTCTCGTGTGTTAGAGGAAAATCTCTTCATGATATGGTGAAGCCGTCGTCTGTTTTGGAAGACGGTGGCCAGGAAAATGGTGGTAAGTTCTCCTGAAAGCAAAC
GCAGAAGGAAATAAAAATGACTGACAATTAGATCGTACAAAGATGAGCATTGTTTTTATTTATTTTTTTTGAGACGAAGTCTTGTTCTGTCACCCAGGCTGGAGTGC
AATGGCACGATCTTGGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGTGATTCTACTGCCTCAGCCTCCTGAGAAGCTAGAATTACAGGCATGCTCCACCACTACCC
AGCTAGTATTTGTAGTTTTAGTAGAGACGGGATTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCT
GGGATGACAGGCGTGAGCCACCGTGCCCCGGCCCCGGATTCTATTTCTTAAACTGCAGGAAATACTAAATACTTTACATGCGTATGTTTCTCATTTCATTTTCCTAC
AGGAATTGGAGAGTGAGTTGCTGAGATTTCATTTTTATTCTCTTGGTCCTCATCATGCAAATGGCTAAATGGTTTTTTTTTGAAACAGGGTCTTGCTCTGTCCCCCA
GGCTGGAGTGCAATGGGGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATCACAGGTGCGCAC
CACCGTGTCCGTCTGATTGTTAAATTTTTCCTAGAGACGAGGTTTTGCCACGTTGCCCAGGCTGGTCTCAAACTCCTGAGCTCAGGTGATCCCCCCACCTCGGCCTC
CGAAGGCGCTGGGCTTCCAGGCGTGCAGAGGGCTCTTTGTCTACATCCGTCACTGGATTTTCTACCCACGGGGACATGTGACCCCAGCGAGGTGGTCGGGGAGTCGT
CGCCGGCCAGCGCTGCAGCCCAGGGATGATGACGGCTCTAGGAGTTTGGTTGTGGAGCACAGATGCCGGCGGTGTGGGCACGAGGCCCTGCACCTGCGTTTCTCCTG
CACGGGCCCCGGAACCCCGGGATGGGATCTCTGTGGAGTGATCACGTCAGTGGGCTCTCGGCGAGGCTGGCGGGTGGGAGTCGGCGGCATCCGGGCTGGGCCCCCGA
GGCTTGGCTCCCTCTCTGCACCCCAGGGGTCCCACAGGCTGGTGACAGCGCGACTGAAGGCCACGCTGCAGTCAGGAGGCCGCCCCACGTGTGACCCCGGGTCTTCC
CACCCCGCCCTGACCCAGGGTCCTGGCCTTCCACGTGTGACCCTGGTTCCGGCACCTGTCCCCGCCCCCCCATCCTGTTCCCCTCCTCTCTCCCTGCCCCCCCTCCC
CTCTCCCCGCCCCCCCTCCCCTCTCCCCGCCCCCCTCCCCTCTCCCCGTCCCCCTCCCCTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCTCTCTCCC
CGTCCCCCCTCCCCTCTCCCCGCCCCCCCGTCCCCCCTCCCCTCTCCCCGTCCCCCCTCCCCTCTCCCCGTCCCCCCTCCCCTCTCCCCGTCCCCCCTCCCTCTCTC
CCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCCTCTCCCCGCCCCCCCGTCCCCCCTCCCC
TCTCCCCGTCCCCCCTCCCCTCTCCCCGTCCCCCCTCCCCTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCTCTCCC
CCGCCCCCCCCATCCCCCCTCCCCTCTCCCCGCCCCCCCTCCCCGTCCCCCTCCCCTCTCCCCCTCCCCCCTCCCCGCCCCCCTCCCCTCTCCCCCTCCCCCCTCCCC
GCCCCCCTCCCCTCTCCCTGCCCTCCCCTCTCCCTGCCCCCCCTCCCTGCCCCCCTCCCCTCTCCGCCCCACCTCTCCCCTCCCCCCTCCCCCTCCTCTCTCCCCCC
TCCCCCTCCTCTCTCCCCCTCCCCCCCTCTCTCCCCCTCCCCCTCCTCCTCCTCTCTCCCCCTCCCCCCTCCCCCTCCCCTCTCCTCTCTCCCCGCCCCCCTCCCTC
TCTCCCCGCCCCCCTCCCTCTCTCCCCGTCCCCCTCTCTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCCCTCTCTCCC
CCGTCCCCCCTCCTCTCCCCGTCCCCCCTACTCTCCCCGTCCCCCCTCCCTCTCTCCCCGTCCCCCCTCTCCCCGTCCCCCCTCCCCCTCTCCCCGTCCCCCCCTCC
CCCTCTCCCCGTCCCCCCTCCCCCTCTCCCCGTCCCCCCCTCCCCCTCTCCCCGTCCCCCCCTCCCCCTCTCCCCGTCCCCCCCTCCCCCTCTCCCCGTCCCCCCC
```

TCCCCCTCTCCCCGTCCCCCCCTCCCCCTCTCCCCGTCCCCCCCTCCCCCTCTCCCCGTCCCCCCCTCCCCCTCTCCCCGTCCCATCTCTCCCTTGAGTCCTTGGGT
GGTCCTGGCCGCACGCCTACCTGGCGGCCCCATCCGTCCCTTCAGGGCAGGGACAGCCGATGCCCCGCACCTCCCCCATCGCTTTTCTGCCTCTGCCCCCTCCCCAC
ATAGGCCTGGCGTCCCCTGGACTCCTGTAGGCCGAGCTGTAGTCTTAAGAAGTGTGTCCCGCAGCCCTAAAAGGTTGGGGTACCCTCCGGCCGTCATGTTCTGGGCA
CCCTATAGTCCTAAAGGGTCAGGGCACCCCATGGCTGTCATGTTTAGAGTACCCCGTAGTCCTAAAGGATTGGGGCACCCCACATCCCCGGTGTTCAGGGCACGTCA
TAGTCCTAAAAGGTCGGGGTACCCCTGGATGAGGGGCGGTCAGGAGGGTCAGGGCCTGGGTGGACATCCTGTCATCCTACCCTGTGGGCTATGCCCTGTGTCCCTGA
CTCTGACATTGGGTGAGGTGAGGACGGGGCAGGTTTCCATGGTACATTTATTTATTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGCCTGGAGTGCAGTGGTGTGA
TCTCGGCTCCCTGCAAGCTCCGCCTCCCGGGTTCATGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATCACAGGCGCCCGCCACCACGCCCGGCTAATTTTT
TGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCTAACTCCTGACCTCGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATGACAGGC
GTGAGCCCCTGTGCCTGGTGGGACCCCAGTTTTCTATTAAAAAAAGGCCGGGCGCGGTGGCTCACACCTGTCATCCCAGCATTTTGGGAGGCCGAGGCGGGTGGATC
ACTTGAGGTCAGTTGTTTGAGACCATCCTGGCCAACATGGAGAAACCCCGTCTCAACTGAAAATACAAAAATTAGCCGGGCGTGGTGGTGGCGGGCGCCTGTGATCC
CAGCTACGCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGAAGGCAGAGGCTGCAGTGAGCCGAGATCGCACCATGGCAGTCCAGCCTGGGCGACAAACAGAAA
ACAGGCACAGCATAGAGTCTGCCTTTCTGGAATCGCTATAGCTAGGGGGCTGCCTCCCTAAGCCCCACTTCTCTGAGTCTCCTGCAAGGTTTGTGTGAGTGCCTGGG
GGCCGAGGCCCTGGCGAGGAGGATGGGGGCTCGGGGGCCTCTGTGACCTGGGGGAGGGGCTGTCGCGTCCCCCTGTGGTCGCGTCCCCCTGTGGTCGCGTCCCCCTG
TGGTCGCACCTGTGGGCTGTCGGCCGTAGGTACCGCTCAGGCCCGTCCTGGTTGCGGGGCAAGGGCCTGTGTGTTGGGGGCTGAGAGACGGGGCCTAGAAGCCCTGT
TGTGGGGAGGCCTGTGGGGGGCTTGTCCAAGCGGGACCTCCTATCCCTGCACCACTGGGGTCGTCCCCCCCGACGGGGCCTCGTCCCCCCCATGGCTGCTCCGTGAG
GTCAGGCAGCGCCGCGCTAGGGAGATGGGGTGGCCCTGTCCCCGCAGAGGGGCTGTTCTCTGGGGCAGAGGCGGGCAGGACCCCCCGCCTTGTGGTCAGGACCCCCC
TCCTTGTGGTCAGGACCCACTCGTGGCGGCCGGGGCTGCTGCGGGAGCTGATACCGGGGTGGAGACAGCTACGCCATGGTGAGGACACGGGGCCGGGGGAGGCTGCA
GCCGACGGCCAGGCCCCCGTGCGGAAGACCCGGGGCAGCCCCCGCCTAGGAGCTGCGCAGACACGTCCTGGGGCCGCTGTTTCACCGAGACCCGAACGGCTGCTCCC
CAGGCTCTGAGTGGCGTGTGCGGCAGCCAAGGCTGGGGGAGCTGTGACTTCTACACGGCGCCAGGGCGCAGGGATCCCTGCTGTGCCAGGGTCTGCCCGTGGGTCAT
GGGGAACGTGGTTCGAGGCCCCTCTGCTGTAGCCACGCAGCAGCGTCTTGCCGGGTGCGTGGCCTTGTCTGCAGGCTTGAGACCCAGCGGGGCCGTCTGTGCCACGT
GGGGTCTGCGCCGGCCCAGGGGCCGCACCGCTGTCAGCCACAGGATGTGTGTCTCGGGGCTGCGGCAGACGGGGCTTCTGTGCCGTGTGGGGTCCGACCCCTCGGCC
TGGGCAGGGCGGCCCGTGGGCACCGTGCTGGGGACGACCTCCGGCCTCTGCTCCTTCCCCGGCTCTGAGTCATAAAGGCCTCCCGCTCCCCCGGGAGCCGTGTCTTC
TCTAGGAACTGCTGAGCGCGGCTGGGTCTGGCTCTGCCTCATTCCTGCATCTGATGCCCGCCCGTTCTCCCCCAACTCCCCTTTCTCCATCTCCTCCCTCCCCTCCG
TCTTGCCCGTCTCCATTTCCTCCATCTCTCCCTTCTCCGTCTCCTCCGTCTCGTGGCTGTCCAGACCCCGTCCCTCCGCAGGCTGGGTCCCCCCTTGTTCTGGGCGG
GGCTCTCGGTGCCGGTCTGGTGCCAGTGCTGGGAACCTGTGGTGCCCGTCACCTCTGCTCCCCGTCACCTCGCTCCACCTCGAGGAGCTGTGTCCTGTCCCACGAG
GAGGCGCCCCGGGACCTCAGGGCTGCAGGCCATGGCTGGGTCCGGCTTTGCCTCATCCCCTGTATCTGAGGCCCCCCAGTTCTCCCCCATCTCCCCCTCTCCATCTC
CTCTGTCTCCCCCGTCTCATGGCTCTCCGGGCCCCGTCCCCCCCACCAGGCTGGGTCCCCTTATTCTGGGTAGCCCGTCTTGCCATCACTGTCCCTGCCGCCCCTCA
GGACAGCGCTGTGGATGCCGGTCTGGTGCCGGCGGTCACCCCTGCTGCGCCTTGAGCAGCTGTGTCTTGTCCCGTGAGGAGGTGCCCCGGGACCTCAGGGCTGTGGC
TGGCACCTTTCATGGTCCTCCCGGAACAGAGCTCATGCTGCTGCTCCTGGGTGGACCCGGGGCTGCCCCTGTGAACACTCTTGGTTCTGACACGAAGGCACCCGTG
AGGGAGGGGAGGGGTGGGCTGGGTCGGGTCCCTGCTCGGGGATCTCCAAGGGGTGACGGCCACGAAGTGCGGGGGACCCACGATTGTCATTCAGAGGGTGAGGCGCT
GAGCCCTGGAACATCCTGGCTTCTGGCGGGACGCTGGCCACGGTCAGGGTGCCCCGAGTGCTTGTTTTGTGGGGTCCCTGGCCCCCTGCACACGTGAGCCCTCGGGG
GAATGGACCCCCAGAACACGTGGAGGGCATGACTGCTGGGCTCTGGGTCTCCCCGCAGCGCCGCCGTCCACGAGGACCCCCGACAGGACCCCCCCACTTCTCCGGCA
AAGGTGTCGGTGTGCACGCCCGGCCCCGGCACTGCCCGGGGCATCTGGGGTTGTGGCTGGGGAAGGACGGGGCACTTGGGGGAATGCATTGGGTACAGGGATGTCGG
AGGATGCTCAGTGGCTGTGCAGGGGCCTCACCAGCGTCGTGGGTGGGGGACGCTGTCTTTCACGGCCCCAGGGAGGCCTCTCATGTCTCAGCCTTGAACCCTAATTG
GGGGGGCTCCCAGTGTCAGTGTGGGGAGGCTGAGGGCGCTGGTCAGAGCCCTCCCATGACCATGGGCTCCGGCTCATAGGTGTGTGTGTGTTGGGGGGTGGTGTCCG
GAGGTGGTGTGGGGCGGACATGGTCTCCCAGACGGGCTCCCCGGAGCTGAGGGCAGGGTGGGCCCTGCGGGCCGTGGCTCCCAGTGTCATTTCCCTGCGGCAGAACC
TCCTGCCCTGCACCTGTCCTGTTGTTGGGGGTCCTGACCCTTCGCAGACCCAGGTGGGGGCATCCGGGAATTCAGGGGAGCACCTGGTGACACTGAGGACACCTTGG
GGACAGCAGCATTCCCCCGGGAGGTGCTGCCTCGTCTACCCACCACCTGGAGCCCCCGCAGTCTCTCCTCAGGGGTCACGGGACCACCCTGTGCGTCCGCCTCCTTC
ACAGCGACGGCTGCCGGTGGATGAACGGGGTGCGTATGCTGCTCTCGGGGTGGTCCCGGCTTGGCCCTGCCCTGCCGCGTGGACGCTGCCACCTCCGTTTCCTGGGG
GAGAGTTTTGTCGTGTGTGAGGACGGGGTTGTTCCCATGGTCACCACCTGTGACCCTCACCTGCGTGCAGAGGGAGCACAGCTGGGCCTGGGCTCATCCACGTCGGG
TGCCCCTGGGGTGCTTGTGGCCTCCCCTGCACAGGGAGCTCCTCCTGGTGGGGCCTCCGACCCCCCTCATTTAGGTGTCCCCCAGCTGTGGACGTCCCCACTCTGGC
GCGGGCTCCTTCCTTCTGGGGTCGTTTGTGGAACGGGGCCTGGGTCCGTGGCTCCGGGAGGGAACGGCCCCGAGGGTGGGGTCCCCCTGGCTGGGGAGGGGCCCGC
AGAGAGGCCCAGGCTGCACCCCGAGGCGTCTGGCGTGGCTGCGGTCCCCCTCCAGCACTGACGCTTGTAAATCCTGGACGGAGCCACGCTGTGGCTGAGGGAGCCGA
GCCCGTGACTGAAAATCCTCCAAACGTACAAGCGTGATCGCCAGAAACGGTTTTGTACGTTTACACAAAACATTCACACAGCCTGTGGTGGAGGCTCCTGTCCAGGA
CTGAGGCGCCCGGGAGCCGCCGGTCACCGTTGTGCGCACACGGACCCTTTCCACAGACGCAGGCCCCGGAACCCAGGCTGGGTCGGGAACGGCAAGCGCCAGAGGGT
GTCCGTGTGGGAACCCGTCCCATTCACGCGCGGCCCTACGTGTCCCCCTGGCACAGAGCTCTGGGCAGGTCCAGCCACGAACCCACAGCGGCAATCAACACGCTTCT
GTGAATAAATAAAAGTTTATCATTCCGTACAAACGCACTCATTTTCCACAACAGTTTTTACACGAGCCGCGGTGGCCCGGTGGTGGCACGTGGGGAGCGGCCCCGCG
GCGGCGTTCTCGCGGGCGGCGTCACAGCGGCTCCAGGTCCTCGTCCCCGCATGCGTACTCGTACAGGTCCACGGCGCCCAGCGGTGAGGGCGCCTCGAAGAAGGGCC
TCTGGGCCAGCGGGGAGCGCAGCGCACTCAGCTTCTGCTCCACAGGGCTGAGCTCGGCCTCGAACCTGCAACGAGGGGATGGCGAAGACGTGGCCAGCAGCGCGGAG
CAGGCCCTGGGCCGTTTTCCGGGAAACACAGGCTGCTCGGCCGCCAACCTCCAGCACGAGACAGTCCCCTGAGCCGACCTTGAGCTCCAGCCGCTGAGCCAGCAACG
CGCTCCTGGCTGAGGACGCCGGCTCCACGAATGCTCCCGGGGGAGGGGCCTGCGGTGCAGGTGAAAACCCACCCAGGACGGGACGGAGCCCCGGGCGTGAGCTACAC
ACAGAGGACCTGGGCGGGCGCGCCTCTTCCCCAAAGCGAGCTGCACGGGACCTGTGTTTACAACTGGGTCGTGGTGTCCACACAGGACTGGAGAACAAACAGGACCA
CAGTTTAGACACATTCTCCCAAAGCACAGAAGCTCCCAGTGGAGGCTACCACTTCCGTTCTGACCAGAACACGCAGCAGTGCGAGGACTGGACGCGCGCCTTCCCTT
AAGGTCTCGGCAGTGACGTCGGGCGCTGCTCCGGGGAACCCAGCGCGGCCACGGACCACAGACTCTCAGAGATCCTGTGGGGACACCTCTGCTTTCCAAGACGCC
CCCGACAGCCTGGCACTGAGCTCCCCGCGGTCACCCCCACGGACCTGAGCTCCCACTGCCGGAGCAGGCGTGGACACGGTGAAGGCAGTCGACCCCAGCGTGCTGCT
GGCTGTGCCGCAGACACAAAGCCGGGCAGCCGTGGGGGTGGACACTGGTCAGGTGGGGAGCAGAGGTGACAGACGGTGGTGGGGACAGGCGCACACCCATCCAGGCT
AGGTCTGGGGGCCACAGCTGCCTTAGGGTGTCCGACAGGAGAAGGCAGAGACTGGGAAGTTGTGGGGACCCAGAGGTTGGCTGCCAGGTCAGAGCCACCGTGCAAAC
AATGACCACGAGGAGGACGGACACACAGGACAGACGGCCCAGGGTGGGCCGCAGGGGCCCTGTGCTGCTTGGCAGGAATCGGACGGGGCGGCTGATGGGGGAGCAGG
GGTTAAAATGCACGTCCATGGAACCCATGGGTCGGGAACCCAACGGTAACCAGAAAGTATTTTGAGCTGCACCAAAATCAGAAGGCCGCGGGCTGCGGTGTCCATGG
CCCACGTCGCTGCCTGAAGATACTCAAGGAAAACAGCAGAACAAAGGGGAGAAACACCCGTGGTTCCCTGGGAAGAGGAGTGAACCTCGTCGGCCCCTCGCCTCACT
GATCACACACAAAGTCAGTATCAGGAGCGTGAGAGACGCTATCACTGGATTCTGCAGACATCAGAGTCACGAGCAGGAAATACCAACAATAACCCACTGATGGGAAC
GGGCAAGTTTCCGAAACACACAAACCAGCAAAGCTCACTCAAGGCTGGGCAGGGGCTCACACCTGTAATTCCGGCACGTTGGGGGACCGAGGCGGGGAGAACCCCTG
AGCTCAGGCATCCAAGACCAACCTGGGCAACACGGTGAAACCCCATCTCTACAAAAAAATACAAAAAAGTAGCCGGGTGTGGTGGTGCGTGCCTGGAAGCCCAGCTA
CTCGGGGGGCTCAGGTGCGAGGATTCCTTAAGCCCAGCAGGTGGAGGGTGCAGTGAGCTGAGATTGCGCCAATGCCCTCCAGCCTGCGTGACACAGCCAGACCCTGT
CTCCAAAAAAATAATAAAATACTACAGATAATCTCTGCAAAGAGAATCCAGTAATATGTGAGGATAACACATGACTCGGCAAGATCTATCCTGGGAACACGAGGCTG
GCTGAAAATGACTCGGTGTAACTCGCTGTATGGACACACAGGAACCACGGAGCGAAAGCATCTGAGAAATTCGGCCCCCACTGCGAGAAAGCCCCCGCAGACCAG
GAAGAGGGAGAACATCCTCGCCAATGAGCGTGCTTACCGCTGCGTGAGGGGAGACGTCCCCCAGGACTGGGAACGAGGACGGCACACAAACCAGCAGGGCTCACTCA
AGGCTGGCCGTGGTCTGCCCTGACATCCACCACATCTCCCTGCGGGTTCCAGGCCATGCAGCAGGCAAGAGAAAGGAAGTAACGTACACAGAGAGGTGAAAAGGAAG
ACGAGAGACGGACCCACTGACAGACATGGGCTGTCCACGCAGAAAACCCCAAAGAACCGACAGAGCTTTGAGATCCAGTAAGTGAGTTTCATGAAGGTGCAAAACAC
AAGGTCAATAAATTCTATATGCTACCAACGAACAACTGGAAACGGGAACCTTAAAAGCACCGCTGACAACAGGAAAACATCAGGAAACACTCAGAAGTCTCACAACA
TTCGCATCAAGCGTACACGGAAAGCTACAAAAATCCTGCCGAGAGGACACGGACGGCGCCCATGGACTGGGCCTCCTGCGGTTATGAAGGCGGCAGCTGGGCCCACAG
GGATCCATTCACTCAGCGAGATCTCACTCAAAAGTTCAGATTGTTTGTTTTTGTTTTGGAAATGGAGTTTTGCTCTCGTTGCCCGGGCTGGAGTGCAGTGGCGCGAT
CTTGGCTCACTGCAACCTGCACCTCCCGAGTAGCTGGGATTACAGACGTGTGCCACCATCCCTGGCTTTTTTTTTTTGGAGACGGAGTCTCGCTCTGTCACCCAGGC
TGGAGTGTGGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCCTCCCAGATTCACGCCATTCTCTTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCGCCAC
CACGCCTGGCTAATTTTTTTTTTGTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCTCCTCAGCCTCCCAA
AGTGCTGGGATGACAGGCGTGAGCCACCGCGCCCGGCCAGAAGGAGCCTATTCTTATGTGTTTTCATACAAAGGCAAAGAAACCAGAATAGCCAACGTAGCGCGGGA
CAGAAGCAGAACTGAGAAGCTCAGCACGGGATGGCAAAAGTGACCACGTCGCTCCAGCGTGCACCACGGGAGGCCGGTGGGGGCCACAGTCCACGTGGTCCCGGGAC
CAGCACTCACCACCACGTCAGAGTCCAACACACGTTCCCACAGGACGGCAGAGCAAGGACGGGCTGAGTCTAAGATTCACACACAGGCAGCCCAACACCAGAATGGC
CCCGACAGTTCCACACGGCGCCCAAGTCGGAGGCTTCTCACCACAGCCAGCTCTGCCGCGGGAACGGACAGACGCGTGCAGCAGTGGGACGGAGGAAGGTTCCGGAA
CGGAGCCCGCACACGGCACAGTCACGACGCCGAGGCAACGCTGCGGCCTGCAAAGCTGATGCTGGGACAGTCGGACACCCGGACACCCACGTACCACAAGGGGAACC
TCAATCCGAGCTTCACACCCCACAGGGAAGGTTACTCAAGAGGACCCCAGACCAAAACACAAAACATCGCAGAGGAAAATCTTCGTGGCTTTGGATTCAGAAATGAA
TTCTTAGATATGACACCAGAAACACCGCTGTAAAAATAACCAGTCAGTGTCTCCTCCTCCAAATTCAGAACATCTCTGAGACACCACGTGAAGAGGGGGTGAAAACA
TACGTCACAAACTGGGAAAAAGGCGAACACGTCGCAAATAAAGGGTTTGTATGCTGAAGACACGAAGAACCCTCACGCTCCGTGAGAAAGCAACCAGCCCAGGGCCT
CTCAGGGCCTCCAAGGCTGGAAGGCCGAAGCGGAGGCTGGAATGGCCACGGGCCCTGCAGCCTGCCCAGCAGGTGCCAGCCGTGGGATAAGGACAGACGTCGTTGGC
GAAACACGACTCGGCCTCCCCGTGCTGGGCCTGATGTGCGAAGCCTCGCGCCCAGGATCACAGAACCCCACGCGGGGAATGACGGGCAGACTGCACCGAGGCCCGGC
CCTGTCATCGGCTGTCCTTACCTCACCGGCCCCGTGTCAGGTCTCACCCCGCCCCTCTGTGTCCGCGCACCCCACCTGACTGACCCCGCATCCCCCGGCTGCACACA
CACCCGTCCTCCCACTCACCCGTCCTCCCCACTCACCCGTCCTCCCACTGACCCGTCCCCCCACTCACCCGTCCTCCCCACTCACCCGTCCTCCCAGGGCTCTCCCG
CAGTCTCCTCGGCCACCAGGATGTCGTACTCCTCGGCCGCGTACTTCTCCCAGTCCGAGAGCTCGGGGCCGCCGCTGTCACCGTCCTGGAGGAAGCACACGGGTTAC

GTACACGGCGTGGCGCGGCCCGGCCCGCGTGCCCGCCGCACTCACCCTGAGCAGGGAGATCTGCTCTTTCTGCTCGTGGTCGAGGTACTTCTCGATGTTGAAGAAGG
TGTCGAAGAAGACGTTAGCCAGCTTGCAGCGCTTCAGGTCCTGCAGCGTGATCTTCCCTGCGGGGAGGGGAGTGCGTCCAAGGCGCGTGAGCCCGGTCTCACCTTCG
GGGCCTGGGTGTGGGGTGCGCGCGTCCTGTCACACGTGCTTAAGGACGCGGCACGAAGCTCCGGGCTCTCACGCCACGTGTTTGACGTGAAAGGCGGACACGCGAGT
GTCCTGGTTCTGGGTGGGGACTGAGGCGTGGGAGGCCGTGATGCGGAGGCGACTAGGGGCCACCGGTGCCCCAAGGATGCGTGGACTGGGACTAGCGCAGGGAGGCG
CGGCCAGCAGGGGCGCTGCTGGATCTGAAATTACAGAGAACGGAGCTCCCCGTGCAGGCAGAGGCCAGAGGGTGTCCCGGGGACACTGAGCAGCAGGAGGCGCCCCA
TGGCCGGGGGGGGCAGGGCTGCAGGGCGGTGCCCGGGCCTCAGAGCCATTTCCAAAGTGAAAGTCCCAGAGCTGGATTCTGAGATGCAGAAGGAAGCCAGGCGGCGT
GGACGGTGGGGCTCCCCTTTGCCTTAGAGGGCACGGGGGTGGGTGGGAGGAGCTGGGGAGATGGTGTGGGGGGAACAGGGGTGAGGGGAGGAGCCGGGGGGGGGTGA
GAGGGGAGGAGCCGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGTGAGAGGGGAGGAG
CCGGGGGGGGATGAGAGGGGAGGAGCCGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGG
TGAGAGGGGAGGAGCCGGGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGTGAGAAGGG
AGGAGCCGGAGGGGGTGAGAGGGGAGGAGCCGGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGGGTGAGAGGGGAGGAGCTG
GGGGGGGTGAGAGGGGAGGAGCCGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGTGAGAGGGGAGGAGCCGGGGGGGTGAGAGGGGAGGAGATGGGGGGGGTGAGAGG
GGAGGAGATGGGGGGGGTGAGAGGGGAGGAGATGGGGGGGGTGAGAGGGGAGGAGATGGGGGGGTGAGAGGGGAGGAGATGGGGGGGGTGAGAGGGGAGGAGATGGG
GGGGGCGAGAGGGGGGGAGGAGCCGGGGGGGGTGAGAAGGGAGGAGACGGGGGGGGCGGGGTGAGAGGGGAGGAGACGGGGGGGGGGGGGGTGAGGGGAGGACCGCC
TTGTGCGGCATCAGCACGAACGTGGCAACTGCTCCCTCAGCCCAGTGCCAACGTCCCCGGTGAGCTCCGGCCCTAGGTTCTCTAGGTCACCCCTCGGGGGGTGTCGA
CGCCCCGAATAGGAGGTTCTTGCACTGTGAGACTAGGCACCTGTGGTGCTGGCCCTCCTCACAAACTCAGTAGGACAGCGCCTCACAGGGACGTCCCCTCACCCTGG
TCCGTCCCCCCTCCCGTCTGTCCCCTCACCCTGGGCCGTCCCCCCTCCCGTCCGTCCCCTCACCCTGGGCCGTCCCCCCTCCCGTCCGTCCCCTCACCCTGGGCCGT
CCCCCCTCCCGTCCGTCCCCTCACCCTGGGCCGTCCTCCCTCCCGTCCGTCCCCTCTCCCTGGGCCGTCCTCCCTCCCGTCCGTCCCCTCTCCCTGGGCTGTCATCC
GTCCCCTCTCCCTGGGCCGTCCTCTCGCCCGTCCGTCCCCTCACCCTGGGCCGTCCTCTCGCCCGTCCGTCCCCTCACCCTGGGCCATGCCCTCACGGGGCATCACC
TTCAGTCCTCGGCTTGACCAGGTCCAGCATCTGGCAGAGGCAGTCCTGGAAGGGCAGGGCCTCGATGGCCATGCTGTCCAGCCTTCGGCACTGCTCCTCGTAGAAGT
ACTCGAGCTCGAACATGGACAGGGCGCCGTCCCCGTCCAGGTCCATGCAGCGGAACCAGTACTCGATGCTGCGGCACGGCGAGCTCTGTCAGCCCCTGCCCTGGGCC
CTCCCAGCCCGTGACCTGCAGCCCCTGAGGCAGCCCCCACCGGGGGTGCACGCGTCCCCGCTGTGCCTTGCAGCCCCCACCGGGCGTGCAGGCATCCCCTGCCCCCT
GCCGCCCCCACCGGGCGCGCACGCGTCCCCCTGTGCCGTGCAGCCCCCACCAGGCGTGCACATGTCCCCCTGTGCCGTGCAGCCCCCACCGGGCGTGCACATGTCCC
CCTGTGCCGTGCAGCCCCCACCAGGCGTGCAGGCATCCGCCTGGGGACACATGTCACATGGGCGGCTCCCGGCCCCTCCACTGGGACAAACGCATGCCGCAGCAGGA
ACCCACCTGGTCGGTGTTTTTTTGTCTTCCTCAGAGATCAAAAACCAGACAAAGTCGGCATAGCTGATCTTCCCTTCCTTCTGCACTTTTCTGCCTCTAGATCGAAA
GCCAGGATGGAGAGACGAAGATGCATGTCAGGGAGAGCTTCACAGGAACGGAGCCCCTGTCCACGCGCCTCGGTGAGGGGAGCCCCCCGGGCCCGGCCCTCCTCCTG
CCCCCCTCCTGCCCCTCCTCCTGCCTCTCCGGGGAGGAGGTGGAGGCCCCGTGGCCAGAGGGTTTTCCCCAGATCCAGGCAGGGTCAGGAGTGCACCTTCGTTACTG
CTCACTCAGGCCCAGCGCCCGACAAGAACCCCCGACCTGGGGCCTGGGCCACCCCCTTCCTCAGACTTCGCGTGACAGTCTTGTGCCACCCCCCCCCACTAGGGATT
CACGTGACAGAGACACGTGCCCCCCTCGCCAGGGCCTGGGGTGACAACCACTCGCTGTCGGGGCACAAAAAGCTCACGTCAGGCAACGATGAGGAGAGGGACCGGGG
TCCTCGCAGGGGCAATGGCTGCCGTCAGGCGCCTGAGCCGTACGTACCGTGTGACTGCTCCTGAGAAGATCCTGTCTATCATCTTGGTAGAAAGGGCTGGAAAGGAA
TGCGGTTGATGGGCAGCCCGCACCGTGCCTCGGCCCCGACGTCACCACCCCCCGGAGCCGAGACTGGATGCGGTGGGGACCGAAAAGCTGAGAGGACGCCTGGGTCT
GGGAGAGCCCCGGGGCCCCGATGCCCCTGCACGGCCCATCCTAGGGGCCCACCACGCTTTCCCGTCGAGCAGAGCCAAGTCCAGCATGAAATCCACAGAGCGCAAAG
CTGACCGCGGCTCCAAGACCGACTTGTAAAGAGCAGAATATTCAGGCCTCAAAGGTACAGCTTTCAGACGGAGAGAGAGACCTCGAGTGTGATCACGGAAACAAACA
CGTTTCAACCAAAGGTTCACCAACGGGAGACGGGAGTGAGACCTCAGCAACGGGAGGCGGGAGTGAGACCTCAGCAACGGGAGGCGGGAGTGAGACCTCAGCAACGG
GAGGCGGGAGTGAGACCTCAGCAACGGGAGGCGGGAGTGAGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCAGCAACG
GGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAAC
GGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAA
CGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCAGCA
ACGGGAGGCGGGAGGGAGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCAGC
AACGGGAGGCGGGAGGGAGACCTCGCCAAGGAGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCAG
CAACGGGAGGCGGGAGTGAGACCTCAGCAACGGGAGGCGGGAGTGAGACCTCGCCAAGGAGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCG
CCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTC
GCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCT
CGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACC
TCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGTGAGAC
CTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCGCCAACGGGAGGCGGGAGTGAGA
CCTCGCCAACGGGAGGCGGGAGGGAGACCTCGCCAACGGGAGGCGGGAGTGAGACCTCAGCAACGGGAGGCGGGAGTGAGACCTCACCAAGGAGACGCGGGAGTGAG
ACCTCAGCAACGGGAGGGGGGGAGGGAGACCTCACCAAGGAGACGCGGGAGTGAGACCTCAGCAACGGGAGGCGGTAGGGAGACCTCACCAAGGAGACGCGGGAGTG
AGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCACCAAGGAGAGGCGGGAGGGAGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCAGCAACGGGAGGCGGGAGG
GAGACCTCAGCAACGGGAGGCGGGAGGGAGACGTCGCCAAGGAGAGGCGGGAGGGAGACGTCGCCAACGGGAGGCGGGAGGGAGACGTCGCCAACGGGAGGCGGGAG
GGAGACCTCACCAACGGGAGGCGGGAGTGAGACCTCACCAACGGGAGGCGGGAGGGAGACCTCAGCAACGGGAGGCGGGAGGGAGACCTCACCAACGGGAGGCGGGA
GTGAGACCTCAGCAACGGGAGGCGGGATTGAGACCTCACCAACGGGAGGCCGGAGTGAGACCTCACCAAGGAGAGGCGGGAGTGAGACCTCACCAACGGGAGGCCGG
AGTGAGACCTCACCAACGGGAGGCGGGAGGGAGACCTCACCAACGGGAGGCAGGAGTGAAAGCACCGTCGCCGTCAGCTTGGGCCACGAGAAGGTCCCGCAGCCTGG
GCGGCCATCCCTGCGGTCACCGGTGTCCCTGGGACGCACGAGCCAAGGTGCCGCCCCCCGCTTCAGGCCGCAGTGCGTGAGAAACAGCGCAGCCCGGCCGCACACGG
CATCCTGCCCTGGGACCGAGAGTGGGCTCCAGAGGAACGCGGAAAGCTGGGGCCGCGCCGGGCCGGGGGTTCACAGGGACAGTGAGCGCCTGGTTCTGGGCCGATTC
CCAGCTGTGCAGGTGCTGGCCCAGGCCACACGGGCCGGCGCTGTGGAGCTGTGCTCTCCTCGTGGCTAGTTAGGGGCCCAGGGCGCAGGTGGCTCCCTGAACTGAGC
GCAGGACAAATCCAGGGAATGCTACGTCTGCTCAGGTGCGCGTCCGACGCAGGGATCTACTGCGGCCACGCGACCCAGGACCAGCCCCGGGAGACGCAGACGGGTGC
AGCTCACGTTCGGGGAACCAGGCGTATTATGAGTCGACCACACACAACAGATGCCACCGTAATTCGGAAACAGTCTTGGTCATATGGTTTTTTGGGGTAAAAAGCAC
ACTTAAAGGAATCCTTTAGAAAACTCAAGCCCCACAGGCCACAGGCCGCTTTGAGTTCCTGCAAGTGTGGATGACTCTTAGAGGGGCCGGGAGGCAGTGGCTGCTGT
GAGGACACCTGGGCGGCACCGGCTCCCGCGGAGGTATATGAACGACCGCTAAGCCTGGGGCTCCACCTAGCCCGGGATTGGATGGAAACTGCTCCAGGCCTTGGGGG
CTGGAACCTGGAGCTCCTGAGGGAAGGCGTGTGGCCTGCCCGCCCTTGCTCGGGTGTTAACAAGGAAGCCCCACAGCGCTGCTGGCCCTCGGGCCAGGAGCTTCAGG
ACCAGCGGCCCACACTGACCCTGTAGACGCCCCCAGGCAGAGGCCTCGGGCAGAGGCGCACGCGGGGACCCAGACACGGGGCAGCGACTTGGGGAAGGAGAGGCAGCT
GCAGACACAGAGCGGCGAGTGCGAAGGAGAGGCAGCTGCAGACACAGAGCTGGGAGTGCGGAAGGAGAGGCAGCTGCAGACCCGCAGGCCCTGAGAGAAGGGTGTGC
TCGGTGTCCGCGCGGCCCGCCCGCCCCTGTGCCCCCACGCACCGTGGTCATTGTGCCGCGCCAGGTCGTCCGCGTCGATGAGCAGGTCGTGGTCCGTGTCCAGCTCC
CAGAACTTGCAGTAGATGACGTAGAAATGCTCGTACGAGAAGAATTCGGTCAGCTGGTTGATGTCCGCCTCCTCCTCCAGCAGCGCCACATTCTGCCAAAGGACCCA
GGCGGCCTGAGCGCGGGGCCTCTGCGGGGATGCCCCAAGTCCGCCTGGCTGCGGGCGGGGCAGGGAAGGCTGGGAGGGTCGGGGCCGCTCCGTGGGTGGGGGGGACT
GGGCAGCTGGGGCCGGGGGGCACTCCGAGTGGGGGCAACTCTGCGGTGGGGGGAGCTCCAGCATGGGGGTGGCTCCCTAGGCGGGTGTGGGGGTGGGGGCAGCTCTG
GGGGAGTCAGGGCTCGGCTCCCCCTCAGGGTGGTGCCCAGTACGGCCCCACCCGCCTCCTAGATCAACCGTGGACCCCCAACCGCAGGCGGCGGGGTTCCAGTGCTC
ACAGGCCTGGGGGTGGGCTGGGAGCGCGGTGGAGGTGGGGGTGGGGGTGGGGGTGGGAGAGGGGGTGGGAGGGGAGGAGGGAGGGGGGAGGAGGGAAGGGAAGGGAG
TGGAGGTAGGAGGGGTAGGGACAAGGCAGGGGGCAGGGGACAGGGGGCCGCTCCGCACCTGCAGGAAGGAGCTCCTCCGCAGCTCGGCGCAGGTGATCCTGCCGGAC
CAGGACCGGTTCACGGCGTAGAAGATCCGCTGGATGACCTGCGGGGGCGCTGTCAGTGCGGTGGGTGCGCAGAGACCCCCAGGAGCCTCGCCCCGCACGGAGACCGG
GAGCCGGGAGAGGGGCGCGCCCTTGGTCTCAGCCGCACGGGGCCGCCAGGGCACAGGCGGGGGCAGAGGGAAGGGCCCTGCGGGAGGCGCCGCCCCAGGCCGCGGAA
AGCGGGGAGGGTCTGGCCGGGGACGCCCCGGAGCTACACGGGCCCAGGACAGGTGGGAAGGGGGTGCGGCCACCCCGGAAAACCAGAGTCCCCCCAACACCGGGCTC
CCGGGCGGGTGGAGACTCTCCCAGAGCGCGGCCGCCTCCTCCGGAAGCTCAGGAACCCCGGGCCCCGCCCGCCCCGTCCGCAGAAGCCCCGCGGCGGCCGCTGCAGA
AACACCCGCGCCCCGCGCTGGAACCGACGGCCCCTCCGCTGGGGACCCACCGTGGTGATGTAGCGCGAGTGGAACTCGGACGCCTCCTTCAGGAACGACAGCCCCGG
GTGCGTGTTCACCACGTCCTGCGGGTGGGAAGACACGAGGCGCGTGGTGTAGACGCCGGCCCTCCCGTGAGGTGTGCGGTGTGGACGCTGCAGACGCGGACCCTCCC
GTGAGCGATGAGGTGTGCGGTGTAGACGCCGGCCCTCCCGTGAGGGATGAGGCGTATGGTGTAGACGCCGGCCCTCCCGTGAGGGATGAGGCATGTGGTGTAGACGC
GGGCCCTCCCGTGAGGTGTGCAGTGTAGACGCGGACCCTCCCGTGAGCGATGAGGTGTGCGGTGTAGACGCGGACCCTCCCATGAGGGATGAGGCATGTGGTGTAGA
CGCGGGCCCTCCCATGAGGTGTGCGGTGTAGACGCGGACCCTCCCGTGAGGGATGAGGCGTGTGGTGTAGACGCCGGCCCTCCCATGAGGCATGCGGTGTAGACGCG
GGCCCTCCCGTGAAGGATAAGGCCTGTGGTGTAGACGCAGGCTCTCACCTGCAAGAAGGGGACAAAGTCCTCCTGCACCAGGTAGTTGCAGCCGGGGCTCATGAGCA
GATGGACGAACTTGGCCGCGTCGTCGTGGCAGTTCTGGAGGATTCTGGAAGGACAGGATGACTGGGCACCACCCTCACAGGGGGGTGGCTTTCGACCCCGCAGACGC
AGGGTGGAACACGTGTGTGGTGTTCCACGTGGTGCGTGGCAGGTGGCCACACACGACGGCCAGGCCTGTGCCCGTACAAGGGTTTCTCCTCTCACTGCGGCAGCCTC
AGGGGGCACCCGTCCTGGCACCACGGGGACCCGGGGACGCCTCCGCCCTCCCGACACAGCCCCCTGCCCAGCCCAGGACTCACTTTCTCCACATGGCGACGAACTTG
TGGACGGACACGGAGCCCGTGCGCTCCCCGCCGGCGCCATAGAAGAGCGGCCCCTTCCAGTAGAGGGGGCAGCCGCAGGCCTGGGGCAGAGAGGGCAGGAGTGGGCA
GTCAGCAGGGCCTGGACGCCGGCGCTCCTGCTGCGTACCTCGCGCCTGACCGACGCCGCCCAGAGACCCTCTGCTGCAGAAAGACACAGCACGCTCAGCGCGGCCTG
TCTGGGCATCTGCAAACTCAACGTGGCTTTCGCTCCAGCCTTCAATCAGTCATGAGAGCTGAGCGTGGGAGTGCCGGGCACGCAGTATCCCCTGACCACAGACCACG
CGTGGGCACTGTAACCAGAGGCGGCTGCATCCCCAGCGTCTGGAAATCAAGCGGCACGTACTAAGAACCAGTAAGTACGGGAATAAATGGCCCCGTCAGCCAGGCAA
CGCGGTGAATCGGACGGCGACAGAAACATCTCATGGAATGTGTGGGACGCACCTACAGCAGCGACCCATGGCAGGTTCACAGCCTTAAATACACGCGTGAGGAAGGA

AAGTTTAAAATCGATCGTCTACACTTCTACCCCAAAGGAACAAATACATGGCAGAGATGGAACGCAAAGTAAATAGTAACAGAAAATCAACAAAATCAAAACTCGAT
GATTTGAAAGATTTCTTAAAAATTAATTAACCCTGGCTGGCTGGGCGTGGTGGCTCATGCCTGTAATCCCAGCACTTTAGGAGGCCGAGGGGGGTGGATCACCTGAG
GTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCGTCTCTACTAAAAACACAAAAATTAGCCTGGCATGGTGGCGGGCGCCTGTAATCCCAGCTACTCG
GGAGGCTGAGGCAGGAGAATCACTTGATCCTGGGAGGTGGAGGTTGCAGTAAGCCACGATCACACCACTGCACTCCAGCCTGGGTGACAGAGAGAGACTCTGTCTCA
AAAAAAGAGCAAATACAGATTAGCCATGTCAGCTGTAACTACAGATCCTACAAACATCAGATAGTGGATGCTCATGAATCACTTTATGCCTGTAAGTCAGATGACTG
CTGAAACAGACTCCCCGATGAACAAGTCAAAAAATAAACTACAATAGAAAATCTGAAAGAGCTGTTTCTGCAAGAAATTGTGTCCATAATTTAAAACTGTCCCACCA
AGAAAATTCCAGCCCCCAAAGGCTGCTCCCATGAGTCGACCAAACATTGAAGGAAGAAATCACACATGCGTTCCCCTGGAGAAGAAAAACGAGGGGCGTTTCCCAGT
GCCTGTTCCGAGGCTGGCACAGCCTGGACGCGGTCAGGGACACAGGAAGTCACAGGCCAGTTATCACTGGAGAGCAGAGGCAGAAATCCCACACAGAACACACAGTA
GCAAGTCCAACTCGGCAATAATAGAACAGGAATATCTAATAATAATAATAGACCAGGAAACCAGAATCAGGAGGGACACTCTTATTCTCATAAAGGGTGCTATGGCC
TGAATGCCTGGGTCCCCAAGATCCACGTGCTGAAATCCTCTCCCCCGAGGCGATGGTGTTGGGAGGTGGGGCCTCATGAATGCGATGAGTCCCCTTCTGAGGGACCC
CAGAGAGCTCCCTCACCCCTTCCACCCCGTGAGGACGCAGCGGGAAGGCGCCGTCGATGAACCAGGCACCACTCTGCCACACCGAGATCTTCCAGCCTGCAGACCTG
TGAACAGTAAATGTCTTGCTGGTGACAAGCTGCCCGGGCTATGGTGTTTCTTGACAGCAGCCTGAATGCTCAGACAAAGGGCCTCTACGAAACACCCCAGCTGCTCA
GACAGCGATGACACATGAACACTTCTTCCCTGTGCTGGGAACAAGAGGACAGTGTCACCACCAAGGCCGCCTGGACGGCCGAGCCAGGGCAGGCAAGGGAAAGA
CCGAAAGATACGACGACGGGTAAGAAGAGGCAGAACCGTCATGATTCACAAGCAACACACTGTGTAGACAATCCAAATGCTTACCAAATACGGGAATTAATAAGTAT
CTTCAGCAAAGTTGCTGGATCCAAGGACTATATGCAAAAATAAATTATTTCAATATACTGGCAACAAACAACTAAAAATAAAATCTTCAATAATGCCATCTATCAAG
AGCATGAAAAATCAACCTACGAAAAATGGAACAAAAATCACACACAGTGTCTACAAAGCAAACTGCAGGGCACGGCTGGGTGAAATCAAAGACGACCTCAACGTGGT
GAGGTGGACCTCACTGCTGAGTTCACAGACTGCAGTTTCCTCCAGACGACTCTGTTGACCCAGAGTCCACGCGATCCCCACCCACAGAGTCCAGCAATCCCCATCCA
CATGCACCGAGGGCCTGTTGTTTTGGGGTTTGTAGTATACACGGACAGGCTCATGTGAAAACTTACCAGGGAACACAAAAGAACTCACAGGGGGACCAAGTTGGTGG
ATATGAGTCCCACTGCCCGGTTTCAACCCTTCCTGTAAAGTCAGAGTAATGAAGACTCAGTCAGTCTCTGTCGGGCACCGGTGGCCTAAGAGAGCAAGGGAAGGAAG
TGAGTTGGGACAGAGCCCCGTGGAATATCCACAGGCAAAGACACTTCACACAAACTGAATACCAGGGGATCCCAGCCCGAAGGGTGAAAGGCCAAACAGCGGCCCCT
CCTGGAAGGAAACGAGGGACGCTGCACGGCTCTGCCGCGGCCAAGGTTTCTTTCACGGGCCACACGTGCGCTCCCTGAAGGAAGGGAGCACCGCTGCACCGCCCCAG
CGCAGGACTCAGGGTCAGCTACCGGCACGGTGAACACCATGACAGGACATGACACAGAGGGAGAAATGTCTCTGCAAGGCGAGATTCTGACGAGGGACTCAGAGGCA
GAATTTTAAGCATCAGTGAGAAAGAGGCCCGAACACGACCGAGTGGACATGCACTGCCCTCCGGAGAGGGGTGCACGGTGGCCGTAGACAAGGGGAAGCGCCCAGCA
CTGGCGCCGGCGAGGGCAGGGTTCACAGGGCATGGAGAGGTGGCAACAAGAACATGACACGCGCGGCTGAGACAAAAAGGACTGAAGACGCCAGGTCCCGCGAGGCT
GCCAGGCAGCTGGGTGTCCTCCGCGTTGCTGGGGGAACTGGGACGCCACGCAGCTGCCGGGAAAGCGGCCCACGGGGCGGCTGAGCGTAAAGCCCCGTGGCCCGGCC
AGGACATCTGTCCTCCCAGAAGCACACCAGTGGGAAGAGGGCCCGTGCCCGTGAGAGCCTGCACAGATGGGGACGTCGGCTGCTTGGTTCACGGTGCCCCAGGCGAG
AGGAGCCGGACGCCAGAGCAGGGACTGGAGGAATAAACTGTGGCAGAGTCGTCCTGACAGGTGCAGGGCGGGAGTGACCACGGGGGGGCGTGGGGGACTGCAGAGGG
GGCTGCGCGCCACAGCGGGGATTGCAGTGCTGCTATGGGAAAGGACGCGCTCCGGCTGGCGCACCTGTGGATGACTTGAAGGCAGCCGTGTGCGGCTGGAGGGGCCG
GGCCCCGCAGGACGCACAGGCACGTGCCGGCTCCCGTGATGAGGATCCTGACGGAGCTCAGGCCTGGCCACCTGGAAGGGCGTCCCGGTGCGGTGGGAGGGAACGAG
CCGGGTGCAGCACCGGGTCCTTGGGCTAGTGCTAGTCCCACGCTTTTGACCTCATGAGACCATTCTGGCTCACAGGCCGTGCAGGTTTCAAACCCCAAGCGCTCAAC
ACCCAGGGATGGGGGGATGTGTGAGCCTTCACATCTGTGCTCCTCACCGCGTGCAAGACACACCGTCATGACCCGGTGGCAGGGATGGAGCATCAGATCTATCCAGC
CGTGCAGGATATAGCGTCTCATGTGTCCAGACCTGCAGGCTGGAGTGCAGTGTGTGATCACAGCGTGATCCATTTTTATTCTACTTGTTTTTCGAGACACGGTCTTG
CTGTTGCCCAGGCTGGAGTGCAGTGTGTGATCACAGCGTGATCCATTTTTATTCTACTTGTTTTTCGAGACACGGTCTTGCTGTTGCCCAGGCTGGAGTGCAGTGTG
TGATCACAGCGTGATCCATTTTTATTCTACTTGTTTTTCGAGACACGGTCTTGCTGTTGCCCAGGCTGGAGTGCAGTGTGTGATCACAGCTCAGTGCAGTCTCAACT
TCACGGGCTCAAGCAATCCTCCCACCTTAGCCTCCCAAGTAGCTGGGACCACAGGTGCGCGCTATCATGCCTGGGTAATTTTTTTAATTTCTACAGTGAAGTCTCAC
TATGTTGCCGAGGCTGGTCTGGAACCCTTGGCCGCAAGCCATCCTCCCGCCTTGGCCTCTCAAAGTGCTGGGACTACAGGGATGACACTGCAACCAGAGACAGCTGC
ACCCCCGGCGTCTGGGGGATGAGAACTGAGCGGGAGAGTGCTGGGCACACAGTCCCCCCACCACACCAAACCCGAAAATTACCTCCAAATGAATTAGCTCATAATTT
AAAATAATCCTAAAGTAAATGGAGATAATCAACTCTAAAATGCAAAAGATACAAAGGAAAGCAGTTTTACCTGAATGTCACAAAATCCTTATGGGAAAGGGAATAAA
ACATCAGATACCCTGATCTGACTTTTAAAAGTCTTTTCGTGATCGACACACAGAAAGACCGACTCCTTGTGGGGCCCGGCAACGCGGCGGCACGCCCACCCGCACGG
CCAGGACAGGGAACGAGGCCGCCAGCCCTGACAGTCCCTTCGGCCCCTCCGCAGTCCGCTACGGCGACGGCCCCCAGGCCTCCTCCCCGTGGCTCCAGTTCACCTTT
TCCAGGACACGGCAACGGGACCCGGCCATCTGCGGACTTTTAGACCTGGCTTCCCTCGCTCAGCACATGCGTGTGGGACTCACACACCTTGCTCTGTGTGGATCGTC
TGCCCACCTTGCTCTGTGTGGATCGTCTGCCTGTGTACTGCTGAACCCTATTCCATTATGGCTCTGCTGACCTCAGAGTAGCTCCCTGAGCAGGAAACACCGCCGGC
CCCAAGGCCCACGTGATGGTGCTGCTGAGGGGGACTCGCGAGGCAGACAAAGTCCCGAGCGTCTACAGTCGCACAGGGCACGCCGACACCAACCGACAGGTGTGAGA
GGAGAAAGGAGGACGCCACAACCACGGCCGGCAAGTTCAGCACGCGTTCCTCGGCGGGTGACGGGACAAGAGGACGATCTGCAGACACGTAAAGATGTGAAAACGCT
GCCAAGCTATGTGAGCGAGTGACATCGCGGAGCCGGAGGAGTCTTCAGGGTCACATGAACCAGTGGTGGGCTGTGGCTCGGAGCCCCCACCCAAGTCTCACGGGGAA
CTGTGATCCTGAGTGTTGGAGCAGGGCTCTGCTGGGGGGTGACTGGGTGATGAGGGTGGACAGCCCCCTTGCTGGTCTCGTAATCCTGAGTTCTTGTGAGATCCGAT
GGTCTAAAAGTGGGTGGCAGTGGCCGGGTGCGGTGGCTCAGGCCTGTCATCCCAGCACTTTGAGACGCCAAGGCGGGTGGATCACCCGAGGTCAAGAGTTCGAGACC
AGCATGGCCAACATGGTGAAACCCCATCTCTCCTAAAAATACACAAATTAGCCGGGTGTGGTGGCGCATGCCTGTAATCCCAGCTACCCAGGAGGCTGAGGCAGGAG
AATGGCTTGAACCCAGGAGGCGGAGGTTGCAGTGAGGTGAGATGGTGCCATTGCACTCCAGCTTGGGTGACAGAGCAAGACTCTGTCTCCAAAAAAAAAAGAGAAAA
GATGATACAGACAGCCTGTGTTTATTGAAACAATTCATGCTGTCATTAAAAACCTTCCACAGAGAAAACTATCCCAAATGGCTTCACTGATGAGTTCTACCAAACA
TTCAAATAAAAATAACACCCATTCTAAACAAATGGCTCCAGAAAAGTGAATTCTAAAGGACATTTTCCAAGCAACTCTCCGAGGCCAGCATCCCCCCGATGCTAAAC
CAGACCAAGCTATCACAAGAAAAGGAGACCGCAGGCCAGTCCTTCACGAGCGCAGACGCAACCCCTGCGAGGCCTGGCAATGTGGACCCAACACATAGGAGAAGTGT
GACACTGTGTGGTGGACCAGGGACCGGGGGCTGGGGGCTCACCCAAAGACCGGGGGCTCACCCAAAGACCGGGGCTCACCCAAAGACTGGGGCTCACCCAAGGAGCA
GGGGCTCGCCCAAAGACCGGGGCTCGCCCAGGGACCGGGGGCTCACCCAAAGACCAGGGCTCACCCAAAGACCGGGGGCTCACCCAAAGACTGGGGCTCGCCCAGGG
ACCGGGGGCTCACCCAGGGAGCAGGGGCTCACCCAAAGACCAGGGCTCACCCAAAGACCAGGGCTCATCCAAAGACCGGGGCTCACCCAGGGAGCAGGGGCTCGCCC
AGGGACCAGGGGCTCACCCAAAGACCGGGGCTCGCCCAAAGACCGGGGCTCACCCAGGGACCAGGGGCTCGCCCAAACACAGGGGGCTCACCCAAAGACTGGGGGCT
CACCCAAACACCGGGGGCTCACCCAGGGACTAGGGGCTCACCCAAACACCGGGGGCTCACCCAGGGACCGGGGGCTCACCCAGGGACCGGGGGCTCACCCAAACACT
GGGGGCTCACCCAAACACCAGGGGCTCACCCAAACACCGGGGCTGCCTTCACACTCAGACGTCAGTCAGGTCAAGACCAGTCCTTCATGACCACAGACGCTACCCCT
GCGTGGCCGGGCAATGCACGCTGCGGTGGAACAGGGACCGGGGGCTCACCCAAGGACCGGGGCTGCCTTCACACTCAGACGTCAGTCAGGTCCATTCACCTTTTACA
GGACCAAAGGAAAACCTCACTGCCCTCTCAATACACGCAGAAATACACATTTAGCATCCCGGACAAACTTCAAGGCACTTTTATAATTTTAAAAAGTAAACTGAGGC
CAGGCGTGATGGCTCATGCCTGTCATCATCCCAGCACCATGAGAGGCTGAGGCAAGAGGATCCTTCCAGCCCAGGAGTTCCAGACCAGCCTGGGCAACATAGCAGAA
CCCTGTCTCTACCAAAAAAGAAAAAATCAGCCAGGTGTGGTGGCGGGTGCCAGTAGTCACGGCTACTCGGGAGGGCTGAGGTGGGTGGATCACTGGAGCCCAGGAGG
TCGAGGCTGCAGTGAGCTGAGATCATGCCACTGCACTCCAGCCCGGGGGCAGAATGAGAACCTGTCTCAAAAAAAACACAAGAGCTCCCAGCAAACAAGAAACCAAA
AGGGAAGGGGCCTCCTGAGTCTGATGAAGGACGCCAGGCAGGGCCGCTGCGTCACGGACGACTTGAGCCCGAGGCGGGAAGGGGCGTGCTCTCCACGGCTCTCCAGC
ACTGCACTGGCAGGCCGAGCCTGTGTACAGAAAGTCAGGGAGGAGACACCGATGCAGACGAGAAGAGAAAGGCTGAACTGTGCTCACCTGAAGATGACCTGACCGTG
CGTTAGAGACTCTGCCGGACTTAAAGGTTGAGCTGGTAGAGCCGATGACTGAGTTAAACTTGGTCAGAAAATGCAAAGTCCACATATAAAACCTAACAGTGCTTCTC
AACATTCACAATTGACAATCATATTAAAGTAAAAAAATACCACGTACAGTGTCATCAGAAACAGGAAATGCTTGGGTATCAATGTGATGAACCGTGAACAAGACCAA
ACTACAAGGCGTACTGAAAGAAGTGAAAGATGCCTGAAATAAGGGAGAGCTGGACCGCGTTCATGAGACGGAAGCCTCAATATGGTCAAGATTCTTCCAAGTTCATC
TGGAGTCAACAGGATGCCCATGTAATCCCAGCAGGCTGTCGTAGCAACGGAGGGACCGGCTCTAAAATTCACACGGTGGGAACGAACAACGGTGCAGCCACTCGGGA
AGACAGCGTGCAGGTTCCCCAGAAAGCCGGAAGTGGAACTGCACGCGCACACGGAACACCACGCGATCCGGCCCTCCACACACGCACATGGAACACCGTGCTGTATA
TGCCACAAACGTCTCCGAGTTCTGTGCTCCAAATGGTTACGGTGGTGAATTCTGTTGTGTGAATTTTCCCATAATAAAAAAAAAGTGCCTGCTGTCTGATTCTGTTTA
CATAAAATTCCAGAAAATTCCGGCTAATCTATATTGACAGAAAGGAGGCCTGAGGTAGCGTGGGATGAGGGCAGGTGCGCGTCAGAAAGCAGCAGAGAAAAGTCCAG
CGCACGTGCCCGCTCAAGCGCAAGGATGCAGGGACGCCTATTGTCTTGGTGCTGTGGATAGTTTCATGGGTGGGTACAGACGTCAAAGCTCATTGAATTGTACCTTC
CACACGCACCATGCATCCCACATCAGTTATATCTCAACACAGCTGCCCCAAAATGAAAAAAAGGCAAACTGATGAGAAAGAAACACGTAAGGTGGTCAACACGTACA
TCAGGGAAATGAGGAGTGCAGCCAGGACGGGACGCCTGCACACGCACAGAACGGCTAAAGGCAAAGCGTGTGGCCACGTGCCCGCAGGGGTCAACGCTGCTGCCTGG
AAATGCGTCACTCAGGGCCACCCCGGAAGCAGGGTTTTTTGTTTTTTGTTTTTTTGCGTAGAGACACAGTCTTGCTCAGTCGCCCAGGCTGGAGTGCAGTGGTGCGA
TCTTGGCTCACTGCAACCTCCACCTCCCAGGTTCAAGCGATTTTCCTGTCTCAGCCTCCCGAGTAGCTGGAATGACAGGTGCCCGCCACCATACCTGGCTAATTTTT
GTATTTTTAGTAGAGATGTGGTTTCATCATGTTGGCCAGGCTGGCCTCGAACTCCTGACCTCAGGTGATCTGCCTACCTCAGCCTCCCGAAGTGCTGGGATTAGTGG
AGTGAGCCACCACTCCTGGGCCCTGGAAGCAGTTTTATAGCTTCTTACAAAACTAAACATGTGGCCACCAGATGAGGCAGAGCTCACCTCTGCGGTATTAATCACCT
CAGGAAAATGCAAGCATGTCTGCGTAAATCAGTGCACAGACACACAGCAGCTTTACTCACAAGAAGCAAAAACCAAAGCAACCAGACGTCTATCAACCCTGGAACAG
GGGAGAGATGGAGCACACACACAGCAGAAGCTACTCAGTGATAGAAAAGAAACTGCCCATACGGCCAGGGACACATAGTGAGCCCCACAGTATCCACACCCTGGCCA
GCCACACGGGAGCCCCACGGTAACCACGCCTTGGCCAGCCACACAGCGAGCCCCACAGTATCCACACCCCGCCAGCCACACAGCGAGCCCCACGGTAACCACGCCTT
CGCCAGCCACACAGCGAGCCCCACAGTATCCACACCACAGGACTTAATCAGTCCTCAACACACACACATAGCGAAACATCTTGCTGTACCCTGTATACACAAATATT
ACTTGTCCATTCAAAATAAA

ÚKOL 5
Analyzujte část genomu mnohobuněčného eukaryotického organismu také pomocí programu **GENSCAN** (http://genes.mit.edu/GENSCAN.html). Porovnejte s výsledky z úkolu 4.