

BIOINFORMATIKA V PRAXI – ZÁVĚREČNÝ TEST JARO 2011

Jméno a příjmení:

Datum:

UČO:

Varianta: A B C D

Studijní obor:

MATERIÁLY A POMŮCKY

Při vypracování testu jsou povoleny jakékoliv vlastní materiály a pomůcky, zejména dříve vypracované protokoly a internetové stránky. Zakázána je komunikace s dalšími osobami (osobní, telefonem, elektronická,...), vyjma zkoušejících a dozoru při testu.

OTÁZKA 1

Analyzujte část genomu *Listeria monocytogenes*. Identifikujte potenciální geny (uved'te čtecí rámec a délku potenciálního genu) a určete, které proteiny mohou být těmito geny kódovány.

Sekvence

```
TGCCAGGAATGAACGCTAAAAGATGGGTTGCACTAGGGATTGTTTTTGCACCTTTTGATAGTTAGTGCCTTAGCAA
AGTTTACTTCAAGTCAAATCGCTTCAACGGAAGAGAGTAGTCCGACATTTGTCGAAAGTCTATTCGCGGATACCG
GCGAACTAACCAGAAACAGTAATTGAAGAAGGTGGGGATGATACAATTGCTGTATTATCTGTAGATGGTACTATTTC
AAGATACGGGAGATTCTGGTTCTTTACTTGGAGGCGGGGGCTATGATCATTTCATTTTTTCATGCAACAATTAGAAC
AAGTGAGAAATGATGACTATATCCAAGGTGTGCTGCTGTATGTTAATTCACCAGGTGGCGGTGTTATGGAGTCCG
CGCAAATTCGTGATAAAATTCTACAAATTCAAAAAGAACGCAATATTCCATTCTATGTTTCCATGGGGAGTATGG
CAGCATCAGGTGGATATTATATTTCCGCACCAGCCGACAAAATTTTTGCAAGTAAAGAAAACATTCAGTGGTTCAC
TTGGGGTTATTATGCAAGGTATGACTATAGTGAACCTCATGAAAAAATTAGGTGTTTCAGATAATACCATTAAAA
GCGGCGAATATAAAGACATTATGAGCGGGACAAGACCAATGACGGAAGACGAGAAGAAAATATGCAGTCAATGA
TTGATGATTTTACAATGAGTTCGTTAAAGTCGTTGCAAAAAGGCCGTGGAATGTCCGCAGAAAAAGTACGTAAAA
TTGCAGATGGACGCATATACGACGGCCCAAGCAAAAAGAGAACGGGTTGATTGATGAGTTTGGTTATCAAGAAG
ATGCCCTAGAAGCTTTGAAAAAAGAACAGGGGTTAGCTGATGCGACAGTCATTCAATATGATGCACCAGAAGACT
TTAGCTCGTTATTCTCTGTAGCTGCACAAAAAATCTCTGGTCAAAAATGCAGATATTTCCCAACTAATCAAATTA
CTGGCACACTTAAAGCGCCAAGAATGATGTATTTATATGGAGAATAATTTGCGACGTGCATGCGTC
```

OTÁZKA 2

Proved'te multiple alignment následujících sekvencí a přiložte jeho výstup. Určete 2 navzájem nejpodobnější sekvence.

Sekvence 1

```
MESPVRLSLLYVVLASLLLPGRSVFINRERANNVLQRIIRANSFFEEIKKGNLERECVVEIICSFEEAREVFEDNE
KTTEFWNKYEDGDQCESSPCQNQGECDGLGYSYTCCTEGFEGKNCELVVRKLCSLDNGDCDQFCREEQNSVVC
CAKGYFLGNDGKSLSTAPFPCKTKNKGRAKRSVALNTSNSEPDPELMPDADILYPTESPSELLNLNKTEPEAN
SDDVIRIVGGQECKRGCEPWQALLFSDEETDGFCCGTTILNEFYILTAACHLHQAKRFKVRVGDNLNTEQEDGEMV
HEVDMI IKHNKFQRDITYDFDIAMLRKTPITFRENVAACLPPQKDWAEATLMTQKTGIVSGFGRTHEKGRQSKVL
KMMEVPYVDRNTCRLSTSFSITQNMFCAGYDAKQEDACQGDSSGGPHVTRFKDITYFVTGIVSWGEGCARKGKYGIY
TKVTAFLKWIDRSMKARVGPTSETPRLTHPPY
```

Sekvence 2

```
MAHVRGLQLPGCLALAAALCSLVHSQHVFLLAPQQARSLQLQVRRANTFLEEVRKGNLERECVEETCSYEEAFEALE
SSTATDVFVAKYTACETARTPRDKLAACLEGNAEGLGTNYRGHVNI TRSGIECQLWRSRYPHKPEINSTTHPGA
DLQENFCRNPDSSTTGWCYTTDPTVRRQECISIPVCGQDQVTVAMTPRSEGSSVNLSPPLEQCVPDRGQQYQGR
AVTTHGLPCLAWASAQAKALSKHQDFNSAVQLVENFCRNPDGDEEGVWCYVAGKPGDFGYCDLNYCEEAVEEETG
DGLDEDSRAIEGRTATSEYQTFNPRFTFGSGEADCGLRPLFEKKSLEDKTERELLESYIDGRIVEGSDAEIGMS
PWQVMLFRKSPQELLCGASLISDRWVLTAAHCLLYPPWDKNFTENDLLVRIKHSRTRYERNIEKISMLEKIYIH
```

PRYNWRENLDRIALMKLKKPVAFSDYIHPVCLPDRETAASLLQAGYKGRVTGWGNLKETWTANVGKGQPSVLQV
VNLPIVERPVCKDSTRIRITDNMFCAGKLA

Sekvence 3

IVGGKDCPKGECPWQVLLLVNGAQLCGGTLLINTIWWVSSAAHCFDKIKNWRNLIAVLGEHDLSEHDGDEQSRVAQ
VIIIPSTYVPGTTNHDIALLRHLHQPVVLTDHVVPLCLPERTFSERTLAFVRFSLVSGWGQLDRGATALVLQVLNV
PRLMTQDCLQQSRKVGDSPNITEYMFCAQYSDGSKDSCGKSDSGGPHATHYRGTWYLTGIVSWGQCATVGHFGVY
TRVSQYIEWLQKLMRSEPRPGVLLRAPFP

Sekvence 4

MTASSTTLAQAAEALAEHRLEEAHGLVRAALAAHPHDAEAWRLGCVARAGGKAEAEQAFFRAIQALAPRHALAH
ADLCGLSLDLDRAGEAIALLDRAAASHNQPAQNPPVWTLSLKAAATWMAERRPHDALPALEALVRQAPHAPVPWIN
LAEALQALGHLDRAVGAIRHALAIDPYCAPAWLGLANLRVIRLEPADVAVIKTALGRATSDLARVQLGYALGKAL
GDQAAYEESFRHFERANALRGALTPHDPQALDHFAQAMERIAAPAPSRESGRDGPFIIVGMPRSGSTLVEQILACH
PQVEALGELFELQATAKRIESAPEALPAIISRLTAEESARFGDHYLRSIQRYRRTGRPFPTDKMPANWQLVPLIR
QILPNARIVDVRDPAPCCLSAFMTYFNRRTPFPANLPDLTRYDYDTCRGLMDAMRRAPAHVHLLRYEALIAQPK
GEVRRLLDFLRDLDFDPACLRPHDSARPIFTPQAQQVRKPMGNKGFEGWRNYERWFRHANGA

OTÁZKA 3

Určete teoretický isoelektrický bod a extinční koeficient následujícího proteinu (předpokládejte, že všechny cysteiny tvoří disulfidické můstky). Nezapomeňte uvést jednotky!:

Protein

MHPGLWLLLVTLCLTEELAAAGEKSYGKPCGGQDCSGSCQCFPEKGARGRPGPIIGIQGPTGPGFTGSGT
LSGLKGERGFPGLLGPIYGPYKDKGPMGVPGFLGINGIPGHGPGQPGPRGPPGLDGCNGTQGA VGFPGPDGY
PGLLGPPLPGQKSGKDPVLAPGSFKGMKGDPLPGLDGITGPGAPGFPAGVPGAGPPGLQGPFP
PLGPDGNMGLGFQGEKGVKGDVGLPGPAGPPPSTGELEFMGFPKGKKSKEPGEKGFPGISGPPGF
GTTGEKGEKGIIPGLPGRPGMSEGVQGGQKGTGLGFPGLNGFQGGIEGQKGDIGLPGPDVFI
IDGAVISGNPGDPGVPGLPGLKGDDEGIQGLRGPSPVPLPALSGVPGALGPQGFPGKGDQGNPGR
TIG AAGLPGRDGLPGPPGPPSPPEFETETLHNKESGFPGLRGEQGPKNLGLKGIKGDSGFCACDGG
VNTGPPGEPGPPGWGLIGLPLGKARGDRGSSGAQGPAGAPLVGPLGSPGPKGKKEPILSTIQGM
PDRGDSGSQGFGRVIGEPGKDGVPGLPGLPGLPGDGGQGFPEKGLPGLPGEKGHGPPGLP
GNGLPGLPGRGLPGDKGKDGLPGQGLPGSKGITLPCIIIPGSYGPSGFPPTGFPKGSRGLP
GTPGQPGSSGSKGEPGS PGLVHLPELPGFPGRGEKGLPGFPGLPGKDGLPGMIGSPGLPGS
KATGDI FGAENGAPGEQGLQGLTG HKGFLGDSGLPGLKGVHKGKPLGPKGERGSPGTPG
QVQPGTPGSSGPGYGIKKSGLPGAPGFPGISGH PGKKGTRKKKPPGSIVKKGLPGLKGL
PGNPGLVGLKGS PGSPGVAGLPALSGPKGEKGSVGFVGFPI PGLPGISGTRGLKGI
PGSTGKMGPSGRAGTPGEKGDGRNPGVPIPSRRPMSNLWLKGDKGSQGSAGSNG
FPGPRGDKGEAGRPGPPGLPGAPLPGI IKGVSGKPGPPGFMGIRGLPGLKGS
SITGFPMPGESGSQG IRGSPGLPGASGLPGLKGDNGQTVEISGSPGKQPGESGFKGT
KGRDGLIGNIGFPGNKGEDGKVGVS
GDVGLPGAPGFPVAGMRGEPGLPGSSGHQGAIGPLGSPGLIGPKGFPGLHGLNGLPGTKG
THGTPG SITGVPAGLPGPKGEKGYPIGIGAPGKPLRGQKDRGFPLQGPAGLPAGPISLPS
LIAGQPGDP GRPGLDGERGRPGPAGPPGPPGSSNQDGTGDPGFPPIPGFSGLPGELGLK
GMRGEPGFMGTPGKVGPPGDPGFPGMKGKAGARSSGLQGDPGQTPTAEAVQVPPG
PLGPIIDGIPGLTGDPGAQGPVGLQGSKGLPIPGKDGPSGLPGPPGALGDPGLPGLQ
GPPGFEGAPGQGFPGMPGMPGQSMRVGYTLVKHSQSEQVPPCP IGMSQLWVGYSLLF
VEGQEKAHNQDLGFAGSCLPRFSTMPFIYCNINEVCHYARRNDKSYWLSTTAPIPM
MPVSQTQIPQYISRCSVCEAPSQAIAVHSQDITIPQCPLGWRSLWIGYSFLMHTAAGA
EGGGQSLVSPGS CLEDFRATPFIECSGARGTCHYFANKYSFWLTTVEERQQFGELP
VSETLKAGQLHTRVSRQCVMKSL

Nepovinná otázka za bonusové body:

Mohli byste protein z předcházející úlohy beze strachu ochutnat? (Předpokládá se, že máte pud sebezáchovy...) Odpověď zdůvodněte.

OTÁZKA 4

Identifikujte v následující sekvenci štěpící místo pro daný restriční enzym a navrhněte jiný enzym, který toto místo rovněž rozpoznává.

Enzym: Eag I

Sekvence:

ATGGCCTACAGGACGGCCATCGACTCAGTATTGTCAGCAACATCCTCAATTCACCATCACACAAAACGCAGCTACA
 GCCGTGCGTGATAACAAGCTTGCAAGAACAGCCATGCACCTTCAACTCCGTCTTTAGCGTCCTTTTGTCTGTGGC
 CTCGCAGCAGCACACATGCAGATGAGCTGGCCTTATCCCATACGCAGTCCACTTGACCCTCAGAACAGAGGATCC
 GACAAGGACTATGACATGGCCAACCCCTCAGCCCGTCTGGCTCCGACTTTCCTTGCAAGGGCTACCACAAGAAC
 ACTCCCTGGCGCGCAACGGCGGAATACAACGGCCGGGCGAGTCATATAACATGACGGTGGCCGGTGTGCAACAC
 ACAGTGGAGGATCATGCCAGCTGTCTCTGAGTTACGACGATGGGAAAACCTTCAAAGTCATCAAGTCCAAGGTGG
 GAGGGTGCCCGCTTGATTTCAAGTACGACTTCACGATGCCGAGCGATGTGGTCAACGGCCATGCTCTGTTTGCTT
 GGTCTGTTTAAATCTAGTGGGCAATCGTGAGATGTACATGAACTGCGCAAACGTTGAGGTGAGCGGAGGGAGCG
 GCAGCAAGCAATCGTTTGAAAACAATTACCCGGATATGTTTCGTGGCCAATGTTGGAAATGGCTGTTCTACGGTGG
 AAGGGAAGCATAACGGTGTGTTGCCCATCCGGGAAAGCAGGTTACCTATGCCGGTGGGCTTGATGCGTCCTCACCTC
 CATTCCCCAACTGCTCTTGAAATCACAACCAATGAATGTTACACACTTTTGTACGTGATACACATTTGTGAGTATC
 GGCAAGGCAGGGCCATACTGAACTCTGGCTGTTGAGTGAAATAAGGATCAGTTGATCGGGATATACATAGTACAA
 ATCGGTTACATATGGCCTACAGGACGGCCATCGACTCAGTATTGTCAGCAACATCCTCAATTCACCATCACACA
 AACGCAGCTACAGCCGTCGTGTATAACAAGCTTGCAAGAACAGCCATGCACCTTCAACTCCGTCTTTAGCGTCCTT
 TTGTCTGTGGCCTCGCAGCAGCACACATGCAGATGAGCTGGCCTTATCCCATACGCAGTCCACTTGACCCTCAG
 AACAGAGGATCCGACAAGGACTATGACATGGCCAACCCCTCAGCCCGTCTGGCTCCGACTTTCCTTGCAAGGGC
 TACCACAAGAACACTCCCTGGCGCGCAACGGCGGAATACAACGGCGGCGAGTCATATAACATGACGGTGGCCGGT
 GCTGCAACACACAGTGGAGGATCATGCCAGCTGTCTCTGAGTTACGACGATGGGAAAACCTTCAAAGTCATCAAG
 TCCAAGGTGGGAGGGTGGCCGCTTGATTTCAAGTACGACTTCACGATGCCGAGCGATGTGGTCAACGGCCATGCT
 CTGTTTGCTTGGTCTGTTTAAATCTAGTGGGCAATCGTGAGATGTACATGAACTGCGCAAACGTTGAGGTGAGC
 GGAGGGAGCGGCAGCAAGCAATCGTTTGAAAACAATTACCCGGATATGTTTCGTGGCCAATGTTGGAAATGGCTGT
 TCTACGGTGGAAAGGGAAGCATAACGGTGTGTTGCCCATCCGGGAAAGCAGGTTACCTATGCCGGTGGGCTTGATGCG
 TCCTCACCTCCATTCCCCAACTGCTCTTGAAATCACAACCAATGAATGTTACACACTTTTGTACGTGATACACAT
 TGTGAGTATCGGCAAGGCAGGGCCATACTGAACTCTGGCTGTTGAGTGAAATAAGGATCAGTTGATCGGGATATA
 CATAGTACAAATCGGTTACAT

Enzym rozpoznávající stejné místo:

OTÁZKA 5

Uveďte základní charakteristiku následujících primerů.

Primer	L	R
Sekvence	CAGTGGAGGATCATGCCAGC	AGGCCATCAGTTGATCGGGATATA
Délka		
T _m		
GC [%]		
Dimer (obr.)		
Vlásenka (obr.)		

OTÁZKA 6

Predikujte lokalizaci následujících proteinů z *Candida albicans* v buňce. Ke každému proteinu uveďte pouze jednu (NEJPRAVDĚPODOBNĚJŠÍ) možnost.

Protein 1

MRQFQIILISLVVSIIRCVDVVDITSPKSGETFSGSSGSASIKITWDDSDSDSPKSLDNAKGYTISLC
TGPTSDGDIQCLDPLVKNEAIAGKSKTVSIPQNSVPNGYFFFQIYVTFFTNGGTTIHYSRPFKLTGMSGPT
ATLDVTETGSVPADQASGFDTATTADSKSFTVPYTLQTGKTRYAPMQMPGTKVTATTWSMKFPTS AVTY
YSTKAGTPNVASTITPGWSYTAESAVNYASVAPYPTYWYPASERVSKATISAATKRRRWLD

Protein 2

MAIKIGINGFGRIGRLVLRVALGRKDIEVVAVNDPFIAPDYAAYMFKYDSTHGGRYKGEVTASGDDLVIDG
HKIKVFQERDPANIPWGKSGVDYVIESTGVFTKLEGAQKHIDAGAKKVIITAPSADAPMFVVGVNEDKYT
PDLKIIISNASCTTNCLAPLAKVVNDTFGIEEGLMTTVHSITATQKTVDGPPSHKDWRGGRTASGNIIPSSST
GAAKAVGKVIPELNGKLTGMSLRVPTT DVSVDLTVRLKKAASYEEIAQAIAKKASEGPLKGVVLGYTEDAV
VSTDFLGSSYSSIFDEKAGILLSPTFVKLISWYDNEYGYSTRVVDLLEHVAKASA