

BIOINFORMATIKA V PRAXI – ZÁVĚREČNÝ TEST JARO 2014

Jméno a příjmení:

Datum:

UČO:

Varianta: **A**

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 *Pseudomonas aeruginosa*. Identifikujte potenciální geny (uved'te ty nejpravděpodobnější, rovněž uveďte čtecí rámec a délku potenciálního genu) a určete, kteří proteiny mohou být těmito geny kódovány.

```
CGTGCATGGGACATTCTGTTGCAGGCGTCTGCCTGGAAGAACCCGCGGTACTGACCGCCTTTCGGTCGCTGCTGC
ACCCGCAGGACCCGCCCAACAGAGGGACCGAATACTCTTCGTCACCGCCGAGCTCAGCGACTTCGTCAAGGTCG
GCGGGCTCGGCGACTTTTCCGCCGCCCTGCCCCGGGTGCTCAGGCGCCAGCATGCGGTACGGGTCTGCTGCCCC
GCTACCGCCAGGTGCTGGAGCGCTGCAGTGACCTGCGAATCCTCGGCAGCCTGCCGGGCCGCGCGGGCGATCCCGC
CGTGCGAGATCGGCCTGGTGACCCTCGACGACGGCCTGGAGGTGATGCTGGTGCTCTGCCCGTTGCTCTACGAGC
GCGAGGGCACCCCTACATGGACGACCAGGGCAACGACTGGGCTGACAACCACCTGCGCTTCGCCCCCTCTGCC
TGGCCGCCGCCGAGATCGCCGGCGGTACGGCGCCCAGGGCTGGCAGCCGGGCTGGTGCACGCCAACGACTGGC
CGTCGGCGCTGACCCCGGCCTACATGGCCTGGAACGGTGTGCGCACGCCGAGCCTGTTACCATCCACAACCTCG
CCTACCAGGGCCTCTGCGACCTGCAGTGACGCGCGAAGTGGGCTGCCCCGACGAAGCGCTCAGCCCCGAGAGCA
TGGAGTTCCACGGGCGCTGTCGTTCCCTCAAGGCGGGCATCGCCCATGCCATCACATACCACGGTGAGCGAGA
CCTACGCGCAGGAGATCACACCCCGGAATACGGCTGCGGCCCTGCACGGCATCCTCAAGTGCAAGGTGAGAAGC
GCCAGCTCAGCGGCATCGTCAACGGTATCGATGACAGCTGGCAGCCGATTGCGATCCGCACCTGGTGCCGGCT
TCAGCGCCCGCAATGGGCCGGCAAGCGCGCAACACACGCTATGTGAGGAGCGTTTTCGGCCTCGAGCCGGGCA
AGGGGCCGCTGTTTCGCGGTGGTCTCGCGGCTGGTGCAGCAAAAGGGTATCGACCTGACCCTGGAGATCAGCGACG
CCCTGCTCCAGGCCGGTGGCCGGCTGGTGCAGCATCGGGCGCGGTGAACCGTCGCTGGAGAAGGCCATGCTCGATC
TCTCGCGGCGTCATCCGGGCCAGGTGGCGGTGCACATCGGCTTCGACGAGACCGACGCGGGCGCATCTATGCCG
GCAGCGACTTCTGCTCATGCCATCGCGCTACGAGCCTTGGCGCCTGAGCCAGCTCTACGCGCAATGCTTCGGCT
CGCTGCCGATCGCCGTTGCACCGCGGGCTCGCCGACACCATCGTCGACGGCGTCACCGGCTTCTGTTCCGCG
AGGAAACCGCGCAGAGCTATCTGGATGCGGTGCTGCGCGGATCAACGTCTACCATTGCCCGCTCTGCTCAATG
CCATGCGCTGCAAGGCGATGGCCGCGCCGATGTTCTGGCGGACTCGGTGGAACCTACAACCGCTTACCGGC
GCCTGCTGAGGAATACCGCGCCGGCCTGGCGCGGGGTGCGCCCGTGA
```

OTÁZKA 2

Pomocí programu ClustalW2 proveďte multiple alignment následujících sekvencí a přiložte jeho výstup. Určete 2 navzájem nejpodobnější sekvence.

Sekvence 1

```
KCRI IREFPDIVFKGLTLVQVSQKFGKAGFEDVKVTEEIVHLNEDCCKGDAVECMERMEATDHI CEAK
DKLSSKLABDCCA KSI LERTPCLLALPNDES DLSKELKNYYEDERVCENYKDKLLFLAHFTHDYARSHQE
SSPQ SCLRVSKGFEGLEKCCASENHA ECLKQAPI LLEAALKEIEELRKQNCGALQLLGF RDYNIQLLFR
YFFKMPQV TAPTLVELAGRM TKVAVYCCGLAENKQQTCAEEKLDILLGEMCEKEKHTFVNDNVRHCCVDS
YANRRKCF TDLQRYPNYVAPK WDESKLHFNE D LCKGSEDDQIKKKLEVLVEYMKMKPDCGPEKLKEVVEA
FRKIDIKCAAEDHQKCFDDEKAGLLQIIEAH
```

Sekvence 2

MMSSKIQPAVPPKSSKSSKQEELICKTKIKGDGQDERVKEKAVQSRLFTNNDTVRAKFTFIDLFAIGIGGF
RIAMQNLGGECVFSSEWDEKAKQTYEANFGEVPPFGDITLEETKQCIPEQFDVLCAGFPCQAFSIAGRGG
FEDTRGTLFFDVAEIIIRHRPKAFFLENVKGLTNHDRGRTLQTIILNTLREDLGYFVPEPEIVNAKDFGVP
QNRERIFIVGFRGDTDIKYFEYPEPTGIQTAFRDIREKDTVPTKYLLSTQYIDTLRKHQRHEEKGNFGF
YEIISDDGIANAIIVVGGMGREERNLVIDKRITDFTPTTNIKGEVNREGIRKMTPREWARLQGFPETFKIPV
ADASAYKQFGNSVAVPAVQAVAGQVLIIRLGWRDA

Sekvence 3

MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNIITYAEYFEMSVRLAEAMKRY
GLNTNHRIVVCSSENSLQFFMPVVLGALFIGVAVAPANDIYNERELLNSMNI SQPTVVVFSKKGLQKILNVQ
KKLP I IQK I IIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVAL
PHRTACVRFSHARDPIFGNQIIPDTAILSVPVFFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRSLQDY
KIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGRQGYGLTETTSAIL
ITPEGDDKPGAVGKVVFFFEAKVVDLDTGKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDKDGWLHS
GDIAYWDEDEHFFIVDRLKSLIKYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHG
KTMTEKEIVDYVASQVTTAKKLRGGVVVDEVVFKGLTGKLDARKIREILIKAKKGGKSKL

Sekvence 4

KWEKHSCGFLKSFGERAFQANKLVQLSQKFPKAPFEEIHKVVLTAATKLQKECCHGDMIECLDDRVEVMAY
ICSKQAVFSSKIKDCCEKPIVDREVCILQADLDEKPADLPSIAGQYIESTEVCCKHYEKGKDVFLAHFVYE
YSRRHPEFSSQMLLRTGKGYQDTLDKCKTENPPECYKGAGEELARHIQESQEILKTHCSFYTSQKDPF
QKMVLVRYTKMPQLPAEELIEISKKLTGVGVKCCPLSEDKRLSCSEKHLMSVLFEICRQHEASPVNNHV
THCCTDSYSEMRPCFTKLGVDSDSYVPEFCPSTFLFDEQLCTAPEEARLKKQLTFLVCLIQLKPIEDEQ
LKKLVTDYHAMEEKCCQAENKQECFSTEGEKLTQEGKALLGVQLNVPGLQ

OTÁZKA 3

Určete teoretický isoelektrický bod a extinční koeficient následujícího proteinu (předpokládejte, že žádný cystein netvoří disulfidické můstky). Nezapomeňte uvést jednotky!:

MVHLTPEEKSAVTALWGKVNVDVGGGALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDG
LAHLNDNLKGTFTATLSELHCDKLVDPENFRLLGNVLVLCVLAHFFGKEFTPPVQAAAYQKVVAGVANALAHKYH

Nepovinná otázka za bonusové body (uplatitelné u kolokvia):

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

OTÁZKA 4

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

Enzym: Xma I

Sekvence:

GCCGCCGCCATCAGCTGAGACCGGCTGCGCCAGGCAGTGGGCCCCAGGGTGACAGCGGCTGGAGGTGGC
CGAGGAAGCCGGCGAGGTCAGGGTAAGAGGAGACTGCAGAGGCGGGGAGAGGGGGCGGGAGGGGACACC
TCAGAACTGCAGCTCGCTAGTGGTTGCTTTGGGGTGCAGAGCGGGGCTCCCCAGGGCGCAGGGCCGGC
GCCTCCCTCTCGGGCTCCCACGCCCGCTCCGGACTCCGGGACGCGACGTGCGAGCACCAGGGGCGCGCGG
CAATGGAGCCTTGGGGACCGGATACTACCGCCAGATCCTGGCAGGCTACGACCGTGTCCCATTCCTCTGT

CAACTTCTCCTCCGAAGAACAGAAAGGGAAACGCGATGGGCCCTGGTCATCTGGCCACAGGTGACTCCTA
 GAAATCCACATAGCCTACCTGCATCTGATGAAATCAGCCAGCCAGCCAACCAGCCATGTCTTACACTCCG
 GGGATTGGTGGGGACTCTGCGCAGCTAGCACAACGCATCTCCTCTAACATCCAAAAGATCACACAATGCT
 CCGTGGAAATACAAAGGACTCTGAATCAACTTGGAAACCTCAAGACTCACCGGAACTGAGGCAACAGCT
 GCAGCAGAAGCAGCAGTATACTAACCACCTTGGCAAAGAAACCGATAAGTACATTAAAAGAGTTTGGCTCT
 CTGCCCACCACCCCAAGTGGAGCAGCGTCAGAGAAAGATACAGAAAGACCGATTAGTGGCAGAGTTCACCA
 CATCGCTGACAACTTCCAGAAGGCCAGAGGCAGGCTGCTGAGAGAGAGAAAGAGTTCGTTGCTCGAGT
 GCGAGCCAGCTCCAGGGTATCGGGTGGTTTTCTGAAGACAGCTCAAAAGAAAAGAATCTTGATCCTGG
 GAAAGCCAAACACAGCCTCAAGTGCAGGTCGAAGATGAAGAAATCACAGAGGATGACCTCCGACTCATT
 ATGAGAGAGAGTCTTCAATCAGGCAGCTGGAAGCTGATATTATGGACATTAATGAAATATTTAAAGACTT
 GGGGATGATGATTCATGAACAAGGCGATATGATTGACAGCATAGAAGCCAATGTAGAAAAGTGCAGGAACTT
 CACGTTCAACAGGCAAACCAGCAGCTGTCAAGAGCGGCAGACTATCAGCGCAAATCCAGGAAAACCTCTCT
 GCATTATCATTTTTTATCCTCGTGGTCGGAATCGTGATCATCTGTCTCATCGTATGGGGACTGAAAGGCTG
 AGCCCCGAGGGCGTGGACGACTGCATGATGCTGTCTGAGCGATGCAGGCAGATTCTTGCGATCACTTTCT
 CTTATCGTTATCTTGAGGCTGTTGTGTAATAATGATGGTTCCATACTTTGCCATTTTTACTAGGGTGGGG
 GATTCTTTTTGGATTGAGTCTGATATTTCTAATACCAAGGCTTTTCTAAACACCCGGGCTGGCACAGC
 TTCCCTGCTTTTTAGTTCAACAACCTGTTAGGTTTTGCTCACGTTGTATATGCCTTTCATTTGTAATTTAT
 TTATTGTAACTTTTTACTGAGTTTTGGGAATTTTATTGATTGGCGTGTATACTCTGCCTGCACCTCTGTA
 ACTGTGTAACATGTGTAACATGATTGCATTATCTTATTTAAAAAGAAAGCCCAAGTTTAAATTTCAATTTG
 TCCAGTAAATGTTCTCAATAGGAAACCTTAACATAACTGTTTGTGGGATGTTTTATCCGCTGTGGTCCA
 CAGCAGTGCAGATCGTTATTGCTCTTTCATTATTGGATTGTTGTTTTCGTCTCTGAATAAAAATCAGTGAA
 TCTTTTTCTGATGTGTATGTATATATGCATTTGCTGTGGATTACCTAGTAAAGACAGTAAGGATAGGTTG
 TTTTTAACATTATGGACTGTTTAAATCCAATGCATACTAAAGAAAGCTTAGTTACTGTATTGTTTCAAGTAA
 AGTTGGAGGAAGTGGCTGAAAGGTCCATACCTGCATCTATTCTGTTTCAAGATGCTAAAGATTCCAGTCTCT
 TACTGCATGAACACTTTCACCATATGGTCTGGCTTTGTGGCCATTATACCTTCCCTCGTGCGCCATGG
 AACACTTCTTAGGCCTGTATCTTTGGGTCCTGTCTTACAAAGTAATGTTCCCATAGTTACAGATTTGGG
 ATTTTCAATTTTATAGTTTTTGGACTTAAAGTTTATAGAGTTTATTTAGATGCCATTTTTTATTTATTC
 TTCATTCAAATAGGAATCAGACGCTGCCAATGAACACACTGTCTTTTTTGAATCAATGTTTTAATGCA
 CTAATATGAATGCTGTAAGAGGGGTTATCTTAGTTTTATCATTTGTGTTTTATAATGTTCTTACACAGCAG
 TTTGCTAATGATGGCACTGTTTAAATGTGGCTCTGAAACCTTTAAATGGGAACAAGTGAAATATTAACAAA
 ATAATTGTATCTTTGCAAGCAGAATAAAGATGATTCTGGAGATGTCTGCT

Enzym rozpoznávající stejné místo:

OTÁZKA 5

**Uveďte základní charakteristiku následujících primerů. Nezapomeňte uvést jednotky.
 Na základě parametrů označte primer vhodnější pro další práci.**

Primer	A	B
Sekvence	GACTGGTATACCCTAGGCGTA	CTCAGACCAATCATTCCGGTTAC
Délka		
T_m		
GC [%]		
Dimer (obr. a volná energie)		
Vlásenka (obr. a volná energie)		

Vhodnější primer:

OTÁZKA 6

Predikujte lokalizaci následujících proteinů ze *Staphylococcus aureus* v buňce. Ke každému proteinu uveďte pouze jednu (NEJPRAVDĚPODOBNĚJŠÍ) možnost.

Protein 1

MKALLLLKTSVWLVLLFSVMGLWQVSNAAEQYTP IKAHVVT TIDKATTDKQQVTP TKEAAHQFGEEAATNVSASAQ
GTADEINNKVTSNAFVSNKPSTAVSTKVNETHDVDTQQASTQKPTQSATFTLSNAKTASLSPRMFAANVPQTTHK
ILHTNDIHGRLAEEKGRVIGMAKLT I KEQEKPDMLDAGDAFQGLPLSNQSKGEEMAKAMNAVGYDAMAVGNHE
FDFGYDQLKKLEGMLDFPMLSTNVYKDGKRAFKPSTIVTKNGIRYGIIGVTTPETKTKTRPEGIKGVEFRDPLQS
VTAEMMRIYKDVDTFVVISHLGIDPSTQETWRGDYLVKQLSQNPQLKKRITVIDGHSHTVLQNGQIYNNDALAQT
GTALANIGKVTFNYRNGEVSNIKPSLINVKDVENVTPNKALAEQINQADQTFRAQTAEVII PNNTIDFKGERDDV
RTRETNLGNAIADAMEAYGVKNFSKKTDFAVTNGGGIRASIAKGVTRYDLISVLPFGNTIAQIDVKGSDVWTA
EHS LGAPT TQKDGKTVLTANGLLHISDSIRVYYDMNKPSGKRINAIQILNKETGKFENIDLKRVYHVTMNDFTA
SGGDGYSMFGGPREEGISLDQVFLASYLKTANI AKYDTTEPQRMLLGKPAVSEQPAKGQQGSKGSESGKDVQPIGD
DKAMNPAKQPATGKVVLLPTHRGTVSSGTEGSGRTLEGATVSSKSGNQLVRMSV PKGSAHEKQLPKTG TNQSSSP
AAMFVVLVAGIGLIATVRRRKAS

Protein 2

MKCLIRFILVLGLLLISSAMVYINPAAYAEQDQKWEKIKERGELRVGLSADYAPMEFEHTVNGKTEYAGVDIDLAK
KIAKDNNLKLKIVNMSFDSL LGALKTGKIDIIISGMTSTPERKKQVDFSDSYMMTKNIMLVKKDKVNDYKDIKDF
NNKKVGAQKGTEQEKIAQTEIENASITSLRSLPDVILALKSGKVEGVVVEKPVAEAYLKQNPKLGISNVKFNEEE
KDTVIAVPKDSPKLLSQINKTIKEVKDKGLIDKYMTNAANAMNDDSGFISKYGSFFLKGKIKITILISLIGVALGS
ILGAFVALMKLSKIKIISWIASIYIEILRGTPMLVQVFIVFFGITAAALGLDISALVCGTIALVINSSAYIAEIR
AGINAVDKQMEARS LGLNRYQTMKSVIMPQAIKNILPALGNEFVTLIKESSIVSTIGVGEIMFNAQVVQGISF
DPFTPLLVAALYFVLTFVLTRIMMIEGRNLASD