C3211 Aplikovaná bioinformatika
Jaro 2013

Produkce rekombinantních proteinů

Úkol 1

Náš známý student dostal za úkol vyzkoušet produkci několika hypotetických/nových proteinů v expresním systému *E. coli*. Je stále stejně zbrklý a začal s experimenty, aniž by si o proteinech cokoliv zjistil. Takže když narazil na problémy, je úplně ztracený. Zkuste mu poradit, co může způsobovat potíže při produkci proteinů.

A) Hypotetický protein z Gram-negativní bakterie

MGSASSLHAQAQITGRPEWIWLALGTALMGLGTLYFLVKGMWIAIGTGLMALGTLYFIVKGMVWLVCN
HQQFQLFMSINPVPTSVMWPHHHMALQGTMGMTIFNNCHAPNWSGVCAFHILIHHSGFWTCTHLLFWM
SCSMVPQGGWIWNGMGHAFFFTNVSLNFLWFTLCVACVWNNACPIIMSVQCFGTMSVLQPWLWHVSPH
LTSQAAMPPIPQQCWNPQPIASSANCTITCMQNAWQIAFLGGSVTGFNHIVTMIATGCWLALGTALMG
LGTLYFLVKGMWIAIGTGLMALGTLYFIVKGM

Problém: buňky málo a pomalu rostou!

B) Nově identifikovaný protein z byliny používané v tradičním čínském léčitelství.

MNIKMRYRCEIKHRAAGIPRYREADFTRRIWRAIMTKTCPKLPIVCSPIDKQLPRWNPEQPYDNLIIN
PLMNSVKFPHRQHHRKLPRMFLITRCDNRAMIYNEPDAQIRRFKVIDFITFYQRPCCNWQYPCLAQSY
VIQPRIRRTIVPRHTVGGEQRIILIEEMCLNFMRPHELWLLRARLCFRHTRHTIVEVCWYRRHHQSGP
QRLEVWGWKINQLPIEYCLCLWVIQVRNGVRLNLELFPIAQRLNEYHWTRYDRTPVWRFISLLRKRQI
CFRWPMSDWRKHDVRYGSPARRSLDGGKPPIRSQPCIPRLGNPYRNRTRRRCDPFAPHSIRGSITALR
PEWEYKCSFEVLLRSLRPIPIRVMKAMSPPGYDIHMQFDIAWMPPARRFRGASMMRRYWGTDRRSGDY
PKTMAWFLRVHTHMMERGSRNTPKRDKPCFN

Problém: protein není produkován! Ani v nerozpustné formě (inkluze)!

C) Hypotetický protein z Gram-negativní bakterie

MVTVNSRVDSARASHITLNVMTVLDVGAILDDIKAQEKRLGETPDTGTSIGHKYIYMSSDDPRGWSVG
NDPGNITINAHVGETLSFACASTSENSEYAAFLYRLTGGDPHLDPSHVEVIKLQNAAQPTPSNGYPFT
TAPVAFSSCDAKVAQQGQAKNFYVWAGLFTLDESGEKEVLAAYVKWEPTLNVGSGGLSGLSWERTNET
YQTLGLGLNDVDGTTEQMFKKPSYSDRTALMDVSQIEASLASFAEIVDFHAIMLRLKPCALRAKLDYL
ARRVSKEIEVPHVVNIGGPFKNHGNGGKEVPIIAQ

Problém: protein je produkován, ale všechen pouze ve formě inkluzí!

D) Nově izolovaný protein z patogenní houby, optimalizované kodony, ověřené exony/introny, studentova kolegyně dosáhla pěkných výtěžků!

MTQPHCICKGHDNAFYLVCRSWQKVDQCNVQRNMWWFNCSSLMMTYAMYANGRFPAHAEHEAKRPQYE
QEHNGMDVCGETMEARWCITWPCCESIRKFINSAWLDIFQYTKDKLLYIDKKWVSLLVFHRQNFNYEG
CLTYVQSFHGYTMETPTDRDPVRLFPFNYCSKGMPIHTGESWPHAMWDAIIPSRVWDMQHCGKVL

Problém: všechny buňky (12 lahví! Student připravoval medium půl dne!) v průběhu kultivace umřely…

Úkol 2

Student dostal za úkol odhadnout, zda se v proteinech, se kterými pracuje, tvoří disulfidické můstky. Pomozte mu!

Protein 1

FDSKLKEDKVAQVASEFRCVEIDVLLPVKFERNMTVGLLKGSPEHHQAQHVIDLAEAKKI
NCVGMYSVKGAEPLEWTPINYVIAVGRFNTRIAYFHYGKSTGELDRLNKLFAALCDQEPG
NTLDDAGLMRYRMISQALLPFEIISLSASSTVGKTSGAIAATDYTQPFLDGPRVSRNWRG
VESTLKRNHQGILIDEPQTM

C3211 Aplikovaná bioinformatika
Jaro 2013

Protein 2

AWKGEVLANNEAGQVTSIIYNPGDVITIVAAGWASYGPTQKWGPQGDREHPDQGLICHDAFC
GALVMKIGNSGTIPVNTGLFRWVAPNNVQGAITLIYNDVPGTYGNNSGSFSVNIGKDQS

Úkol 3

Student, poučen předchozími nezdary, se rozhodl *vybrat si* proteiny, se kterými bude pracovat. Využijte nástroj **PROSO** a pomozte mu…

Protein 1

AWKGEVLANNEAGQVTSIIYNPGDVITIVAAGWASYGPTQKWGPQGDREHPDQGLICHDAFCGALVMK
IGNSGTIPVNTGLFRWVAPNNVQGAITLIYNDVPGTYGNNSGSFSVNIGKDQS

Protein 2

LVIVDAVTLLSAYPEASRDPAAPTVIDGRHLYVVSPGDAAQLGHNDSRLFTGLSPGDQLHLRETALAL
RAEVSVLFIRFALKDAGIVAPIELEVRDAATAVPDADDLLHPSCRPLKDHYWRSDVLAAGATTCTADF
AVCDRDGTVSGYFRWETSIEIAGSQPDTKQPGFKPSSDRNGNFSLPPNTAFKAIFYANAADRQDLKLF
IDDAPEPAATFVGNSEDGVRLFTLNSKGGKIRIEASANGRQSATDARLAPLSAGDTVWLGWLGAEDGA
DADYNDGIVILQWPIT

Protein 3

PLLSASIVSAPVVTSETYVDIPGLYLDVAKAGIRDGKLQVILNVPTPYATGNNFPGIYFAIATNQGVV
ADGCFTYSSKVPESTGRMPFTLVATIDVGSGVTFVKGQWKSVRGSAMHIDSYASLSAIWGTAAPSSQG
SGNQGAETGGTGAGNIGGGGERDGTFNLPPHIKFGVTALTHAANDQTIDIYIDDDPKPAATFKGAGAQ
DQNLGTKVLDSGNGRVRVIVMANGRPSRLGSRQVDIFKKSYFGIIGSEDGADDDYNDGIVFLNWPLG

Protein 4

MGSASSLHAQAQITGRPEWIWLALGTALMGLGTLYFLVKGMWIAIGTGLMALGTLYFIVKGMVWLVCN
HQQFQLFMSINPVPTSVMWPHHHMALQGTMGMTIFNNCHAPNWSGVCAFHILIHHSGFWTCTHLLFWM
SCSMVPQGGWIWNGMGHAFFFTNVSLNFLWFTLCVACVWNNACPIIMSVQCFGTMSVLQPWLWHVSPH
LTSQAAMPPIPQQCWNPQPIASSANCTITCMQNAWQIAFLGGSVTGFNHIVTMIATGCWLALGTALMG
LGTLYFLVKGMWIAIGTGLMALGTLYFIVKGM

Úkol 4

Student se rozhodl výsledky ověřit pomocí jiného predikčního nástroje. Predikujte rozpustnost proteinů z úkolu 3 s využitím dalšího programu. Výsledky porovnejte.

Úkol 5

Analyzujte následující geny a určete četnost využitých kodonů v *E. coli* a *Saccharomyces cerevisiae.*

Gen 1

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTC
TATCGAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTAACCGTTTCACCT
CTAACATTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATCTCCGTTGCGCAGACCACCGAAGGCGCGCTGTCC
GAAATCAACAACAACTTACAGCGTGTGCGTGAACTGACGGTACAGGCCACTACCGGTACTAACTCTGAGTCTGATCTGTCTTC
TATCCAGGACGAAATTAAATCCCGTCTGGATGAAATTGACCGCGTATCTGGTCAGACCCAGTTCAACGGCGTGAACGTGCTGG
CAAAAAATGGCTCCATGAAAATCCAGGTTGGCGCAAATGATAACCAGACTATCACTATCGATCTGAAGCAGATTGATGCTAAA
ACTCTTGGCCTTGATGGTTTTAGCGTTAAAAATAACGATACAGTTACCACTAGTGCTCCAGTAACTGCTTTTGGTGCTACCAC
CACAAACAATATTAAACTTACTGGAATTACCCTTTCTACGGAAGCAGCCACTGATACTGGCGGAACTAACCCAGCTTCAATTG
AGGGTGTTTATACTGATAATGGTAATGATTACTATGCGAAAATCACCGGTGGTGATAACGATGGGAAGTATTACGCAGTAACA
GTTGCTAATGATGGTACAGTGACAATGGCGACTGGAGCAACGGCAAATGCAACTGTAACTGATGCAAATACTACTAAAGCTAC
AACTATCACTTCAGGCGGTACACCTGTTCAGATTGATAATACTGCAGGTTCCGCAACTGCCAACCTTGGTGCTGTTAGCTTAG
TAAAACTGCAGGATTCCAAGGGTAATGATACCGATACATATGCGCTTAAAGATACAAATGGCAATCTTTACGCTGCGGATGTG
AATGAAACTACTGGTGCTGTTTCTGTTAAAACTATTACCTATACTGACTCTTCCGGTGCCGCCAGTTCTCCAACCGCGGTCAA
ACTGGGCGGAGATGATGGCAAAACAGAAGTGGTCGATATTGATGGTAAAACATACGATTCTGCCGATTTAAATGGCGGTAATC

TGCAAACAGGTTTGACTGCTGGTGGTGAGGCTCTGACTGCTGTTGCAAATGGTAAAACCACGGATCCGCTGAAAGCGCTGGAC
GATGCTATCGCATCTGTAGACAAATTCCGTTCTTCCCTCGGTGCGGTGCAAAACCGTCTGGATTCCGCGGTTACCAACCTGAA
CAACACCACTACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGACCGAAGTGTCCAATATGTCGAAAG
CGCAGATCATCCAGCAGGCCGGTAACTCCGTGTTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGT
TAA

Gen 2

ATGGATACCTCTCACGAAATTCATGATAAAATACCCGATACATTAAGAGAGCAGCAGCAGCATTTGCGCCAAAAAGAGTCGGA
AGGGTGCATAACAACATTGAAAGATCTTAATGTACCTGAGACGAAGAAACTTTCCTCCGTCTTACATGGCAGAAAGGCGAGTA
CATATTTAAGAATTTTTAGAGATGATGAATGTTTAGCCGATAACAACAATGGTGTAGATAGCAATAATGGCGGGTCCGTGACA
TGCGCGGACAAAATAACAAGGTCAGAAGCAACTCCCAAATCGGTACCAGAAGGACTACAAGTTTCTGAAAAAAAGAATAACCC
AGATACTTTGTCCTCTTCATTGTCAAGTTTCATTTTATCTAACCACGAGGAGCCGGCAATTAAGCCCAATAAACACGTGGCGC
ATAGAAATAATATTACTGAGACAGGTCAAGGCTCTGGTGAGGATATCGCAAAACAACAATCACACCAACCACAGGTGTTACAC
CACCAGACTTCTTTAAAGCCAATTCAGAATGTTGATGAGGGATGTATTTCGCCTAAGTCCACTTATCAGGAGAGCTTACATGG
AATATCGGAGGACTTAACGTTAAAGCCTGTTTCTTCTGCAACATACTATCCTCACAAAAGTAAAGCAGATTCTGGTTATGAAG
AAAAGGATAAAATGGAAAATGATATTGATACTATTCAGCCTGCCACCATAAATTGTGCCTCTGGTATAGCGACACTTCCTAGT
TCATATAACCGACATACTTTCAAAGTTAAGACATATTCAACTTTATCACAATCCCTAAGGCAAGAAAACGTCAACAATCGTAG
TAACGAAAAAAAACCTCAACAGTTTGTGCCACATAGCGAATCAATTAAAGAAAAACCCAATACGTTTGAACAAGATAAAGAAG
GCGAGCAAGCGGACGAAGAAGAGGATGAAGGTGATAATGAACATAGGGAATATCCCCTAGCAGTTGAACTAAAACCATTTACG
AATAGAGTTGGAGGTCACACTGCAATATTCAGATTTTCGAAGAGAGCGGTTTGTAAGGCTCTAGTGAATAGAGAAAACAGGTG
GTATGAGAATATAGAGCTGTGCCATAAAGAATTATTGCAATTCATGCCACGTTATATTGGTGTCCTGAATGTTAGGCAACATT
TTCAGTCAAAAGATGACTTTTTAAGCGATTTGGATCAAGAAAACAACGGTAAAAATGATACCAGTAATGAAAACAAAGACATA
GAAGTCAATCATAACAATAACGACGATATTGCGCTTAATACAGAACCCACGGGGACCCCCTTAACACACATACATTCCTTCCC
CTTGGAACATTCGTCGAGACAAGTACTTGAAAAAGAGCATCCTGAAATTGAGTCTGTCCATCCACACGTGAAAAGATCACTGT
CAAGTTCCAACCAACCTTCCTTGTTGCCCGAGGTTGTATTAAATGATAACAGGCATATTTAA

Úkol 6

Analyzujte geny z různých organismů a určete, které z nich je možné použít pro expresi v *E. coli* bez optimalizace kodonů. Za jakých podmínek?

Gen 1 *(Saccharomyces cerevisiae)*

ATGTCATTCGACGACTTACACAAAGCCACTGAGAGAGCGGTCATCCAGGCCGTGGACCAGATCTGCGACGATTTCGAGGTTAC
CCCCGAGAAGCTGGACGAATTAACTGCTTACTTCATCGAACAAATGGAAAAAGGTCTAGCTCCACCAAAGGAAGGCCACACAT
TGGCCTCGGACAAAGGTCTTCCTATGATTCCGGCGTTCGTCACCGGGTCACCCAACGGGACGGAGCGCGGTGTTTTACTAGCC
GCCGACCTGGGTGGTACCAATTTCCGTATATGTTCTGTTAACTTGCATGGAGATCATACTTTCTCCATGGAGCAAATGAAGTC
CAAGATTCCCGATGATTTGCTAGACGATGAGAACGTCACATCTGACGACCTGTTTGGGTTTCTAGCACGTCGTACACTGGCCT
TTATGAAGAAGTATCACCCGGACGAGTTGGCCAAGGGTAAAGACGCCAAGCCCATGAAACTGGGGTTCACTTTCTCATACCCT
GTAGACCAGACCTCTCTAAACTCCGGGACATTGATCCGTTGGACCAAGGGTTTCCGCATCGCGGACACCGTCGGAAAGGATGT
CGTGCAATTGTACCAGGAGCAATTAAGCGCTCAGGGTATGCCTATGATCAAGGTTGTTGCATTAACCAACGACACCGTCGGAA
CGTACCTATCGCATTGCTACACGTCCGATAACACGGACTCAATGACGTCCGGAGAAATCTCGGAGCCGGTCATCGGATGTATT
TTCGGTACCGGTACCAATGGGTGCTATATGGAGGAGATCAACAAGATCACGAAGTTGCCACAGGAGTTGCGTGACAAGTTGAT
AAAGGAGGGTAAGACACACATGATCATCAATGTCGAATGGGGGTCCTTCGATAATGAGCTCAAGCACTTGCCTACTACTAAGT
ATGACGTCGTAATTGACCAGAAACTGTCAACGAACCCGGGATTTCACTTGTTTGAAAAACGTGTCTCAGGGATGTTCTTGGGT
GAGGTGTTGCGTAACATTTTAGTGGACTTGCACTCGCAAGGCTTGCTTTTGCAACAGTACAGGTCCAAGGAACAACTTCCTCG
CCACTTGACTACACCTTTCCAGTTGTCATCCGAAGTGCTGTCGCATATTGAAATTGACGACTCGACAGGTCTACGTGAAACAG
AGTTGTCATTATTACAGAGTCTCAGACTGCCCACCACTCCAACAGAGCGTGTTCAAATTCAAAAATTGGTGCGCGCGATTTCT
AGGAGATCTGCGTATTTAGCCGCCGTGCCGCTTGCCGCGATATTGATCAAGACAAATGCTTTGAACAAGAGATATCATGGTGA
AGTCGAGATCGGTTGTGATGGTTCCGTTGTGGAATACTACCCCGGTTTCAGATCTATGCTGAGACACGCCTTAGCCTTGTCAC
CCTTGGGTGCCGAGGGTGAGAGGAAGGTGCACTTGAAGATTGCCAAGGATGGTTCCGGAGTGGGTGCCGCCTTGTGTGCGCTT
GTAGCATGA

Gen 2 *(Burkholderia sp.)*

ATGTCTACTGGTGTGCAAACTAAAGCTGTGCCGGGCGCGGCCCAGCACGCCGATGGACCGAGGCTGCTGGCCGACATCGGCGG
TACCAATGCGCGTTTCGCGCTCGAGACGGGCCCCGGCGAAATCGGCTCGGTGCAGGTCTATCCCTGCGCCGAGTATCCGAGCG
TCGCCGACGTCATCAAGAAGTACCTGAAGGACACGAAGATCGGCCGCGTCAATCACGCGGCGATCGCGATTGCGAACCCGGTC
GACGGCGACCAGGTCAGCATGACCAATCACGACTGGACCTTTTCGATCGAAGCGACGCGCCGCACGCTCGGCTTCGACACGCT
GCTCGTCGTCAACGACTTCACCGCGCTCGCGATGGCGCTGCCCGGCCTCACCGACACGCAGCGCGTGCAGGTCGGCGGCGGCC
AGCGCCGGCCGAACAGCGTGATCGGCCTGCTCGGCCCCGGCACCGGCATGGGCGTCTCGGGCCTGATCCCCGCCGACGACCGC
TGGATCGCGCTCGGCAGCGAAGGCGGCCACGCCACCTTCGCGCCCGCCGACGAGCGCGAAGAGATCGTGCTGCACTACGCGCG
CAAGAAGTGGTCGCACGTTTCGTTCGAACGGGTGGCCGCGGGCCCGGGCCTCGAGGTGATCTATCGCGCGCTCGCGGGCCGCG
ACAAGAAACGCGTGCCGGCCAGCGTCAACCCGGCCGATGTGGTCAAGCGCGCGCACGAAGGCGAGCCGCTCGCGGCCGAATCG
GTCGACGTGTTCTGCGGCATTCTCGGCACCTTCGCCGGCAACATCGCGGTGACGCTCGGCGCGCTCGGCGGCATCTATATCGG
CGGCGGTGTCGTGCCGCGTCTCGGCGAGTTGTTCGCGCGTTCGTCGTTCCGCCAGCGCTTCGAGGCGAAGGGCCGCTTCGAGG
CGTATCTGCAGAATGTGCCGACCTATGTGATCACCGCCGAATACCCGGCGTTCCTCGGCGTATCGGCGATTCTCGCGGAGCAG
CTGTCGAACCGCGCGGGCGGCAGCTCGTCGGCGGTGTTCGAGCGGATTCGCCAGATGCGCGACGCGCTGACGCCGGCCGAGCG

CCGCGTCGCCGATCTCGCGCTGAACCATCCGCGTTCGATCATCAACGATCCGATCGTCGACATCGCGCGCAAGGCCGACGTCA
GTCAGCCGACCGTGATCCGCTTCTGCCGCTCGCTCGGCTGCCAGGGGTTGTCGGATTTCAAGCTGAAGCTCGCGACCGGTTTG
ACCGGCACGATTCCGGTCAGCCACAGCCAGGTGCATCTCGGGGACACCGCGACCGACTTCGGCGCGAAGGTGCTCGACAACAC
CGTGTCGGCGATCCTGCAGTTGCGCGAGCATCTGAACTTCGAGCACGTGGAACGCGCGATCGACCTGCTGAACGGAGCACGGC
GCATCGAGTTCTACGGGCTCGGCAATTCGAACATCGTCGCGCAGGACGCGCACTACAAGTTCTTCCGTTTCGGTATTCCGACC
ATCGCGTACGGTGACCTGTACATGCAGGCAGCCTCGGCCGCGCTGCTCGGCAAGGGCGACGTGATCGTCGCGGTGTCGAAGTC
GGGGCGCGCGCCCGAGCTGCTGCGCGTGCTCGACGTCGCGATGCAGGCCGGCGCGCAGGTGATCGCGATCACCTCGAGCAACA
CGCCGCTCGCGAAGCGCGCGACGGTCGCGCTGGAAACCGATCACATCGAGATTCGCGAGTCGCAGCTGTCGATGATCTCGCGC
ATCCTGCACCTCGTGATGATCGACATTCTCGCGGTCGGTGTGGCGATTCGTCGCGCGGTGCCGAGCGATGAAGTGGCCGAGAC
CGTCGAGAAAGCGCGCGAAGGCGCCGACGACGACGCGACTGCCGTACTCGACTGGCTGAGCCACGGCGCGGCTTCGTCGGCGC
GCGACTGA

Gen 3 *(nezmámý organismus)*

ATGAATATTAAAATGAGATATAGATGTGAAATTAAACATAGAGCTGCTGGAATTCCAAGATATAGAGAAGCTGATTTCACTAG
AAGAATTTGGAGAGCTATTATGACTAAAACTTGTCCAAAACTTCCAATTGTTTGTTCTCCAATTGATAAACAACTTCCAAGAT
GGAATCCAGAACAACCATATGATAATCTTATTATTAATCCACTTATGAATTCTGTTAAATTCCCACATAGACAACATCATAGA
AAACTTCCAAGAATGTTCCTTATTACTAGATGTGATAATAGAGCTATGATTTATAATGAACCAGATGCTCAAATTAGAAGATT
CAAAGTTATTGATTTCATTACTTTCTATCAAAGACCATGTTGTAATTGGCAATATCCATGTCTTGCTCAATCTTATGTTATTC
AACCAAGAATTAGAAGAACTATTGTTCCAAGACATACTGTTGGAGGAGAACAAAGAATTATTCTTATTGAAGAAATGTGTCTT
AATTTCATGAGACCACATGAACTTTGGCTTCTTAGAGCTAGACTTTGTTTCAGACATACTAGACATACTATTGTTGAAGTTTG
TTGGTATAGAAGACATCATCAATCTGGACCACAAAGACTTGAAGTTTGGGGATGGAAAATTAATCAACTTCCAATTGAATATT
GTCTTTGTCTTTGGGTTATTCAAGTTAGAAATGGAGTTAGACTTAATCTTGAACTTTTCCCAATTGCTCAAAGACTTAATGAA
TATCATTGGACTAGATATGATAGAACTCCAGTTTGGAGATTCATTTCTCTTCTTAGAAAAAGACAAATTTGTTTCAGATGGCC
AATGTCTGATTGGAGAAAACATGATGTTAGATATGGATCTCCAGCTAGAAGATCTCTTGATGGAGGAAAACCACCAATTAGAT
CTCAACCATGTATTCCAAGACTTGGAAATCCATATAGAAATAGAACTAGAAGAAGATGTGATCCATTCGCTCCACATTCTATT
AGAGGATCTATTACTGCTCTTAGACCAGAATGGGAATATAAATGTTCTTTCGAAGTTCTTCTTAGATCTCTTAGACCAATTCC
AATTAGAGTTATGAAAGCTATGTCTCCACCAGGATATGATATTCATATGCAATTCGATATTGCTTGGATGCCACCAGCTAGAA
GATTCAGAGGAGCTTCTATGATGAGAAGATATTGGGGAACTGATAGAAGATCTGGAGATTATCCAAAAACTATGGCTTGGTTC
CTTAGAGTTCATACTCATATGATGGAAAGAGGATCTAGAAATACTCCAAAAAGAGATAAACCATGTTTCAAT

Úkol 7

Optimalizujte následující geny/proteiny pro produkci v hostitelských organismech.

1

MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNANDGISVAQTTEGALS
EINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKNGSMKIQVGANDNQTITIDLKQIDAK
TLGLDGFSVKNNDTVTTSAPVTAFGATTTNNIKLTGITLSTEAATDTGGTNPASIEGVYTDNGNDYYAKITGGDNDGKYYAVT
VANDGTVTMATGATANATVTDANTTKATTITSGGTPVQIDNTAGSATANLGAVSLVKLQDSKGNDTDTYALKDTNGNLYAADV
NETTGAVSVKTITYTDSSGAASSPTAVKLGGDDGKTEVVDIDGKTYDSADLNGGNLQTGLTAGGEALTAVANGKTTDPLKALD
DAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG

Pro produkci v *E. coli.*

2

MDTSHEIHDKIPDTLREQQQHLRQKESEGCITTLKDLNVPETKKLSSVLHGRKASTYLRIFRDDECLADNNNGVDSNNGGSVT
CADKITRSEATPKSVPEGLQVSEKKNNPDTLSSSLSSFILSNHEEPAIKPNKHVAHRNNITETGQGSGEDIAKQQSHQPQVLH
HQTSLKPIQNVDEGCISPKSTYQESLHGISEDLTLKPVSSATYYPHKSKADSGYEEKDKMENDIDTIQPATINCASGIATLPS
SYNRHTFKVKTYSTLSQSLRQENVNNRSNEKKPQQFVPHSESIKEKPNTFEQDKEGEQADEEEDEGDNEHREYPLAVELKPFT
NRVGGHTAIFRFSKRAVCKALVNRENRWYENIELCHKELLQFMPRYIGVLNVRQHFQSKDDFLSDLDQENNGKNDTSNENKDI
EVNHNNNDDIALNTEPTGTPLTHIHSFPLEHSSRQVLEKEHPEIESVHPHVKRSLSSSNQPSLLPEVVLNDNRHI

Pro produkci v *E. coli*. Nesmí obsahovat štěpící místa pro *HindIII* a *NdeI*!

3

ATGGCAACACAAGGAGTGTTCACCCTTCCCGCCAACACCCGGTTCGGCGTCACCGCCTTCGCCAACTCGTCCGGAACCCAGAC
GGTGAACGTGCTGGTCAACAACGAGACGGCCGCGACCTTCAGCGGGCAAAGCACCAATAACGCCGTCATCGGCACCCAGGTGC
TCAACTCCGGCAGCAGTGGCAAGGTACAGGTCCAGGTCAGCGTCAACGGCCGCCCCTCGGATCTGGTCTCGGCACAGGTAATC
CTGACCAACGAGCTGAACTTCGCCCTGGTCGGCTCTGAAGACGGCACCGACAACGACTACAACGACGCCGTCGTGGTGATCAA
CTGGCCGCTCGGCTAG

Pro produkci v HEK 293.

C3211 Aplikovaná bioinformatika
Jaro 2013

Úkol 8

Analyzujte část genomu z *G. gallus*. Predikované geny optimalizujte pro produkci v *E. coli*.