

5. vaja: Naloge in odgovori

V05-01

Je pri kvasovki *Saccharomyces cerevisiae* opazna optimizacija rabe kodonov? Kateri kodoni za katere aminokislinske ostanke so najredkejši?

Uporabite Codon Usage Database (<http://www.kazusa.or.jp/codon/>).

Odgovor

Pri organizmu *S. cerevisiae* je opazna optimizacija rabe kodonov in sicer predvsem pri naslednjih aminokislinskih ostankih (uporabljen kriterij: <10% pogostost za določen aminokislinski ostanek):

- arginin (redko se uporabljajo kodoni CGC, CGA in CGG)
- levcin (redko se uporablja kodon CUC)

Podani rezultati so za *Saccharomyces cerevisiae* [gbpln]: 14411 CDS's (6534504 codons).

V05-02

Gen za človeški histatin 3 (histatin-3) želimo izraziti v bakteriji *Escherichia coli* K12. Kodoni za katere aminokislinske ostanke bi lahko bili pri tem problematični? V primeru, da bi bila količina tRNA za redke kodone izredno kritična – pri katerem aminokislinskem ostanku bi se translacija ustavila?

Potek dela: poiščite cDNA zapis za histatin 3 --> identificirajte CDS --> zaporedje CDS uporabite za analizo v Graphical Codon Usage Analyser: <http://gcua.schoedl.de> (obe možnosti: “each triplet position vs. usage table” in “each codon vs. usage table”). POZOR: obvezno vpišite ime zaporedja (polje “sequencename” na vrhu spletnega obrazca), sicer vam bo program javil napako!

Odgovor

1. Zapis v bazi Nucleotide (NCBI): <http://www.ncbi.nlm.nih.gov/nuccore/BC095438.1>

Koda zapisa: BC095438

CDS (od start do stop kodona, vključno z njima): 75-230

Zaporedje:

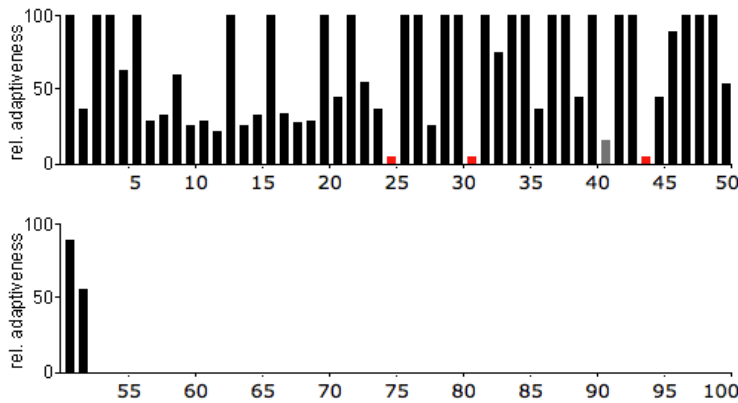
```
>histatin-3_Hs_cds
```

```
ATGAAGTTTTTTGTTTTTGCTTTAATCTTGCTCTCATGCTTTCCATGACTGGAGCTGATTCACATGCA
```

```
AAGAGACATCATGGGTATAAAAAGAAAATTCCATGAAAAGCATCATTCACATCGAGGCTATAGATCAAAT
```

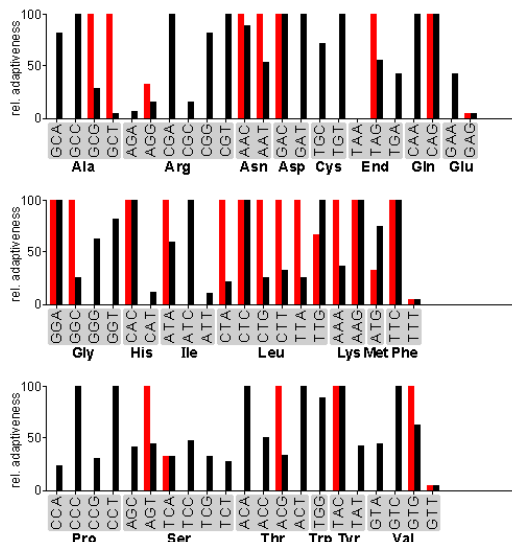
```
TATCTGTATGACAATTGA
```

2. Analiza “each triplet position vs. usage table” (meje: 20% sivo, 10%rdeče – privzete nastavitve):



Opazimo, da bi problem lahko nastal pri 25., 31., in 44., potencialno pa tudi pri 41. aminokislinskem ostanku (vsi so Arg), saj so kodoni, ki so uporabljeni na teh mestih, v izbranem organizmu zelo redko uporabljeni. Na teh mestih bi se translacija lahko ustavila.

3. Analiza “each codon vs. usage table”:



Rdeči stolpci označujejo pogostost uporabe kodona v vstavljenem zaporedju ob upoštevanju izhodnega organizma (*Homo Sapiens*), črni pa v tarčnem organizmu (*Escherichia coli* K12). Neskladja so lahko problematična v primerih, ko imamo visok rdeč in nizek črn stolpec. V teh primerih bi bilo smiselno razmisliti o zamenjavi kodona za katerega izmed tistih, ki kodirajo za isto aminokislino, vendar imajo visok črn stolpec.

V05-03

Iz plazmidnega konstrukta piMP0A-Ep123TC želimo z pomočjo restrikcijskih endonukleaz izrezati najmanjši fragment, ki bo vseboval najdaljši odprti bralni okvir v celoti, pri tem pa želimo, da pri rezanju zraven tega fragmenta nastane le še en fragment, ki ustreza preostanku plazmida; le-tega bomo uporabili za nadaljnje delo. Kateri restrikcijski nukleazi lahko uporabimo, da dosežemo navedeno? Zaporedje plazmidnega konstrukta je v datoteki "V05 - piMP0A-Ep123TC.txt". Za analizo uporabite le tiste restriktaze, katerih prepoznavno mesto je dolgo 6 ali več baznih parov.

Potek dela: iz datoteke skopirajte nukleotidno zaporedje --> identificirajte najdaljši odprti bralni okvir (ORF Finder) --> celotno zaporedje konstrukta prilepite v obrazec orodja Webcutter (<http://rna.lundberg.gu.se/cutter2/>) --> nastavite ustrezne parametre analize --> analizirajte rezultate. POZOR: vse podatke (odprti bralni okvir – lokacija, ...) si sproti zapisujte, saj jih boste potrebovali v nadaljevanju.

Odgovor

Analiza zaporedja s programom ORF Finder najde najdaljši bralni okvir med 352 in 1296 nukleotidom, dolžine 945 nukleotidov.

Frame	from	to	Length
+1	352	1296	945

Restrikcijski endonukleazi moramo izbrati tako, da ne bodo rezale v območju znotraj najdaljšega odprtega bralnega okvirja, hkrati pa je zaželeno, da režejo čim bližje pred začetkom oz. za koncem okvirja. Zaporedje nato skopiramo v orodje Webcutter in ustrezno nastavimo parametre.

Please select the type of analysis you would like

- Linear sequence analysis
- Circular sequence analysis
- Find sites which may be introduced by silent mutagenesis

Please indicate how you would like the restriction sites displayed

- Map of restriction sites
- Table of sites, sorted alphabetically by enzyme name
- Table of sites, sorted sequentially by base pair number

Please indicate which enzymes to include in the display

- All enzymes
- Enzymes not cutting
- Enzymes cutting once
- Enzymes cutting exactly times
- Enzymes cutting at least times, and at most times
- Rainbow highlights for enzymes from the Standard polylinker

Please indicate which enzymes to include in the analysis

- All enzymes in the database
- Only enzymes with recognition sites equal to or greater than bases long

Glede na našo nalogo si bomo najlažje pomagali z razporeditvijo po položaju nukleinskih kislin, ki jih prepoznajo (»Table by Site Position«). Najprej izberemo restriktazo, ki reže najbližje pred mestom 352. Temu kriteriju ustrezajo naslednje:

Position	Enzyme	Cuts	Recognition Site	More info
333	XmaIII	1	C/GGCCG	More info
333	NotI	1	GC/GGCCG	More info
333	Eco52I	1	C/GGCCG	More info
350	Bpu1102I	1	GC/TNAGC	More info
350	BlpI	1	GC/TNAGC	More info
350	Bsp1720I	1	GC/TNAGC	More info
350	CelII	1	GC/TNAGC	More info
358	KasI	1	G/GCGCC	More info
359	NarI	1	GG/CGCC	More info
360	EcoRV	1	GAATTC	More info

Nato izberemo še restriktazo, ki reže čimprej za 1296.

Position	Enzyme	Cuts	Recognition Site	More info
1127	EcoRI	1	G/AATTC	More info
1301	AvaI	1	C/ycgrg	More info
1301	BsoBI	1	C/ycgrg	More info
1301	XhoI	1	C/tcgag	More info
1301	PaeR7I	1	C/tcgag	More info
1301	Sfr274I	1	C/tcgag	More info
1301	Ama87I	1	C/ycgrg	More info
1301	Eco88I	1	C/ycgrg	More info
1301	BcoI	1	C/ycgrg	More info
1307	SalI	1	G/tcgac	More info

Ustrezna je izbira kateregakoli para izmed zgoraj izbranih restrikcijskih endonukleaz.

V05-04

Načrtajte par začetnih oligonukleotidov, s katerima bi s PCR lahko pomnožili najdaljši odprti bralni okvir v celoti (točno od START do STOP kodona). Upoštevajte splošne smernice za načrtovanje oligonukleotidov (temperatura tališča, dimeri, sekundarne strukture, ...), pri načrtovanju pa si pomagajte s spletnim orodjem OligoCalc (<http://www.basic.northwestern.edu/biotools/oligo.html>). Kot končen rezultat navedite zaporedji obeh oligonukleotidov (dopišite, kateri je smeren in kateri protismeren), seveda v smeri 5'-->3'.

Odgovor

Začetni oligonukleotidi naj bi bili dolgi med 20 in 25 nt. Za začetek si zato izpišemo prvih 30 in zadnjih 30 nukleotidov v najdaljšem bralnem okvirju.

Prvih 30: ATGGCGCCCCCGCAGGTCCTCGCGTTCGGG

Zadnjih 30: GTGAGATGCATAGGGAAGCAATGCATAA

Sedaj želimo dobiti dva oligonukleotida s čimbolj podobnima T_m, pri čemer želimo da je T_m v območju od 55 do 70 °C. Pri napovedovanju T_m si pomagamo z orodjem OligoCalc.

Začnemo s prvimi 20 nt in analiziramo T_m (upoštevamo »Salt Adjusted«!). Rezultat je 68,7 °C kar je ravno še sprejemljivo. To je sedaj naše izhodišče za začetni smerni oligonukleotid:

ATGGCGCCCCCGCAGGTCCT

Nato analiziramo zadnjih 20 nt (POZOR, kopirajte od desne proti levi!). T_m je 52,3 °C, kar je bistveno manj od izhodišča za začetni smerni oligonukleotid. Zato poskusimo kar z vsemi 30 nt. Rezultat je 69,9 °C, kar pomeni da lahko kakšen nukleotid spustimo. Zato zberemo po en nukleotid z leve in spremljamo spremembo T_m. Izkaze se, da je dovolj če izbrišemo že enega in dobimo T_m 67,4 °C. Izhodišče za protismerni začetni oligonukleotid je torej:

GTGAGATGCATAGGGAAGCAATGCATAA

Za še boljše ujemanje T_m začetnemu smernemu oligonukleotidu zberemo še en nukleotid na desnem koncu. Končni izhodišči sta torej

Smerni: ATGGCGCCCCCGCAGGTCCT

Protismerni: GTGAGATGCATAGGGAAGCAATGCATAA

Edino kar nam še preostane je protismerni oligonukleotid zapisati v pravilni orientaciji. Ker gre za protismernega, mora biti usmerjen tako, da bo 3'-konec obrnjen proti zaporedju, ki ga želimo prepisovati in ne proti koncu. V primeru, da smo se priprave oligonukleotidov lotili tako, kot je opisano v tem postopku, to preprosto dosežemo tako, da naredimo obratno komplementarno zaporedje že pripravljene izhodišča za protismerni oligonukleotid (lastnosti se pri tem ne spremenijo). Končni rezultat je torej:

Smerni oligonukleotid: 5'- ATGGCGCCCCCGCAGGTCCT-3'

Protismerni oligonukleotid: 5'- TTATGCATTTAGTTCCTATGCATCTCAC-3'

Analiza nastanka dsDNA:

Pri protismernem oligonukleotidu nam analiza pokaže možnost nastanka lasnic in dimerizacije. Znotraj oligonukleotida se namreč zaporedje ATGCAT dvakrat ponovi. Popolna rešitev bi seveda bila, da eno izmed ponovitev ne vključimo v oligonukleotid, vendar to v tem primeru ni mogoče, saj bi bil potem oligonukleotid bistveno prekratek. Ker gre le za 6 nt bodo take strukture pri ustrezni temperaturi, kot je temperatura 5 °C pod T_m najverjetneje razpadle, vseeno pa je nastanek dsDNA lahko ena izmed razlag v primeru neuspešne rekacije.

Postopek se vam sprva morda zdi zelo zapleten, vendar ga boste z vajo začeli razumeti. V kolikor sami najdete drugačen pristop za pripravo oligonukleotidov je to tudi vredno, dokler bodo rezultati pravilni (načinov konstruiranja je namreč več)!

V05-05

Podobno kot pri prejšnji nalogi načrtajte par začetnih oligonukleotidov, s katerima bi s PCR lahko pomnožili najdaljši odprti bralni okvir v celoti (točno od START do STOP kodona), le da tokrat uporabite spletno orodje Primer3 (<http://primer3.ut.ee/>). Kot matrico uporabite celotno zaporedje plazmidnega konstrukta. Končen rezultat načrtovanja naj bosta zaporedji obeh oligonukleotidov (dopišite, kateri je smeren in kateri protismeren), seveda v smeri 5'-->3'.

Skopiramo celotno zaporedje v ustrezno okno

Paste source sequence below (5'→3', string of ACGTNacgtn -- other letters treated as N -- numbers and blanks ignored). FAST.

NONE

```
CTCATGGTTATGGCAGCACTGCATAATTCTTACTGTGCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG
AGAATAGTGATGCGGGCAGCCAGTTGCTCTTGCCCGGGTCAATACGGGATAATACCGGCCACATAGCAGAACCTTAAAAGTGCTCATCATTG
GAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTGATCTTCAGCA
TCTTTTACTTTTACCACGCGTTTCTGGGTGAGCAAAAACAGGAAGGCCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACT
CATACTCTTCCTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAA
TAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACTATAAAAAATAGGCATATCAGG
AGGCC
```

Nato pa izberemo od kje do kje naj oligonukleotida pomnožita. Odločimo se da morata pomnožiti celoten najdaljši odprti bralni okvir. Med katerima mestoma se nahaja smo že ugotovili!

[Force Left Primer Start](#) [Force Right Primer Start](#)
[Force Left Primer End](#) [Force Right Primer End](#)

Za pridobitev optimalnih rezultatov se nekoliko poigramo še z omejitvami za lastnosti oligonukleotidov:

[Primer Size](#) Min Opt Max
[Primer Tm](#) Min Opt Max [Max Tm Difference](#) [Table of thermodynamic parameters](#) SantaLucia 1998
[Product Tm](#) Min Opt Max
[Primer GC%](#) Min Opt Max

Ko zaključimo stisnemo »Pick Primers«. Primer rezultata:

```
PRIMER PICKING RESULTS FOR piMPOA-Ep123TC

No mispriming library specified
Using 1-based sequence positions
OLIGO      start  len  tm    gc%  any th  3' th  hairpin seg
LEFT PRIMER  352   18   67.90 77.78 26.77 0.00  60.93 ATGGCGCCCCCGCAGGTC
RIGHT PRIMER 1296   30   65.06 43.33 10.62 1.52  51.45 TTATGCATTGAGTTCCTATGCATCTCACC
SEQUENCE SIZE: 3900
INCLUDED REGION SIZE: 3900
```

Vidimo, da sta oligonukleotida zelo podobna tistim, ki smo jih načrtali sami. Zaradi drugačnega algoritma se nekoliko razlikuje izračunana Tm. Če z rezultatom nismo zadovoljni, se lahko vrnemo na prejšnjo stran, spremenimo parametre in poskusimo znova.