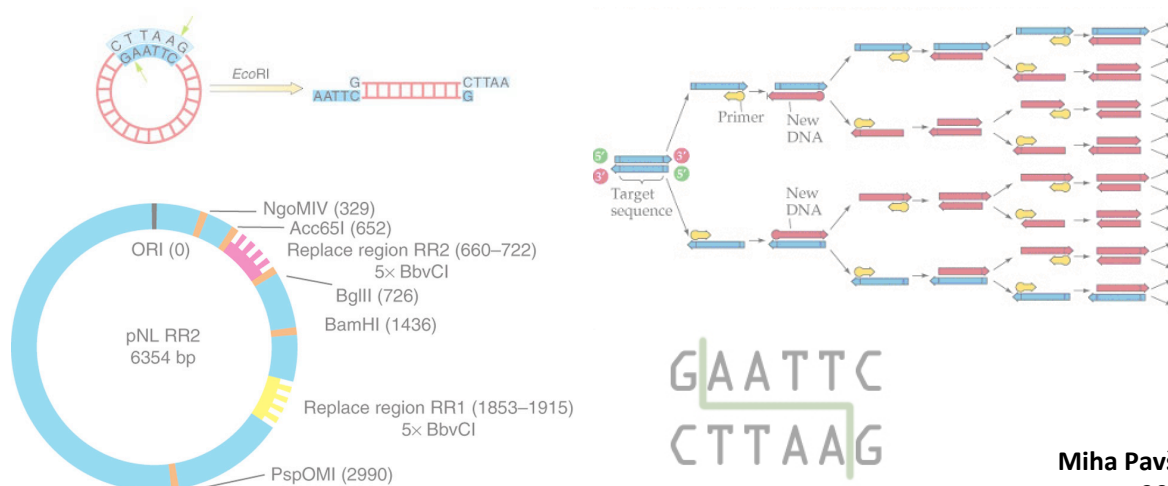


5. vaja

Analiza nukleotidnih zaporedij in molekulsko kloniranje



Miha Pavšič
marec 2014

Pregled vaje

- **raba kodonov**
 - aplikacija tabel rabe kodonov na heterologe gene
- **restriksijske endonukleaze (*restriction endonucleases, restriction enzymes*)**
 - iskanje mest (linearna, krožna dsDNA)
- **plazmidne karte (*plasmid maps*)**
- **verižna reakcija s polimerazo (PCR, *polymerase chain reaction*)**
 - matrica
 - načrtovanje začetnih oligonukleotidov
 - ohranitev bralnega okvirja

v obliki kviza v
spletni učilnici

Raba kodonov

Pri nekaterih organizmih se nekateri kodoni za določen ak-ostanek pojavljajo pogosteje kot drugi kodoni za isti ak-ostanek.

- velja predvsem za **hitro rastoče organizme** (→ optimizacija rabe kodonov)
 - višja hitrost in natančnost translacije
 - najbolj izrazita je v genih, ki se močno izražajo
 - raba kodonov povezava s količino tRNA za določen ak-ostanek!
- pri počasi rastočih organizmih (npr. *Homo sapiens*) ali pri organizmih z majhnimi genomi optimizacija rabe kodonov ni izrazita

Problem v praksi:

- pri organizmu Y je raba kodonov optimirana, pri organizmu X pa ne (oz. je drugačna)
- v organizmu Y želimo izraziti gen iz organizma X (*heterologno izražanje*)

↓ pomanjkanje določene/ih tRNA

- počasnejša translacija
- napake pri translaciji
- prehitra prekinitvev translacije → skrajšane oblike proteina

Raba kodonov

Primer: raba kodonov pri *Escherichia coli* K12.

Od šestih kodonov za Arg se le dva (CGU in CGC) uporabljata pogosto.

UUU F 0.57 19.7	UCU S 0.11 5.7	UAU Y 0.53 16.8	UGU C 0.42 5.9
UUC F 0.43 15.0	UCC S 0.11 5.5	UAC Y 0.47 14.6	UGC C 0.58 8.0
UUA L 0.15 15.2	UCA S 0.15 7.8	UAA * 0.64 1.8	UGA * 0.36 1.0
UUG L 0.12 11.9	UCG S 0.16 8.0	UAG * 0.00 0.0	UGG W 1.00 10.7
CUU L 0.12 11.9	CCU P 0.17 8.4	CAU H 0.55 15.8	CGU R 0.36 21.1
CUC L 0.10 10.5	CCC P 0.13 6.4	CAC H 0.45 13.1	CGC R 0.44 26.0
CUA L 0.05 5.3	CCA P 0.14 6.6	CAA Q 0.30 12.1	CGA R 0.07 4.3
CUG L 0.46 46.9	CCG P 0.55 26.7	CAG Q 0.70 27.7	CGG R 0.07 4.1
AUU I 0.58 30.5	ACU T 0.16 8.0	AAU N 0.47 21.9	AGU S 0.14 7.2
AUC I 0.35 18.2	ACC T 0.47 22.8	AAC N 0.53 24.4	AGC S 0.33 16.6
AUA I 0.07 3.7	ACA T 0.13 6.4	AAA K 0.73 33.2	AGA R 0.02 1.4
AUG M 1.00 24.8	ACG T 0.24 11.5	AAG K 0.27 12.1	AGG R 0.03 1.6
GUU V 0.25 16.8	GCU A 0.11 10.7	GAU D 0.65 37.9	GGU G 0.29 21.3
GUC V 0.18 11.7	GCC A 0.31 31.6	GAC D 0.35 20.5	GGC G 0.46 33.4
GUA V 0.17 11.5	GCA A 0.21 21.1	GAA E 0.70 43.7	GGA G 0.13 9.2
GUG V 0.40 26.4	GCG A 0.38 38.5	GAG E 0.30 18.4	GGG G 0.12 8.6

kodon za posamezni ak-ostanek kodon ak **XXX** **YYY** frekvenca na 1000 kodonov (za vse ak)
 enočrkovna koda ak-ostanka frekvenca rabe tega kodona (za to ak)

Raba kodonov

Primerjava: raba kodonov za Arg pri *Escherichia coli* K12 in človeku.

CGU	R	0.36	21.1
CGC	R	0.44	26.0
CGA	R	0.07	4.3
CGG	R	0.07	4.1

AGU	S	0.14	7.2
AGC	S	0.33	16.6
AGA	R	0.02	1.4
AGG	R	0.03	1.6

E. coli K12

Izrazita preferenca za dva kodona!

CGU	R	0.08	4.5
CGC	R	0.18	10.4
CGA	R	0.11	6.2
CGG	R	0.20	11.4

AGU	S	0.15	12.1
AGC	S	0.24	19.5
AGA	R	0.21	12.2
AGG	R	0.21	12.0

Homo sapiens
(standardni kod)

Ni izrazite preference!

kodon za posamezni ak-ostanek kodon ak XXX YYY frekvenca na 1000 kodonov (za vse ak)
 enočrkovna koda ak-ostanka frekvenca rabe tega kodona (za to ak)

Raba kodonov

Vaja:

- ogled rabe kodonov pri različnih organizmih
- primerjava rabe kodonov v določenem genu iz nekega organizma z rabo kodonov v drugem organizmu

Codon Usage Database

<http://www.kazusa.or.jp/codon/>

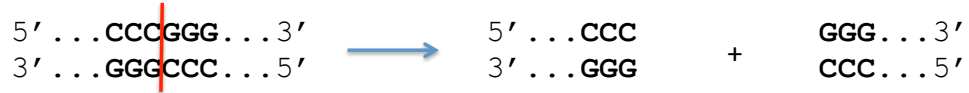
g raphical
c odon
u sage
a nalyser

<http://gcua.schoedl.de>

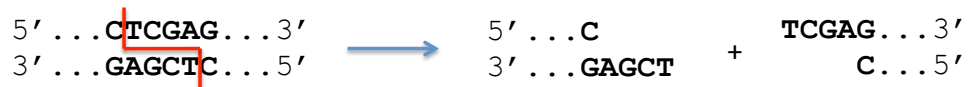
Restriksijske endonukleaze

- prepoznajo določeno mesto na DNA (**restriksijsko mesto**, *restriction site*) in ga razrežejo

- **topi konci** (*blunt ends*), npr. *SmaI*



- **lepljivi koncu** (*sticky ends*), npr. *XhoI*



- zaporedja pogosto **palindromna**
- imena restriksijskih endonukleaz izvirajo iz imen (mikro)organizmov, iz katerih so bile izolirane, npr. *EcoRI* – *Escherichia coli*
- ti encimi so izredno uporabni pri molekularnem kloniranju – usmerjeno sestavljanje fragmentov DNA (npr. vstavljanje vključka v plazmidni vektor)

Restriksijske endonukleaze

Vaja:

- restriksijska analiza (*restriction analysis*) plazmidnega konstrukta

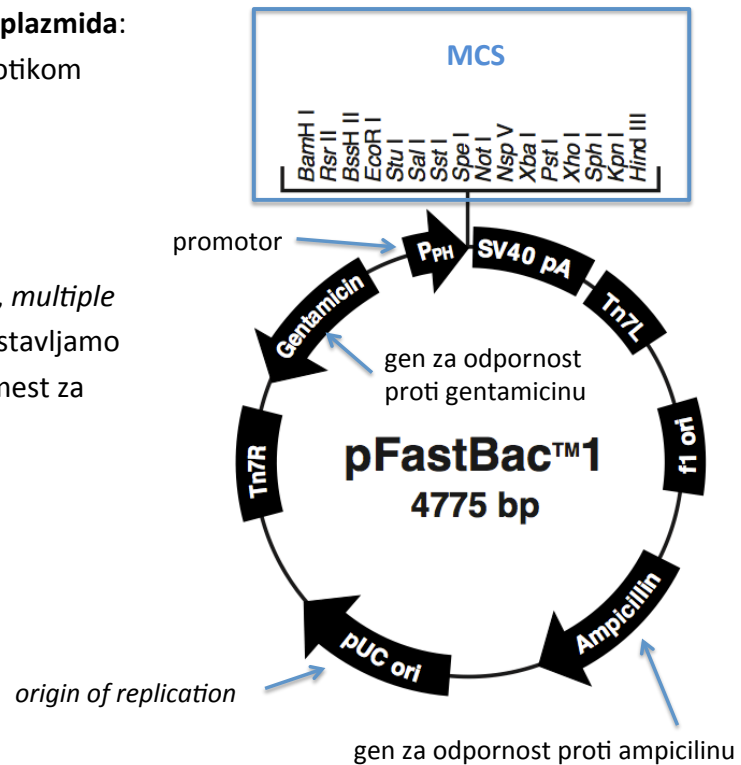


<http://rna.lundberg.gu.se/cutter2/>

Plazmidne karte

S plazmidnimi kartami **grafično ponazorimo lokacijo in obseg določenih regij plazmida:**

- geni za odpornost proti antibiotikom
- ori (*origin of replication*)
- promotorji
- ...
- vključki
- multiplo klonirno mesto (MCS, *multiple cloning site*) – sem ponavadi vstavljamo vključke preko prepoznavnih mest za restrikcijske endonukleaze

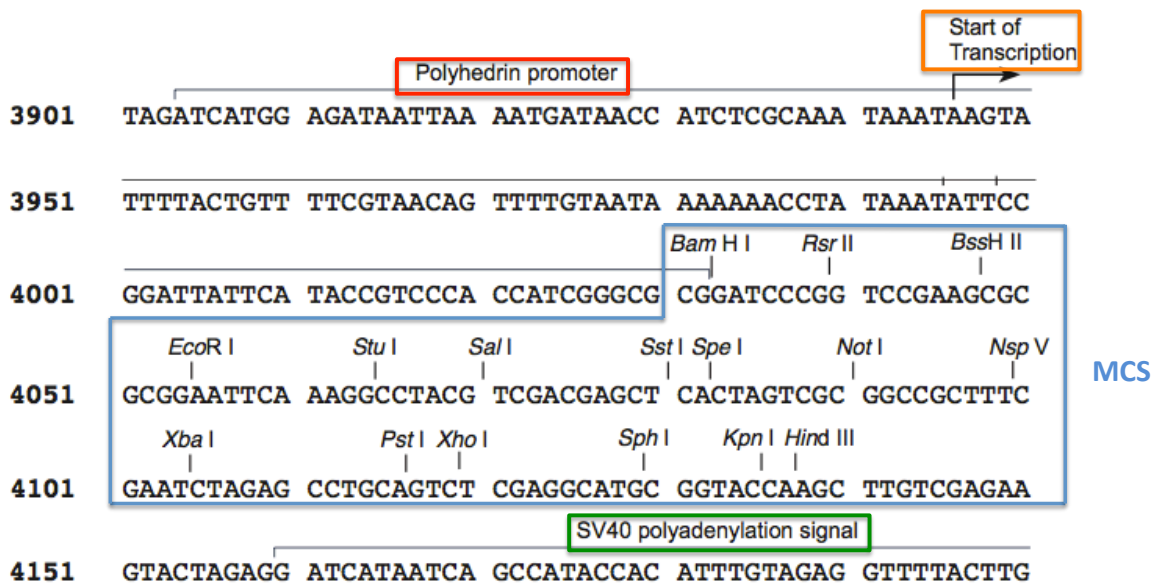


Plazmidne karte

Multiplo klonirno mesto

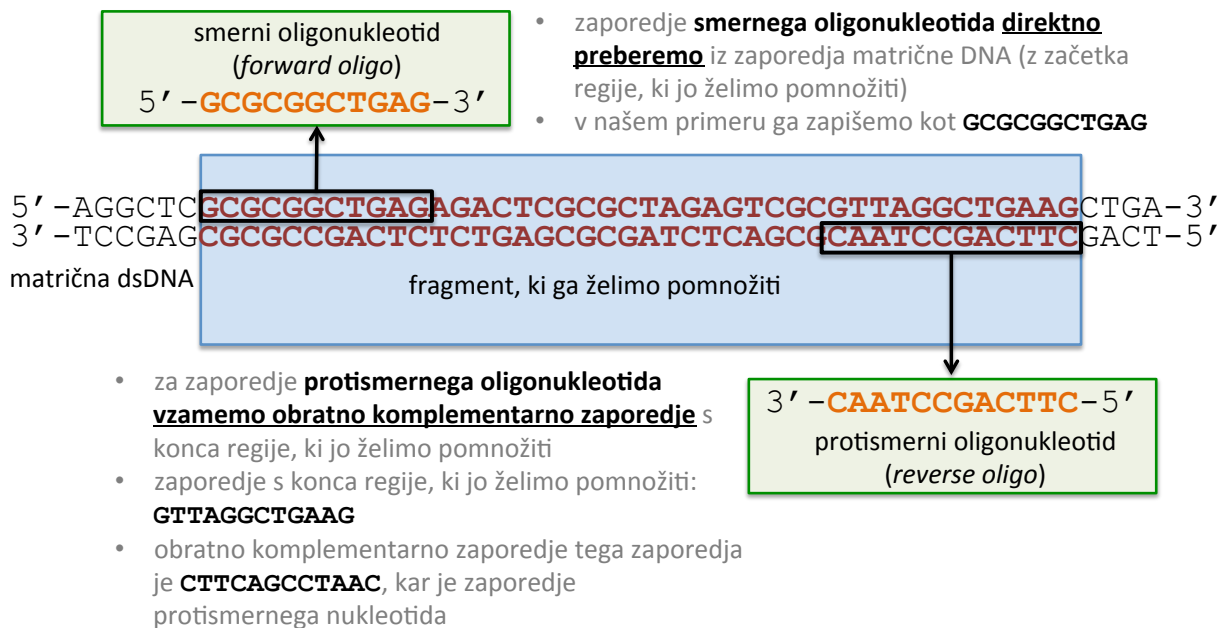
Primer vektorja za izražanje proteinov v insektnih celicah: pFastBac 1.

Vključek z genom, ki ga želimo izraziti, vstavimo preko mest od *Bam*HI do *Hind*III (vključek mora imeti START in STOP kodon!).



Verižna reakcija s polimerazo (PCR)

Načrtovanje začetnih oligonukleotidov



Na 5'-konca oligonukleotidov lahko dodamo druga zaporedja, ki niso komplementarna matrici (npr. restrikcijska mesta za usmerjeno vstavljanje v plazmidne vektorje) → ta zaporedja potem dobimo v končnem PCR produktu.

Verižna reakcija s polimerazo (PCR)

Načrtovanje začetnih oligonukleotidov

Splošne smernice za izbiro optimalnih začetnih oligonukleotidov:

- dolžina 20-25 nt
- temperatura tališča (T_m) 55-70°C → pomembno za prileganje (*annealing*), na splošno velja $T_a = T_m - 5^\circ\text{C}$
- večji delež GC → višja T_m
- oligonukleotidi naj med seboj ne tvorijo dsDNA:
 - sekundarne strukture (regije dsDNA znotraj istega oligonukleotida)
 - dimeri oligonukleotidov (regije dsDNA med dvema oligonukleotidoma)
- na 3'-koncu naj bo G ali C (tri vodikove vezi v bp – stabilnejše!)
 - 3'-konec mora tvoriti par z matrico, sicer ne pride do polimerizacije
- specifičnost samo za izbrano DNA (uravnavamo tudi s T_a)

Izračun T_m

Za kratke oligonukleotide, < 14 nukleotidov, v prisotnosti 50 mM soli

$$T_m = 4^\circ\text{C} \times (\text{število G in C v oligonukleotidu}) + 2^\circ\text{C} \times (\text{število A in T v oligonukleotidu})$$

Za daljše oligonukleotide (N je dolžina oligonukleotida)

$$T_m = 64.9^\circ\text{C} + 41^\circ\text{C} \times (\text{število G in C v oligonukleotidu} - 16.4)/N$$

Za daljše oligonukleotide ob upoštevanju koncentracije soli

$$T_m = 81.5^\circ\text{C} + 16.6^\circ\text{C} \times (\log_{10}[\text{Na}^+] + [\text{K}^+]) + 0.41^\circ\text{C} \times (\%GC) - 675/N$$

Verižna reakcija s polimerazo (PCR)

Potek PCR (poenostavljen primer): vezava začetnih oligonukleotidov

5' -AGGCTC**GCGCGGCTGAGAGACTCGCGCTAGAGTCGCGTTAGGCTGAAG**CTGA-3'
 3' -TCCGAG**CGCGCCGACTCTCTGAGCGCGATCTCAGCGCAATCCGACTTC**GACT-5'

matrična dsDNA



95°C - denaturacija
 (verigi matrične dsDNA se ločita)

5' -AGGCTC**GCGCGGCTGAGAGACTCGCGCTAGAGTCGCGTTAGGCTGAAG**CTGA-3'

3' -TCCGAG**CGCGCCGACTCTCTGAGCGCGATCTCAGCGCAATCCGACTTC**GACT-5'



~55°C (T_a) - vezava začetnih oligonukleotidov
 (*annealing*, prileganje; optimalno T_a izračunamo iz T_m)
 [temperaturo hitro znižamo s 95°C, da ne nastane matrična dsDNA]

5' -AGGCTC**GCGCGGCTGAGAGACTCGCGCTAGAGTCGCGTTAGGCTGAAG**CTGA-3'
 3' -**CAATCCGACTTC**-5'

smerni oligonukleotid
 (*forward oligo*)

protismerni oligonukleotid
 (*reverse oligo*)

5' -**GCGCGGCTGAG**-3'
 3' -TCCGAG**CGCGCCGACTCTCTGAGCGCGATCTCAGCGCAATCCGACTTC**GACT-5'



(se nadaljuje)

Verižna reakcija s polimerazo (PCR)

Potek PCR (poenostavljen primer): polimerizacija



~72°C - polimerizacija
 (ta temperatura je optimalna za delovanje termostabilnih
 DNA-polimeraz)

5' -AGGCTC**GCGCGGCTGAGAGACTCGCGCTAGAGTCGCGTTAGGCTGAAG**CTGA-3'
 ← **CTCAGCGCAATCCGACTTC**-5'

smerni oligonukleotid
 (*forward oligo*)

protismerni oligonukleotid
 (*reverse oligo*)

5' -**GCGCGGCTGAGAGACTCG** →
 3' -TCCGAG**CGCGCCGACTCTCTGAGCGCGATCTCAGCGCAATCCGACTTC**GACT-5'



5' -AGGCTC**GCGCGGCTGAGAGACTCGCGCTAGAGTCGCGTTAGGCTGAAG**CTGA-3'
 ..AG**GCGCGGCTGAGAGACTCTCTGAGCGCGATCTCAGCGCAATCCGACTTC**-5'

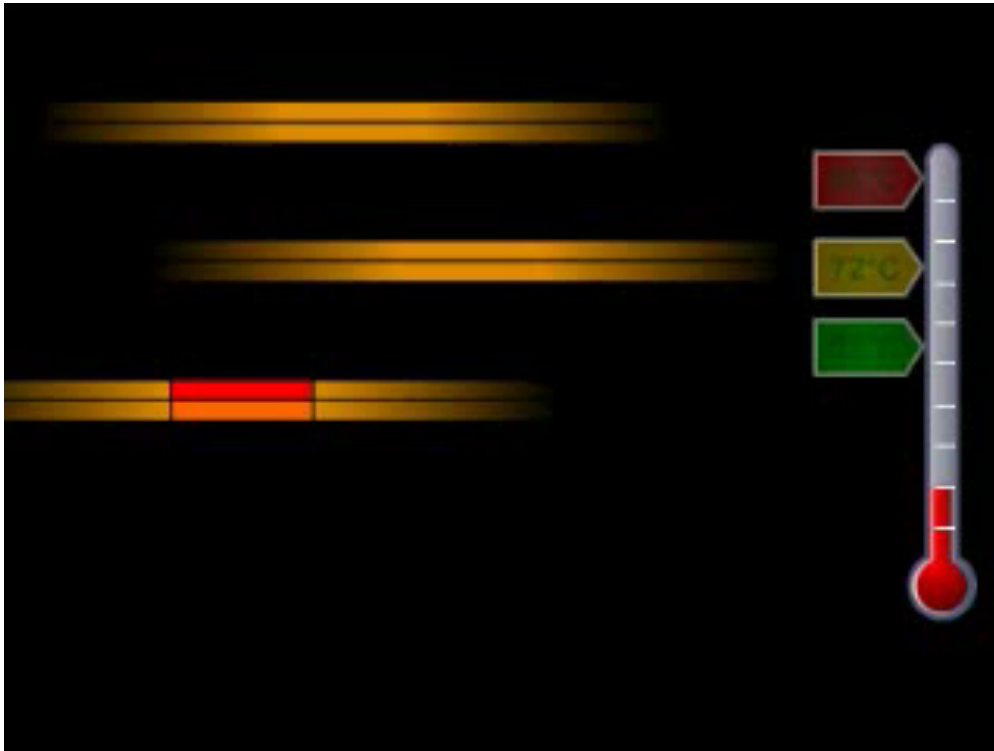
5' -**GCGCGGCTGAGAGACTCGCGCTAGAGTCGCGTTAGGCTGAAG**CT.
 3' -TCCGAG**CGCGCCGACTCTCTGAGCGCGATCTCAGCGCAATCCGACTTC**GACT-5'



ciklično ponavljanje (ponavadi 20-30x):

- ponovna denaturacija (95°C)
- vezava začetnih oligonukleotidov (T_a)
- polimerizacija (72°C)

Verižna reakcija s polimerazo (PCR)



Vir: YouTube (https://www.youtube.com/watch?v=eEcy9k_KsDI)

Verižna reakcija s polimerazo (PCR)

Vaja:

- načrtovanje začetnih oligonukleotidov – samostojno
 - identifikacija regije, ki jo želimo pomnožiti
 - izpis zaporedij oligonukleotidov in analiza njihovih lastnosti (T_m , dimeri, ...)
- načrtovanje začetnih oligonukleotidov z orodji (Primer3, Primer-BLAST)
 - identifikacija regije, ki jo želimo pomnožiti
 - ustreznost nastavitve parametrov → program sam predlaga optimalne oligonukleotide

<http://www.basic.northwestern.edu/biotools/oligocalc.html>

<http://primer3.ut.ee>

<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>
(primerjanje zaporedij oligonukleotidov z bazami zaporedij nt – za zmanjšanje nespecifične vezave)