

Primer Designer

Tanja Gošnjak

Univerza v Ljubljani, Fakulteta za kemijo in kemijsko tehnologijo

Povzetek

V sklopu natečaja za študentski IKT projekt sem obravnavala težavo, povezano z načrtovanjem začetnih oligonukleotidov za vstavljanje genov v vektorje, zato predlagam izdelavo aplikacije, ki lahko opravlja vse potrebne naloge hkrati. Aplikacija bi vključevala funkcionalnosti za preverjanje združljivosti restrikcijskih encimov, prisotnost START in STOP kodonov, prilagajanje bralnega okvirja, izračun temperature tališča in preverjanje morebitnih sekundarnih struktur DNA. Uporabnik bi vnesel zaporedje gena, zaporedje MCS izbranega vektorja in restrikcijske encime, program pa bi izpisal seznam ustreznih encimov in zasnovanih začetnih oligonukleotidov na podlagi uporabnikovih zahtev. Ta aplikacija bi lahko bila uporabna tako za začetnike kot za izkušene raziskovalce.

Motivacija

Študentje se z načrtovanjem začetnih oligonukleotidov (angl. primer) za vstavljanje gena v vektor srečamo pri študiju, kar pa se marsikomu lahko zdi zapleteno. Pri tem je potrebno upoštevati številne dejavnike, kot so načrtovanje previsnih koncev (angl. overhang), preverjanje kompatibilnosti encimov, prisotnost START in STOP kodona ter prilagajanje bralnega okvirja. Poleg tega je potrebno izračunati temperaturo tališča začetnih oligonukleotidov in preveriti morebitne tvorbe sekundarnih struktur DNA, kar lahko predstavlja dodatne težave.

Čeprav so na voljo aplikacije, ki omogočajo izvedbo posameznih korakov pri načrtovanju začetnih oligonukleotidov (preverjanje temperature tališča, kompatibilnosti encimov...), pa nisem našla aplikacije, ki bi omogočala izvedbo vseh nalog hkrati. Zato sem se odločila ustvariti aplikacijo, ki bi bila zmožna izvesti vse naloge hkrati ter s tem olajšati in pospešiti postopek načrtovanja.

Moja aplikacija bi omogočala hitro in natančno načrtovanje začetnih oligonukleotidov na podlagi vseh omenjenih zahtev. Poleg tega vključuje tudi funkcionalnosti za preverjanje kompatibilnosti encimov, prisotnosti START in STOP kodona, prilagajanje bralnega okvirja, izračun temperature tališča in preverjanje tvorbe morebitnih sekundarnih struktur DNA.

Aplikacija

Uporabnik sledi navodilom, ki jih aplikacija od njega zahteva (Slika 1). Najprej vnese zaporedje gena, ki ga želi vstaviti v izbrani vektor, in zaporedje območja z več mesti za kloniranje (angl. multiple cloning site, MCS) iz vektorja, kjer se nahaja več različnih cepitvenih mest za restrikcijske endonukleaze, ki služijo kot mesta, kamor vstavimo tujo DNA. Program izpiše nabor vseh encimov, ki cepijo MCS enkrat, in ne cepijo zaporedja za naš gen, kar nam je lahko v pomoč pri iskanju ustreznih encimov. Nato vpišemo imena encimov, ki jih imamo na voljo v laboratoriju.

```

Enter the gene sequence: ATGACTACGGTCGACAACATCGTCGGGTGGTGATCGCGGTGGCGCTAATGGCGTTCCCTATTGCGGGCGCTGCTGTTCCGGAGAAAGTTCTGA
Enter the multicloning site (MCS) sequence:
TGCCTGCAGGTGCGACTCTAGAGGATCCCGGGTACCGAGCTCGAAATCaTGCCGCTGTTTTACAACGTCGTCGACTGGGAAAACCCCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTT
CGCCAGCTGGCGTAATAGCGAAGAGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTCTCCTTACGCATCTGTCGGGATTTTCACACCGCAT
ATGGTGACACTCTCAGTACAATCTGCTGTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
CGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACCCTCATCACGAAACGCGGAGACGAAAGGGCCTCGTGTATACGCCCTATTTTATtgAGGTTAATGTACGATAAATAATGGTTCTTAGACGTC
GGTGGCACTTTTCGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCGTATAAATGCTTCAATAATTTGAAAAAGGAAAGT
ATGAGTATTC AACATTTCCGCTGTCGCCCT
Enzymes that cut MCS once and never cut the gene: ['AccBSI', 'PspGI', 'ZraI', 'BstSLI', 'SfcI', 'PstI', 'FspI', 'AvaI', 'MbiI',
'PvuII', 'BglI', 'BsrBI', 'BstNI', 'PleI', 'Acc65I', 'EcoRI', 'Psp6I', 'NspI', 'SfoI', 'TspRI', 'Eco53kI', 'BfaI', 'FriOI',
'BciI130I', 'BstNSI', 'HinfI', 'BamHI', 'Alw44I', 'BstAPI', 'BmuI', 'Psp124BI', 'XbaI', 'FauNDI', 'HpyAV', 'AatII', 'BseBI',
'SchI', 'BaeGI', 'BstSFI', 'MvaI', 'ApaLI', 'Asp718I', 'BsuI', 'Bst2BI', 'NdeI', 'TasI', 'BstX2I', 'BssSI', 'XapI', 'Bst2UI',
'XspI', 'BseMII', 'TatI', 'SdaI', 'SspDI', 'AcsI', 'SacI', 'Ple19I', 'BspMAI', 'HgaI', 'BsoBI', 'Mly113I', 'BseSI', 'EcoT38I',
'DinI', 'BstYI', 'SstI', 'CseI', 'BfuI', 'BmeT110I', 'TspGWI', 'Sse9I', 'BciVI', 'BceAI', 'XmaI', 'EheI', 'Ama87I', 'FspBI',
'BtsIMutI', 'SspMI', 'PfoI', 'Eco24I', 'BauI', 'BanII', 'SbfI', 'VneI', 'SmiMI', 'EcoRII', 'EaeI', 'Eco88I', 'SspI', 'BsiHKCI',
'Acc16I', 'BspCNI', 'Ecl136II', 'Sse8387I', 'PvuI', 'MlyI', 'EgeI', 'RseI', 'PsuI', 'TscAI', 'NsbI', 'ApoI', 'MaeI', 'BfmI',
'BmrI', 'Mfli', 'MslI', 'TspMI', 'MluCI', 'NarI', 'SmaI', 'KasI', 'PpsI', 'AcoI', 'Eco109I', 'KpnI', 'Cfr9I', 'EcoICRI',
'PluII', 'AjnI', 'XceI'].
Enter available enzymes (separated by a comma):

```

Slika 1: Primer izpisa kompatibilnih restriktaz za vstavitev gena za protein KdpF *M. tuberculosis* v MCS vektorja pUC19.

Program izpiše seznam primernih encimov in jih razvrsti od 5' proti 3' koncu MCS glede na restriksijsko mesto, ki je zapisano v kvadratnem oklepaju (Slika 2). Na podlagi tega se odločimo kateri encim bomo uporabili za cepitev na 5' in 3' koncu MCS. Program izpiše vse START kodone, ki se nahajajo pred restriksijskim mestom za izbran encim na 5' koncu in vse STOP kodone, vključno z njihovim zaporedjem, ki se nahajajo za restriksijskim mestom za izbrani encim na 3' koncu. Program nato izpiše prepoznavna mesta za izbrana encima.

```

Enter available enzymes (separated by a comma): FspI, ApaLI, TatI, AatII, SspI, ZraI, XbaI, HincII, EaeI, FspI, SmaI, BmrI, EcoRI,
NdeI, SacI
The appropriate enzymes are: ['FspI', 'ZraI', 'SspI', 'XbaI', 'AatII', 'TatI', 'EaeI', 'SmaI', 'BmrI', 'NdeI', 'SacI', 'ApaLI',
'EcoRI']
The available enzymes cleave the MCS from the 5' end to the 3' end (enzyme: cleavage site): {XbaI: [17], SmaI: [30], SacI: [42],
EcoRI: [44], EaeI: [52], BmrI: [86], FspI: [186], NdeI: [258], ApaLI: [263], TatI: [273], ZraI: [513], AatII: [515], SspI: [629]}
At the 5' end, I will use a restriction enzyme: XbaI
At the 3' end, I will use a restriction enzyme: BmrI
The sequence before the restriction site for XbaI does not contain the start codon (ATG).
The sequence after the restriction site for BmrI contains stop codons at the following positions:
TAA at position 102
TAA at position 141
TAG at position 144
TGA at position 191
TGA at position 209
TGA at position 285
TAG at position 295
TAA at position 299
TGA at position 330
TGA at position 340
TGA at position 383
TGA at position 458
TGA at position 476
TAA at position 482
TGA at position 498
TAA at position 493
TAA at position 496
TAG at position 507
TAA at position 567
TGA at position 595
TAA at position 603
TGA at position 609
TAA at position 612
TAA at position 624
TGA at position 630
TGA at position 646
The recognition sequence for XbaI is TCTAGA.
The recognition sequence for BmrI is ACTGGG.

```

Slika 2: Izpis programa po izbiri restriktaz za vstavitev gena v MCS.

Uporabnik se odloči ali bo uporabil START in/ali STOP kodon iz vektorja in katerega (smiselna je uporaba najbližjega restriksijskemu mestu). V primeru da teh ne izbere, so ti samodejno dodani kot previsni konci na začetni oligonukleotid. Če pred oziroma po restriksijskem mestu ni START oziroma STOP kodona, je tudi takrat ta samodejno dodan na zaporedje načrtovanega začetnega oligonukleotida.

```

Since there is no start codon before the restriction site, it will be added to the primer.
Will you use the given STOP codon? (yes/no): yes
Enter the sequence position of the STOP codon: 102
Forward primer: 5'-TATTCTAGAATGATGACTACGGTCGACAACATC-3'
Tm of forward primer: 62 °C
Length of the complementary primer: 21
Reverse primer: 5'-TATCCCAGTTATCAGAACTTCTCCGGAAACAG-3'
Tm of reverse primer: 62 °C
Length of the complementary primer: 21

```

Slika 3: Izpis programa po izbiri STOP kodona. START kodon je bil dodan samodejno na previsni konec zaradi njegove odsotnosti pred restrikcijskim mestom izbranega encima (zelen pravokotnik). Moder pravokotnik – dodano zaporedje pred restrikcijskim mestom, da je omogočena restrikcija encimu. Oranžen pravokotnik – zaporedje, ki ga prepozna restriktaza Bmrl. Siv pravokotnik – prilagajanje bralnega okvirja. Bel pravokotnik – komplementarno zaporedje gena.

Program izpiše načrtovane oligonukleotide na podlagi uporabnikovih zahtev glede uporabe restriktaz in START ter STOP kodona (Slika 3). Začetni oligonukleotidi imajo omejitev temperature tališča (Tm) med 60 in 65°C ter tvorbe lasnic ali dimerov z največ štirimi zaporednimi nukleotidi. V primeru, ko START oziroma STOP kodon ne ustreza bralnemu okvirju, je na previsni konec dodanih toliko nukleotidov, da bo vstavljen gen ustrezal bralnemu okvirju. S pomočjo programa je omogočeno enostavno in hitro načrtovanje ustreznih začetnih oligonukleotidov za vstavljanje genov v vektorje. Poleg tega bi z uporabo aplikacije prihranili čas, saj vključuje vse korake, ki bi sicer bili ločeni in bi jih morali opraviti s pomočjo več različnih orodij.

Nadaljnji razvoj

V prihodnosti bi ustvarili spletno različico aplikacije. Uporabno bi bilo dodati še anotacijo zaporedja MCS, ker ta lahko že v osnovi vsebuje določena vstavljena zaporedja oznak. Na podlagi tega bi lažje izbrali ustrezne restriktaze za vstavev gena v vektor. Aplikacijo bi lahko naredili uporabniku prijaznejšo, s tem da bi dodali grafični prikaz vektorja z ustreznimi restrikcijskimi mesti in START ter STOP kodoni. Izbira encimov in START/STOP kodonov bi potekala s klikom na izbran element. Dodali bi lahko tudi, da bi uporabnik sam izbiral med območjem temperature tališča in dovoljeno dolžino zaporednih nukleotidov, ki lahko tvorijo sekundarne strukture. Aplikacija Primer Designer je v osnovi namenjena kot pomoč, vendar ima tudi potencial za prodajo na trgu.

Spletna aplikacija bi bila uporabna za začetnike - študente, da lahko preverijo svoje rezultate, kot tudi za znanstvenike in raziskovalce, ki načrtovanje oligonukleotidov uporabljajo pri vsakdanjem delu.