

Molekuláris biológia házi feladat

Kurdi Bogdán, CGP6A5

2016. 06. 07.

Megadott kód:

```
GGGCGAGTTTATCGCCTCAGCCTGACGTCAGCGGGGAGCTAAACACCGAAGCATTGACCTGTATTGTTGA
CGTCTGGAAAGTGAATTATAGATGGCAGCTGGATTCATATCAGTTACGAGGAAACCGATCTCGAAATG
ATGCCTTTTCTTGTGTCACAGGCCAATAAGAAGTATCCAGAGTTAAATCTTAAATTTGTTATGTCAGTCC
ATGAGCTTGTTCCTCTATAAAGGAGACCAGAATGGAAGGCGTTGAATCTGCCCGATTTCTCGTAAATAT
GGGAAGTTCAGGTATCCATATTTTCAGTCGTCGATTTTATGAGTTATGGACGGAAAGACATCGGTGATTTTG
TTCGAACCAGCAGCGTGTAGCGCTTTTGGACCTGCTTTACTGGCGTTGAGGACCAAAGCAGCTCTTGAAC
GTGAACAACCTGCCTGATTGTTATTTGCTATGGTCGAGCTGGACATTCAACGAAGCTCTTCTGAATGCGG
TATTTTGTAGCCTGGCGCTCGCCAAAAAACTTCAGCTTGAATTTATGAACTTAGTAAAAATTCATGAAGAT
AATATTTGTGAACGCTCTGTGTGGTGAAGAACCTTTTCTCCCGTCCGATAAAGCAGACCGCTATCTGCCGG
TGAGTTTTTACAAACATACTCAAGGCGCACAAACGATTAATGAATATGTGGAGGCCAATCCGGCGCGGG
```

Primer tervezés: (<http://bioinfo.ut.ee/primer3-0.4.0/>)

Nem változtattam a beállításokat. A primerek tervezésénél nem volt kijelölt keresendő pontmutáció, vagy rövidebb szekvencia, amit vizsgálnánk, ezért automatikusra állítottam a bal és jobb primer keresését.

Eredmény:

	start	len	tm	gc%	any	3'	seq
1 LEFT PRIMER	242	20	60.07	45.00	4.00	1.00	AATGGAAGGCGTTGAATCTG
RIGHT PRIMER	414	20	60.00	55.00	5.00	1.00	GAGCTGCTTTGGTCTCTCAAC

SEQUENCE SIZE: 700

INCLUDED REGION SIZE: 700

Alternatív primerek:

	start	len	tm	gc%	any	3'	seq
2 LEFT PRIMER	65	20	60.02	50.00	2.00	0.00	TGTTGAGCGTCTGGAAAGTG
RIGHT PRIMER	258	20	60.07	45.00	4.00	1.00	ATTCAACGCCTTCCATTCTG
PRODUCT SIZE: 194, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 3.00							
3 LEFT PRIMER	65	20	60.02	50.00	2.00	0.00	TGTTGAGCGTCTGGAAAGTG
RIGHT PRIMER	261	20	60.07	45.00	4.00	2.00	CAGATTCAACGCCTTCCATT
PRODUCT SIZE: 197, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 3.00							
4 LEFT PRIMER	253	20	60.16	45.00	7.00	1.00	TTGAATCTGCCCGATTTCTC
RIGHT PRIMER	414	20	60.00	55.00	5.00	1.00	GAGCTGCTTTGGTCTCTCAAC
PRODUCT SIZE: 162, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 0.00							
5 LEFT PRIMER	122	20	60.01	45.00	6.00	2.00	GGAAACCGATCTCGAAATGA
RIGHT PRIMER	272	20	60.16	45.00	7.00	3.00	GAGAAATCGGCAGATTCAA
PRODUCT SIZE: 151, PAIR ANY COMPL: 6.00, PAIR 3' COMPL: 2.00							

start = primer kezdőpontjának sorszáma (melyik nukleotidnál kezdődik)

len = primer hossza (minél rövidebb, annál könnyebb előállítani)

tm = olvadáspont (legyen minél kisebb)

gc% = G-C bázisok aránya (összefügg az olvadási hőmérséklettel)

any = primer önkomplementaritása (legyen minél kevesebb)

3' = 3' önkomplementaritás (legyen minél kevesebb)

seq = primer szekvencia 5'3' irányban (a jobb primer fordítva van)

Konklúzió:

- A megtervezett primerek 20 nukleotid hosszúak. Egy primernek legalább 18 hosszúnak kell lennie, és minél hosszabb, annál drágább a szintézis. Ez a 20 egy átlagosan jó hosszúnak tűnik.
- A primerek olvadáspontja 60.00 és 60.16 között változik. Az 5 primerpár közül az elsőnek a legalacsonyabb a hőmérséklete (max. 60.07 °C), a program valószínűleg ez alapján választotta ki ezt a megoldást.
- Az ön-komplementaritási érték miatt az utolsó kettő primerpár elmarad a többtől (nagy a komplementaritás), ezek kizáró feltételek lehetnek a primertervezéskor. A 2. megoldásnál a legideálisabb a komplementaritás, ezek is elég jó primernek tűnnek.

NCBI Blast: (<http://blast.ncbi.nlm.nih.gov>)

A kezdőképernyőn kiválasztjuk a „microbes”-t, mert baktériumok / vírusok genomját keressük (betegség meghatározása céljából). A lekérdezési szekvenciának – ne finomkodjunk – beadtam az egészet (700 nukleotid). A lenti opciók közül a „Highly similar sequences (megablast)”-ot választottam, hogy minél pontosabb eredményt kapjunk.

blastn blastp blastx tblastn

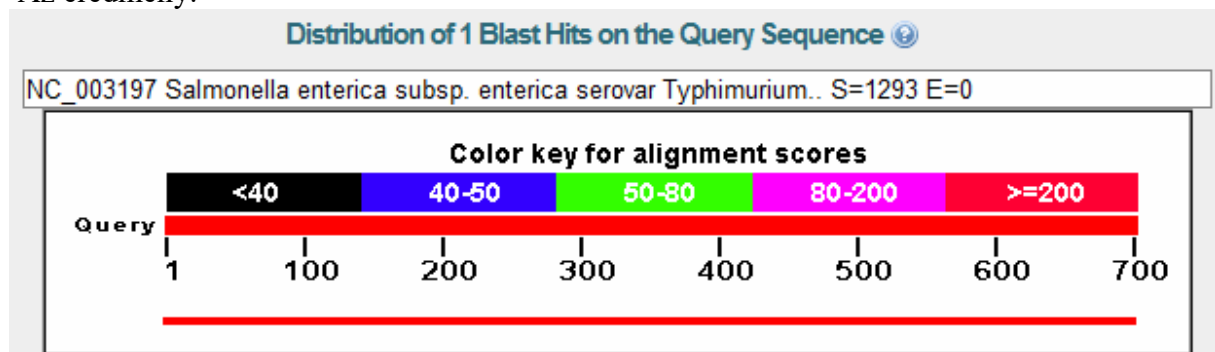
BLASTN programs search nu

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

```
TTCGAACCAGCAGCGTGTAGCGCTTTTGGACCTGCTTACTGGCGTTGAGGACCAAAGCAGCTCTTGAAC
GTGAACAACCTGCCTGATTGTTATTTTGCATGGTCGAGCTGGACATTCAACGAAGCTCTTCTGAATGCGG
TATTTTTAGCCTGGCGCTCGCCAAAAAAGCTTCAAGCTTGAATTTATGAACCTTAGTAAAAATTCATGAAGAT
AATATTTGTGAACGCTGTGTGGTGAAGAACCTTTTCTCCGTCGGATAAAGCAGACCGCTATCTGCCGG
TGAGTTTTTACAAACATACTCAAGGGCGACAACGATTAATGAATATGTGGAGGCCAATCCGGCGCGCGG
```

Az eredmény:



Mindössze egyetlen baktériumot talált, ami egyébként 100%-osan lefedi az adott szekvenciát, nincsenek lyukak (gaps). Az adott szekvencia a baktérium genomjának 3010037. nukleotidjánál kezdődik, 700 nukleotid hosszú.

„Salmonella enterica subsp. enterica serovar Typhimurium str. LT2 chromosome, complete genome”

max score: 1293
total score: 1293
query cover: 100%
E. value: 0.0
Ident: 100%
Accession: NC_003197.1
Identites: 700/700(100%)
Gaps: 0/700(0%)

Range 1: 3010037 to 3010736 [GenBank](#) [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1293 bits(700)	0.0	700/700(100%)	0/700(0%)	Plus/Minus

Query	1	GGGCGAGTTTATCGCCTCAGCCTGACGTGACGCGGGGAGCTAAACACCGAAGCATTGACCT	60
Sbjct	3010736	GGGCGAGTTTATCGCCTCAGCCTGACGTGACGCGGGGAGCTAAACACCGAAGCATTGACCT	3010677
Query	61	GTATTGTTGAGCGTCTGSAAGTGAAATTATAGATGGCAGCTGGATTATATCAGTTACG	120
Sbjct	3010676	GTATTGTTGAGCGTCTGSAAGTGAAATTATAGATGGCAGCTGGATTATATCAGTTACG	3010617
Query	121	AGGAAACCGATCTCGAAATGATGCCTTTTCTTGTTCACAGGCCAATAAGAAATATCCAG	180
Sbjct	3010616	AGGAAACCGATCTCGAAATGATGCCTTTTCTTGTTCACAGGCCAATAAGAAATATCCAG	3010557
Query	181	AGTTAAATCTTAAATTTGTTATGTCAGTCCATGAGCTTGTTCCTCTATAAAGSAGACCA	240
Sbjct	3010556	AGTTAAATCTTAAATTTGTTATGTCAGTCCATGAGCTTGTTCCTCTATAAAGSAGACCA	3010497
Query	241	GAATGGAAAGCGTGAATCTGCCCGATTTCTCGTAAATATGGSAAGTTCAGGTATCCATA	300
Sbjct	3010496	GAATGGAAAGCGTGAATCTGCCCGATTTCTCGTAAATATGGSAAGTTCAGGTATCCATA	3010437
Query	301	TTTCAGTCGTCGATTTTAGAGTTATGSAACGAAAGACATCGGTGATTTTGTTCGAACCAG	360
Sbjct	3010436	TTTCAGTCGTCGATTTTAGAGTTATGSAACGAAAGACATCGGTGATTTTGTTCGAACCAG	3010377
Query	361	CAGCGTGTAGCGCTTTTGGACCTGCTTTACTGGCGTTGAGGACCAAGCAGCTCTTGAAC	420
Sbjct	3010376	CAGCGTGTAGCGCTTTTGGACCTGCTTTACTGGCGTTGAGGACCAAGCAGCTCTTGAAC	3010317
Query	421	GTGAACAACCTGCCTGATTGTTATTTTGTCTATGGTCGAGCTGGACATTCACGAAAGCTCTT	480
Sbjct	3010316	GTGAACAACCTGCCTGATTGTTATTTTGTCTATGGTCGAGCTGGACATTCACGAAAGCTCTT	3010257
Query	481	CTGAATGCGGATTTTATGCTGCGCTCGCCAAAAAATTCAGCTTGAATTTATGAACT	540
Sbjct	3010256	CTGAATGCGGATTTTATGCTGCGCTCGCCAAAAAATTCAGCTTGAATTTATGAACT	3010197
Query	541	TAGTAAAAATTCATGAAGATAATATTTGTGAACGTCTGTGTGGTGAAGAACCTTTTCTCC	600
Sbjct	3010196	TAGTAAAAATTCATGAAGATAATATTTGTGAACGTCTGTGTGGTGAAGAACCTTTTCTCC	3010137
Query	601	CGTCCGATAAAGCAGACCGCTATCTGCCGCTGAGTTTTTACAACATACTCAAGGCGCAC	660
Sbjct	3010136	CGTCCGATAAAGCAGACCGCTATCTGCCGCTGAGTTTTTACAACATACTCAAGGCGCAC	3010077
Query	661	AACGATTAATGAATATGTGSAGGCCAATCCGGCGGCGGG	700
Sbjct	3010076	AACGATTAATGAATATGTGSAGGCCAATCCGGCGGCGGG	3010017

100%-os egyezés a baktérium genomjának 700 hosszú szakaszával.

Az NCBI adatbázisában megkerestem az adott baktériumot (a Blast felajánlotta a linket):

- [Salmonella enterica subsp. enterica serovar Typhimurium str. LT2](#)

Submitter: Washington University Genome Sequencing Center

Human Pathogen

Morphology: Gram:Negative, Shape:Bailli, Motility:Yes

Environment: OxygenReq:Facultative, OptimumTemperature:37, TemperatureRange:Mesophilic, Habitat:HostAssociated

Phenotype: Disease:Gastroenteritis and food poisoning

Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Other RNA	Gene	Pseudogene
Chr	-	NC_003197.1	AE008468.1	4.86	52.2	4,451	22	86	10	4,631	54
Plsm	pSLT	NC_003277.1	AE008471.1	0.093939	53.1	102	-	-	1	112	-

([http://www.ncbi.nlm.nih.gov/genome?term=16763390\[gi\]&RID=NDWED53T015&log\\$=mapalign&blast_rank=1](http://www.ncbi.nlm.nih.gov/genome?term=16763390[gi]&RID=NDWED53T015&log$=mapalign&blast_rank=1))

A fenotípus információknál olvashatunk a betegségről: gastroenteritis (gyomor-bélhurut), ételmérgezés. Erről a betegségről gyűjtöttem néhány információt:

Gyomor-bélhurut (gastroenteritisz):

Begyullad a gyomor-bél rendszer, hasmenés, hányás, hasi görcsök. Magyarországon 87 halálozás 2011-ben, az egész világon átlag 1,4 millió évente. A WHO szerint 2000-ben a szalmonella baktérium miatt 216510 ember halt meg, tehát a bélhurutok 15%-a. /Wikipedia/