Advanced Cell Biology & Biotechnology

Biotechnology Project Lab

Giovanna Gambarotta
& Isabella Tarulli

The lecture of November 8th 2021 is about to begin….

Summary of the previous lesson activity

1-find the length of the intron between domain EGF-like and domains alpha and beta

2-find coding exons (number and length) and introns (length) of TBP and of GAPDH to decide where to design primer.

3-project overview

PROJECT OVERVIEW

Genebank sequence analysis

↓

primer design

↓

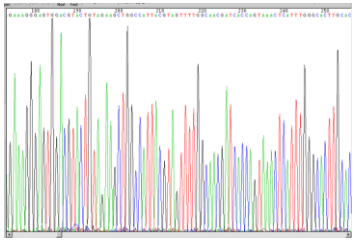
RT-PCR

↓

cloning in the vector pCRII-blunt

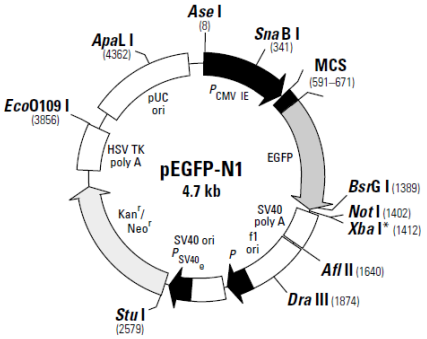
↓

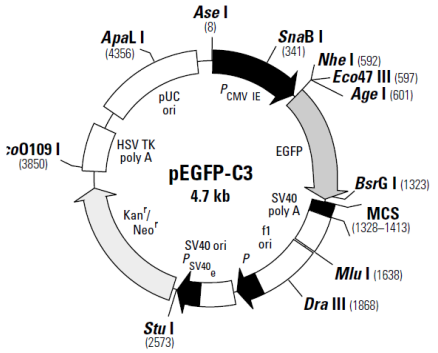
sequence analysis

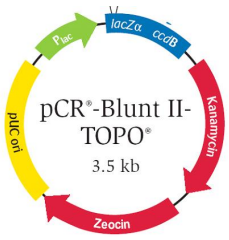


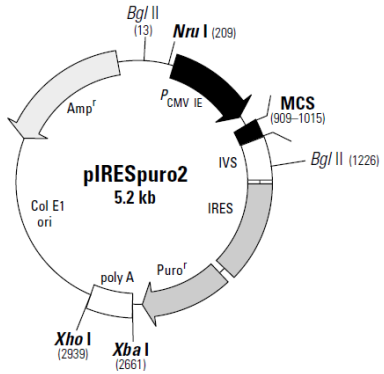
- real time PCR
- protein quantification

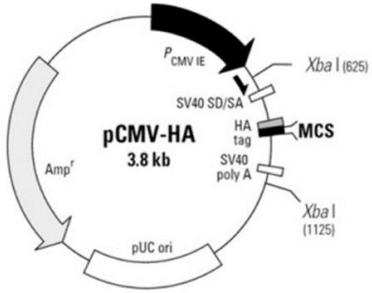
subcloning in different expression vectors

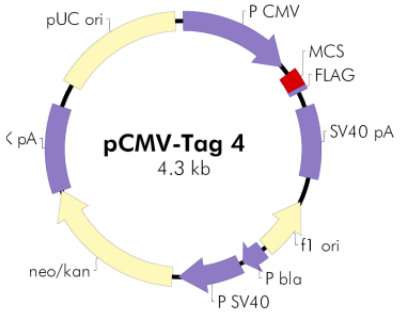




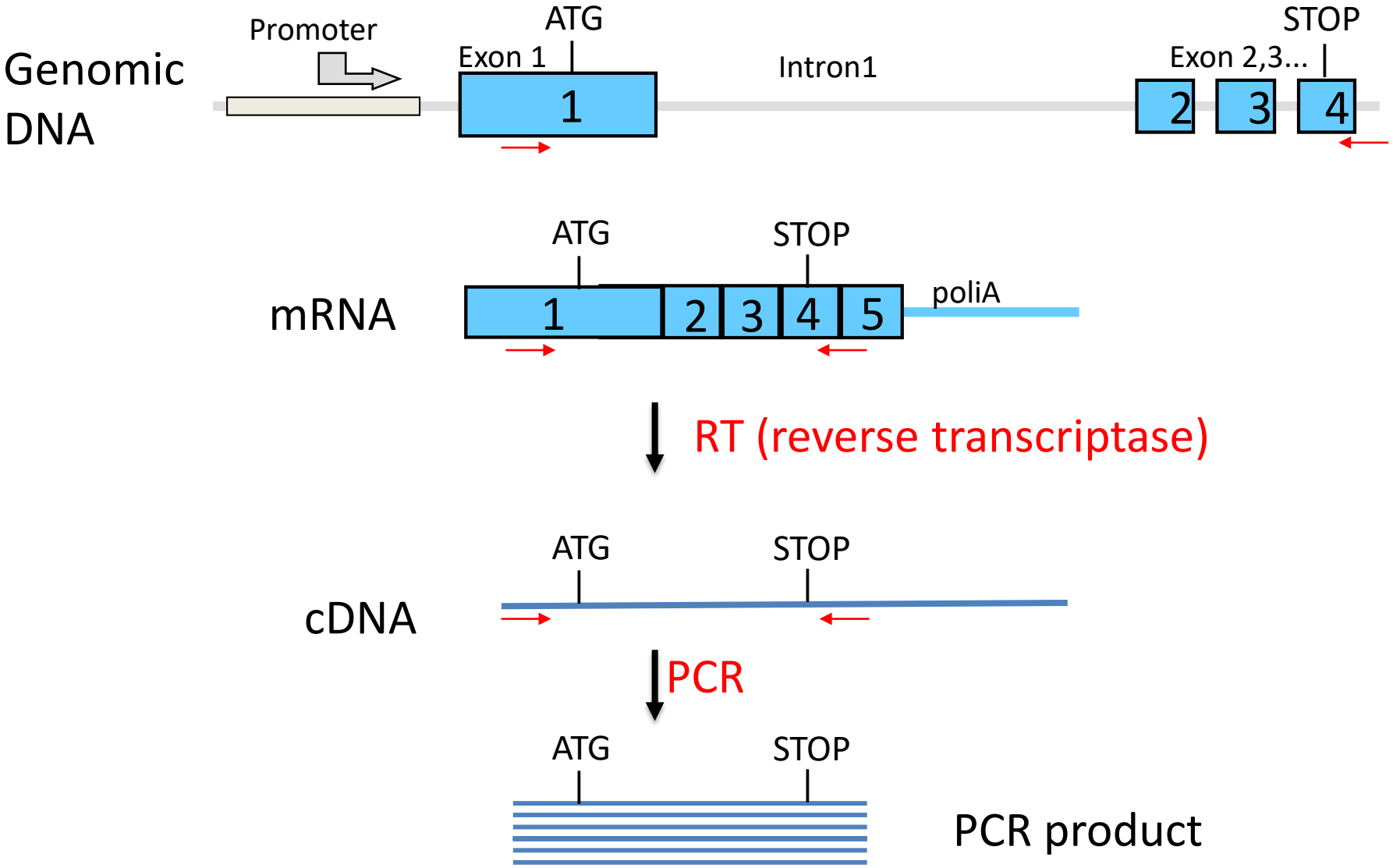




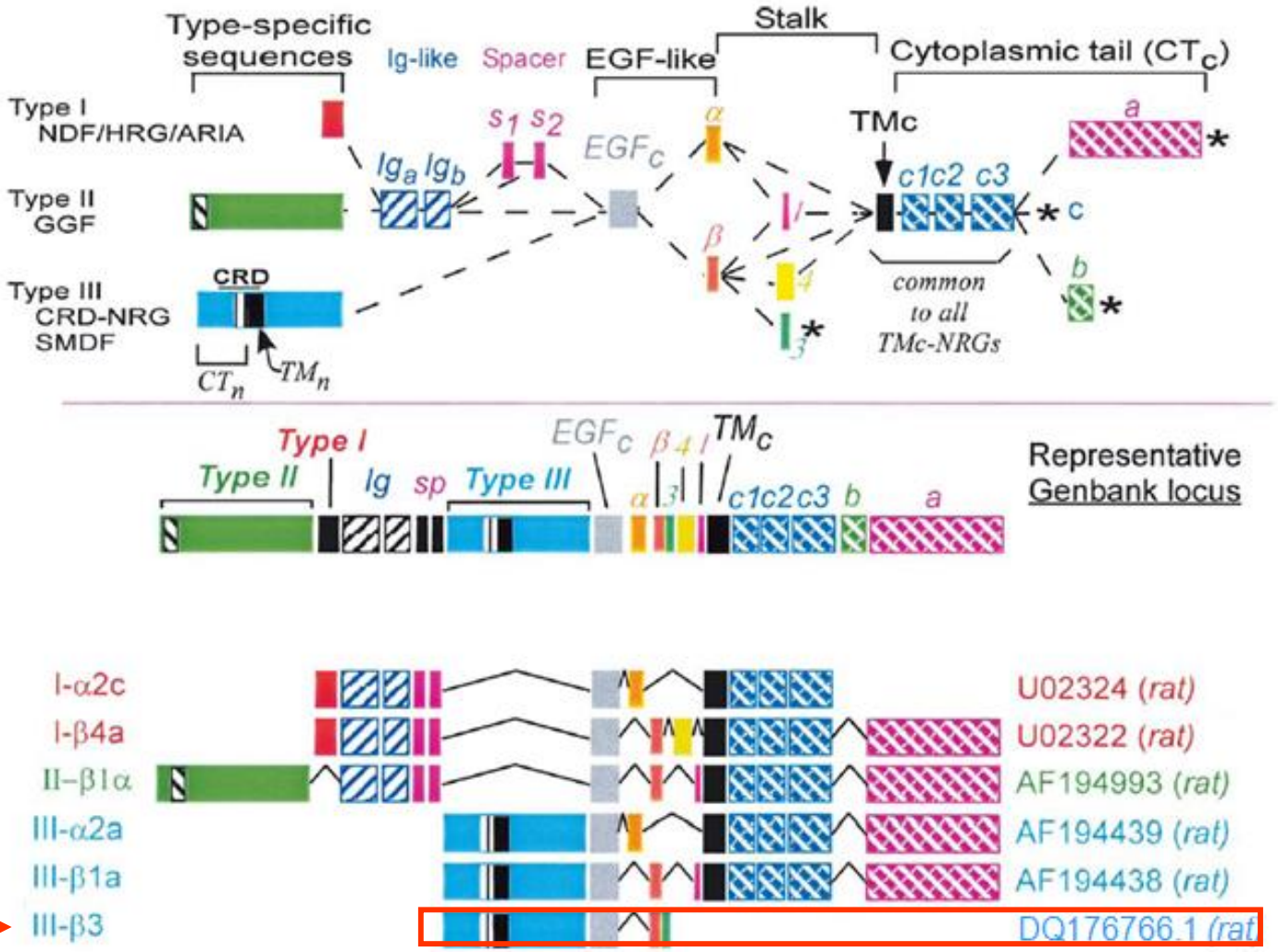


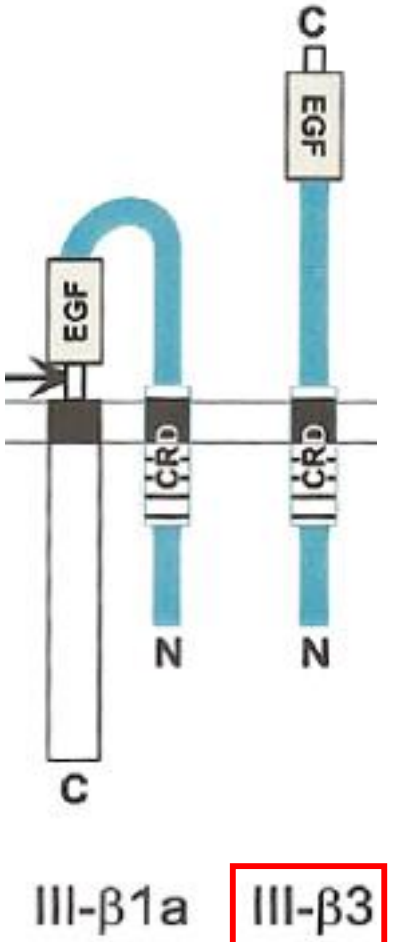


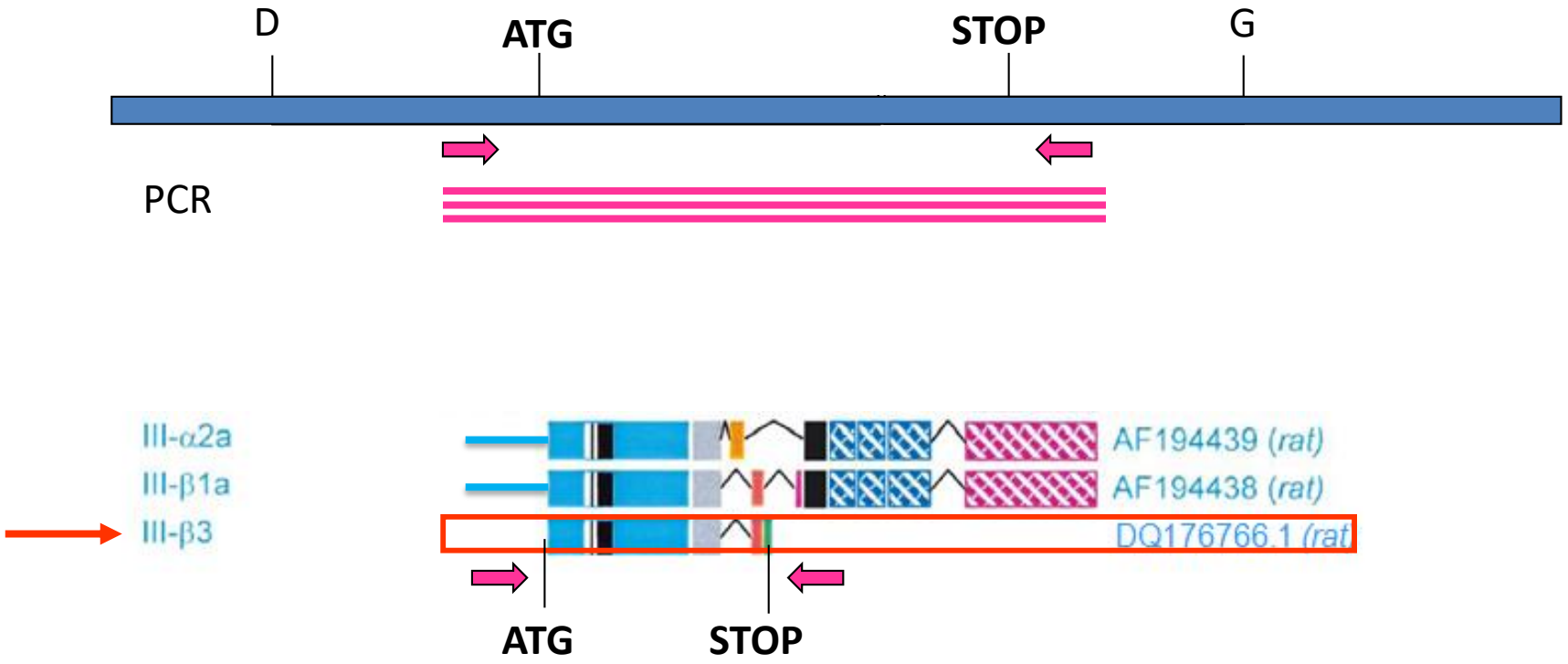
Cloning the **full length cDNA** to express NRG1 type III beta3



Cloning of rat NRG1-typeIIIβ3







The sequence DQ176766 is starting from the ATG and is ending after the STOP. Therefore I suggest you to copy 30 bases upstream ATG from AF194438 or AF194439 to have 30 bp sequence before ATG, to be able to design a good primer before or near ATG.

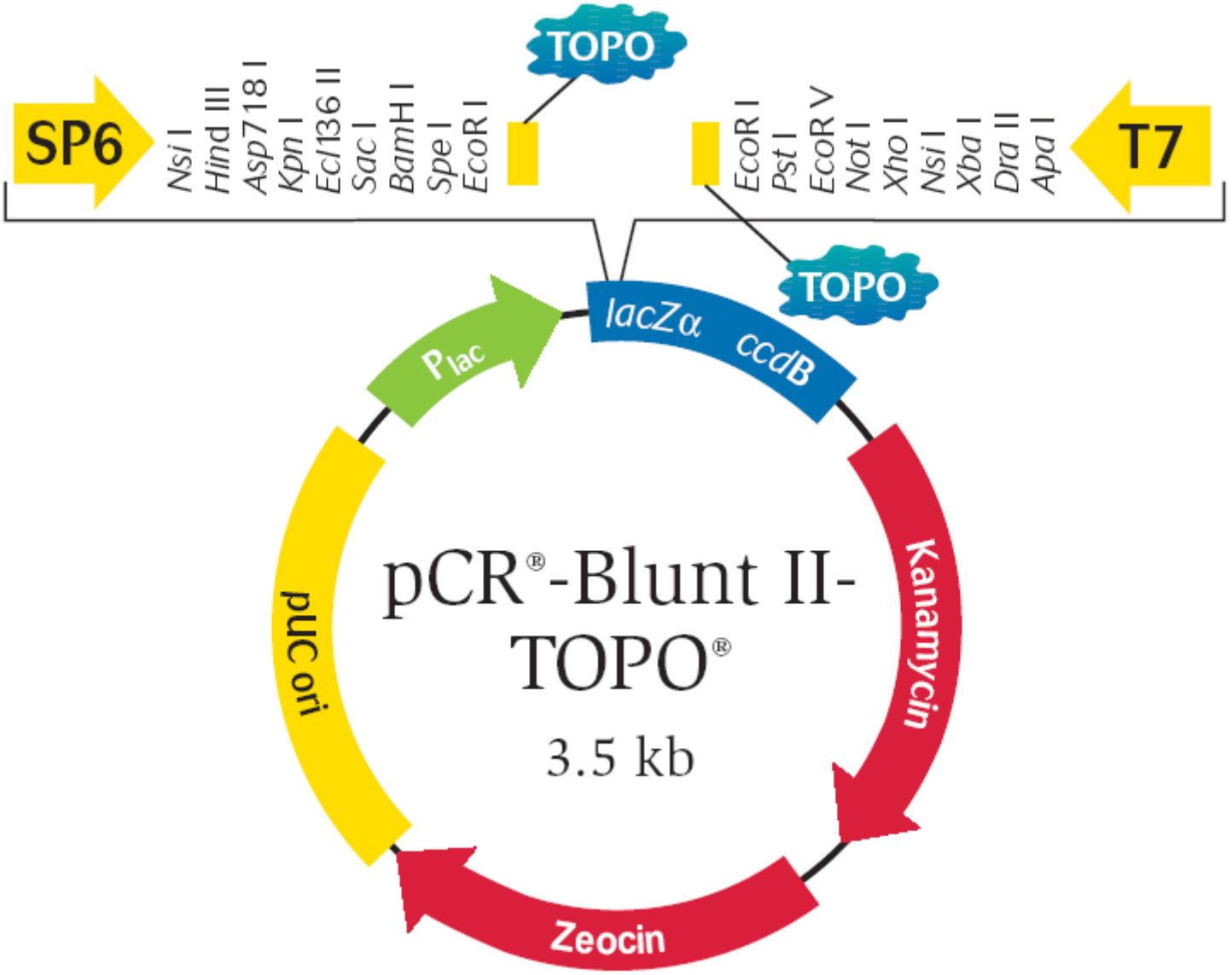
You can copy 30 bases before ATG and copy them before the sequence DQ176766, as shown in the next slide

AGAGAGGCCA GGCCTTCTGG AGGTGAGCCG

(Remember that ATG and STOP position will be shifted 30 bases!).

AGAGAGGCCA GGCCTTCTGG AGGTGAGCCG

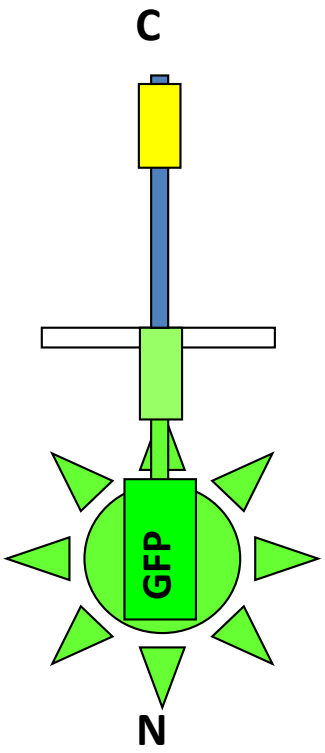
```
  1 atggagattt attccccaga catgtctgag gtagctggcg ggaggtcctc cagcccctcc
 61 actcagctga gtgcagcccc atctcttgat gggcttccgg cagcggagga acatatacca
121 gacacccaca cagaagatga gagaagccct ggactcctgg gcctggcggt gccctgctgt
181 gtgtgcctgg aagctgagcg cctgagaggg tgtctcaact ccgagaagat ctgcattgtt
241 cccattctgg cttgcctagt cagcctctgc ctctgcattg ctggcctgaa gtgggtattt
301 gtggacaaga tatttgaata cgactctcct acccaccttg accctggggg gttaggccag
361 gaccctgtga tttctctgga tccaactgct gccccagcca ttttggtatc atctgaggca
421 tacacttcac ctgtctctaa ggctcagtct gaagctgggg ctcatgttac agtacaaggt
481 gaccatgctg ctgtggcctc tgaaccttca gcagtaccga cccggaagaa ccggctgtct
541 gcttttcctc cctttcactc tactgcaccg cccttccctt ctccagctcg gacccctgag
601 gtgagaacac ccaagtcagg aactcagcca caaacaacag aaactaacct gcaaactgct
661 cctaaacttt ccacatcaac atccacgact gggaccagcc atctcataaa gtgtgcggag
721 aaggagaaaa ctttctgtgt gaatgggggc gagtgcttca cggtgaagga cctgtcaaac
781 ccgtcaagat acttgtgcaa gtgcccaaat gagtttactg gtgatcgttg ccaaaactac
841 gtaatggcca gcttctacag tacgtccact ccctttctgt ctctgcctga gtag GAGCAT

901 GCTCAGTCGA TGCT
```

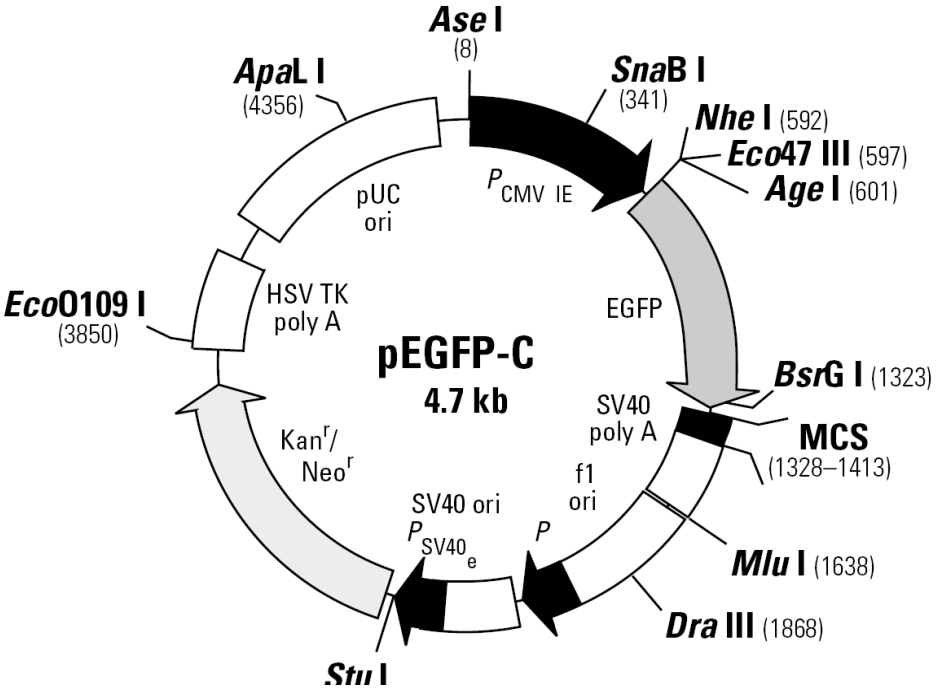


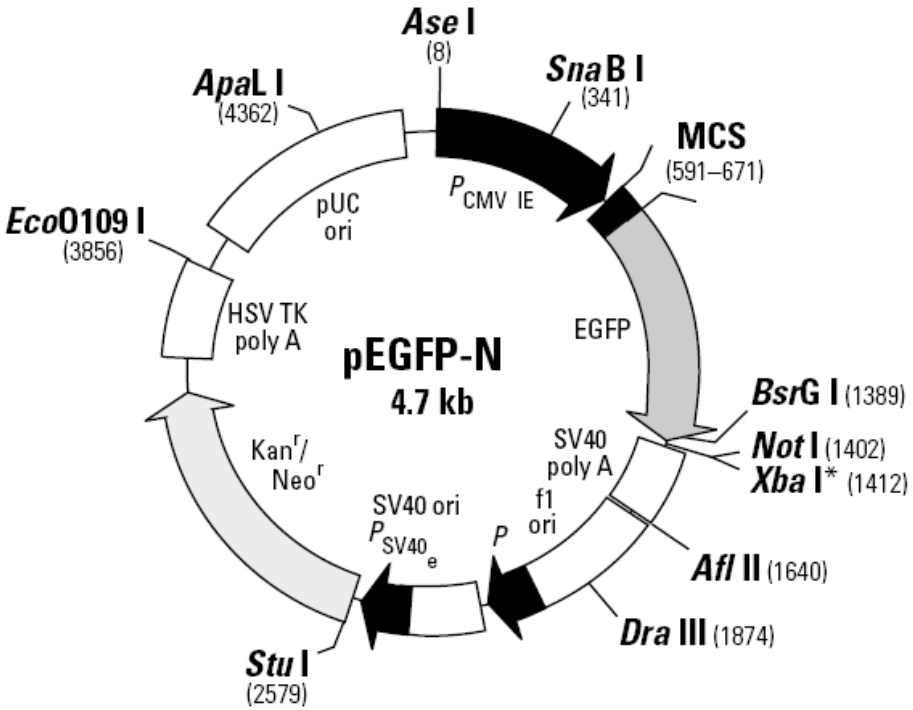
http://www.invitrogen.com/

Cloning vector for the RT-PCR amplification product

Which expression vector will you use to have GFP in the cytoplasmic region, considering that N terminus is in the cytoplasm and C terminus is extracellular??

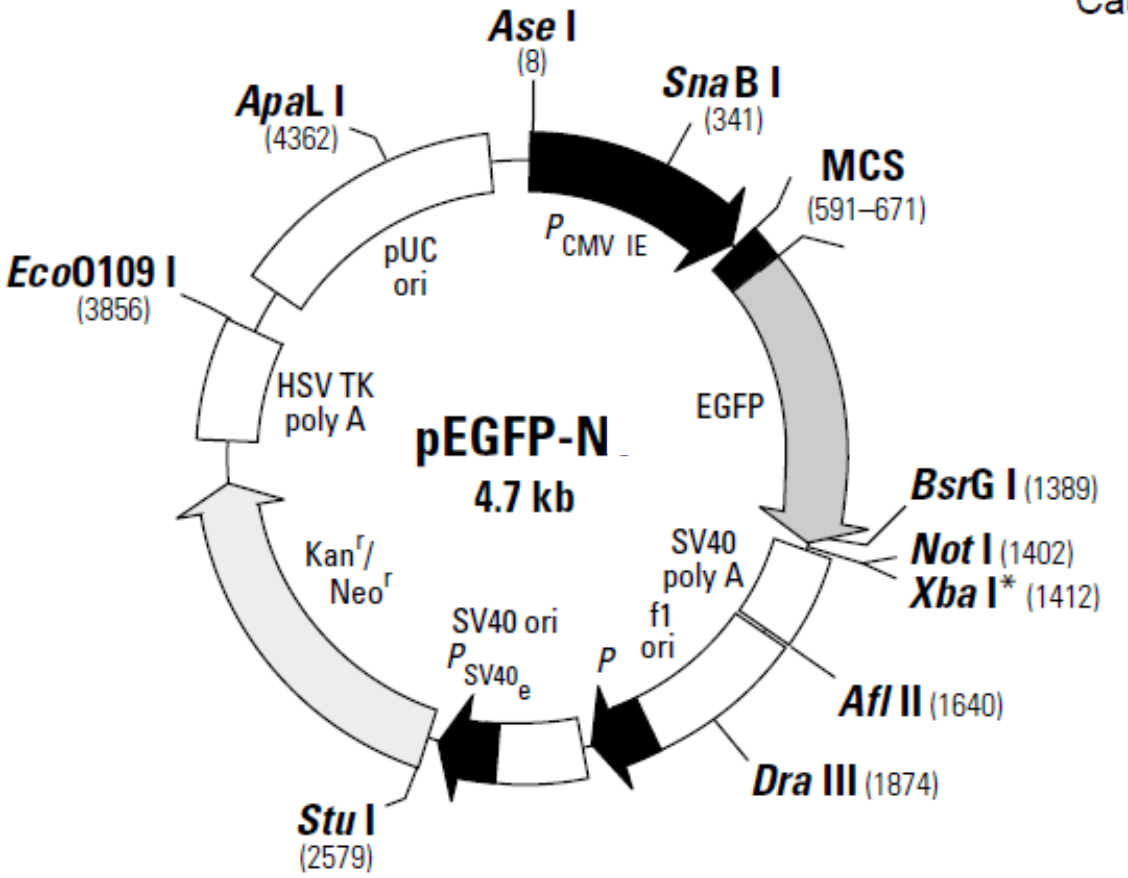






PT3027-5

Catalog #6085-1



591 601 611 621 631 641 651 661 671

G CTA GCG CTA CCG GAC TCA GAT CTC GAG CTC AAG CTT CGA ATT CTG CAG TCG ACG GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG

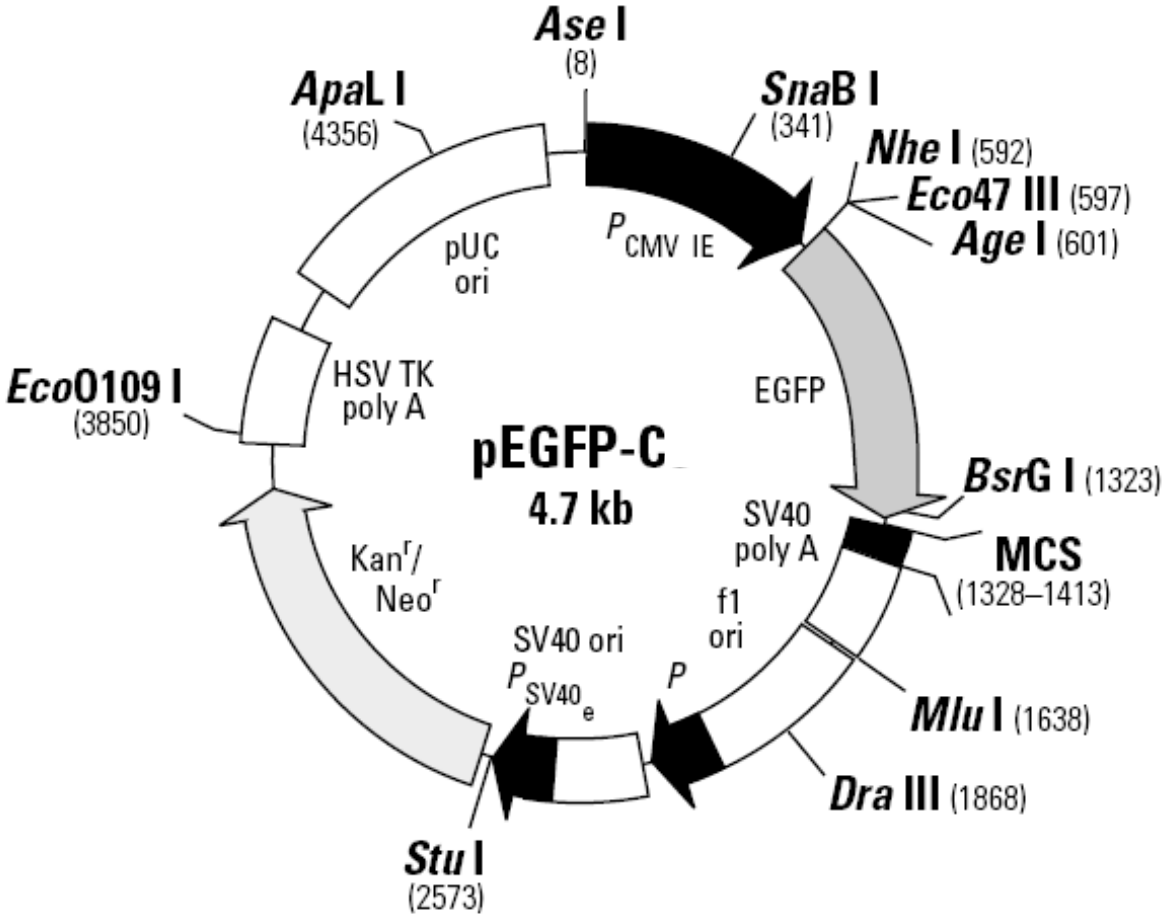
EGFP

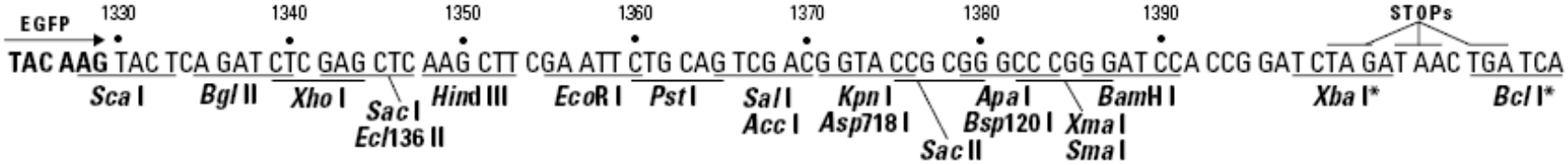
Nhe I Eco47 III Bgl II Xho I Sac I Ecl136 II Hind III EcoR I Pst I Sal I Acc I Kpn I Asp718 I Sac II Apa I Bsp120 I Xma I Sma I BamH I Age I

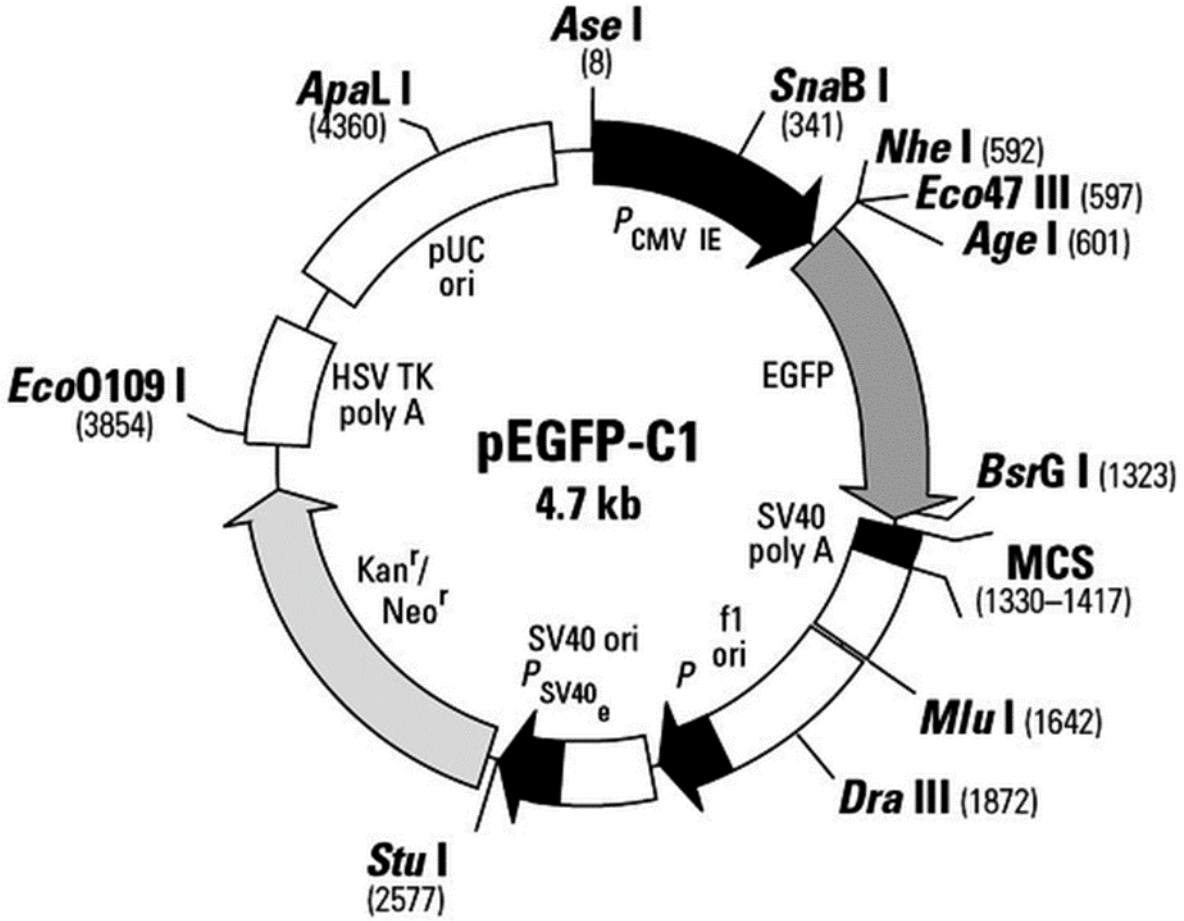
Clontech

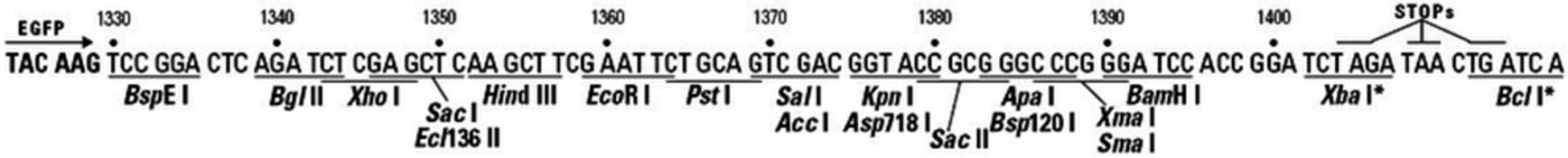
PT3052-5

Catalog #6082-1

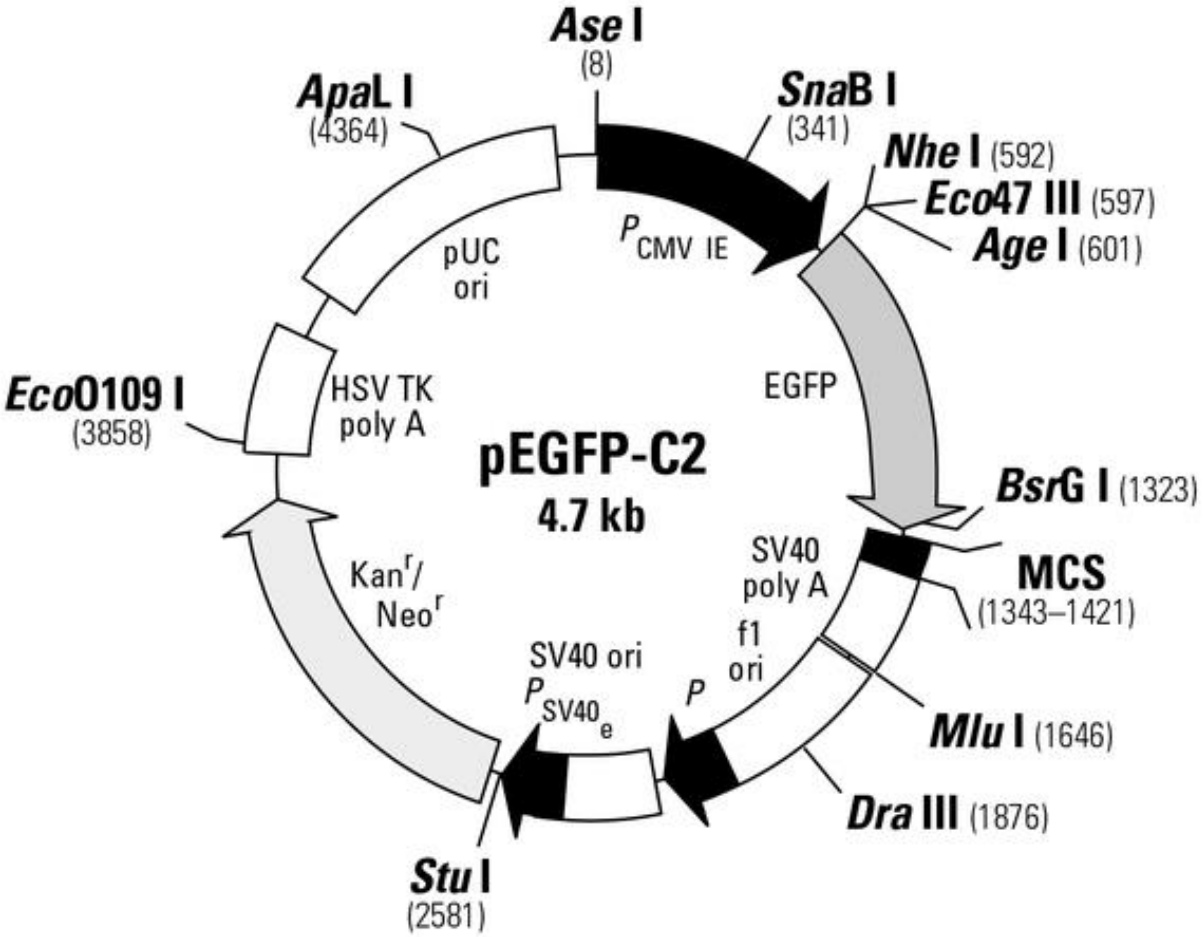


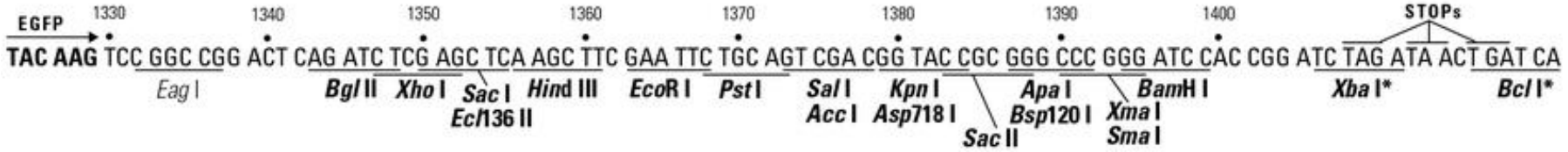


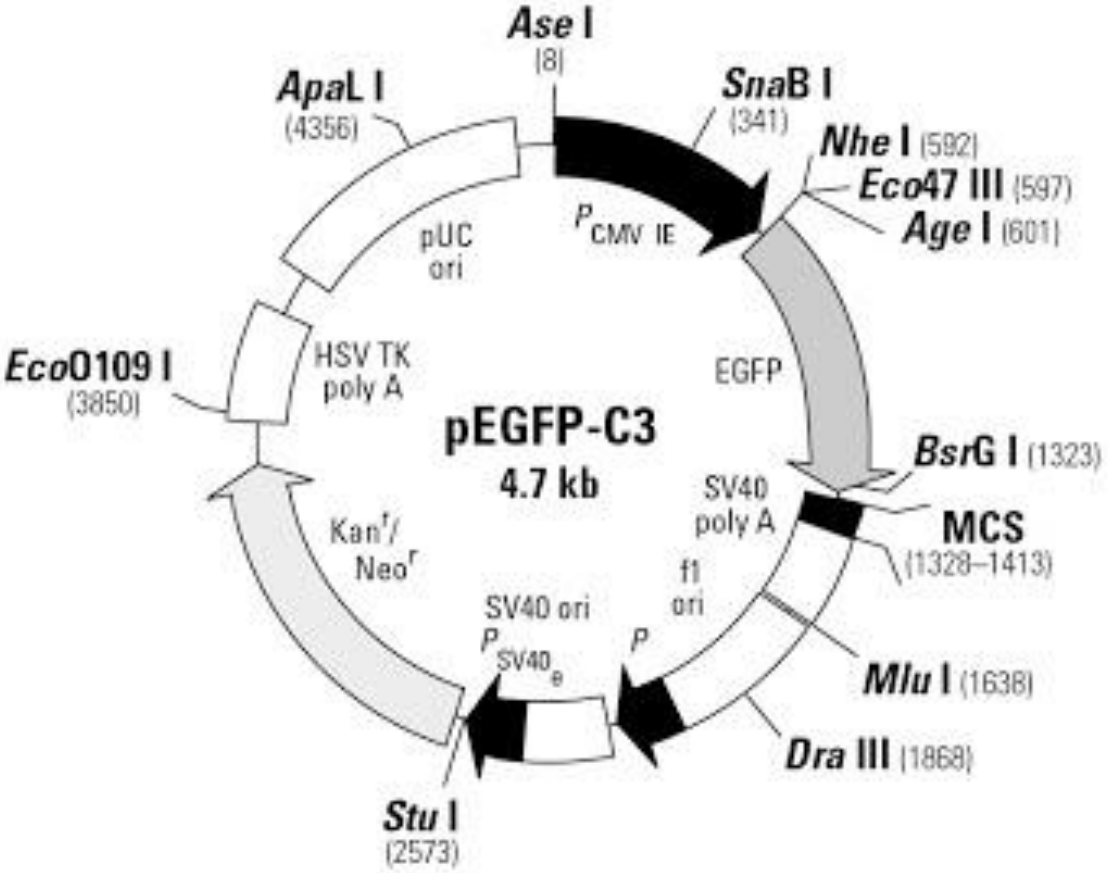




Restriction Map and Multiple Cloning Site (MCS) of pEGFP-C1. (Unique restriction sites are in bold). The *Xba* I and *Bcl* I sites (*) are methylated in the DNA provided by CLONTECH. If you wish to digest the vector with these enzymes, you will need to transform the vector into a *dam*⁻ host and make fresh DNA.


Ase I
(8)
SnaB I
(341)
Nhe I (592)
Eco47 III (597)
Age I (601)
BsrG I (1323)
MCS
(1343–1421)
Mlu I (1646)
Dra III (1876)
Stu I
(2581)
ApaL I
(4364)
EcoO109 I
(3858)
pUC
ori
P CMV IE
EGFP
HSV TK
poly A
pEGFP-C2
4.7 kb
SV40
poly A
f1
ori
P
SV40 ori
P SV40e
Kanr/
Neor


EGFP
1330
1340
1350
1360
1370
1380
1390
1400
STOPs
TAC AAG TCC GGC CGG ACT CAG ATC TCG AGC TCA AGC TTC GAA TTC TGC AGT CGA CGG TAC CGC GGG CCC GGG ATC CAC CGG ATC TAG ATA ACT GAT CA
Eag I
Bgl II
Xho I
Sac I
Ecl136 II
Hind III
EcoR I
Pst I
Sal I
Acc I
Kpn I
Asp718 I
Sac II
Apa I
Bsp120 I
Xma I
Sma I
BamH I
Xba I*
Bcl I*



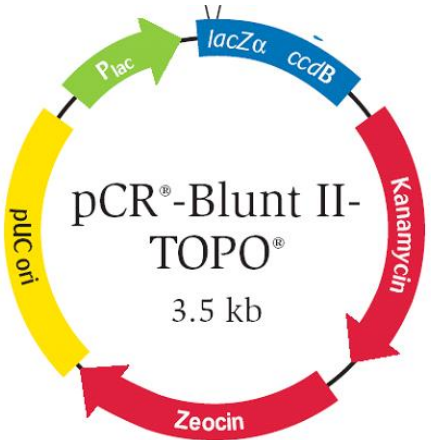
Restriction Map and Multiple Cloning Site (MCS) of pEGFP-C3. (Unique restriction sites are in bold). The *Bcl* I site cannot be used for fusions since it contains an in-frame stop codon. The *Xba* I and *Bcl* I sites (*) are methylated in the DNA provided by CLONTECH. If you wish to digest the vector with these enzymes, you will need to transform the vector into a *dam*⁻ host and make fresh DNA.

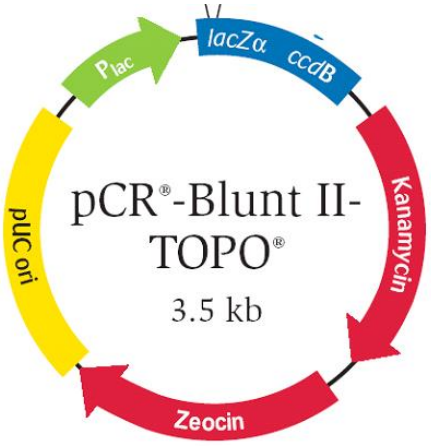
First AIM

* you have to prepare two primers, sense and antisense, to clone the full length NRG1 type III beta 3. One primer immediately before ATG, one after STOP.

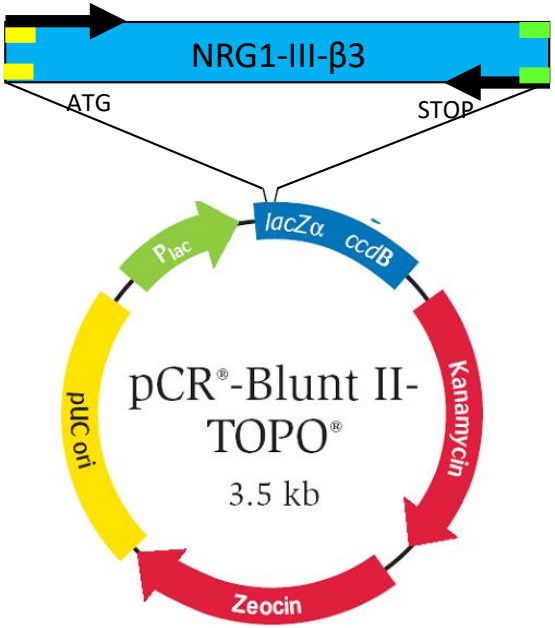
* pairs of primers should meet the following criteria:

1- Tm similar (about 60° C calculated by the Allawi's method)
2- finish with G or C
3- have a content of G and C ≥ 50% (if possible!)
4- not form secondary structures
5-18-28 base long

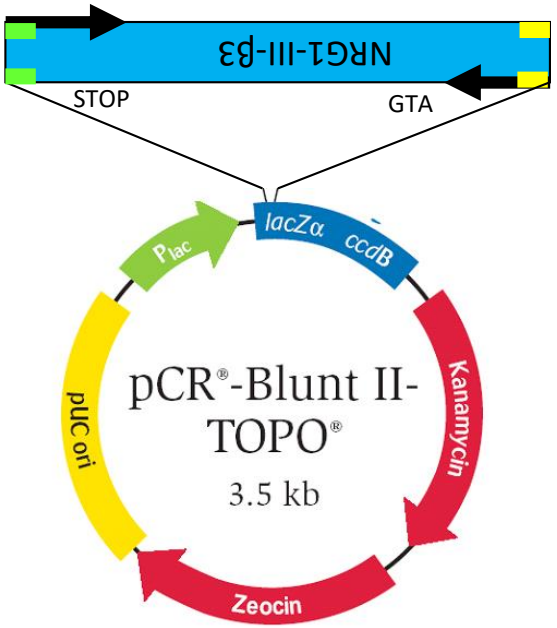


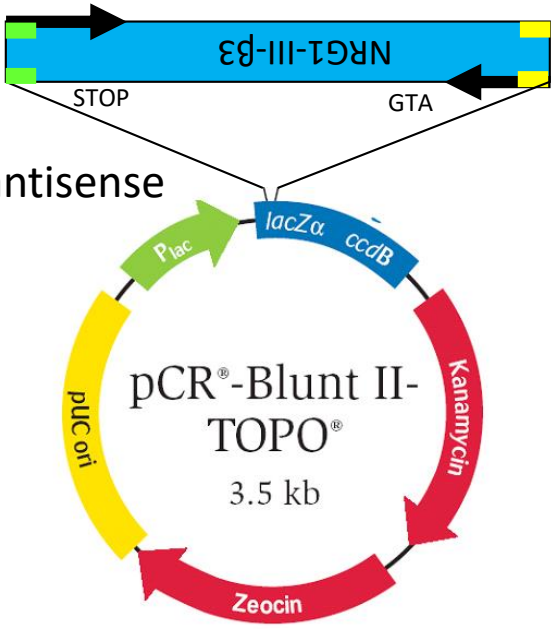


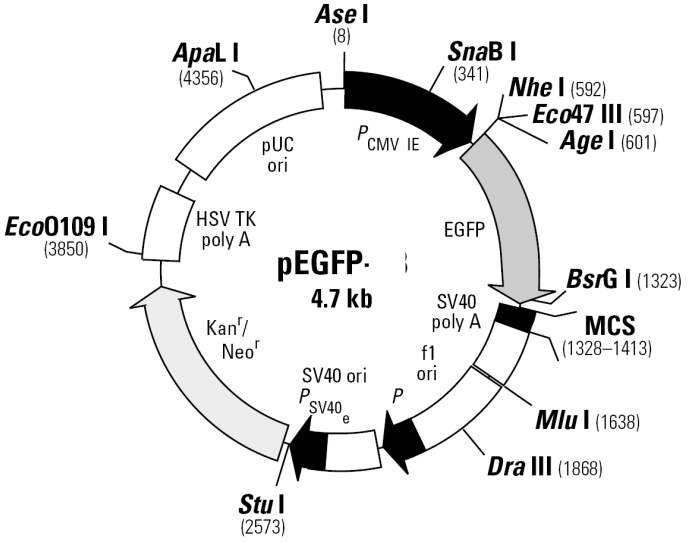
insert sense

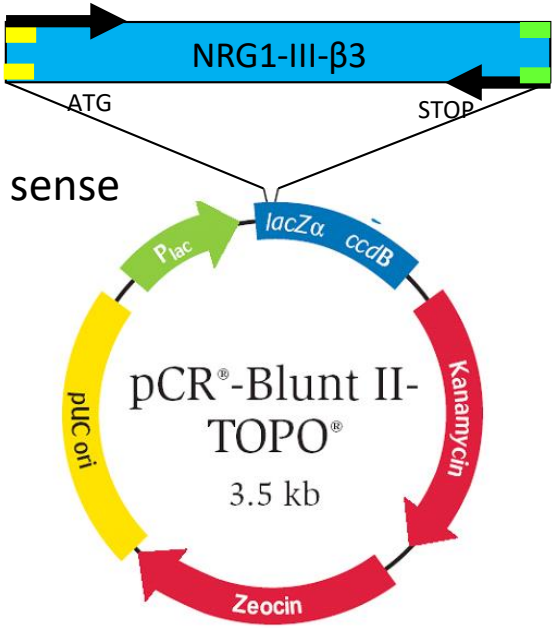


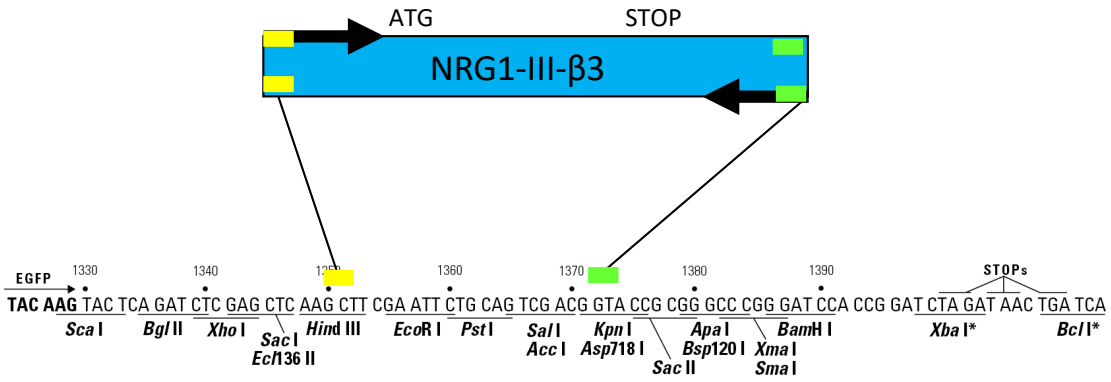
insert antisense











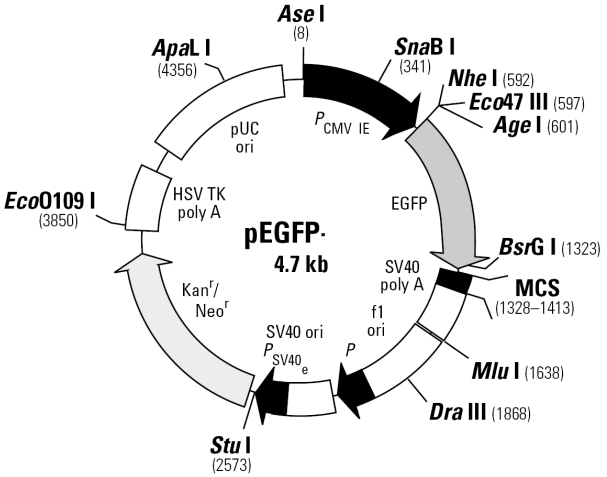
Second AIM

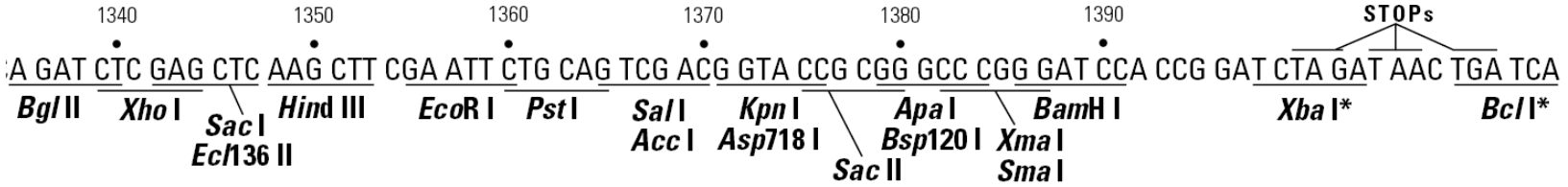
❋ because NRG1typeIIIβ3 will be cloned in frame downstream EGFP in one vector pEGFP: insert restriction sites in the 5‘ of the primers to facilitate following subcloning from pCR-bluntII-TOPO

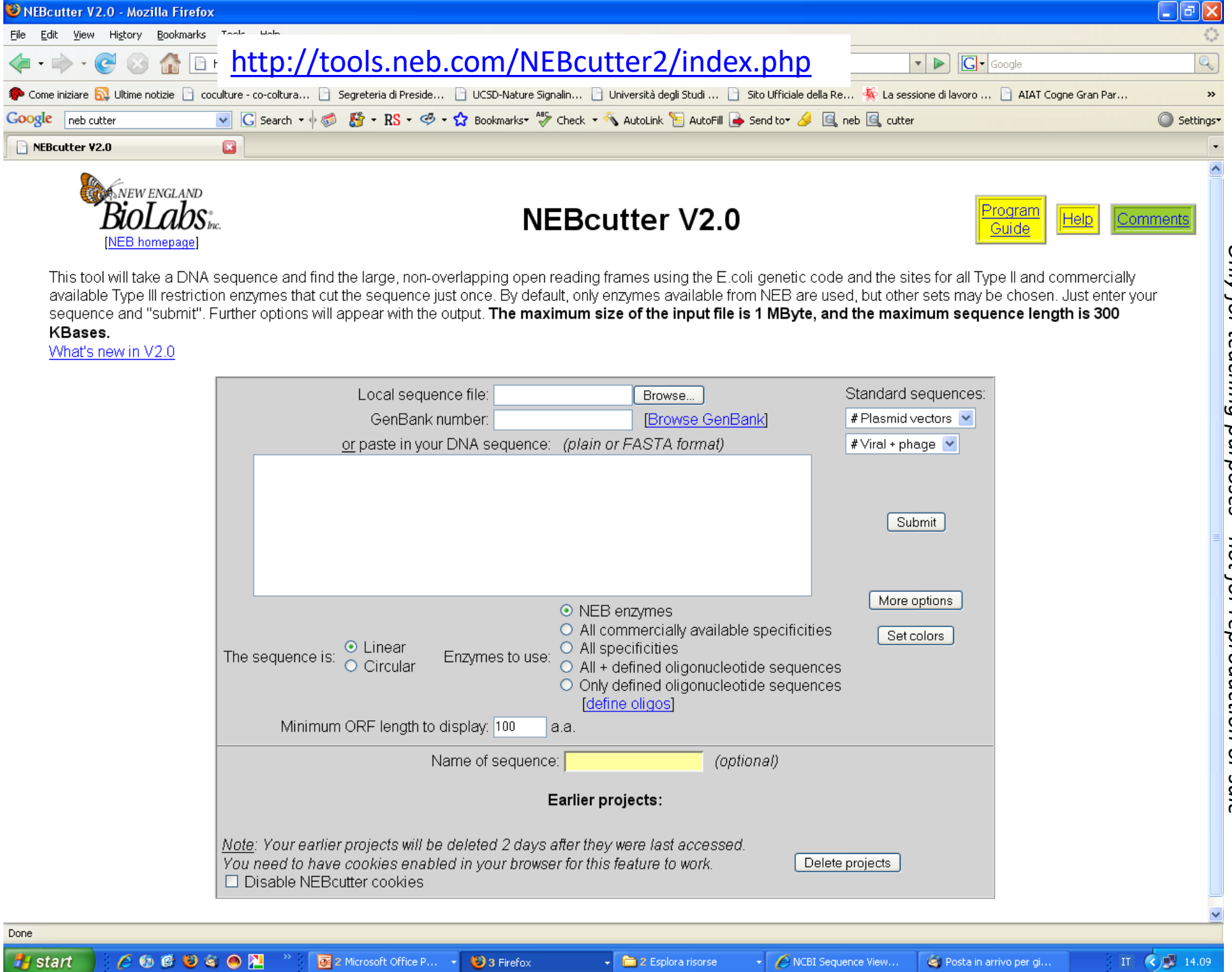
❋ restriction sites must be compatible with the multiple cloning site of the vector pEGFP (from BglII to BamHI)

❋ restriction sites do not have to cut NRG1

-> Make the list of enzymes in the MCS of the vector and investigate which of them do not cut NRG1







NEBcutter V2.0

nce and find the large, non-overlapping open reading frames using the E.coli genetic code and the sites for all Type II and commer
equence just once. By default, only enzymes available from NEB are used, but other sets may be chosen. Just enter your sequence
maximum size of the input file is 1 MByte, and the maximum sequence length is 300 KBases.
Bcutter

Local sequence file: Sfoglia... Nessun file selezionato.

GenBank number: [Browse GenBank]

or paste in your DNA sequence: *(plain or FASTA format)*

```
atttcaaagt
     2161 ctcactttta ttgataaaat aaaaatcatt ctactgaaca gtccatcttc
tttatacaat
     2221 gaccacatcc tgaaaagggt gttgctaagc tgtaaccgat atgcacttga
aatgatggta
     2281 agttaatttt gattcagaat gtgttatttg tcacaaataa acataataaa
aggagttcag
     2341 atgtttttct tcattaacca aaaaaaaaaa aaaa
//
```

Standard sequences:

Plasmid vectors

Viral + phage

Submit

More options

Set colors

The sequence is: Linear / Circular

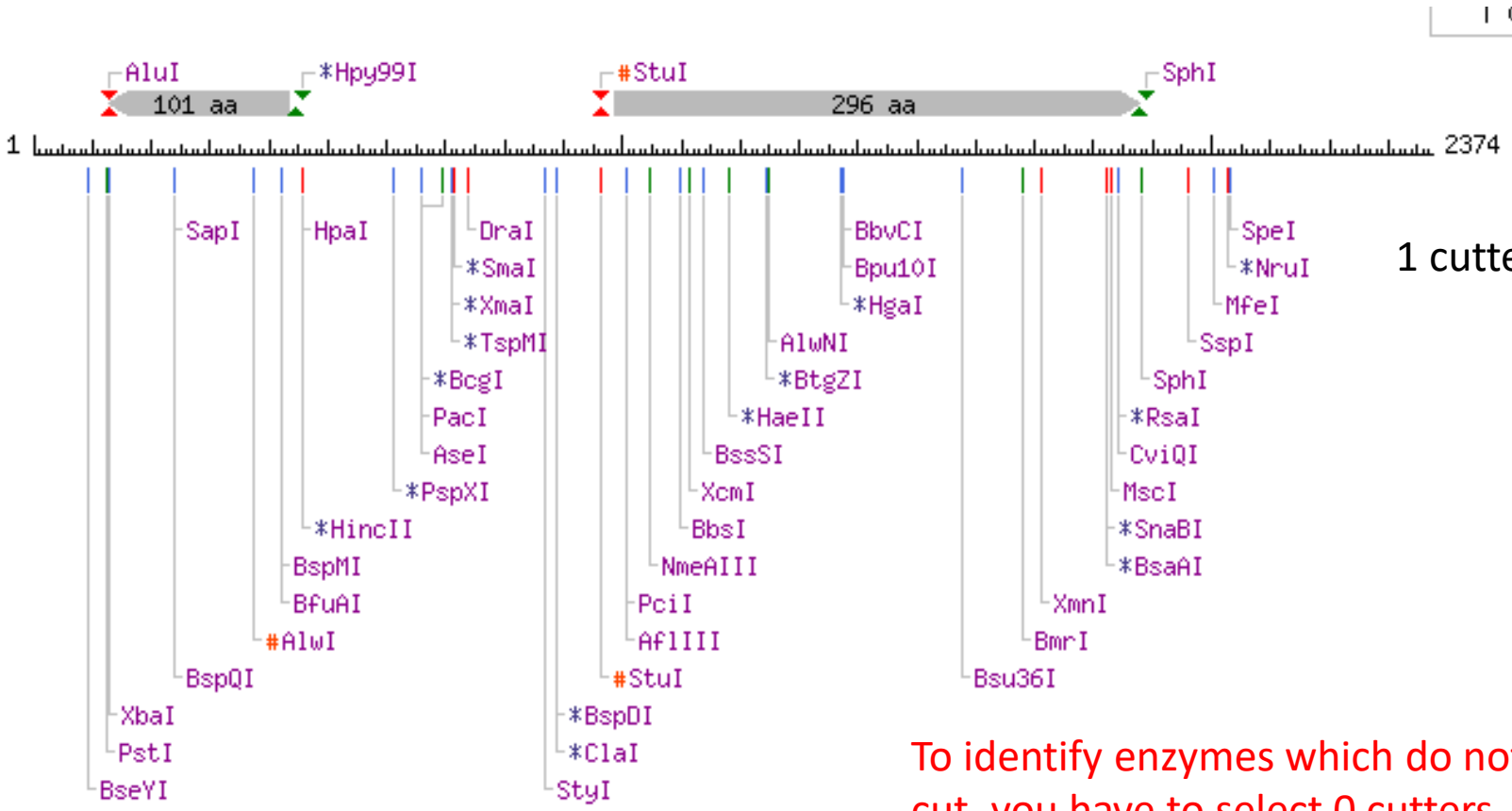
Enzymes to use:
- NEB enzymes
- All commercially available specificities
- All specificities
- All + defined oligonucleotide sequences
- Only defined oligonucleotide sequences

[define oligos]

Minimum ORF length to display: 100 a.a.

Name of sequence: *(optional)*

- The system automatically shows 1 cutter enzymes



1 cutters

To identify enzymes which do not cut, you have to select 0 cutters

Main options	Availability	Display	Zoom	List
New DNA	All commercial	2 cutters	Zoom in	0 cutters
Custom digest	All	3 cutters	More...	1 cutters
View sequence				All sites
ORF summary				Save all sites
Save project				Flanking enzymes
Print				

Minimum ORF length to display: 100 aa. OK

NEW ENGLAND BioLabs Inc.
NEBcutter

[Back to main display]

Enzymes that don't cut

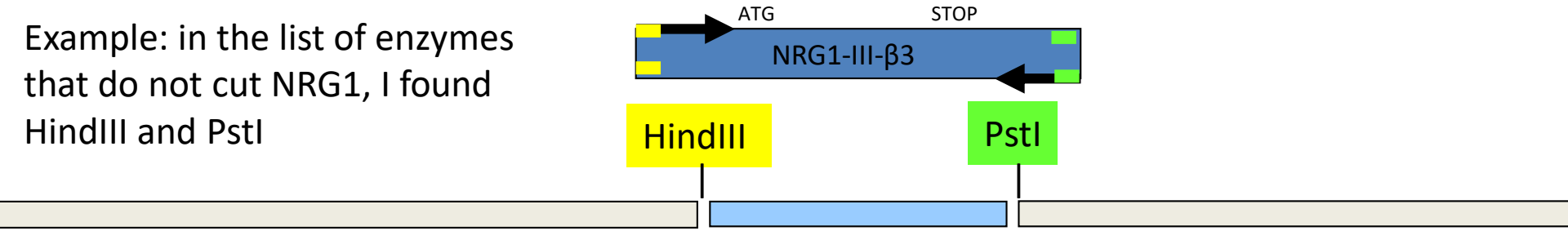
Help | Comments

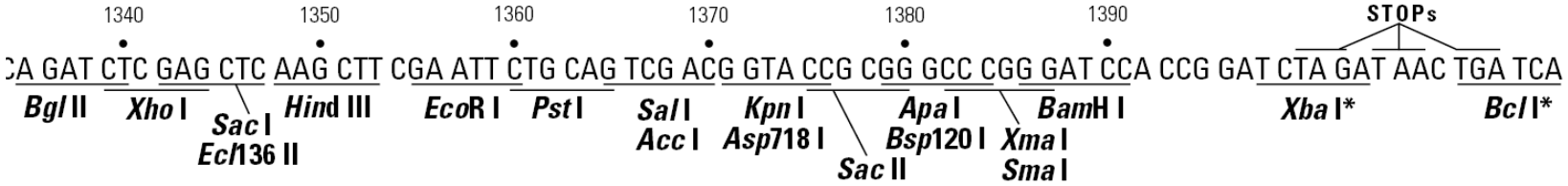
unnamed sequence

Number of cuts = 0 OK Save as text file

#	Enzyme	Specificity
1	Acc65I	G^GTAC_C
2	AccI	GT^MK_AC
3	AclI	AA^CG_TT
4	AfeI	AGC^GCT
5	AflII	C^TTAA_G
6	AgeI	A^CCGG_T
7	AleI	CACNN^NNGTG
8	ApaLI	G^TGCA_C
9	AscI	GG^CGCG_CC
10	AseI	AT^TA_AT
11	AsiSI	GCG_AT^CGC
12	AvaI	C^YCGR_G
13	AvrII	C^CTAG_G
14	BaeI	_(N)5^(N)10ACNNNNGTAYC(N)7_(N)5^
15	BbvCI	CC^TCA_GC
16	BceAI	ACGGC(N)12^NN_
17	BciVI	GTATCC(N)5_N^
18	BmgBI	CAC^GTC
19	BmtI	G_CTAG^C
20	BsaBI	GATNN^NNATC

Example: in the list of enzymes that do not cut NRG1, I found HindIII and PstI





5′-CGTTAACTTGACC**ATG**TGCATC**TAG**CTCCATGGCATGC-3′

5′-**CGTTAACTTG**ACC**ATG**TGCATC**TAG**CTCCATGGCATGC-3′
3′-GCAATTGAACTGGTACACGTAGATCGAG**GTACCGTACG**-5′

First AIM

Primer sense:	5′-**CGTTAACTTG**-3′
Primer antisense:	5′-**GCATGCCATG**-3′

5′-CGTTAACTTGACC**ATG**TGCATC**TAG**CTCCATGGCATGC-3′

5′-CGTTAACTTGACC**ATG**TGCATC**TAG**CTCCATGGCATGC-3′
3′-GCAATTGAACTGGTACACGTAGATCGAGGTACCGTACG-5′

First AIM

Primer sense: 5′-CGTTAACTTG-3′
Primer antisense: 5′-GCATGCCATG-3′

Add the **HindIII site** to the 5′ of the primer sense

5′-AAGCTT-3′
3′-TTCGAA-5′

Add the **PstI site** to the 5′ of the primer antisense

5′-CTGCAG-3′
3′-GACGTC-5′

Second AIM

Primer sense: 5′-AAGCTTCGTTAACTTG-3′
Primer antisense: 5′-CTGCAGGCATGCCATG-3′

5′-CGTTAACTTGACC**ATG**TGCATC**TAG**CTCCATGGCATGC-3′

5′-CGTTAACTTGACC**ATG**TGCATC**TAG**CTCCATGGCATGC-3′
3′-GCAATTGAACTGGTACACGTAGATCGAGGTACCGTACG-5′

First AIM

Primer sense: 5′-CGTTAACTTG-3′
Primer antisense: 5′-GCATGCCATG-3′

Added the **HindIII site** to the 5′ of the primer sense

5′-AAGCTT-3′
3′-TTCGAA-5′

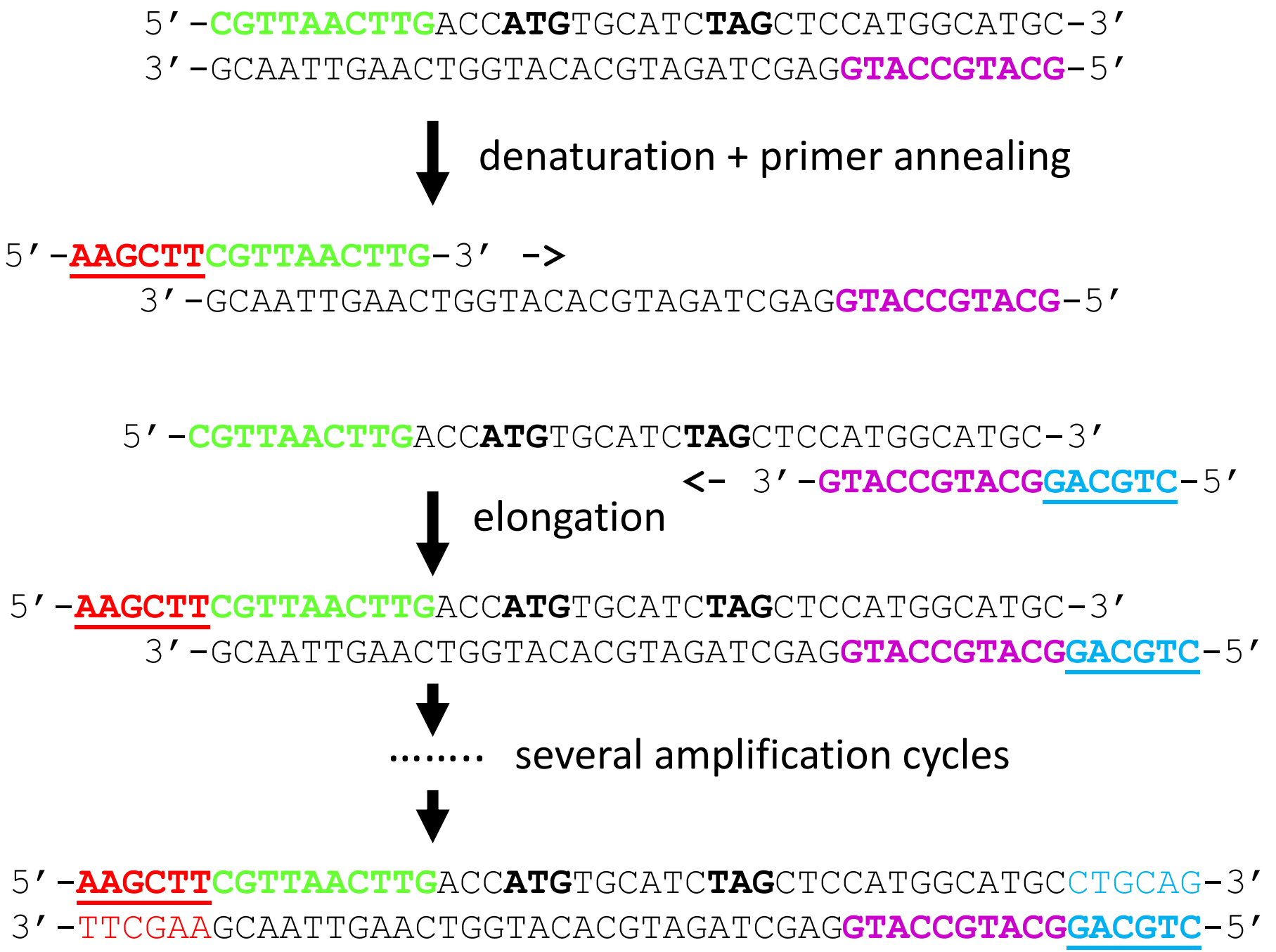
Added the **PstI site** to the 5′ of the primer antisense

5′-CTGCAG-3′
3′-GACGTC-5′

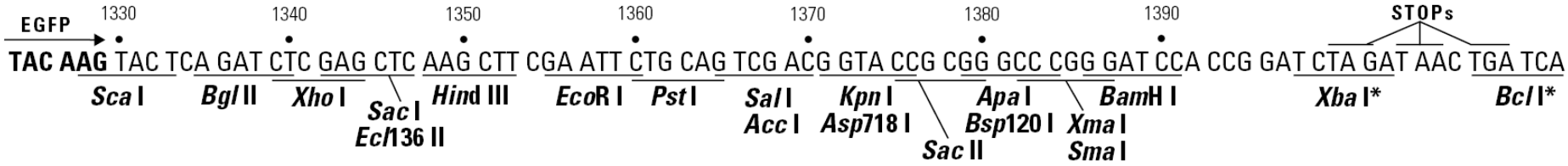
Second AIM

Primer sense: 5′-AAGCTTCGTTAACTTG-3′
Primer antisense: 5′-CTGCAGGCATGCCATG-3′

5′-CGTTAACTTGACC**ATG**TGCATC**TAG**CTCCATGGCATGC-3′
3′-GCAATTGAACTGGTACACGTAGATCGAGGTACCGTACG-5′


5′-CGTTAACTTGACCATGTGCATCTAGCTCCATGGCATGC-3′
3′-GCAATTGAACTGGTACACGTAGATCGAGGTACCGTACG-5′
denaturation + primer annealing
5′-AAGCTTCGTTAACTTG-3′ ->
3′-GCAATTGAACTGGTACACGTAGATCGAGGTACCGTACG-5′
5′-CGTTAACTTGACCATGTGCATCTAGCTCCATGGCATGC-3′
<- 3′-GTACCGTACGGACGTC-5′
elongation
5′-AAGCTTCGTTAACTTGACCATGTGCATCTAGCTCCATGGCATGC-3′
3′-GCAATTGAACTGGTACACGTAGATCGAGGTACCGTACGGACGTC-5′
..... several amplification cycles
5′-AAGCTTCGTTAACTTGACCATGTGCATCTAGCTCCATGGCATGCCTGCAG-3′
3′-TTCGAAGCAATTGAACTGGTACACGTAGATCGAGGTACCGTACGGACGTC-5′

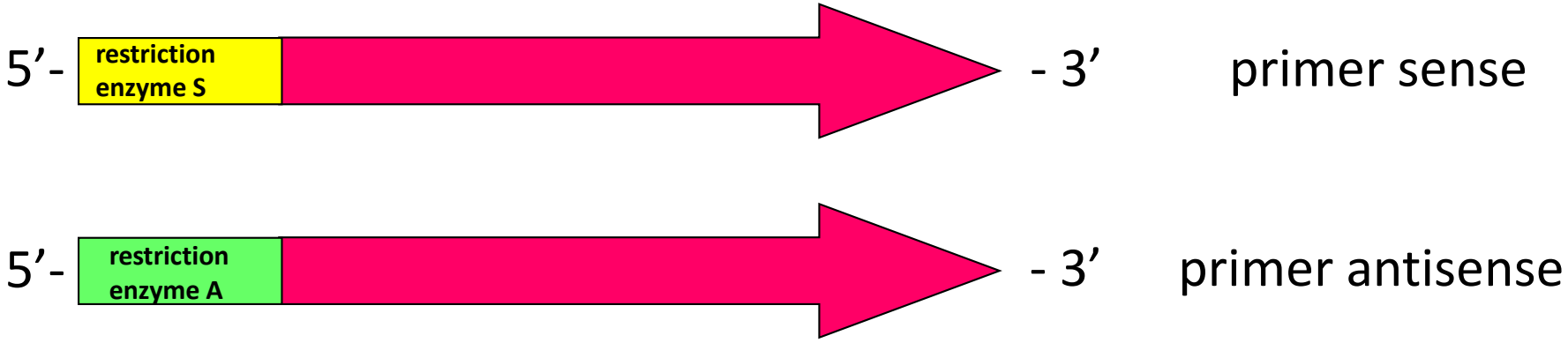
Pay attention to the reading frame!



Example: you decide to include HindIII site upstream the primer:

EGFP

TAC AAG TAC TCA GAT CTC GAG CTC **AAG CTT** CGA ATT CTG CAG TCG



TAC AAG TAC TCA GAT CTC GAG CTC **AAG CTT** **N**NN NNN NNN N**AT G**NN N

EGFP linker NRG1

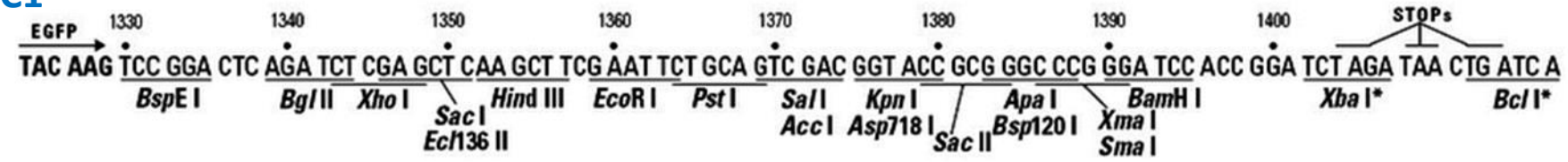
TAC AAG TAC TCA GAT CTC GAG CTC **AAG CTT** NNN NNN NNN NNN NNN NNN **ATG** NNN

EGFP linker NRG1

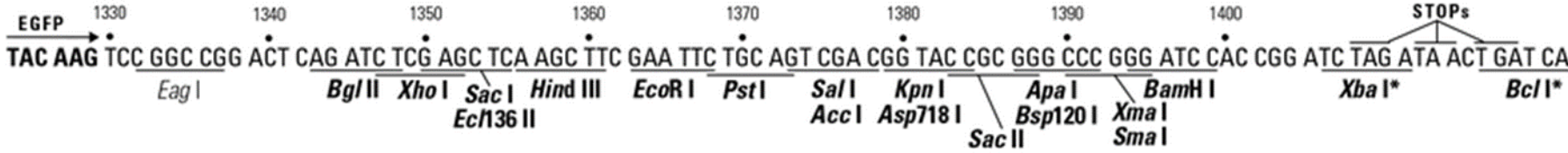
✻ translate the triplets between GFP and NRG1 into amino acids : there must be no STOP

✻ the triplets must be in frame: coding from EGFP to NRG1

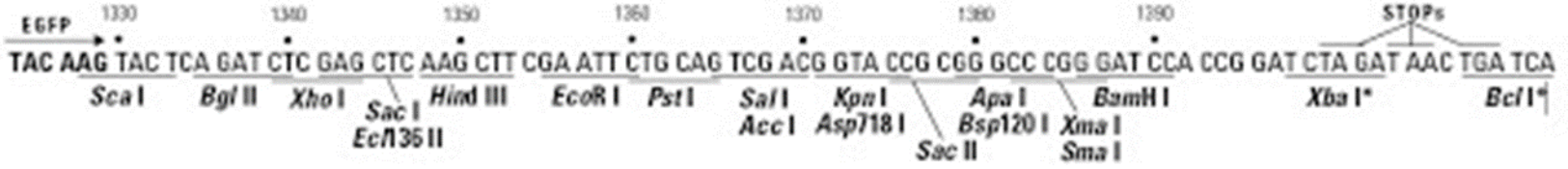
EGFP-C1



EGFP-C2



EGFP-C3



❋ insert your sequence in vector pEGFP-C1

❋ translate the triplets between the GFP and the NRG1 into amino acids : there must be no STOP & the triplets must be in frame: coding from EGFP to NRG1

❋ if it does not work try with pEGFP-C2

❋ if it does not work try with pEGFP-C3

Primer: 5'-**AAGCTT**NNNNNNNNNNATGNNN-3'

EGFP-C1

TAC AAG TCC GGA CTC AGA TCT CGA GCT C**AA GCT T**NN NNN NNN NN**A TG**N NN

EGFP — linker — HindIII — NRG1

EGFP-C2

TAC AAG TCC GGC CGG ACT CAG ATC TCG AGC TC**A AGC TT**N NNN NNN NNN **ATG** NNN

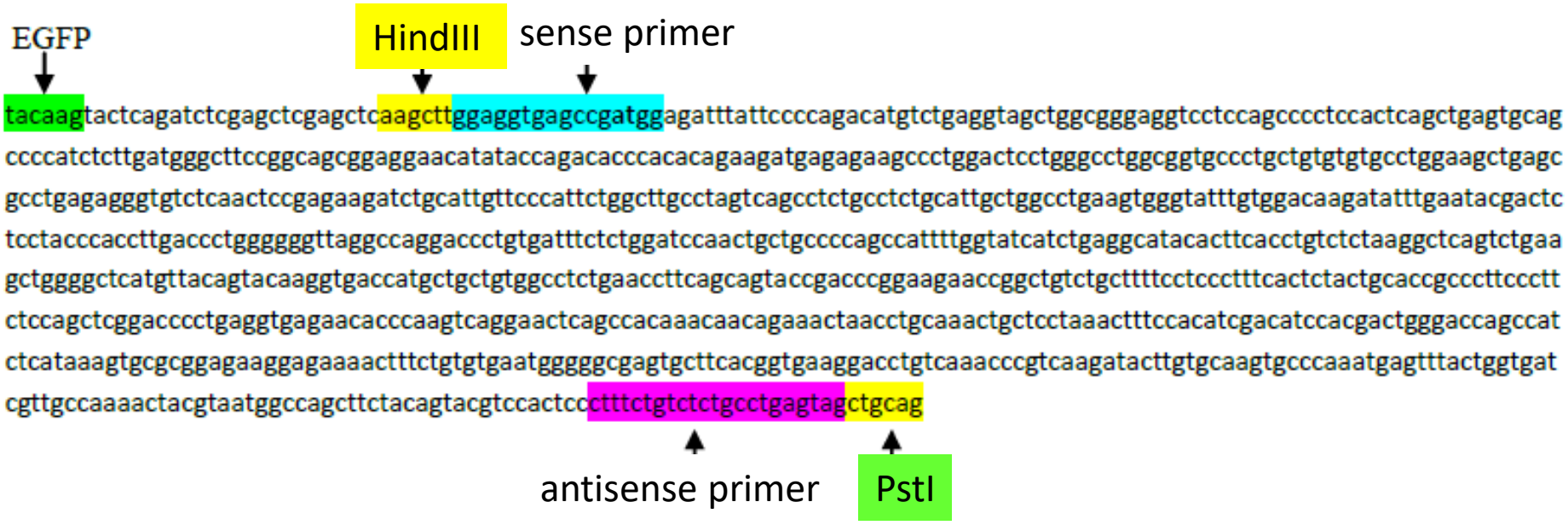
EGFP — linker — HindIII — NRG1

EGFP-C3

TAC AAG TAC TCA GAT CTC GAG CTC **AAG CTT** NNN NNN NNN N**AT G**NN N

EGFP — linker — HindIII — NRG1

1-prepare primers
2-add restriction sites (not cutting the NRG1 cDNA insert!)
3-verify the correct frame to choose the suitable expression vector
4-insert your construct into the expression vector, to obtain the final map and verify the correct frame.



pEGFP-C3

```
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAA
ATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA
GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC
AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTC
CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGA
TAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA
CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC
CACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC
CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGA
GTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCC
ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG
CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGA
GTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTACTCAGATCTCGAGCTCAAGCTTCGAATTC
TGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAGG
TTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTG
TTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTC
TAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAACGCGTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTA
AATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA
GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAA
CCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA...
```

AAGCTT=HindIII
CTGCAG=PstI

pEGFP-C3

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAA
ATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA
GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC
AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTC
CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGA
TAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA
CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC
CACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC
CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGA
GTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCC
ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG
CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGA
GTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTACTCAGATCTCGAGCTC**AAGCTT**

HindIII

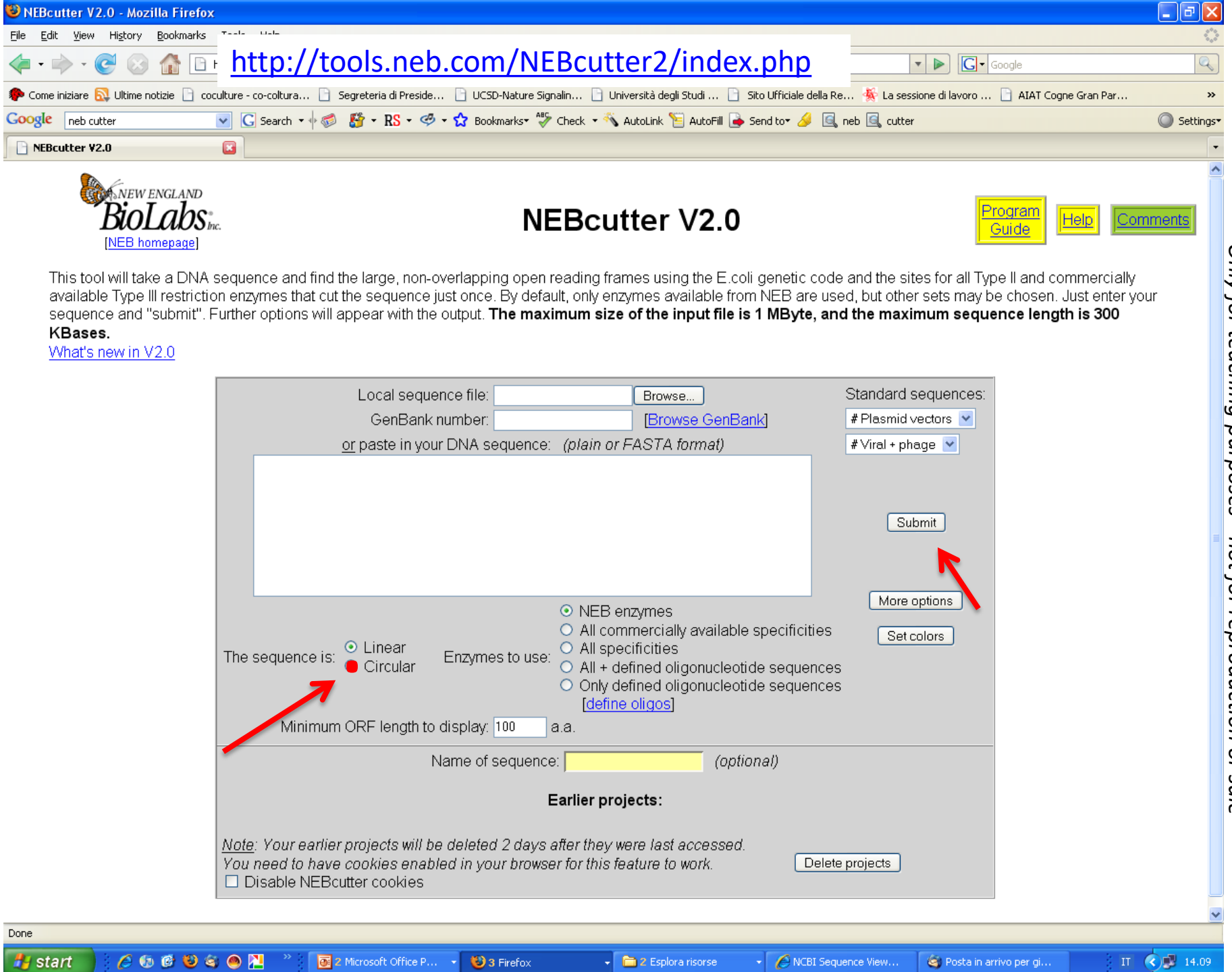
PstI

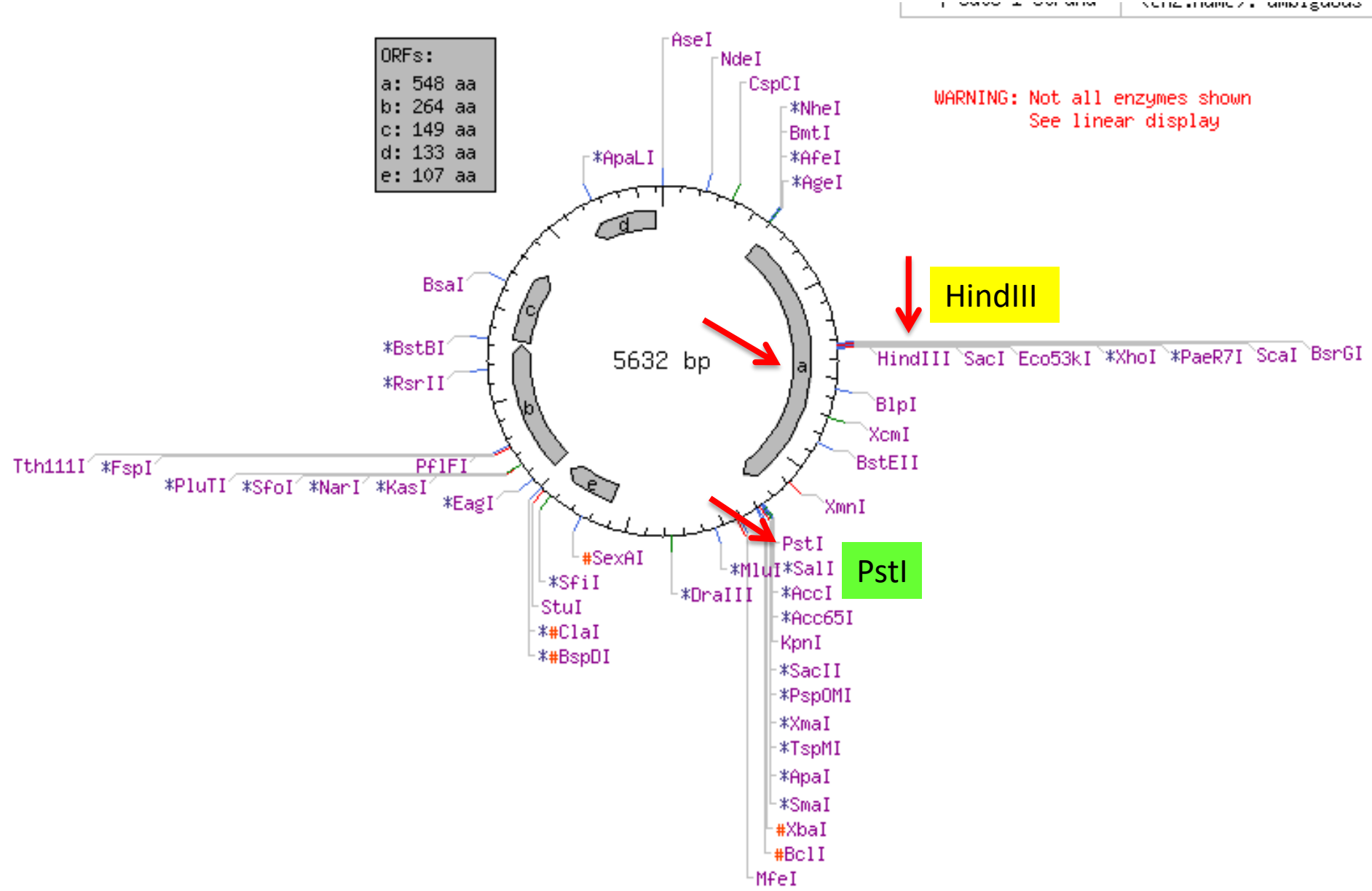
CTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAGG
TTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTG
TTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTC
TAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAACGCGTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTA
AATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA
GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAA
CCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA...


TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAA
ATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA
GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC
AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTC
CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGA
TAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA
CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC
CACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC
CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGA
GTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCC
ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG
CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGA
GTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTACTCAGATCTCGAGCTCAAGCTT
HindIII
aagcttggaggtgagccgatggagatttattccccagacatgtctgaggtagctggcgggaggtcctccagcccctccact
Cagctgagtgcagccccatctcttgatgggcttccggcagcggaggaacatataccagacacccacacagaagatgagaga
Agccctggactcctgggcctggcggtgccctgctgtgtgtgcctggaagctgagcgcctgagagggtgtctcaactccgag
Aagatctgcattgttcccattctggcttgcctagtcagcctctgcctctgcattgctggcctgaagtgggtatttgtggac
Aagatatttgaatacgactctcctacccaccttgaccctgggggttaggccaggaccctgtgatttctctggatccaact
Gctgccccagccattttggtatcatctgaggcatacacttcacctgtctctaaggctcagtctgaagctggggctcatgtt
Acagtacaaggtgaccatgctgctgtggcctctgaaccttcagcagtaccgacccggaagaaccggctgtctgcttttcct
Ccctttcactctactgcaccgcccttcccttctccagctcggacccctgaggtgagaacacccaagtcaggaactcagcca
Caaacaacagaaactaacctgcaaactgctcctaaactttccacatcgacatccacgactgggaccagccatctcataaag
Tgcgcggagaaggagaaaactttctgtgtgaatgggggcgagtgcttcacggtgaaggacctgtcaaacccgtcaagatac
Ttgtgcaagtgcccaaatgagtttactggtgatcgttgccaaaactacgtaatggccagcttctacagtacgtccactccc
tttctgtctctgcctgagtagctgcag
PstI
CTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAGG
TTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTG
TTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTC
TAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAACGCGTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTA
AATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA
GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAA
CCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA ...
AAGCTT=HindIII CTGCAG=PstI

For restriction analysis and identification of ORF and maps, you can use NEB cutter:

http://tools.neb.com/NEBcutter2/index.php




ORFs:
a: 548 aa
b: 264 aa
c: 149 aa
d: 133 aa
e: 107 aa
WARNING: Not all enzymes shown
See linear display
5632 bp
HindIII
PstI
AseI
NdeI
CspCI
*NheI
BmtI
*AfeI
*AgeI
*ApaLI
BsaI
*BstBI
*RsrII
Tth111I
*FspI
PflFI
*PluTI
*SfoI
*NarI
*KasI
*EagI
#SexAI
*SfiI
StuI
*#ClaI
*#BspDI
*DraIII
*MluI
PstI
*SalI
*AccI
*Acc65I
KpnI
*SacII
*PspOMI
*XmaI
*TspMI
*ApaI
*SmaI
#XbaI
#BclI
MfeI
XmnI
BstEII
BlpI
XcmI
HindIII
SacI
Eco53kI
*XhoI
*PaeR7I
ScaI
BsrGI

NEW ENGLAND BioLabs Inc.
NEBcutter

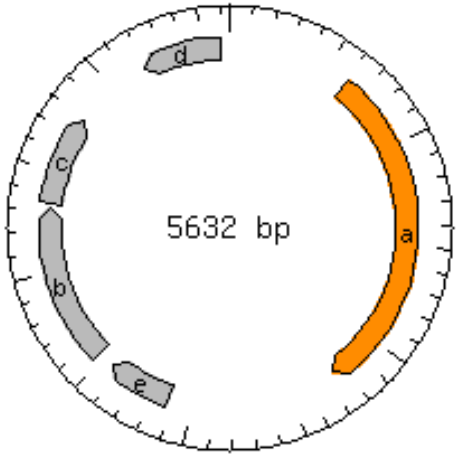
ORF Sequence

unnamed sequence

ck to main display]

Coding region: 613..2259

ORFs:
a: 548 aa
b: 264 aa
c: 149 aa
d: 133 aa
e: 107 aa



[Edit] - [Delete] - [Add new ORF] - [Locate multiple cutters that excise this ORF] - [Silent Mutagenesis]

Protein sequence:

```
> 548 aa
MVSKGEELFT GVVPILVELD GDVNGHKFSV SGEGEGDATY GKLTLKFICT
TGKLPVPWPT LVTTLTYGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF
FKDDGNYKTR AEVKFEGDTL VNRIELKGID FKEDGNILGH KLEYNYNSHN
VYIMADKQKN GIKVNFKIRH NIEDGSVQLA DHYQQNTPIG DGPVLLPDNH
YLSTQSALSK DPNEKRDHMV LLEFVTAAGI TLGMDELYKY SDLELKLGGE
PMEIYSPDMS EVAGGRSSSP STQLSAAPSL DGLPAAEEHI PDTHTEDERS
PGLLGLAVPC CVCLEAERLR GCLNSEKICI VPILACLVSL CLCIAGLKWV
FVDKIFEYDS PTHLDPGGLG QDPVISLDPT AAPAILVSSE AYTSPVSKAQ
SEAGAHVTVQ GDHAAVASEP SAVPTRKNRL SAFPPFHSTA PPFPSPARTP
EVRTPKSGTQ PQTTETNLQT APKLSTSTST TGTSHLIKCA EKEKTFCVNG
GECFTVKDLS NPSRYLCKCP NEFTGDRCQN YVMASFYSTS TPFLSLPE
```

EGFP

NRG1

Blast this sequence at NCBI

Exercize: prepare a **single slide** like the following one, containing all necessary information:

1-the entire sequence from EGFP to the antisense primer, highlighting the primers, the restriction sites, the ATG, the STOP codon)
2- the vector used (1, 2 or 3)
3- restriction sites used (do not use HindIII!!)
4- the primers (written correctly, with restriction sites in **bold**)
5- the protein translation
6- the full map showing the fusion protein

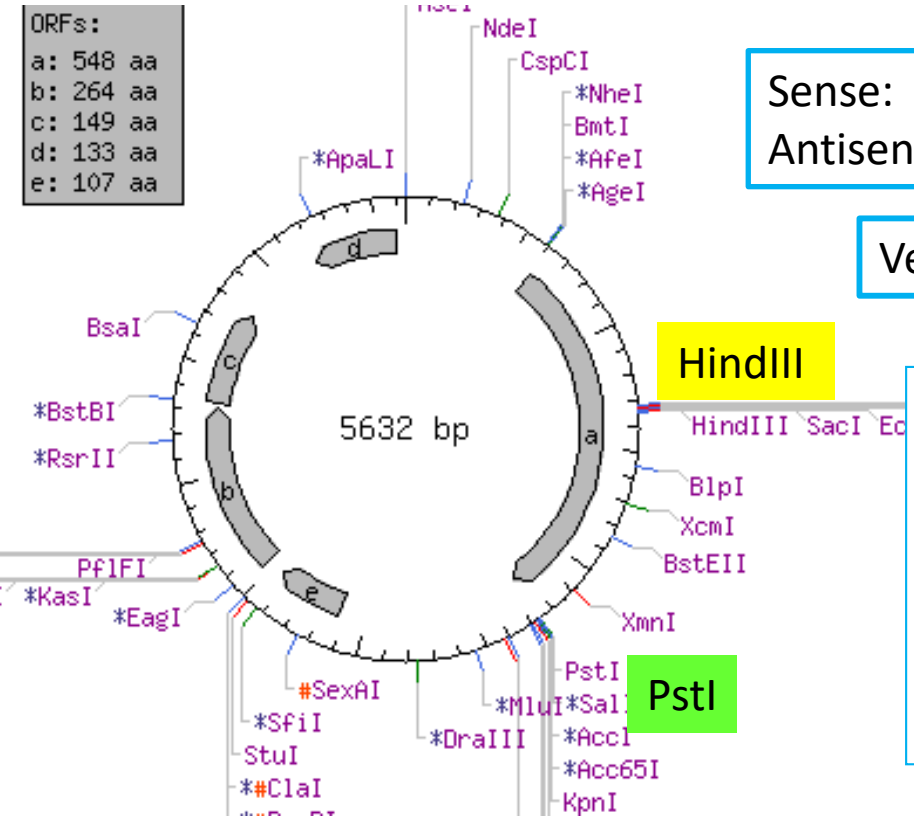
DO NOT use the enzyme HindIII!

EGFP

HindIII sense primer

```
tacaagtactcagatctcgagctcgagctcaagcttggaggtgagccgatggagatttattccccagacatgtctgaggtagctggcgggaggtcctccagcccctccactcagctgagtgcag
ccccatctcttgatgggcttccggcagcggaggaacatataccagacacccacacagaagatgagagaagccctggactcctgggcctggcggtgccctgctgtgtgtgcctggaagctgagc
gcctgagagggtgtctcaactccgagaagatctgcattgttcccattctggcttgcctagtcagcctctgcctctgcattgctggcctgaagtgggtatttgtggacaagatatttgaatacgactc
tcctacccaccttgaccctgggggttaggccaggaccctgtgatttctctggatccaactgctgccccagccattttggtatcatctgaggcatacacttcacctgtctctaaggctcagtctgaa
gctggggctcatgttacagtacaaggtgaccatgctgctggcctctgaaccttcagcagtaccgacccggaagaaccggctgtctgcttttcctccctttcactctactgcaccgcccttccctt
ctccagctcggacccctgaggtgagaacacccaagtcaggaactcagccacaaacaacagaaactaacctgcaaactgctcctaaactttccacatcgacatccacgactgggaccagccat
ctcataaagtgcgcggagaaggagaaaactttctgtgtgaatgggggcgagtgcttcacggtgaaggacctgtcaaacccgtcaagatacttgtgcaagtgcccaaatgagtttactggtgat
cgttgccaaaactacgtaatggccagcttctacagtacgtccactccctttctgtctctgcctgagtagctgcag
```

antisense primer PstI



PRIMERS

Sense: 5’-**AAGCTT**GGAGGTGAGCCGATGG-3’

Antisense: 5’-**CTGCAG**CTACTCAGGCAGAGACAGAAAG-3’

Vector: pEGFP-C3

TRANSLATION

```
> 548 aa
MVSKGEELFT GVVPILVELD GDVNGHKFSV SGEGEGDATY GKLTLKFICT
TGKLPVPWPT LVTTLTYGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF
FKDDGNYKTR AEVKFEGDTL VNRIELKGID FKEDGNILGH KLEYNYNSHN
VYIMADKQKN GIKVNFKIRH NIEDGSVQLA DHYQQNTPIG DGPVLLPDNH
YLSTQSALSK DPNEKRDHMV LLEFVTAAGI TLGMDELYKY SDLELKLGGE
PMEIYSPDMS EVAGGRSSSP STQLSAAPSL DGLPAAEEHI PDTHTEDERS
PGLLGLAVPC CVCLEAERLR GCLNSEKICI VPILACLVSL CLCIAGLKWV
FVDKIFEYDS PTHLDPGGLG QDPVISLDPT AAPAILVSSE AYTSPVSKAQ
SEAGAHVTVQ GDHAAVASEP SAVPTRKNRL SAFPPFHSTA PPFPSPARTP
EVRTPKSGTQ PQTTETNLQT APKLSTSTST TGTSHLIKCA EKEKTFCVNG
GECFTVKDLS NPSRYLCKCP NEFTGDRCQN YVMASFYSTS TPFLSLPE
```