Esercizio 1

Clonaggio di una sequenza nucleotidica codificante

In **Allegato 1** è fornita la sequenza del cDNA (più regioni fiancheggianti) codificante per la **proteina prionica (PRNC)** umana. La struttura primaria della proteina è la seguente:

[1]MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGGGWGQPHGGGWGQPHGGGWGQPHGG
GWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHRYPN
QVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPP
VILLISFLIFLIVG[253]

In **Allegato 1** trovate anche la mappa di restrizione della sequenza di DNA generata dal programma Webcutter 2.0. Sulla mappa in allegato sono evidenziati il codone di inizio della regione codificante e il codone di stop. Verificate di trovarvi nella regione codificante traducendo i primi 5 codoni e gli ultimi 5, utilizzando il *codice genetico* (**Allegato 2**).

Disegnate degli oligonucleotidi (della lunghezza di 25-28 nucleotidi) che servano da *primer* per una reazione di PCR per l'amplificazione del cDNA della proteina prionica, in modo da poter poi clonare tale cDNA nel plasmide pET22b(+) (mappa in **Allegato 3**), tra i siti di restrizione *Nco*I e *Xho*I. È possibile utilizzare questi due enzimi di restrizione per il clonaggio? Altrimenti utilizzate *Nde*I come sito di clonaggio a monte. Ipotizzate due diversi clonaggi: nel primo dovrete aggiungere alla proteina prionica, in posizione C-terminale, la coda di istidine presente nel plasmide (attenzione a non andare fuori *frame*!!!). Nel secondo clonaggio invece dovete fare in modo che la coda di istidine non venga tradotta.

ALLEGATO 1

>NM_000311.5 Homo sapiens prion protein (PRNP), transcript variant 1, mRNA 2435 bp

```
   1 gccagtcgct gacagccgcg gcgccgcgag cttctcctct cctcacgacc gaggcagagc
  61 agtcatatg gcgaaccttg gctgctggat gctggttctc tttgtggcca catggagtga
 121 cctgggcctc tgcaagaagc gcccgaagcc tggaggatgg aacactgggg gcagccgata
 181 cccggggcag ggcagccctg gaggcaaccg ctacccacct cagggcggtg gtggctgggg
 241 gcagcctcat ggtggtggct gggggcagcc tcatggtggt ggctgggggc agccccatgg
 301 tggtggctgg ggacagcctc atggtggtgg ctggggtcaa ggaggtggca cccacagtca
 361 gtggaacaag ccgagtaagc caaaaaccaa catgaagcac atggctggtg ctgcagcagc
 421 tggggcagtg gtgggggggcc ttggcggcta catgctggga agtgccatga gcaggcccat
 481 catacatttc ggcagtgact atgaggaccg ttactatcgt gaaaacatgc accgttaccc
 541 caaccaagtg tactacaggc ccatggatga gtacagcaac cagaacaact ttgtgcacga
 601 ctgcgtcaat atcacaatca agcagcacac ggtcaccaca accaccaagg gggagaactt
 661 caccgagacc gacgttaaga tgatggagcg cgtggttgag cagatgtgta tcacccagta
 721 cgagagggaa tctcaggcct attaccagag aggatcgagc atggtcctct tctcctctcc
 781 acctgtgatc ctcctgatct ctttcctcat cttcctgata gtgggatgag gaaggtcttc
 841 ctgttttcac catctttcta atctttttcc agcttgaggg aggcggtatc cacctgcagc
 901 ccttttagtg gtggtgtctc actctttctt ctctctttgt cccggatagg ctaatcaata
 961 cccttggcac tgatgggcac tggaaaacat agagtagacc tgagatgctg gtcaagcccc
1021 ctttgattga gttcatcatg agccgttgct aatgccaggc cagtaaaagt ataacagcaa
1081 ataaccattg gttaatctgg acttattttt ggacttagtg caacaggttg aggctaaaac
1141 aaatctcaga acagtctgaa ataccttgc ctggatacct ctggctcctt cagcagctag
1201 agctcagtat actaatgccc tatcttagta gagatttcat agctatttag agatattttc
1261 cattttaaga aaacccgaca acatttctgc caggtttgtt aggaggccac atgatactta
1321 ttcaaaaaaa tcctagagat tcttagctct tgggatgcag gctcagcccg ctggagcatg
1381 agctctgtgt gtaccgagaa ctggggtgat gttttacttt tcacagtatg ggctacacag
1441 cagctgttca acaagagtaa atattgtcac aacactgaac ctctggctag aggacatatt
1501 cacagtgaac ataactgtaa catatatgaa aggcttctgg gacttgaaat caaatgtttg
1561 ggaatggtgc ccttggaggc aacctcccat tttagatgtt taaaggaccc tatatgtggc
1621 attcctttct ttaaactata ggtaattaag gcagctgaaa agtaaattgc cttctagaca
1681 ctgaaggcaa atctcctttg tccatttacc tggaaaccag aatgattttg acatacagga
1741 gagctgcagt tgtgaaagca ccatcatcat agaggatgat gtaattaaaa aatggtcagt
1801 gtgcaaagaa aagaactgct tgcatttctt tatttctgtc tcataattgt caaaaaccag
1861 aattaggtca agttcatagt ttctgtaatt ggcttttgaa tcaaagaata gggagacaat
1921 ctaaaaaata tcttaggttg gagatgacag aaatatgatt gatttgaagt ggaaaaagaa
1981 attctgttaa tgttaattaa agtaaaatta ttccctgaat tgtttgatat tgtcacctag
2041 cagatatgta ttacttttct gcaatgttat tattggcttg cactttgtga gtattctatg
2101 taaaaatata tatgtatata aaatatatat tgcataggac agacttagga gttttgttta
2161 gagcagttaa catctgaagt gtctaatgca ttaacttttg taaggtactg aatacttaat
2221 atgtgggaaa ccctttgcg tggtccttag gcttacaatg tgcactgaat cgtttcatgt
2281 aagaatccaa agtggacacc attaacaggt ctttgaaata tgcatgtact ttatattttc
2341 tatatttgta actttgcatg ttcttgtttt gttatataaa aaaattgtaa atgtttaata
2401 tctgactgaa attaaacgag cgaagatgag cacca
```

Table by Enzyme Name

Enzyme name	No. cuts	Positions of sites	Recognition sequence
AatI	1	737	agg/cct
AccB1I	3	20 347 1566	g/gyrcc
AccI	2	995 1208	gt/mkac
AcsI	1	1979	r/aatty
AcyI	1	21	gr/cgyc

Alw44I	2	593 2260	g/tgcac
AlwNI	3	8 419 1443	cagnnn/ctg
Ama87I	1	181	c/ycgrg
AocI	2	219 2246	cc/tnagg
ApaLI	2	593 2260	g/tgcac
ApoI	1	1979	r/aatty
Asp700I	3	835 1532 2271	gaann/nnttc
AspI	1	602	gacn/nngtc
AtsI	1	602	gacn/nngtc
AvaI	1	181	c/ycgrg
BalI	1	107	tgg/cca
BanI	3	20 347 1566	g/gyrcc
BanII	2	1204 1384	grgcy/c
BbeI	1	24	ggcgc/c
BbiII	1	21	gr/cgyc
BbsI	1	840	gaagac
Bbv16II	1	840	gaagac
BcoI	1	181	c/ycgrg
BglI	2	201 444	gccnnnn/nggc
BlpI	1	1362	gc/tnagc
BpiI	1	840	gaagac
BpmI	3	155 203 1376	ctggag
Bpu1102I	1	1362	gc/tnagc
BpuAI	1	840	gaagac
BsaBI	1	707	gatnn/nnatc
BsaHI	1	21	gr/cgyc
BsaI	1	670	ggtctc
BsaMI	1	1624	gaatgc
BsaOI	1	49	cgry/cg
Bse21I	2	219 2246	cc/tnagg
Bse8I	1	707	gatnn/nnatc
BseRI	2	39 777	gaggag
Bsh1285I	1	49	cgry/cg
Bsh1365I	1	707	gatnn/nnatc
BshNI	3	20 347 1566	g/gyrcc
BsiEI	1	49	cgry/cg
BsmI	1	1624	gaatgc
BsoBI	1	181	c/ycgrg
Bsp143II	2	24 142	rgcgc/y
Bsp1720I	1	1362	gc/tnagc
Bsp19I	2	295 561	c/catgg
BspHI	1	1036	t/catga
BspMI	1	897	acctgc
BsrBRI	1	707	gatnn/nnatc
BsrDI	1	2066	gcaatg
Bst1107I	1	1209	gta/tac
BstDSI	3	16 295 561	c/crygg
BstEII	1	631	g/gtnacc
BstH2I	2	24 142	rgcgc/y
BstMCI	1	49	cgry/cg
BstPI	1	631	g/gtnacc
BstXI	2	302 1709	ccannnnn/ntgg
Bsu36I	2	219 2246	cc/tnagg
CelII	1	1362	gc/tnagc
Cfr42I	1	19	ccgc/gg
Cfr9I	1	181	c/ccggg
CfrI	1	105	y/ggccr
CvnI	2	219 2246	cc/tnagg
DraI	2	1601 1632	ttt/aaa
DraII	2	436 1605	rg/gnccy
DraIII	1	2086	cacnnn/gtg
DsaI	3	16 295 561	c/crygg
EaeI	1	105	y/ggccr

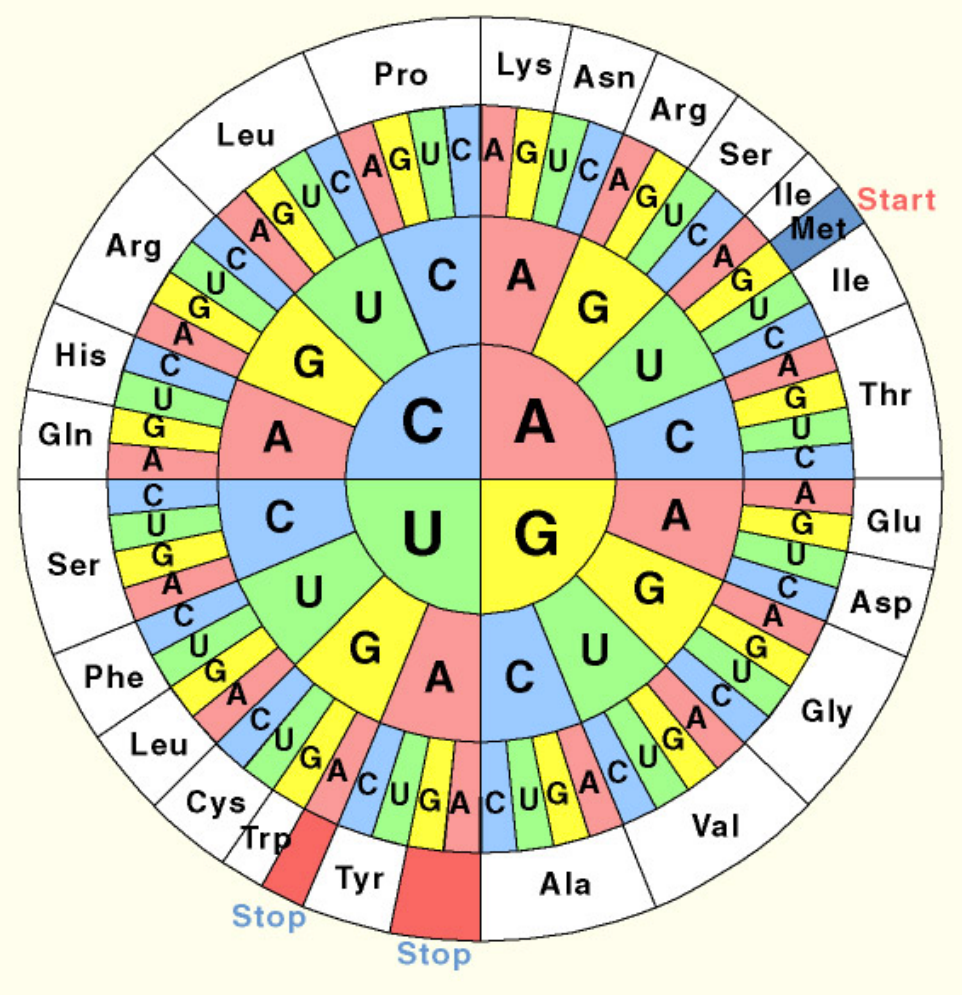
Eam1104I	1	772	ctcttc
EarI	1	772	ctcttc
Ecl136II	2	1202 1382	gag/ctc
Eco147I	1	737	agg/cct
Eco24I	2	1204 1384	grgcy/c
Eco31I	1	670	ggtctc
Eco57I	3	1193 1686 2179	ctgaag
Eco64I	3	20 347 1566	g/gyrcc
Eco81I	2	219 2246	cc/tnagg
Eco88I	1	181	c/ycgrg
Eco91I	1	631	g/gtnacc
EcoICRI	2	1202 1382	gag/ctc
EcoO109I	2	436 1605	rg/gnccy
EcoO65I	1	631	g/gtnacc
EcoT22I	2	2190 2324	atgca/t
EheI	1	22	ggc/gcc
FriOI	2	1204 1384	grgcy/c
GsuI	3	155 203 1376	ctggag
HaeII	2	24 142	rgcgc/y
Hin1I	1	21	gr/cgyc
HincII	1	2168	gty/rac
HindII	1	2168	gty/rac
HpaI	1	2168	gtt/aac
Hsp92I	1	21	gr/cgyc
KasI	1	20	g/gcgcc
Ksp632I	1	772	ctcttc
KspI	1	19	ccgc/gg
MamI	1	707	gatnn/nnatc
MluNI	1	107	tgg/cca
Mph1103I	2	2190 2324	atgca/t
MscI	1	107	tgg/cca
Msp17I	1	21	gr/cgyc
Mva1269I	1	1624	gaatgc
NarI	1	21	gg/cgcc
NcoI	2	295 561	c/catgg
NsiI	2	2190 2324	atgca/t
PacI	1	1997	ttaat/taa
Pme55I	1	737	agg/cct
Ppu10I	2	2186 2320	a/tgcat
PpuMI	1	1605	rg/gwccy
Psp124BI	2	1204 1384	gagct/c
Psp5II	1	1605	rg/gwccy
PspAI	1	181	c/ccggg
PspALI	1	183	ccc/ggg
PspEI	1	631	g/gtnacc
PstI	3	415 898 1748	ctgca/g
PvuII	3	419 1443 1654	cag/ctg
RcaI	1	1036	t/catga
SacI	2	1204 1384	gagct/c
SacII	1	19	ccgc/gg
Sfr303I	1	19	ccgc/gg
SmaI	1	183	ccc/ggg
SseBI	1	737	agg/cct
SspI	1	1462	aat/att
SstI	2	1204 1384	gagct/c
SstII	1	19	ccgc/gg
StuI	1	737	agg/cct
Tth111I	1	602	gacn/nngtc
VneI	2	593 2260	g/tgcac
XbaI	1	1673	t/ctaga
XmaI	1	181	c/ccggg
XmnI	3	835 1532 2271	gaann/nnttc
Zsp2I	2	2190 2324	atgca/t

The following endonucleases were selected but don't cut this sequence:

AatII, Acc113I, Acc16I, Acc65I, AccB7I, AccBSI, AccIII, AclNI, AfeI, AflII, AflIII, AgeI, AhdI, Aor51HI, ApaI, AscI, AseI, AsnI, Asp718I, AspEI, AviII, AvrII, BamHI, BanIII, BbrPI, BbuI, BcgI, BclI, BfrI, BglII, BlnI, Bpu14I, Bsa29I, BsaAI, BsaWI, BscI, Bse118I, BseAI, BseCI, BsePI, BsgI, BsiI, BsiMI, BsiWI, BsmBI, Bsp106I, Bsp119I, Bsp120I, Bsp13I, Bsp1407I, Bsp68I, BspCI, BspDI, BspEI, BspLU11I, BspTI, BspXI, BsrBI, BsrFI, BsrGI, BssAI, BssHII, BssSI, Bst98I, BstBI, BstD102I, BstI, BstSNI, BstX2I, BstYI, BstZI, Bsu15I, CciNI, Cfr10I, ClaI, CpoI, Csp45I, CspI, DrdI, EagI, Eam1105I, EclHKI, EclXI, Eco105I, Eco255I, Eco32I, Eco47III, Eco52I, Eco72I, EcoNI, EcoRI, EcoRV, Esp1396I, Esp3I, FauNDI, FbaI, FseI, FspI, HindIII, Kpn2I, KpnI, Ksp22I, LspI, MfeI, MflI, MluI, MroI, MroNI, MspCI, MunI, NaeI, NdeI, NgoAIV, NgoMI, NheI, NotI, NruI, NspV, PaeI, PaeR7I, Pfl23II, PflMI, PinAI, Ple19I, PmaCI, PmeI, PmlI, PshAI, PshBI, Psp1406I, PspLI, PspOMI, PstNHI, PvuI, RsrII, SalI, SapI, SbfI, ScaI, SexAI, SfiI, Sfr274I, SfuI, SgfI, SgrAI, SmiI, SnaBI, SpeI, SphI, SplI, SrfI, Sse8387I, SspBI, SunI, SwaI, Van91I, Vha464I, VspI, XcmI, XhoI, XhoII, XmaIII

ALLEGATO 2

Il codice genetico



ALLEGATO 3

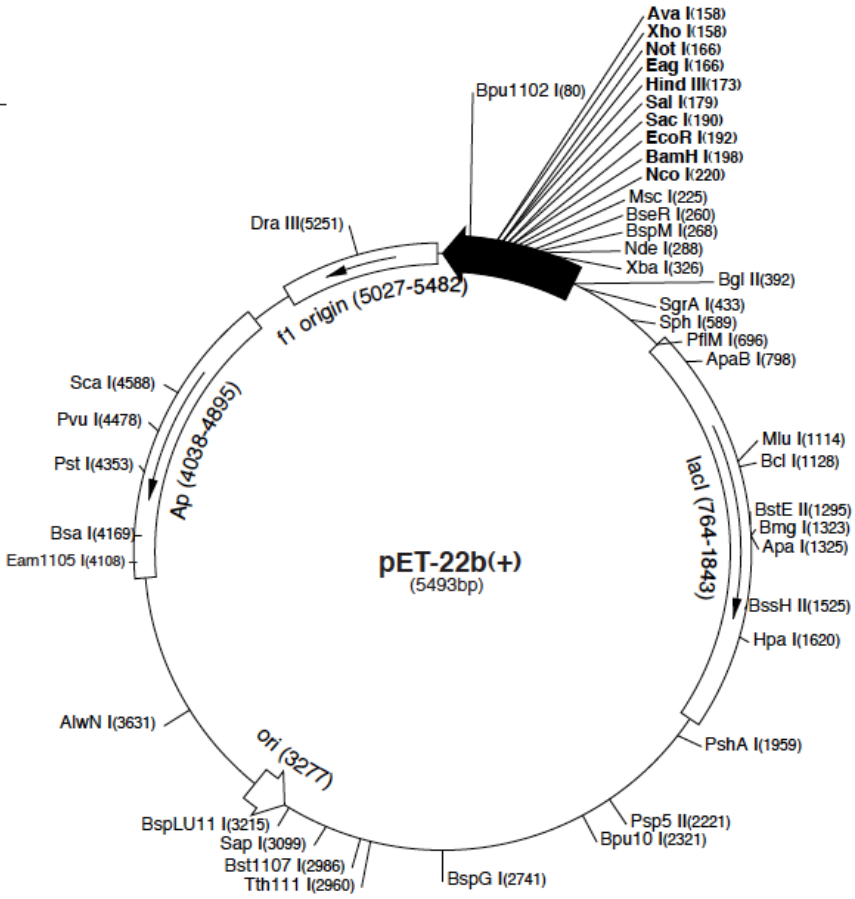
Mappa del vettore pET-22b

pET-22b(+) Vector

TB038 12/98

The pET-22b(+) vector (Cat. No. 69744-3) carries an N-terminal *pelB* signal sequence for potential periplasmic localization, plus optional C-terminal His•Tag® sequence. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer (Cat. No. 69337-3).

pET-22b(+) sequence landmarks	
T7 promoter	361-377
T7 transcription start	360
pelB coding sequence	224-289
Multiple cloning sites	
(*Nco* I - *Xho* I)	158-225
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	764-1843
pBR322 origin	3277
bla coding sequence	4038-4895
f1 origin	5027-5482



Bgl II — T7 promoter — lac operator — *Xba* I — rbs

--- AGA TCT CGA TCC CGC GAA ATT AAT ACG ACT CAC TAT AGG GGA ATT GTG AGC GGA TAA CAA TTC CCC TCT AGA AAT AAT TTT GTT TAA CTT TAA GAA GGA GAT

Nde I — pelB leader — *Nco* I — *Bam*H I — *Eco*R I — *Sac* I

ATA CAT ATG AAA TAC CTG CTG CCG ACC GCT GCT GCT GGT CTG CTG CTC CTC GCT GCC CAG CCG GCG ATG GCC ATG GAT ATC GGA ATT AAT TCG GAT CCG AAT TCG AGC TCC

M K Y L L P T A A A G L L L A A Q P A M A M D I G I N S D P N S S S

Sal I — *Hind* III — *Not* I — *Xho* I — His tag — *Blp* I

GTC GAC AAG CTT GCG GCC GCA CTC GAG CAC CAC CAC CAC CAC CAC TGA GAT CCG GCT GCT AAC AAA GCC CGA AAG GAA GCT GAG TTG GCT GCT GCC ACC GCT GAG CAA

V D K L A A A L E H H H H H H Stop

T7 terminator

TAA CTA GCA TAA CCC CTT GGG GCC TCT AAA CGG GTC TTG AGG GGT TTT TTG ---

Esercizio 2

Clonaggio e analisi di una sequenza nucleotidica codificante

Scaricare dal sito https://www.ncbi.nlm.nih.gov/nuccore la sequenza nucleotidica codificante per Ftr1 del lievito *Pichia pastoris* (accession number AJ937309).

Utilizzare il programma Webcutter 2.0 http://heimanlab.com/cut2.html per generare la mappa di restrizione di Ftr1.

Disegnare due oligo per il clonaggio in pET22b includendo il leader pelB e aggiungendo la coda di istidine al C-terminale.

Utilizzare il programma Translate https://web.expasy.org/translate/ per ottenere la traduzione della sequenza nucleotidica in amminoacidica.

Disegnare un oligo per sostituire il residuo Glu 156 con Ala.

Utilizzare il programma ProtParam https://web.expasy.org/protparam/ per ottenere i parametri biochimici della proteina.