



small t intron

ataaaatTTTTAAGTGTATAATGTGTTAACTACTGATTCTAATTGTTTGTGATTTTAGATTCCAACCTATGGAACGTATGAATGGGAGCAGTG
TATTTAAAAATTCACATATTACACAATTTGATGACTAAGATTAACAAACACATAAAATCTAAGGTTGGATACCTTGACTACTTACCCTCGTCAC

small t intron

BseRI

GTGGAATGCCTTAAATGAGGAAAACCTGTTTGTCTAGAAGAAATGCCATCTAGTGTATGAGGCTACTGCTGACTCTCAACATTCTACTCCTC
CACCTTACGGAATTACTCCTTTTGGACAAAACGAGTCTTCTTTACGGTAGATCACTACTACTCCGATGACGACTGAGAGTTGTAAGATGAGGAG

SV40 NLS

BbsI

CAAAAAAGAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTGTAGTCATGCTGTGTTTAGTAATAGAACTCTTGCT
GTTTTTCTTCTCTTTCCATCTTCTGGGGTTCCTGAAAGGAGTCTTAACGATTCAAAAAATCAGTACGACACAAATCATTATCTTGAGAACGA

Pro Lys Lys Lys Arg Lys Val Glu Asp Pro Lys Asp Phe Pro Ser Glu Leu Leu Ser Phe Leu Ser His Ala Val Phe Ser Asn Arg Thr Leu Ala
SV40 NLS (in frame with SV40 NLS)

BsgI

TGCTTTGCTATTTACACCACAAAGGAAAAGCTGCCTGCTATACAAGAAAATATGGAAAAATATTCTGTAACCTTTATAAGTAGGCATAACAG
ACGAAACGATAAATGTGGTGTCTTCTTTCGACGTGACGATATGTTCTTTAATACCTTTTATAAGACATTGGAAATATTCATCCGTATTGTC

Cys Phe Ala Ile Tyr Thr Thr Lys Glu Lys Ala Ala Leu Leu Tyr Lys Lys Ile Met Glu Lys Tyr Ser Val Thr Phe Ile Ser Arg His Asn Ser
(in frame with SV40 NLS)

TTATAATCATAACATACTGTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATGTGTACTTTAGCTTTT
AATATTAGTATGTATGACAAAAAAGATGAGGTGTGTCGATCTCACAGACGATAATTATGATACGAGTTTTTAACACATGGAAATCGAAAA

Tyr Asn His Asn Ile Leu Phe Phe Leu Thr Pro His Arg His Arg Val Ser Ala Ile Asn Asn Tyr Ala Gln Lys Leu Cys Thr Phe Ser Phe
(in frame with SV40 NLS)

BsaBI*

TAATTTGTAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGC
ATTAACATTTCCCAATTTCTTATAAACTACATATCACGGAACGTATCTCTAGTATTAGTCGGTATGGTGTAACATCTCCAAAATGAACG

Leu Ile Cys Lys Gly Val Asn Lys Glu Tyr Leu Met Tyr Ser Ala Leu Thr Arg Asp His Asn Gln Pro Tyr His Ile Cys Arg Gly Phe Thr Cys
(in frame with SV40 NLS)

MfeI

HpaI

TTAAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTGTTTATTGCACTTATAATGGTTACA
AAATTTTTGGAGGGTGTGGAGGGGACTGGACTTTGTATTTACTTACGTTAACAACAACAATTGAACAATAACGTCGAATATTACCAATGT

Phe Lys Lys Pro Pro Thr Pro Pro Pro Glu Pro Glu Thr
(in frame with SV40 NLS)

SV40 poly(A) signal

AATAAGCAATAGCATCACAATTTCAAAAATAAAGCATTTTTTCACTGCATCTAGTTGTGGTTGTCCAAACTCATCAATGTATCTTATCAT
TTATTTGTTATCGTAGTGTAAAGTGTATTTTCGTAAAAAAAGTGACGTAAGATCAACACCAACAGGTTTGAGTAGTTACATAGAATAGTA

SV40 poly(A) signal

Acc65I

KpnI

Eco53kI

SacI

BanII

GCTGGATCCCCGGTACCAGACTCGAATTCACTGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCAACTTAATCGCCTT
CAGACCTAGGGGCCATGGCTCGAGCTTAAGTGACCGGCAGCAAAATGTTGACGACTGACCCTTTGGGACCGCAATGGGTGAATTAGCGGAA

SV40 poly(A) signal

M13 fwd

KasI

NarI

SfoI

GCAGCACATCCCCCTTTCGCACTGGCGTAATAGCGAAGAGGCCGCACCGATCGCCCTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCG
CGTCGTGTAGGGGAAAGCGGTGACCCGCAATTATCGCTTCTCCGGGCGTGGCTAGCGGGAAGGGTTGTCAACGCGTCGGACTTACCCTTACC

PluTI

NdeI

BstAPI

CCTGATGCGGTATTTTCTCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GGACTACGCCATAAAGAGGAATGCGTAGACACGCCATAAAGTGTGGCGTATACCACGTGAGAGTCAATGTAGACGAGACTACGGCGTATCAATT

PfoI

g ccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccggg
cggtcggggctgtggggcgtgtggggcactgctgcgggactgcccgaacagacgagggccgtaggcgaatgtctgttcgacactggcagaggccc

2090

EcoO109I

agctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgccattttttataggttaatgtcatgat
tcgacgtacacagctctccaaaagtggcagtagtggtttgctgcttcccgagcactatgaggataaaaaatatccaattacagtacta

2185

ZraI AatII

aataatggtttcttagacgtcagggtggcacttttcggggaaatgtgcggaaccctattgtttattttctaaatacattcaaatatgtatc
ttattacaaagaatctgcagtcaccggtgaaaagccctttacacgcgccttggggataaacaataaaaagatttatgtaagttatacatag

2280

AmpR promoter

cgctcatgagacaataaccctgataaatgcttcaataatattgaaaaggaagagtatgagtattcaacatttccggtgctgcccttattccctt
gcgagctactctgttattgggactatttacgaagtattataacttttccctctcatactcataagttgtaaggcacagcgggataagggaaa

2375

AmpR promoter

1 5 10
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe
signal sequence
AmpR

tttggggcattttgccttctgtttttgctcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagttgggttacat
aaacgccgtaaaacggaaggacaaaaacgagtggtctttgcgaccactttcattttctacgacttctagtcacccacgctgctcacccaatgta

2470

15 20 25 30 35 40 45
Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile
signal sequence
AmpR

XmnI

cgaaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaagttctgctatgtggcg
gcttgacctagagttgtcgccattctaggaactctcaaaagcggggcttcttgcaaaaggttactactcgtgaaaattcaagacgatacaccgc

2565

50 55 60 65 70 75
Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly
AmpR

cggtattatcccgatttgacgcccgggcaagagcaactcggctgcgcgcatacactattctcagaatgacttgggtgagctaccagtcacagaa
gccataatagggcataactgcccgccttctcgttgagccagcggcgtatgtgataagagcttactgaaaccaactcatgagttggtcagtgctt

2660

80 85 90 95 100 105 110
Ala Val Leu Ser Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu
AmpR

aagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataaacactgcccgaacttacttctgacaacgat
ttcgtagaatgacctaccgtactgtcatctcttaaatcgtcacgacgggtatgggtactcactattgtgacgccgggtgaaatgaagactgttgcta

2755

110 115 120 125 130 135 140
Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile
AmpR

cgaggaccgaaggagctaaccgctttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaaatgaagccatac
gctcctggcttctcgattggcgaaaaaacgtttgtacccttagtacattgagcggaaactagcaacccttggcctcgacttacttcgggtatg

2850



























145 150 155 160 165 170
Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile
AmpR

caaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgtttgcgcaactattaactggcgaactacttactctagcttccggcaa
gtttgctgctcgcactgtgggtctacggacatcgttaccggttttgcaacgcgtttgataatgaccgcttgatgaaatgagatcgaaggccgtt

2945

175 180 185 190 195 200
Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln
AmpR

Enzymes	Sites	
AatII	1	2205
AccI	1	43
Acc65I	1	1724
AflIII	1	4012
AhdI	1	3124
AlwNI	1	3603
BanII	1	1734
BbsI	1	1168
Bpu10I	1	106
BsaI	1	3058
BsaBI	1*	1483*
BseRI	1	1126
BseYI	1	3708
BsgI	1	1252
BspEI	1	329
BspQI	1	4129
BsrFI	1	3039
BstAPI	1	1954
BtgI	1	634
Eco53kI	1	1732
EcoO109I	1	2144
HpaI	1	1584
KasI	1	1897
KpnI	1	1728
MfeI	1	1571
MscI	1	600
NarI	1	1898
NcoI	1	634
NdeI	1	1950
NmeAIII	1	2977
PaeR7I	1	1
PasI	1	565
PciI	1	4012
PfoI	1	2085
PluTI	1	1901
PspFI	1	3712
SacI	1	1734
SalI	1	42
SapI	1	4129
SfoI	1	1899
TliI	1	1
XbaI	1	48
XhoI	1	1
XmnI	1	2524
ZraI	1	2203

Feature	Location	Size (bp)			Type
✓ CmR	120 .. 779	660			CDS
/gene	= cat				
/product	= chloramphenicol acetyltransferase				
/note	= confers resistance to chloramphenicol				
/translation	= MEKKITGYTTVDISQWHRKEHFEAFQSVQAQCTYNQTVQLDITAFLLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFFSSLWSEYHDDFRQFLH IYSQDVACYGENLAYFPKGFIEENMFFVSNPWSFSTSFDLNVANMDFPAPVFTMGKYYTQGDKVLMLPLAIQVHHAVCDGFHVRMLNELQYQCDEWQGGGA*				
	219 amino acids = 25.7 kDa				
✓ small t intron	945 .. 1010	66			intron
/note	= simian virus 40 (SV40) small t antigen intron				
✓ SV40 NLS	1140 .. 1160	21			CDS
/product	= nuclear localization signal of SV40 large T antigen				
/translation	= PKKKRKV				
	7 amino acids = 883.1 Da				
✓ SV40 poly(A) signal	1585 .. 1719	135			polyA_signal
/note	= SV40 polyadenylation signal				
✓ M13 fwd	1742 .. 1758	17			primer_bind
/note	= common sequencing primer, one of multiple similar variants				
✓ AmpR promoter	2232 .. 2336	105			promoter
/gene	= bla				
✓ AmpR	2337 .. 3197	861			CDS
▶ 2 segments					
/gene	= bla				
/product	= β-lactamase				
/note	= confers resistance to ampicillin, carbenicillin, and related antibiotics				
/translation	= MSIQHFRVALIPFFAAFLPVFA,HPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMSTFKVLLCGAVLSRIDAGQEQLGRRRIHYSQNDLVEYSPVTEKHLTDGMTVRELC SAAITMSDNTAANLLLTITIGGPKELTAFLEHMGDHDVTRLDRWEPENEAIPNDERDTTTPVAMATTLRKLLTGELTLASRQQLIDWMEADK VAGPLLR SALPAGWFIADKSGAGERGSR GIIAALGPDGKPSRIWIYTTGSQATMDERNRQIAEIGASLIKHW*				
	286 amino acids = 31.6 kDa				
✓ ori	3368 .. 3956	589			rep_origin
/direction	= RIGHT				
/note	= high-copy-number ColE1/pMB1/pBR322/pUC origin of replication				
✓ CAP binding site	4244 .. 4265	22			protein_bind
/bound_moiety	= E. coli catabolite activator protein				
/note	= CAP binding activates transcription in the presence of cAMP.				
✓ lac promoter	4280 .. 4310	31			promoter
▶ 3 segments					
/note	= promoter for the E. coli lac operon				
✓ lac operator	4318 .. 4334	17			protein_bind
/bound_moiety	= lac repressor encoded by lacI				
/note	= The lac repressor binds to the lac operator to inhibit transcription in E. coli. This inhibition can be relieved by adding lactose or isopropyl-β-D-thiogalactopyranoside (IPTG).				
✓ M13 rev	4342 .. 4358	17			primer_bind
/note	= common sequencing primer, one of multiple similar variants				

Description:

Created: 2014/05/11

Last Modified: 2014/05/11

Accession Number:

Code Number:

Sequence Author:

DNA Type: Natural DNA

Source Organism:

Sequence Class: UNA

Comments:

References: