

1045

PST

HR-ePST

ctatttgtgtcatcaaaggatagagataaaagacaccaggaaagcttagacaag  
+-----+-----+-----+-----+-----+-----+-----+  
qataaacacacactaqtttcctatctctatttctgtqttccttcgaaatctgttc

1100

PST

HR-ePSI

atagaggaagagcaaaacaaaagtaagaccaccgcacagcaagcggccggccgt  
tatctcccttcgcgttttgtttcattctqqtqacatqtcqttcqccqaccqacqa

1155

PST

HR-ePSI

The diagram shows a DNA sequence with two restriction enzyme sites highlighted. The sequence is represented by two rows of letters: gatcttcagacctggaggaggatatgagggacaattggagaagtgaattata and ctaqaqaagtctqqacacctccctctataactccctgttaaccctttcacttaatata. Two vertical lines with labels above them indicate cleavage sites: 'EcoNI' is at position 10, and 'MfeI' is at position 25.

1210

EcoNI

MfeI

aaatataaagttagtaaaaattgaaccatttaggagtagcacccaccaaggcaaaga  
tttatatttcatcattttaacttgatatactcatcatggatattccatatttt

1265

pHAGE-A1050-0550

gaagagtggtgcagagaaaaaaaagagcagtggaaataggagctttgttccttgg  
|||||

1205

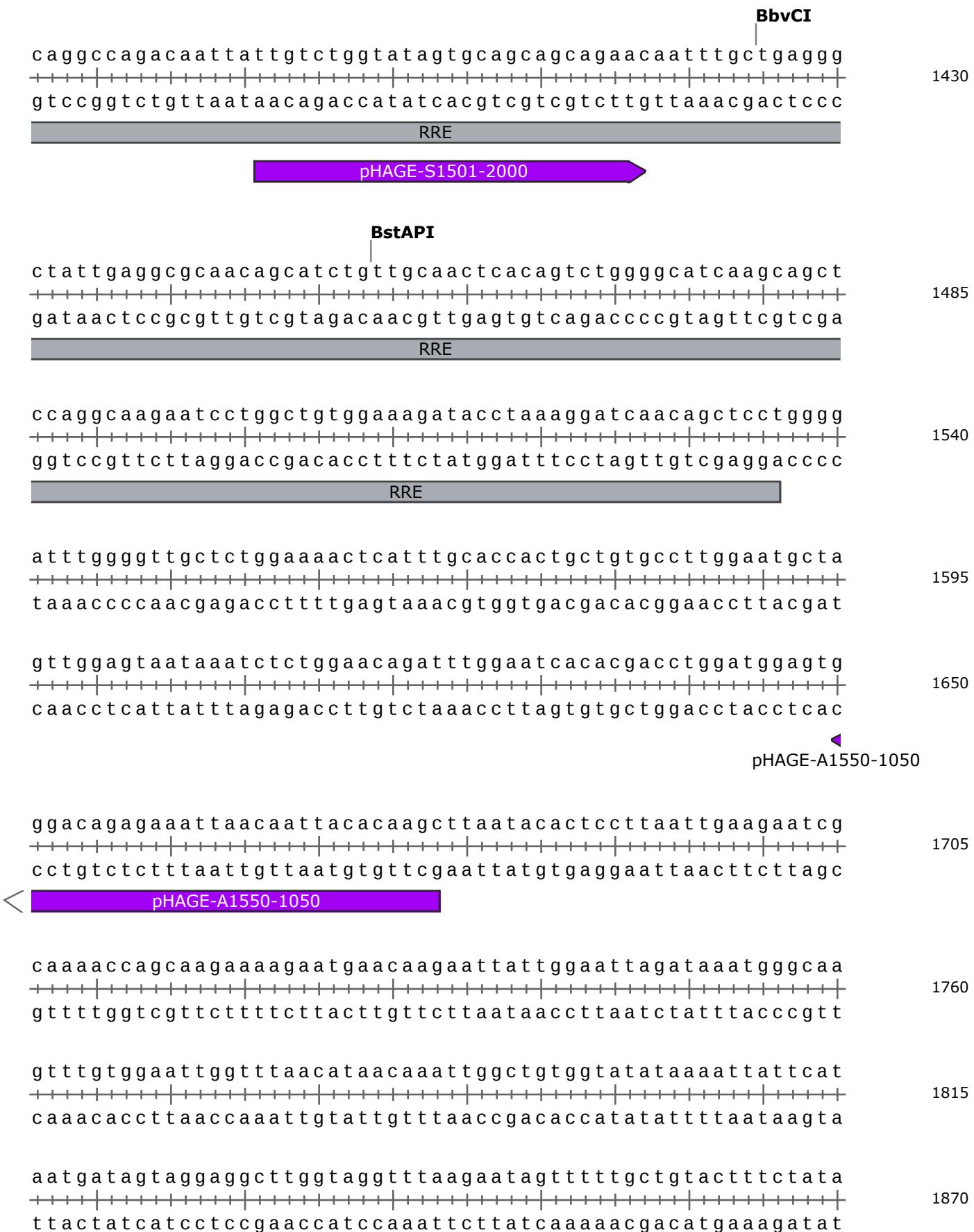
gaagagtggtgcagagagaaaaagagcagtggaaataggagcttgg  
+-----+-----+-----+-----+-----+-----+-----+  
cttctcaccacgtctctttttctcgtaacccttatcctcgaaacaaggacc

RRE

g t t c t t g g g a g c a g c a g g a a g c a c t a t g g g c g c a g c g t c a a t g a c g c t g a c g g t a  
+  
c a a g a a c c c t c g t c g t c t t c g t g a t a c c c g c g t c g c a g t t a c t g c g a c t g c c a t  
**RRE**

1375

RRF



gtgaatagagttaggcaggatattcaccattatcgttcagacccacctccaa  
+  
cacttatctcaatccqtcctataaqtqqtaataqcaaagtctqqqtggqgtt

1925

pHAGE-S2001-2500

KfI  
PpuMI

ccccgaggggacccgacaggcccgaaggaatagaagaagaaggtggagagagaga  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
qqqqctccctqqqctqtccqqqcttccttatcttcttcttccacctctctct

1980

2035

HIV cpPu (Trip)

aaatggcagtattcatccacaatttaaaaagaaaaaggggggattgggggtacag  
|||||  
tttaccgtataaqtqaqgtqtaaaattttctttccccccctaaccgggatqtc

2090

HIV cpPu (Trip)

pHAGE-A2050-1550

tgcaggaaagaatagtagacataatagcaacagacatacaaactaagaatt  
+  
acgtcccctttcttatcatctqtattatcgttgtctqtatqttgattcttaat

2145

HIV cpPu (Trip)

pHAGE-A2050-1550

caaaaacaattacaaaaattcaaaattttcggttattacagggacagcagag  
+  
qtttttgttaatqttttaaqtttaaaqcccaaataatqccctgtcgtc

2200

HIV cpPu (Trip)

SpEl

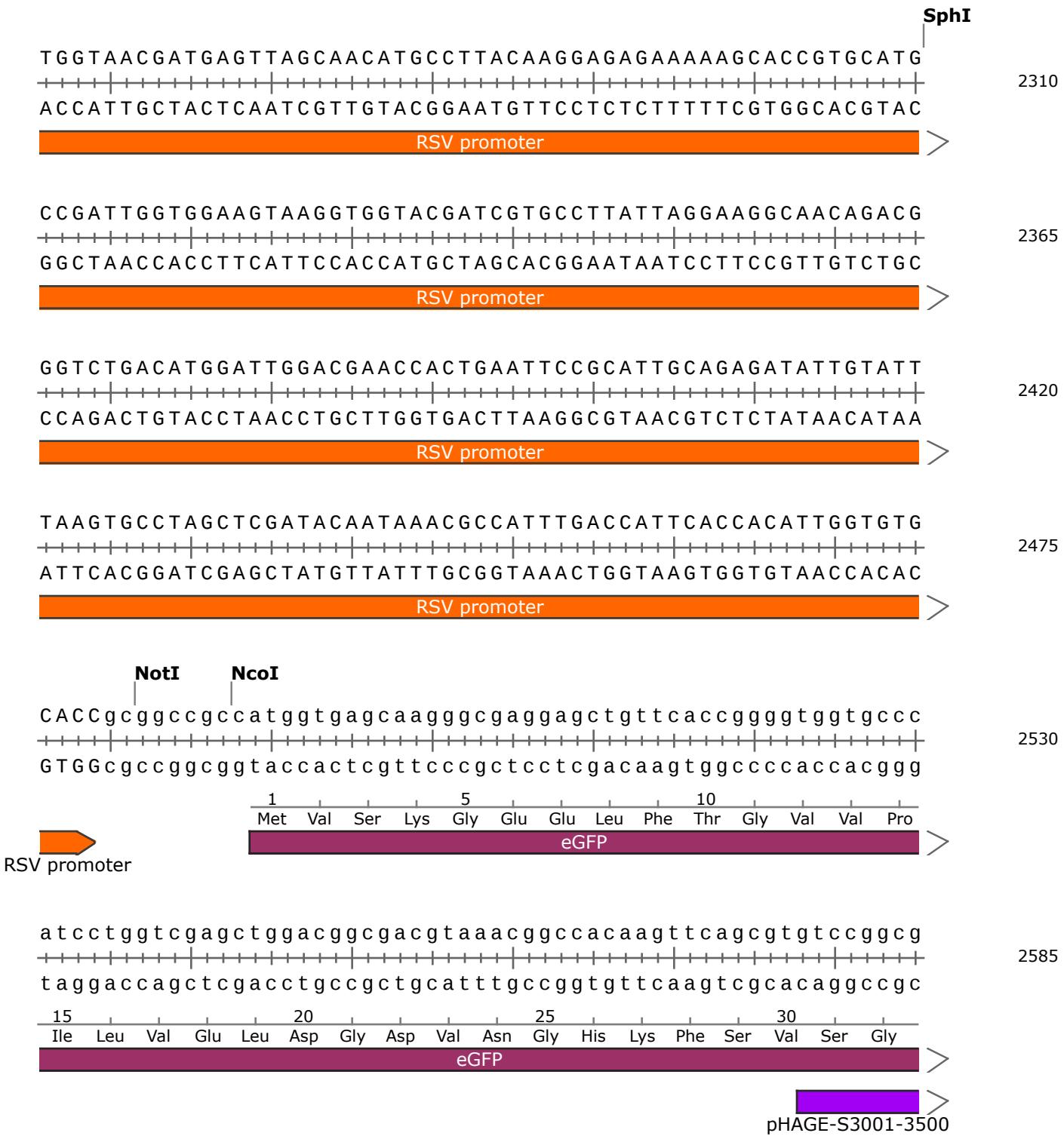
atccagttggactagtAATGTAGTCTTATGCAATACTCTTAGTCTTGCAACA  
+-----+-----+-----+-----+-----+-----+-----+-----+  
taggtcaaacctqatcattACATCAGAATACGTTATGAGAACATCAGAACGTTGT

2255

HIV cpPu (Trip)

RSV promoter

1



aggcgaggcgatgccacccatggcaagctgaccctgaagttcatctgcaccac  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 tcccgctcccgctacggtgatgccgttcgactgggacttcaagtagacgtggtg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 Glu Gly 35 Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr  


pHAGE-S3001-3500

2640

cggcaagctgcccgtgccctggccaccctcggtaccaccctgaccatacggcgtg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 gccgttgcacgggcacgggaccgggtggagcactgggtggactggatgccgcac  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val  


2695

cagtgcattcagccgtacccgaccacatgaagcagcacgacttcttcaagtcgg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 gtcacgaagtccccgtggatgggctgggtacttcgtcgtaactggatggcaggc  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser  


2750

ccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacggcaa  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ggtacgggcttccgtatgcagggtcccgatggatggtaacttcgtcgatggc  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn  


2805

ctacaagaccggcgccgagggtgaagttcgaggggcgacaccctggtaaccgcac  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 gatgttctggcgccgtccacttcaagctcccgctgtggaccacttggcgttag  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Asp Thr Leu Val Asn Arg Ile  


2860

pHAGE-A3050-2550

gagctgaaggcatcgacttcaaggaggacggcaacatcctggggcacaagctgg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ctcgacttccgtagctgaagttccctgtccgttgttaggaccaccgttgcacc  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 125 Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp 130 Gly Asn Ile Leu Gly His Lys Leu  


2915

agtacaactacaacagccacaacgtctatcatggccgacaaggcagaagaacgg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 tcatgttgcgtgtcggtgtgcagatatagtaccggctttcgcttcttgcc  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 Glu Tyr 145 Asn Tyr Asn Ser His 150 Asn Val Tyr Ile 155 Met Ala 160 Asp Lys Gln Lys Asn Gly  


2970

catcaaggtaacttcaagatccgcacaacatcgaggacggcagcgtgcagtc  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 gtagttccacttgaagttctaggcggtgttagctcctgccgtcgacgtcgag  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 Ile Lys 165 Val Asn Phe Lys Ile Arg 170 His Asn Ile Glu 175 Asp Gly Ser Val Gln Leu  


3025

gccgaccactaccaggcagaaacaccccatcgccgacggccccgtgctgccc  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 cggctggatggactcggtggactcggtggactcggtggactcggtggactcg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 180 Ala Asp His Tyr Gln 185 Gln Asn Thr Pro Ile 190 Gly Asp Gly Pro Val 195 Leu Leu Pro  
  


3080

acaaccactacctgagcaccagtcgcctgagcaaagacccaaacgagaagcg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ttttgtatggactcggtggactcggtggactcggtggactcggtggactcg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 200 Asp Asn His Tyr Leu Ser Thr 205 Gln Ser Ala Leu Ser 210 Lys Asp Pro Asn Glu 215 Lys Arg  

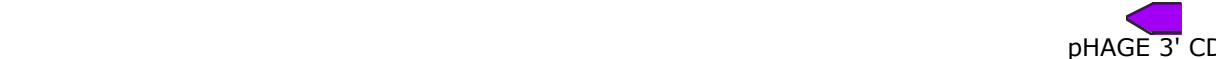

3135

cgcacatggcctgctggagttcggtggccggatcactctcgccatg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 gctagtgtaccaggacgacctaaggactcggtggcccttagtgagagccgtac  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 220 Asp His Met Val Leu Leu Glu Phe 225 Val Thr Ala Ala Gly 230 Ile Thr Leu Gly Met  


3190

BsrGI BamHI BsaBI\*

gacgagctgtacaagtaaggatcctaattcacctctggattacaaaatttgtaa  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ctgctcgacatgttcatccatggattatggagacctaattttaaacactt  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 235 Asp Glu Leu Tyr Lys 240 \*

3245

agattgactggtattcttaactatgttgctccctttacgctatgtggatacgcgt  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 tctaactgaccataagaattgatacaacgaggaaaatgcgatacacctatgcgac  
 WPRE

< pHAGE 3' CDS

cttaaatgccttgtatcatgctattgtctccgtatggcttcatttctcctc  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gaaattacggaaacatagtagataaacgaaggcataccgaaagtaaaagaggag  
 WPRE

pHAGE-A3550-3050

cttgtataaattcctggttgtgtctctttatgaggagttgtggcccggttcagg  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gaacatatttaggaccaacgacagagaaaatactcctcaacaccgggcaacagtcc  
 WPRE

PfIMI

caacgtggcgtggtgtgcactgtgttgcgtacgcacccccactggttggggca  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gttgcacccgacccacacgtgacacaaacgactgcgttgggggtgaccaacccgt  
 WPRE

tgcacccacctgtcagctccttcggactttcgcttccccctccattgc  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 aacggtgtggacagtgcgaggaaaggccctgaaagcgaagggggaggataacg  
 WPRE

pHAGE-S4001-4500

cacggcggaaactcatgcggcctgccttgcggctgtggacaggggctggctg  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gtgcgcgcctttagtgcggggacggAACGGGCGACGACCTGTCCCCGAGCCGAC  
 WPRE

ttggcactgacaattccgtggtgtgtcgaaaaatcatgcgccttcgttggc  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 aaccgtgactgttaaggcaccacaacagccccttagtagcaggaaaggaaccg  
 WPRE

tgctcgccctgtgttgcacccgtggatctgcgcggacgtccctctgtacgtccc  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 acgagcggacacaacgggtggacctaagacgcgcctgcaggaagacgatgcagg  
 WPRE

3300

3355

3410

3465

3520

3575

3630

3685

t tcggccctcaatccagcggaccttccttcccggcgtgcgtccggctctgcgg  
+-----+-----+-----+-----+-----+-----+-----+-----+  
a a g c c g g g a g t t a g g t c g c c t g g a a g g a a g g g c g c c g a c g a c g c c  
+-----+-----+-----+-----+-----+-----+-----+-----+  
**WPRE**

3740

c c t c t t c c g c g t c t t c g c c t c g c c t c a g a c g a g t c g g a t c t c c t t t g g c c g  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
g g a g a a g g c g a g a a g c g g a a g c g g g a g t c t g c t c a g c c t a g a g g g a a a c c c g g c  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**WPRE**

oligo #91 sequencing sense →

c c t c c c c g c c t g a g a t c c t t a a g a c c a a t g a c t t a c a a g g c a g c t g t a g a t c t t  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
g g a g g g g c g g a c t c t a g g a a t t c t g g t t a c t g a a t g t t c c g t c g a c a t c t a g a a  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**WPRE**   **HIV nef**

← pHAGE-A4050-3550 →

a g c c a c t t t t a a a g a a a a g g g g g a c t g g a a g g g c t a a t t c a c t c c c a a c g a a  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
t c g g t g a a a a a t t t c t t t c c c c t g a c c t t c c g a t t a a g t g a g g g t t g c t t  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**HIV PPT**   **dU3 LTR**  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**HIV nef**

3850

3905

g a c a a g a t c t t t t g c t t g t a c t g g g t c t c t c t g g t t a g a c c a g a t c t g a g c  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
c t g t t c t a g a c g a a a a a c g a a c a t g a c c c a g a g a g a c c a a t c t g g t c t a g a c t c g  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**dU3 LTR**  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**HIV R**

Difference from 5' LTR

4015

c t g g g a g c t c t c t g g c t a a c t a g g g a a c c c a c t g c t t a a g c c t c a a t a a g c t t g  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
g a c c t c g a g a g a c c g a t t g a t c c c t t g g g t g a c g a a t t c g g a g t t a t t c g a a c  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**dU3 LTR**  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**HIV R**

4070

c c t t g a g t g c t t c a a g t a g t g t g c c c g t c t g t g t g a c t c t g g t a a c t a g a  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
g g a a c t c a c g a a g t t c a t c a c a c a c g g g c a g a c a a c a c a c t g a g a c c a t t g a t c  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**dU3 LTR**  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
**HIV R**   **HIV U5**

The diagram illustrates the pHAGE-S4501-5000 vector genome. The top sequence shows the DNA sequence: gatccctcagacccttttagtcagtggtggaaaatctctagcagtagtagttcatg and ctagggagtctggaaaatcagtcacaccttttagagatcgcatcatcaagtagac. Below the sequence, three horizontal bars indicate specific regions: a grey bar labeled "dU3 LTR" covers the first 100bp, a grey bar labeled "HIV U5" covers the next 100bp, and a grey bar labeled "3' Flank" covers the remaining sequence. A purple bar at the bottom right is labeled "pHAGE-S4501-5000".

**pGK-PsiI-pHAGE-S4501-5000**

**5' Flank**

**3' Flank**

**pHAGE-S4501-5000**

**XmaI**

**TspMI**

**SmaI**

**PacI**

**5' Flank**

**pHAGE-S6001-6500**

gtgatacgcctatttataggtaatgtcatgataataatggttcttagacgt  
|  
cactatgcggataaaaatatccaattacagtactattattaccaaagaatctgca  
  
pHAGE-S6001-6500

pHAGE-S6001-6500

cagg tgg cact tt cggg aa at gt gc gc gg aacc ct attt gt tt at tt tcta  
 +  
 4345  
 gt cc acc cgt gaaa agcc cttt ac ac gc gc ct tg gg at aa ac aa at aa aa agat  
  
 aata ca tt ca aa at at gt at cc gct cat ga ga ca aata acc ct gata aat gct tcaa  
 +  
 4400  
 tt at gt ta ag tt tata ca ta ca tagg gc gag t a c t ct gt tt at tt gg act at tt ac ga ag tt  
 P-Bla

The diagram illustrates the bla gene construct. It features two horizontal lines representing DNA strands. The top strand has a vertical arrow labeled "SspI" pointing to a restriction site. The bottom strand has a horizontal arrow labeled "P-Bla" pointing to its promoter region. A vertical arrow labeled "AMPr" points to the start site of the bla gene. Below the strands, a scale bar shows positions 1 and 5, with intermediate tick marks at intervals of 5 units.

ccttttttgcggcatttgccttcgtttgctcacccagaaacgctggtgaa  
+-----+-----+-----+-----+-----+-----+-----+-----+  
gaaaaaacgcgtaaaacggaaggacaaaaacgagtggtcttgcgaccactt

4510

Pro Phe Phe Ala Ala 10 Phe Cys Leu Pro 15 Val Phe Ala His Pro 20 Glu Thr Leu Val Lys  
AMPr >

agtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaaactggat  
+-----+-----+-----+-----+-----+-----+-----+-----+  
tcattttctacgacttcttagtcaacccacgtgctcacccaatgttagcttgcaccta

4565

25 Val Lys Asp Ala Glu 30 Asp Gln Leu Gly Ala 35 Arg Val Gly Tyr Ile 40 Glu Leu Asp  
AMPr >



pHAGE-A6050-5550

ctcaacagcgtaagatccttgagagtttgcggccaaagaacgtttccaatga  
+-----+-----+-----+-----+-----+-----+-----+-----+  
gagttgtcgccattcttaggaactctcaaaagcggggcttctgcaaaaggttact

4620

45 Leu Asn Ser Gly Lys Ile Leu 50 Glu Ser Phe Arg Pro 55 Glu Arg Phe Pro 60 Met  
AMPr >

**HincII**

tgagcactttaaagttctgtatgtggcgccgtattatccgtgttgacgcccgg  
+-----+-----+-----+-----+-----+-----+-----+-----+  
actcgtaaaaattcaagacataccgcgccataatagggcacaactgcggcc

4675

65 Met Ser Thr Phe Lys Val Leu Leu Cys 70 Gly Ala Val Leu Ser 75 Arg Val Asp Ala Gly  
AMPr >

gcaagagcaactcggtcgccgcatacactattctcagaatgacttggtgagta  
+-----+-----+-----+-----+-----+-----+-----+-----+  
cgttctcggtgagccagcgggtatgtgataagagtcttactgaaccaactcatg

4730

80 Gln Glu Gln Leu Gly 85 Arg Ile His Tyr 90 Ser Gln Asn Asp Leu 95 Val Glu Tyr  
AMPr >

tcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgca  
+-----+-----+-----+-----+-----+-----+-----+-----+  
agtggtcagtgtctttcgtagaatgcctaccgtactgtcattctcttaatacgt

4785

100 Ser Pro Val Thr Glu Lys His 105 Leu Thr Asp Gly Met 110 Val Arg Glu Leu 115 Cys  
AMPr >

gtgctgccataaccatgagtgataaacactgcggccaacttacttctgacaacgat  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 cacgacggtattggtactcactattgtgacgcccgtgaatgaagactgttgcta

Ser Ala Ala Ile 120 Thr Met Ser Asp Asn 125 Thr Ala Ala Asn Leu 130 Leu Leu Thr Thr Ile

AMP<sub>r</sub>

4840

cggaggaccgaaggagctaaccgcttttgacaaacatggggatcatgtact  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gcctccctggcttcctcgattggcgaaaaacgtgttacccctagtagattga  
 135 Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr  
 140  
 145  
 150  
**AMPr**

4895

cgccttgcgaa

155 160 165 170

Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg

AMP

4950

The diagram illustrates the *FspI* gene structure. The top part shows the DNA sequence: acaccacagatgcctgttagcaatggcaacaacgttgcgcaaactattaactggcga and tgtggtgctacggacatcgtaaccgttggcaacgcgtttgataattgaccgct. Below the sequence, a horizontal line with vertical tick marks indicates restriction enzyme cleavage sites. The bottom part shows the amino acid translation starting at position 175: Asp, Thr, Thr, Met, Pro, Val, Ala, Met, Ala, Thr, Thr, Leu, Arg, Lys, Leu, Leu, Thr, Gly, Glu. A red shaded box labeled 'AMPPr' covers the region from approximately position 180 to 195.

5005

act actt tactctagttccggcaacaattaaataga  
tgatgaatgagatcgaaaggcccgtttaattatctgaccta

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 190 |     |     |     |     | 195 |     |     |     |     | 200 |     |     |     | 205 |     |     |     |
| Leu | Leu | Thr | Leu | Ala | Ser | Arg | Gln | Gln | Leu | Ile | Asp | Trp | Met | Glu | Ala | Asp | Lys |

AMP >

5060

Sequence alignment diagram showing the amino acid sequence of AMPr. The sequence starts with Val at position 210, followed by Ala, Gly, Pro, Leu, Leu, Arg, Ser at position 215, Ala, Leu, Pro, Ala, Gly, Trp, Phe, Ile, Ala, and Asp at position 225. A purple bar highlights the sequence from Val to Asp.

5115

aatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccaga  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ttagacacctcgccactcgcacccagagcgcatacgtaacgtcgtgaccgggtct  
 |-----|-----|-----|-----|-----|-----|-----|-----|-----|  
 Lys Ser Gly Ala 230 Gly Glu Arg Gly Ser 235 Arg Gly Ile Ile Ala 240 Ala Leu Gly Pro Asp  


5170

tggtaaggccctcccgatatcgtagttatctacacgacggggagtcaggcaactatg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 accattcgggagggcatacgatcaatagatgtgctgcccctcagtcgttatac  
 |-----|-----|-----|-----|-----|-----|-----|-----|-----|  
 245 Gly Lys Pro Ser Arg 250 Ile Val Val Ile Tyr 255 Thr Thr Gly Ser Gln 260 Ala Thr Met  
  


5225

gatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattgg  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ctacttgcttatctgtctagcgactctatccacggagtgactaattcgtaacca  
 |-----|-----|-----|-----|-----|-----|-----|-----|-----|  
 265 Asp Glu Arg Asn Arg Gln Ile 270 Ala Glu Ile Gly Ala 275 Ser Leu Ile Lys His 280 Trp  


5280

pHAGE-S7001-7500

aactgtcagaccaagttactcatataacttttagattgattaaaacttcattt  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ttgacagtctggttcaaattgagtatataatgaaatctaactaaatttgaagtaaa  
 \*  
 ->

5335

ttaattaaaaggatctaggtaagatccttttataatctcatgacaaaatc  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 aattaaatttcctagatccacttcttagaaaaactattagagtactggtttag

5390

ccttaacgtgagtttcgttccactgagcgtcagacccgtagaaaagatcaaag  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ggaattgcactcaaaagcaaggtgactcgcagtcgtggggatctttctagttc

5445

gatcttctttagatccttttctgcgcgtaatctgctgctgcaaacaaaaaa  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 ctagaagaactcttagaaaaaaagacgcgcattagacgacgaaacgttttttt

5500



accacccgctaccagcggtggttgttgccggatcaagagactaccaactctttt  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
tggtggcgatggtcgccaccaaacaacggcctagttctcgatggttgagaaaaaa  
ORI

5555

5610

ccgaaggtaactggcttcagcagagcgagataccaaatactgttcttagtgt  
+-----+-----+-----+-----+-----+-----+-----+-----+  
ggcttcattgaccgaagtcgctctcggttatggttatgacaagaagatcaca  
ORI

5665

tccgttagttaggccaccacttcaagaactctgttagcaccgcctacatacctcg  
+-----+-----+-----+-----+-----+-----+-----+-----+  
tcggcatcaatccgggtggtaagttcttgagacatcggtggcgatgtatggagcg  
ORI

5720

tctgctaattcctgttaccagtggtgtgcgcagtggtggcgataagtcgtgttacc  
+-----+-----+-----+-----+-----+-----+-----+-----+  
agacgattaggacaatggtcaccgacgacggtaccgttattcagcacagaatgg  
ORI

pHAGE-S7501-8000 >

5775

gggttggactcaagacgatagttaccggataaggcgcagcggtcgggtgaacgg  
+-----+-----+-----+-----+-----+-----+-----+-----+  
cccaacctgagttctgtatcaatggcctattccggcgtgcgcagccccacttgcc  
ORI

pHAGE-S7501-8000 >

5830

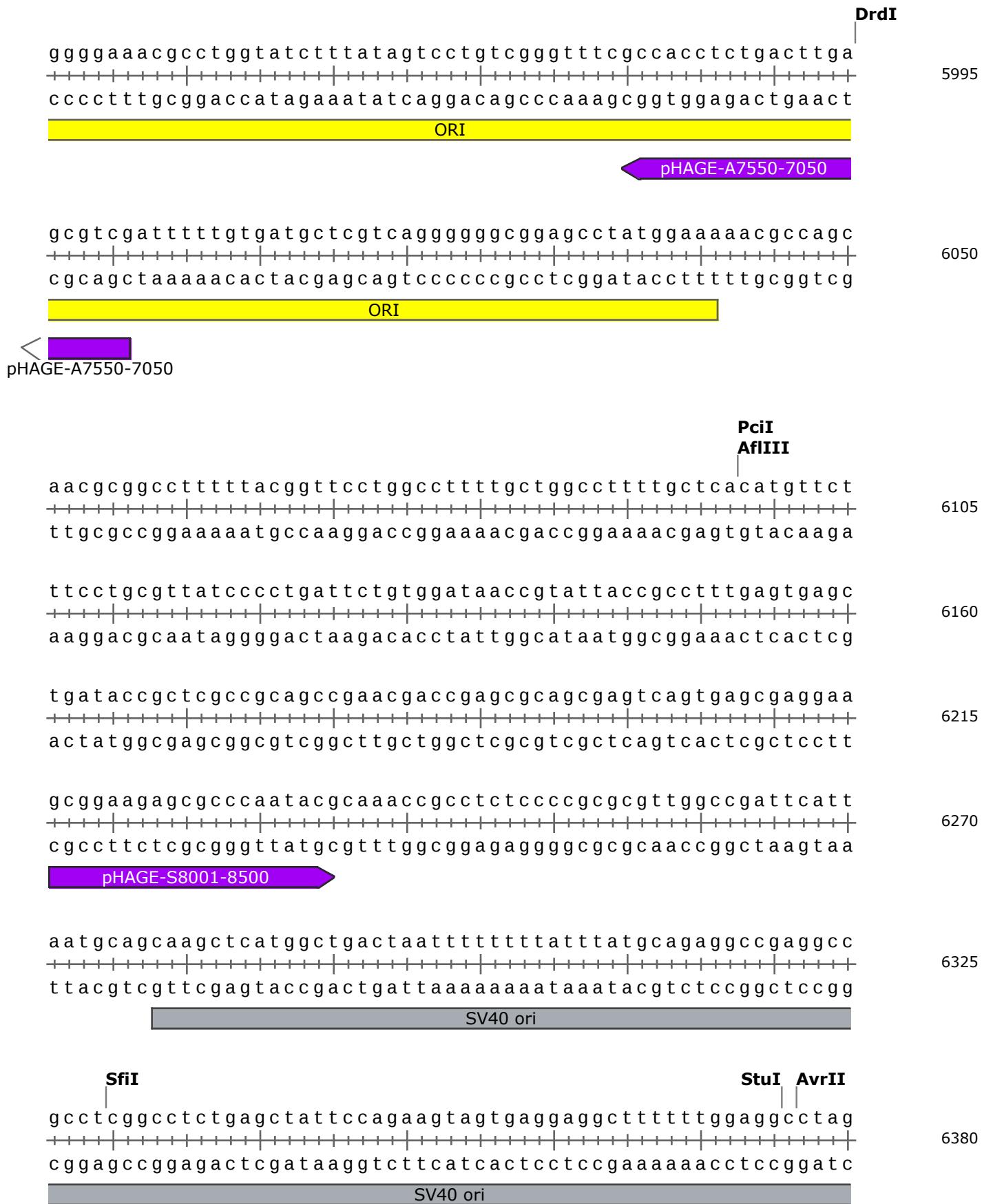
ggggttcgtcacacacagcccagttggagcgaacgacctacaccgaactgagata  
+-----+-----+-----+-----+-----+-----+-----+-----+  
ccccaaagcacgtgtcggtcgaaacctcgcttgcgtggatgtggcttactctat  
ORI

5885

cctacagcgtgagctatgagaaaggcgccacgcttccgaagggagaaaggcgac  
+-----+-----+-----+-----+-----+-----+-----+-----+  
ggatgtcgactcgataactcttcgggtgcgaagggtttccctttccgcctg  
ORI

5940

aggatatccggtaagcggcagggtcggaacaggagagcgcacgaggagcttccag  
+-----+-----+-----+-----+-----+-----+-----+-----+  
tccataggccattcgccgtcccagccttgcgtgcgtccctcgaagggtc  
ORI



gctttgcaaaaagctccccgtggcacgacaggtttccgactggaaagcgggca  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 caaaaacgttttcgaggggcaccgtgtccaaagggctgaccttcgcccgt  
**SV40 ori**6435

gtgagcgcaacgcaattaatgtgagtttagctcactcattaggcacccaggctt  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 cactcgcttgcgttaattacactcaatcgagttagtaatccgtgggtccgaaa  
**pHAGE-A8050-7550**6490

acactttatgctccggctcgatgttgtgtggaatgtgagcggataacaattt  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 tgtgaaatacgaaggccgagcatacaacacacaccttaacactcgctattgttaaa6545

cacacagggaaacagctatgacatgattacgaatttacaaataaagcattttt  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gtgtgtccttgcatactgtactaatgcttaaggtagtttattcgtaaaaaaaa6600

cactgcattctagttgtggttgtccaaactcatcaatgtatcttatcatgtctg  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gtgacgtaagatcaacaccaaacaggtttgagtagttacatagaatagtacagac6655

gatcaactggataactcaagctaaccaaaatcatccaaacttcccacccatac  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 cttagttgacctattgagttcgattggtttagtagggttgaagggtgggtatg6710

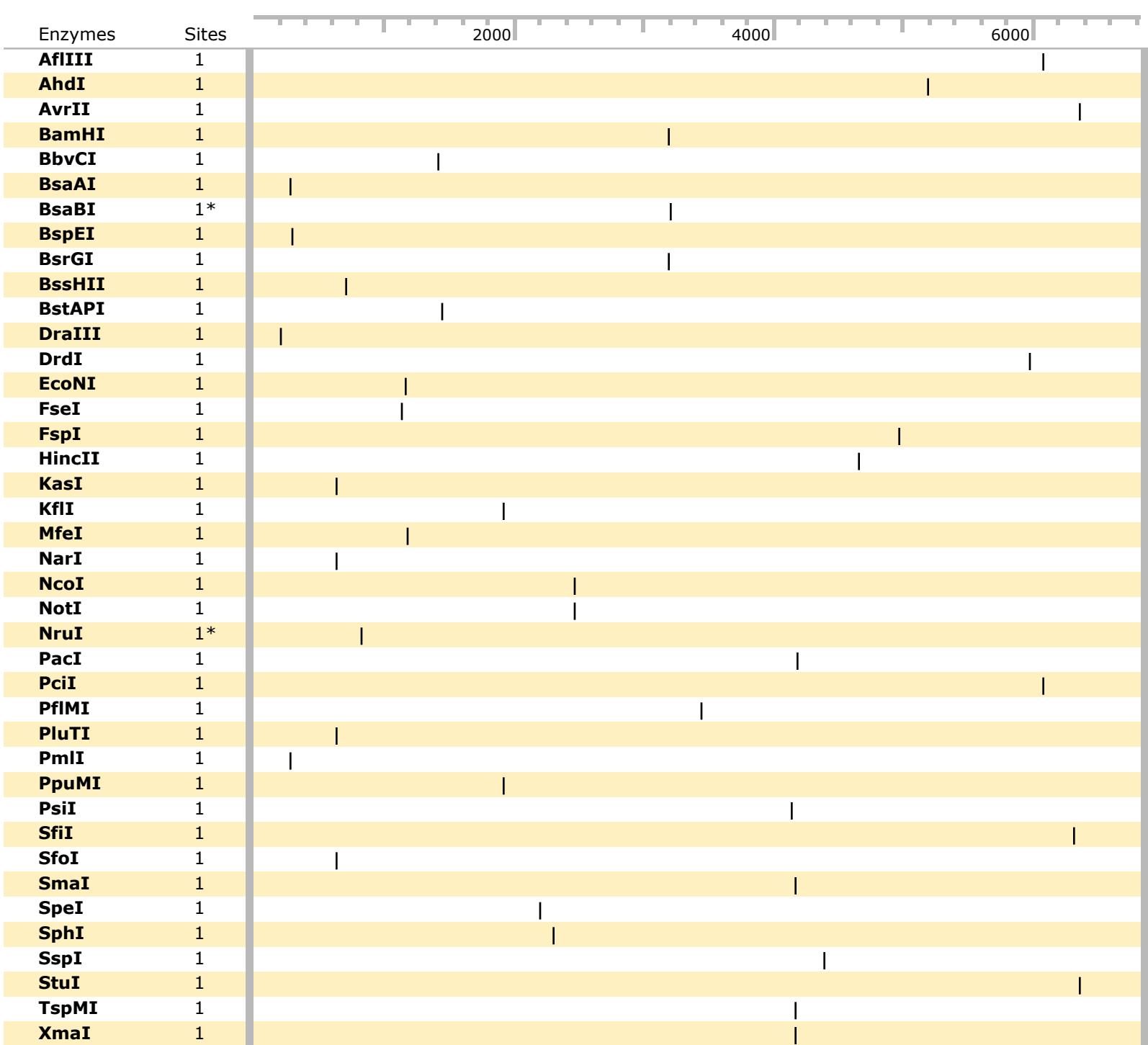
cctattaccactgccaattacctgtggtttcatctactctaaacctgtgattcct  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 ggataatggtgacggtaatggacaccaaagtaatgagatttggacactaagga6765

**pHAGE-S1-500** 

ctgaattatttcatttaaagaaattgtattgttaaatatgtactacaaactt  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 gacttaataaaagtaaaattcttaacataaacaatttatacatgtatgtttgaa6820

agtagt 3'  
 +-----+-----+-----+-----+-----+-----+-----+-----+  
 tcatca 5'6826

| Enzymes       | Sites |       |
|---------------|-------|-------|
| <b>AflIII</b> | 1     | 6097  |
| <b>AhdI</b>   | 1     | 5209  |
| <b>AvrII</b>  | 1     | 6376  |
| <b>BamHI</b>  | 1     | 3209  |
| <b>BbvCI</b>  | 1     | 1424  |
| <b>BsaAI</b>  | 1     | 290   |
| <b>BsaBI</b>  | 1*    | 3214* |
| <b>BspEI</b>  | 1     | 308   |
| <b>BsrGI</b>  | 1     | 3198  |
| <b>BssHII</b> | 1     | 711   |
| <b>BstAPI</b> | 1     | 1453  |
| <b>DraIII</b> | 1     | 212   |
| <b>DrdI</b>   | 1     | 5995  |
| <b>EcoNI</b>  | 1     | 1170  |
| <b>FseI</b>   | 1     | 1150  |
| <b>FspI</b>   | 1     | 4986  |
| <b>HincII</b> | 1     | 4667  |
| <b>KasI</b>   | 1     | 637   |
| <b>KflI</b>   | 1     | 1934  |
| <b>MfeI</b>   | 1     | 1189  |
| <b>NarI</b>   | 1     | 638   |
| <b>NcoI</b>   | 1     | 2487  |
| <b>NotI</b>   | 1     | 2481  |
| <b>NruI</b>   | 1*    | 833*  |
| <b>PacI</b>   | 1     | 4194  |
| <b>PciI</b>   | 1     | 6097  |
| <b>PfIMI</b>  | 1     | 3457  |
| <b>PluTI</b>  | 1     | 641   |
| <b>PmI</b>    | 1     | 290   |
| <b>PpuMI</b>  | 1     | 1934  |
| <b>PsiI</b>   | 1     | 4147  |
| <b>SfiI</b>   | 1     | 6329  |
| <b>SfoI</b>   | 1     | 639   |
| <b>SmaI</b>   | 1     | 4186  |
| <b>SpeI</b>   | 1     | 2212  |
| <b>SphI</b>   | 1     | 2310  |
| <b>SspI</b>   | 1     | 4404  |
| <b>StuI</b>   | 1     | 6375  |
| <b>TspMI</b>  | 1     | 4184  |
| <b>XmaI</b>   | 1     | 4184  |



| Feature                   |   | Location                          | Size (bp) |  |  | Type         |
|---------------------------|---|-----------------------------------|-----------|--|--|--------------|
| ✓ <b>LTR</b>              |   | 1 .. 636                          | 636       |  |  | misc_feature |
| /gene                     | = | LTR                               |           |  |  |              |
| /note                     | = | Long Terminal Repeat              |           |  |  |              |
| ✓ <b>HIV U3</b>           |   | 1 .. 453                          | 453       |  |  | misc_feature |
| /note                     | = | Full Length HIV U3                |           |  |  |              |
| ✓ <b>pHAGE-A0050-8050</b> |   | 158 .. 179                        | 22        |  |  | primer_bind  |
| ✓ <b>pHAGE-S501-1000</b>  |   | 401 .. 421                        | 21        |  |  | primer_bind  |
| ✓ <b>HIV R</b>            |   | 454 .. 551                        | 98        |  |  | misc_feature |
| /note                     | = | HIV Repeat Region                 |           |  |  |              |
| ✓ <b>HIV U5</b>           |   | 552 .. 636                        | 85        |  |  | misc_feature |
| /note                     | = | HIV U5                            |           |  |  |              |
| ✓ <b>PBS-K</b>            |   | 636 .. 658                        | 23        |  |  | misc_feature |
| /note                     | = | tRNA binding site for Lysine tRNA |           |  |  |              |
| ✓ <b>PSI</b>              |   | 637 .. 1155                       | 519       |  |  | misc_feature |
| /gene                     | = | psi                               |           |  |  |              |
| /product                  | = | pbs-gag                           |           |  |  |              |
| /note                     | = | PSI Packaging Sequence            |           |  |  |              |
| ✓ <b>pHAGE-A0550-0050</b> |   | 667 .. 690                        | 24        |  |  | primer_bind  |
| ✓ <b>HIV-PSI</b>          |   | 697 .. 806                        | 110       |  |  | misc_feature |
| ✓ <b>HR-ePSI</b>          |   | 807 .. 1144                       | 338       |  |  | misc_feature |
| ✓ <b>pHAGE-S1001-1500</b> |   | 898 .. 919                        | 22        |  |  | primer_bind  |
| ✓ <b>pHAGE-A1050-0550</b> |   | 1166 .. 1189                      | 24        |  |  | primer_bind  |
| ✓ <b>RRE</b>              |   | 1303 .. 1536                      | 234       |  |  | misc_feature |
| /gene                     | = | rre                               |           |  |  |              |
| /product                  | = | minimal RRE                       |           |  |  |              |
| /note                     | = | Rev Responsive Element            |           |  |  |              |
| ✓ <b>pHAGE-S1501-2000</b> |   | 1391 .. 1417                      | 27        |  |  | primer_bind  |
| ✓ <b>pHAGE-A1550-1050</b> |   | 1650 .. 1678                      | 29        |  |  | primer_bind  |

| Feature                             | Location  | Size (bp) |  |  | Type         |
|-------------------------------------|---|-----------|--|--|--------------|
| ✓ <b>pHAGE-S2001-2500</b>           | 1898 .. 1922  | 25        |  |  | primer_bind  |
| ✓ <b>HIV cpPu (Trip)</b>            | 2034 .. 2211  | 178       |  |  | misc_feature |
| /gene                               | = cppt  |           |  |  |              |
| /note                               | = Central Poly Purine Track   |           |  |  |              |
| ✓ <b>pHAGE-A2050-1550</b>           | 2084 .. 2112  | 29        |  |  | primer_bind  |
| ✓ <b>RSV promoter</b>               | 2218 .. 2479  | 262       |  |  | promoter     |
| /note                               | = Rous sarcoma virus enhancer/promoter  |           |  |  |              |
| ✓ <b>eGFP</b>                       | 2489 .. 3208  | 720       |  |  | CDS          |
| /gene                               | = eGFP  |           |  |  |              |
| /note                               | = Enhanced Green Fluorescent Protein (Mammalian Codon Optimized)  |           |  |  |              |
| /translation                        | = MVSKGEEELFTGVVPILVLDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTGYVQCFSRYPDI<br>DFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQI<br>NFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK* |           |  |  |              |
|                                     | 239 amino acids = 26.9 kDa  |           |  |  |              |
| ✓ <b>pHAGE-S3001-3500</b>           | 2578 .. 2595  | 18        |  |  | primer_bind  |
| ✓ <b>pHAGE-A3050-2550</b>           | 2821 .. 2841  | 21        |  |  | primer_bind  |
| ✓ <b>pHAGE-S3501-4000</b>           | 3076 .. 3097  | 22        |  |  | primer_bind  |
| ✓ <b>WPRE</b>                       | 3215 .. 3806  | 592       |  |  | misc_feature |
| /gene                               | = WPRE  |           |  |  |              |
| /note                               | = Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element  |           |  |  |              |
| ✓ <b>pHAGE 3' CDS</b>               | 3243 .. 3277  | 35        |  |  | primer_bind  |
| ✓ <b>pHAGE-A3550-3050</b>           | 3321 .. 3344  | 24        |  |  | primer_bind  |
| ✓ <b>pHAGE-S4001-4500</b>           | 3474 .. 3494  | 21        |  |  | primer_bind  |
| ✓ <b>oligo #91 sequencing sense</b> | 3767 .. 3790  | 24        |  |  | primer_bind  |
| ✓ <b>pHAGE-A4050-3550</b>           | 3802 .. 3824  | 23        |  |  | primer_bind  |
| ✓ <b>HIV nef</b>                    | 3812 .. 3900  | 89        |  |  | misc_feature |
| /note                               | = HIV nef   |           |  |  |              |
| ✓ <b>HIV PPT</b>                    | 3862 .. 3876  | 15        |  |  | misc_feature |
| /note                               | = Polypurine Tract (Not degraded by RNaseH so it can serve as a primer for + strand DNA synthesis during life cycle)  |           |  |  |              |

| Feature                         | Location   | Size (bp) |  |  | Type          |
|---------------------------------|--|-----------|--|--|---------------|
| ✓ <b>dU3 LTR</b>                | 3877 .. 4112   | 236       |  |  | misc_feature  |
| /gene                           | = dU3 LTR  |           |  |  |               |
| /note                           | = Deleted U3 Long-Terminal Repeat  |           |  |  |               |
| ✓ <b>HIV R</b>                  | 3914 .. 4029   | 116       |  |  | misc_feature  |
| /note                           | = HIV Repeat Region  |           |  |  |               |
| ✓ <b>Difference from 5' LTR</b> | 3925 .. 3925   | 1         |  |  | modified_base |
| /note                           | = This T is a C in the 5' LTR  |           |  |  |               |
| ✓ <b>HIV U5</b>                 | 4030 .. 4112   | 83        |  |  | misc_feature  |
| /note                           | = HIV U5   |           |  |  |               |
| ✓ <b>pHAGE-S4501-5000</b>       | 4110 .. 4141   | 32        |  |  | primer_bind   |
| ✓ <b>3' Flank</b>               | 4113 .. 4183   | 71        |  |  | misc_feature  |
| /note                           | = Chromosomal 3' Flanking Sequence carried over from original HIV integration site   |           |  |  |               |
| ✓ <b>pHAGE-S6001-6500</b>       | 4216 .. 4238   | 23        |  |  | primer_bind   |
| ✓ <b>P-Bla</b>                  | 4387 .. 4421   | 35        |  |  | promoter      |
| /note                           | = Beta Lactamase Promoter  |           |  |  |               |
| ✓ <b>AMPr</b>                   | 4440 .. 5279   | 840       |  |  | CDS           |
| /gene                           | = AMPr   |           |  |  |               |
| /product                        | = beta-lactamase (mature form)   |           |  |  |               |
| /note                           | = Ampicillin Resistance Gene   |           |  |  |               |
| /translation                    | = RVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSSTFKVLLCGAVLSRVDAGC<br>RIHYSQNDLVEYSPVTEKHLDGMTVRELCASAITMSDNTAANLLTTIGGPKELTAFLHNMGDHVTLDRWEPELI<br>DERDTTMRPAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAGERGSRGIIAALGPD<br>VVIYTTGSQATMDERNRQIAEIGASLIKHW |           |  |  |               |
|                                 | 280 amino acids = 30.8 kDa   |           |  |  |               |
| ✓ <b>pHAGE-A6050-5550</b>       | 4489 .. 4512   | 24        |  |  | primer_bind   |
| ✓ <b>pHAGE-S6501-7000</b>       | 4689 .. 4713   | 25        |  |  | primer_bind   |
| ✓ <b>pHAGE-A6550-6050</b>       | 4958 .. 4981   | 24        |  |  | primer_bind   |
| ✓ <b>pHAGE-S7001-7500</b>       | 5209 .. 5233   | 25        |  |  | primer_bind   |
| ✓ <b>pHAGE-A7050-6550</b>       | 5469 .. 5490   | 22        |  |  | primer_bind   |

| Feature                   |   | Location                        | Size (bp) |  |  | Type         |
|---------------------------|---|---------------------------------|-----------|--|--|--------------|
| ✓ <b>ORI</b>              |   | 5508 .. 6041                    | 534       |  |  | rep_origin   |
| /gene                     | = | ORI                             |           |  |  |              |
| /note                     | = | Bacterial Origin of Replication |           |  |  |              |
| ✓ <b>pHAGE-S7501-8000</b> |   | 5713 .. 5737                    | 25        |  |  | primer_bind  |
| ✓ <b>pHAGE-A7550-7050</b> |   | 5980 .. 6001                    | 22        |  |  | primer_bind  |
| ✓ <b>pHAGE-S8001-8500</b> |   | 6216 .. 6235                    | 20        |  |  | primer_bind  |
| ✓ <b>SV40 ori</b>         |   | 6278 .. 6401                    | 124       |  |  | misc_feature |
| /gene                     | = | SV40 ori                        |           |  |  |              |
| ✓ <b>pHAGE-A8050-7550</b> |   | 6456 .. 6482                    | 27        |  |  | primer_bind  |
| ✓ <b>pHAGE-S1-500</b>     |   | 6721 .. 6749                    | 29        |  |  | primer_bind  |