



General Description

DNA pHAGE EF1a1-a1AT-W
Entire molecule length: 8285 bp

Standard Fields

User Fields

Author(s)

Author

Original author

Darrell Kotton
EMAIL: dkotton@bu.edu

References

Comments

EF1a1 promotor cloned into CMV-AAT-W by Not1/Spel ligation by LWK on 8/05.
EF1a1 promotor including both junctions confirmed by sequencing.

Feature Map

CDS (2 total)

human a1ATT CDS

Start: 3411 End: 4664

AMPr

Start: 5899 End: 6738

Ampicillin Resistance Gene

Original Location Description:

6153..6992

Qualifiers:

/gene="AMPr"

/product="beta-lactamase (mature form)"

Intron (1 total)

Intron

Start: 2454 End: 3045

Human EF1a Intron

Misc. Feature (19 total)

LTR

Start: 1 End: 636

Long Terminal Repeat

Original Location Description:

1..636

Qualifiers:

/gene="LTR"

HIV U3

Start: 1 End: 453

Full Length HIV U3

HIV R

Start: 454 End: 551
HIV Repeat Region

HIV U5

Start: 552 End: 636
HIV U5

PBS-K

Start: 636 End: 658
tRNA binding site for Lysine tRNA

PSI

Start: 637 End: 1155
PSI Packaging Sequence
Original Location Description:
637..1155
Qualifiers:
/gene="psi"
/product="pbs-gag"

HIV-PSI

Start: 697 End: 806

HR-ePSI

Start: 807 End: 1144

RRE

Start: 1303 End: 1536
Rev Responsive Element
Original Location Description:
1303..1536
Qualifiers:
/gene="rre"
/product="minimal RRE"

HIV cpPu (Trip)

Start: 2034 End: 2211
Central Poly Purine Track
Original Location Description:
2034..2211
Qualifiers:
/gene="cppt"

Human EF1a Promoter

Start: 2218 End: 3402
Human Elongation Factor 1 alpha promoter

WPRE

Start: 4674 End: 5265
Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element
Original Location Description:
3563..4154
Qualifiers:
/gene="WPRE"

HIV nef

Start: 5271 End: 5359
HIV nef

HIV PPT

Start: 5321 End: 5335
Polypurine Tract (Not degraded by RNAseH so it can serve as a primer for + strand DNA

dU3 LTR

Start: 5336 End: 5571
Deleted U3 Long-Terminal Repeat
Original Location Description:
4225..4460
Qualifiers:
/gene="dU3 LTR"

HIV R

Start: 5373 End: 5488
HIV Repeat Region

HIV U5

Start: 5489 End: 5571
HIV U5

3' Flank

Start: 5572 End: 5642
Chromosomal 3' Flanking Sequence carried over from original HIV integration site

SV40 ori

Start: 7737 End: 7860
Original Location Description:
7991..8114
Qualifiers:
/gene="SV40 ori"

Modified Base (1 total)

Difference from 5' LTR
Start: 5384 End: 5384
This T is a C in the 5' LTR

Primer Binding Site (26 total)

pHAGE-A0050-8050
Start: 158 End: 179 (Complementary)
pHAGE-S501-1000
Start: 401 End: 421
pHAGE-A0550-0050
Start: 667 End: 690 (Complementary)
pHAGE-S1001-1500
Start: 898 End: 919
pHAGE-A1050-0550
Start: 1166 End: 1189 (Complementary)
pHAGE-S1501-2000
Start: 1391 End: 1417
pHAGE-A1550-1050
Start: 1650 End: 1678 (Complementary)
pHAGE-S2001-2500
Start: 1898 End: 1922
pHAGE-A2050-1550
Start: 2084 End: 2112 (Complementary)
oligo #96
Start: 3503 End: 3527
antisense
pHAGE 3' CDS

Start: 4702 End: 4736 (Complementary)
pHAGE-A3550-3050
Start: 4780 End: 4803 (Complementary)
pHAGE-S4001-4500
Start: 4933 End: 4953
pHAGE-A4050-3550
Start: 5261 End: 5283 (Complementary)
pHAGE-S4501-5000
Start: 5569 End: 5600
pHAGE-S6001-6500
Start: 5675 End: 5697
pHAGE-A6050-5550
Start: 5948 End: 5971 (Complementary)
pHAGE-S6501-7000
Start: 6148 End: 6172
pHAGE-A6550-6050
Start: 6417 End: 6440 (Complementary)
pHAGE-S7001-7500
Start: 6668 End: 6692
pHAGE-A7050-6550
Start: 6928 End: 6949 (Complementary)
pHAGE-S7501-8000
Start: 7172 End: 7196
pHAGE-A7550-7050
Start: 7439 End: 7460 (Complementary)
pHAGE-S8001-8500
Start: 7675 End: 7694
pHAGE-A8050-7550
Start: 7915 End: 7941 (Complementary)
pHAGE-S1-500
Start: 8180 End: 8208

Promoter Prokaryotic (1 total)

P-Bla

Start: 5846 End: 5880
Beta Lactamase Promoter

Replication Origin (1 total)

ORI

Start: 6967 End: 7500
Bacterial Origin of Replication
Original Location Description:
7221..7754
Qualifiers:
/gene="ORI"
/product="ColE1 origin of replication"

Mutation (1 total)

EF1a Mutation

Start: 2531 End: 2531

Missing GCCC according to Genbank and Celera sequences

Restriction/Methylation Map

Enzyme	# of cuts	Positions
AfIII	3	518 2603 5455
AgeI	1	2302
ApaLI	3	4885 5997 7243
AvaI	9	296 1928 2269 2371 2557 3072 3189 4567 5644
BamHI	1	3486
BclI	1	3531
BglII	5	474 2790 5304 5370 5411
BssHII	1	712
DraI	8	2062 2725 4053 5321 6091 6783 6802 8243
EagI	3	1145 1149 3405
EcoRI	1	2029
EcoRV	4	36 115 334 4293
FseI	2	1151 2928
FspI	1	6446
HindIII	4	532 1088 1676 5469
NaeI	3	1149 2926 5190
NotI	1	3405
PacI	1	5654
PmlI	2	291 4110
PstI	2	2540 3045
PvuI	1	6300
PvuII	5	436 3603 3816 4193 5299
SacI	4	492 683 3052 5429
SacII	3	2478 2825 5179
SfiI	1	7789
SmaI	1	5646
SpeI	1	2213
StuI	4	2595 2638 3452 7835
XhoI	1	3189
XmaI	1	5644

No cuts: Acc65I, AfeI, AscI, BsrGI, ClaI, EgeI, HpaI, KpnI, MluI, MscI, NcoI, NdeI, NheI, PmeI, Sall,

Sequence

```
1  tggagggtc aattcactcc caaagaagac aagatatacct tgatctgtgg atctaccaca
cacaaggcta cttccctgat tagcagaact acacaccagg
101 gccaggggtc agatataccac tgaccttgg atggtgctac aagctagtagc cagttgagcc
agataaggta gaagaggcca ataaaggaga gaacaccagc
201 ttgttacacc ctgtgagcct gcatgggatg gatgaccgg agagagaagt gttagagtgg
aggtttgaca gccgcctagc atttcatcac gtggcccag
301 agctgcatcc ggagtacttc aagaactgct gatatacgagc ttgctacaag ggactttccg
ctggggactt tccagggagg cgtggcctgg gcgggactgg
```

401 ggagtggcga gccctcagat cctgcatata agcagctgct ttttgccctgt actgggtctc
tctgggttaga ccagatctga gcttgggagc tctctggcta
501 actagggAAC cactgctta agcctcaata agccttgctt tgagtgtctc aagtagtggt
tgcccgtctg ttgtgtgact ctgtaacta gagatccctc
601 agaccctttt agtcagtgtg gaaaatctct agcagtggcg cccgaacagg gacttgaaag
cgaaagggaa accagaggag ctctctcgac gcaggactcg
701 gcttgctgaa gcgcgacgg caagaggcga gggcgggcga ctgggtgagta cgccaaaaat
tttgactagc ggaggctaga aggagagaga tgggtgagc
801 agcgtcagta ttaagcgggg gagaattaga tcgcatggg aaaaaattcg gtttaaggcca
gggggaaaga aaaaatataa attaaacat atagtatggg
901 caagcaggga gctagaacga ttcgcatgta atcctggcct gttagaaca tcagaaggct
gtagacaaat actgggacag ctacaacct cccttcagac
1001 aggatcagaa gaacttagat cattatataa tacagtagca accctctatt gtgtgcatca
aaggatagag ataaaagaca ccaaggaagc tttagacaag
1101 atagagggaag agcaaaaca aagtaagacc accgcacagc aagcggcccg ccgctgatct
tcagacctgg aggaggagat atgagggaca attggagaag
1201 tgaattatat aatataaag tagtaaaaat tgaaccatta ggagttagc ccaccaaggc
aaagagaaga gtgggtcaga gaaaaaaag agcagtggga
1301 ataggagctt tgttccttgg gttcttggga gcagcaggaa gcactatggg cgcagcgtca
atgacgctga cggtagcagg cagacaatta ttgtctggta
1401 tagtgcagca gcagaacaat ttgtgaggg ctattgaggc gcaacagcat ctgttgcaac
tcacagtctg gggcatcaag cagctccagg caagaatcct
1501 ggctgtggaa agatacctaa aggatcaaca gtcctggggg atttgggggt gctctggaaa
actcatttgc accactgctg tgccttggaa tgctagtgtg
1601 agtaataaat ctctggaaca gatttggat cacacgacct ggatggagtg ggacagagaa
attaacaatt acacaagctt aatacactcc ttaattgaag
1701 aatcgcaaaa ccagcaagaa aagaatgaac agaattatt ggaattagat aaatgggcaa
gtttgtggaa ttggtttaac ataacaatt ggctgtggta
1801 tataaaatta ttcataatga tagtaggagg ctggttaggt ttaagaatag tttttgctgt
actttctata gtgaatagag ttaggcaggg atattcacca
1901 ttatcgtttc agaccacct cccaaccccg aggggacccg acaggcccg aggaatagaa
gaagaaggtg gagagagaga cagagacaga tccattcgat
2001 tagtgaacgg atctcgacgg tatcgccgaa tccacaaatg gcagtattca tccacaattt
taaaagaaaa ggggggattg gggggtacag tgcaggggaa
2101 agaatagtag acataatagc aacagacata caaactaaag aattacaaaa acaaattaca
aaaattcaaa attttggggg ttattacagg gacagcagag
2201 atccagtttg gactagtcgt gaggtccgg tgcccgtcag tgggcagagc gcacatcgcc
cacagtcacc gagaagttgg ggggaggggt cggcaattga
2301 accggtgctt agagaaggtg gcgcggggta aactgggaaa gtgatgtcgt gtactggctc
cgcctttttc ccgaggggtg gggagaaccg tatataagt
2401 cagtgtgctc cgtgaacgtt ctttttcgca acgggttgc cgccagaaca caggtaagt
ccgtgtgtgg tccccgccc cctggcctct ttacgggtta
2501 tggcccttgc gtgcctttaa ttaactccac ctggctgcag tacgtgattc ttgatccga
gcttcgggtt ggaagtgggt gggagagttc gaggccttgc
2601 gcttaaggag ccccttcgcc tctgtgctga gttgagcct ggcctgggag ctggggccgc
cgcgtgcgaa tctgggtggc ccttcgccc tgtctcgtg
2701 ctttcgataa gtctctagcc atttaaaatt ttgatgacc tgctgcgacg ctttttttct
ggcaagatag tcttgtaaat gcggccaag atctgcacac
2801 tggatattcg gtttttgggg ccgcgggcgg cgacggggcc cgtgctccc agcgcacatg
ttcggcgagg cggggcctgc gagcgcggcc accgagaatc
2901 ggacgggggt agtctcaagc tggcggcct gctctggtg ctggcctcgc gccgcggtg
atgcccccgc cctgggccc aaggctggcc cggctggc
3001 cagttgcgtg agcggaaaga tggccgcttc ccggccctgc tgcagggagc tcaaatgga
ggacgcggcg ctcgggagag cgggagggtg agtcacccac
3101 acaaaggaaa agggccttcc cgtcctcagc gctcgttca tgtgactcca cggagtaccg
ggcgccttc aggcacctcg attagtctc gagcttttgg
3201 agtacgtcgt ctttaggtt ggggggggg tttatgca tggagtctc ccacactgag
tgggtggaga ctgaagttag gccagcttgg cacttgatg
3301 aattctcctt ggaatttgc ctttttgagt ttgatctt gttcattctc aagcctcaga
cagtggttca aagttttttt cttccatttc aggtgtcgtg
3401 aagcggccgc atgcccgtct ctgtctcgtg gggcctcctc ctgctggcag gcctgtgctg
cctggctcct gtctccctgg ctgaggatcc ccagggagat
3501 gctgcccaga agacagatac atcccacct gatcaggatc acccaacct caacaagatc
acccccacc tggctgagtt cgccttcagc ctataccgc

3601 agctggcaca ccagtcaca agcaccata tcttcttctc cccagtgagc atcgctacag
cctttgcaat gctctccctg gggaccaagg ctgacactca
3701 cgatgaaatc ctggagggcc tgaatttcaa cctcacggag attccggagg ctcagatcca
tgaaggcttc caggaactcc tccgtaccct caaccagcca
3801 gacagccagc tccagctgac caccggcaat ggctgttcc tcagcgaggg cctgaagcta
gtggataagt ttttgaggga tgttaaaaag ttgtaccact
3901 cagaagcctt cactgtcaac ttcggggaca cgaagaggc caagaaacag atcaacgatt
acgtggagaa gggactcaaa gggaaaattg tggatttggg
4001 caaggagctt gacagagaca cagtttttgc tctggtgaat tacatcttct ttaaaggcaa
atgggagaga ccttttgaag tcaaggacac cgaggaagag
4101 gacttccacg tggaccaggt gaccaccgtg aaggtgccta tgatgaagcg tttaggcatg
ttaaactctc agcactgtaa gaagctgtcc agctgggtgc
4201 tgcgatgaa atacctggc aatgccaccg ccatcttctt cctgcctgat gaggggaaac
tacagcacct ggaaaatgaa ctcaccacg atatcatcac
4301 caagttcctg gaaaatgaag acagaaggtc tgccagctta catttaccca aactgtccat
tactggaacc tatgatctga agagcgtcct gggtaactg
4401 ggcatcacta aggtcttcag caatggggct gacctctccg gggtcacaga ggaggcacc
ctgaagctct ccaaggcctg gcataaggct gtgctgacca
4501 tcgacgagaa agggactgaa gctgctgggg ccatgttttt agaggccata cccatgtcta
tccccccga ggtcaagttc acaaacctt ttgtcttctt
4601 aatgattgaa caaaatacca agtctccctt ctcatggga aaagtgggta atccccacca
aaaataaaga tctaatacaa cctctggatt acaaaattg
4701 tgaagattg actggtattc ttaactatgt tgctcctttt acgctatgtg gatacgctgc
ttaaagctt ttgtatcatg ctattgcttc ccgataggct
4801 ttcattttct cctccttgta taaatcctgg ttgctgtctc tttatgagga gttgtggccc
gttgctcagg aacgtggcgt ggtgtgcact gtgtttgctg
4901 acgcaacccc cactggttgg ggcattgcca ccactgtca gctcctttcc gggactttcg
ctttccccct cctattgccc acggcggaac tcatcgccgc
5001 ctgcttgccc cgtgctgga caggggctcg gctgttgggc actgacaatt ccgtgggtgt
gtcggggaaa tcatcgtcct ttccttggct gctgcctgt
5101 gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc cctcaatcca
gcggacctt cttcccgcg cctgtgccg gctctgcggc
5201 ctcttcgcg tcttcgcctt cgcctcaga cgagtcggat ctcccttttg gccgcctccc
cgcctgagat cctttaagac caatgactta caaggcagct
5301 gtagatctta gccacttttt aaaagaaaag ggggactgg aagggctaata tcaactccaa
cgaagacaag atctgctttt tgcttgact gggctctct
5401 ggtagacca gatctgagcc tgggagctct ctggctaact agggaaacca ctgcttaagc
ctcaataaag cttgccttga gtgctcaag tagtgtgtgc
5501 cgtctgttg tgtgactctg gtaactagag atccctcaga cccttttagt cagtgtggaa
aatctctagc agtagtagtt catgtcatct tattattcag
5601 tatttataac ttgcaaagaa atgaatatca gagagtgaga ggcccgggtt aattaaggaa
agggctagat cattcttgaa gacgaaaggg cctcgtgata
5701 cgcctatttt tataggttaa tgatcatgata ataatggttt cttagacgtc aggtggcact
tttcgggaa atgtgcgcg aaccctatt tgtttatttt
5801 tctaaataca tcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat
aatattgaaa aaggaagagt atgagtattc aacatttccg
5901 tgcgcccctt attccccttt ttgocggcatt ttgccttctt gtttttgctc acccagaaac
gctgggtaaa gtaaaagatg ctgaagatca gttgggtgca
6001 cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccctgagag ttttcgcccc
gaagaacgtt ttccaatgat gacactttt aaagtctgc
6101 tatgtggcgc ggtattatcc cgtgttgacg cgggcaaga gcaactcggc gcgcgcatac
actattctca gaatgacttg gttgagtact caccagtcac
6201 agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgctc ccataacat
gagtgataac actgcggcca acttacttct gacaacgatc
6301 ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt
gatcgttggg aaccggagct gaatgaagcc ataccaaacg
6401 acgagctga caccagatg cctgtagcaa tggcaacaac gttgcgcaa ctattaactg
gcgaactact tactctagct tcccggcaac aattaataga
6501 ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggtggctg
gtttattgct gataaatctg gagccggatg gcgtgggtct
6601 cgcggtatca ttgcagcact ggggcccagat ggtaagccct cccgtatcgt agttatctac
acgacgggga gtcaggcaac tatggatgaa cgaaatagac
6701 agatcgtga gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact
catatatact ttagattgat ttaaaacttc atttttaatt

6801 taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc cttaacgtga
 gttttcggtc cactgagcgt cagaccccg agaaaagatc
 6901 aaaggatcctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaaa
 ccaccgctac cagcgggtgg ttgtttgccg gatcaagagc
 7001 taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc
 ttctagtgtg gccgtagtta gccaccact tcaagaactc
 7101 tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg
 cgataagtcg tgtcttaccg ggttgactc aagacgatag
 7201 ttaccggata aggcgcagcg gtcgggctga acgggggggt cgtgcacaca gccagcttg
 gagcgaacga cctacaccga actgagatac ctacagcgtg
 7301 agctatgaga aagcgcaccg cttcccgaag ggagaaaggc ggacaggtat ccggtaaagc
 gcagggtcgg aacaggagag cgcacgaggg agcttccagg
 7401 ggaaacgcc tggatcttt atagctctgt cgggttctgc cacctctgac ttgagcgtcg
 atttttgtga tgctcgtcag gggggcggag cctatggaaa
 7501 aacgccagca acgcggcctt tttacgggtc ctggcctttt gctggccttt tgctcacatg
 ttctttcctg cgttatcccc tgattctgtg gataaccgta
 7601 ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcaggt
 cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc
 7701 gcctctcccc gcgcggttggc cgattcatta atgcagcaag ctcatggctg actaattttt
 tttatattatg cagaggccga gccgcctcg gcctctgagc
 7801 tattccagaa gtagtgagga ggcttttttg gaggcctagg cttttgcaaa aagctccccg
 tggcagcaca gttttcccga ctggaaagcg ggcagtgagc
 7901 gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta cactttatgc
 ttccggctcg tatgttgtgt ggaattgtga gcggataaca
 8001 atttcacaca ggaacagct atgacatgat tacgaatttc acaataaag catttttttc
 actgcattct agttgtgggt tgtccaaact catcaatgta
 8101 tcttatcatg tctggatcaa ctggataact caagctaacc aaaatcatcc caaacttccc
 accccatacc ctattaccac tgccaattac ctgtggtttc
 8201 atttactcta aacctgtgat tcctctgaat tttttcatt ttaaagaaat tgtatattgtt

Component Fragments

#1: FRAGMENT of p201 pHAGE CMV-a1AT-W

parent position: from 2805 to 2212
 length: 7093
 molecule position: from 3405 to 2212

Left Terminus

NotI site #1

Right Terminus

SpeI site #1

#2: FRAGMENT of p302 pHAGE EF1a long-GFP-W

parent position: from 2213 to 3404
 length: 1192
 molecule position: from 2213 to 3404

Left Terminus

SpeI site #1

Right Terminus

NotI site #1