



General Description

DNA p201 pHAGE CMV-a1AT-W
Entire molecule length: 7685 bp

Standard Fields

User Fields

Author(s)

Author

Original author

Darrell Kotton
EMAIL: dkotton@bu.edu

References

Comments

AAW: human a1AT CDS insert generated by PCR from p200 C-AT template adding NotI to 5' end and 11/04 ligated into pHAGE CMV-hHoxB4-W backbone that had been cut NotI/BamH1 (this removes the

p201 pHAGE CMV-ha1AT-W plasmid then confirmed perfect a1AT insert sequence including plasmid maxiprep and glycerol bacterial stock prepared by AAW and stored.

Feature Map

CDS (2 total)

human a1ATT CDS

Start: 2811 End: 4064

AMPr

Start: 5299 End: 6138

Ampicillin Resistance Gene

Original Location Description:

6153..6992

Qualifiers:

/gene="AMPr"

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

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"

CMVp

Start: 2218 End: 2802

Cytomegalovirus Promoter

Original Location Description:

2225..2741

Qualifiers:

/gene="CMVp"

/product="CMV IE promoter"

WPRE

Start: 4074 End: 4665

Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element

Original Location Description:

3563..4154

Qualifiers:

/gene="WPRE"

HIV nef

Start: 4671 End: 4759

HIV nef

HIV PPT

Start: 4721 End: 4735

Polypurine Tract (Not degraded by RNaseH so it can serve as a primer for + strand DNA)

dU3 LTR

Start: 4736 End: 4971
Deleted U3 Long-Terminal Repeat
Original Location Description:
4225..4460
Qualifiers:
/gene="dU3 LTR"

HIV R

Start: 4773 End: 4888
HIV Repeat Region

HIV U5

Start: 4889 End: 4971
HIV U5

3' Flank

Start: 4972 End: 5042
Chromosomal 3' Flanking Sequence carried over from original HIV integration site

SV40 ori

Start: 7137 End: 7260
Original Location Description:
7991..8114
Qualifiers:
/gene="SV40 ori"

Modified Base (1 total)

Difference from 5' LTR

Start: 4784 End: 4784
This T is a C in the 5' LTR

Primer Binding Site (29 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)

pHAGE-S2501-3000

Start: 2393 End: 2417

pHAGE-A2550-2050

Start: 2632 End: 2655 (Complementary)

pHAGE 5' CDS

Start: 2746 End: 2766

oligo #96

Start: 2903 End: 2927

antisense

pHAGE 3' CDS

Start: 4102 End: 4136 (Complementary)

pHAGE-A3550-3050

Start: 4180 End: 4203 (Complementary)

pHAGE-S4001-4500

Start: 4333 End: 4353

pHAGE-A4050-3550

Start: 4661 End: 4683 (Complementary)

pHAGE-S4501-5000

Start: 4969 End: 5000

pHAGE-S6001-6500

Start: 5075 End: 5097

pHAGE-A6050-5550

Start: 5348 End: 5371 (Complementary)

pHAGE-S6501-7000

Start: 5548 End: 5572

pHAGE-A6550-6050

Start: 5817 End: 5840 (Complementary)

pHAGE-S7001-7500

Start: 6068 End: 6092

pHAGE-A7050-6550

Start: 6328 End: 6349 (Complementary)

pHAGE-S7501-8000

Start: 6572 End: 6596

pHAGE-A7550-7050

Start: 6839 End: 6860 (Complementary)

pHAGE-S8001-8500

Start: 7075 End: 7094

pHAGE-A8050-7550

Start: 7315 End: 7341 (Complementary)

pHAGE-S1-500

Start: 7580 End: 7608

Promoter Prokaryotic (1 total)

P-Bla

Start: 5246 End: 5280

Beta Lactamase Promoter

Replication Origin (1 total)

ORI

Start: 6367 End: 6900
Bacterial Origin of Replication
Original Location Description:
7221..7754
Qualifiers:
/gene="ORI"
/product="ColE1 origin of replication"

TATA Signal (1 total)

TATA Box

Start: 2722 End: 2726
CMV Promoter TATA Box

Primary Transcript (1 total)

Transcriptional Start

Start: 2754 End: 2754
CMV Transcriptional Start Site

Restriction/Methylation Map

Enzyme	# of cuts	Positions
AfIII	2	518 4855
ApaLI	3	4285 5397 6643
AvaI	4	296 1928 3967 5044
BamHI	1	2886
BclI	1	2931
BglII	4	474 4704 4770 4811
BssHII	1	712
DraI	7	2062 3453 4721 5491 6183 6202 7643
EagI	3	1145 1149 2805
EcoRI	1	2029
EcoRV	4	36 115 334 3693
FseI	1	1151
FspI	1	5846
HindIII	4	532 1088 1676 4869
NaeI	2	1149 4590
NcoI	1	2527
NdeI	1	2401
NotI	1	2805
PacI	1	5054
PmlI	2	291 3510
PvuI	1	5700
PvuII	5	436 3003 3216 3593 4699
SacI	4	492 683 2735 4829
SacII	1	4579
SfiI	1	7189
SmaI	1	5046

Spel	1	2213
Stul	2	2852 7235
Xmal	1	5044

No cuts: Acc65I, Afel, AgeI, Ascl, BsrGI, ClaI, EgeI, HpaI, KpnI, MluI, MscI, NheI, PmeI, PstI, Sall,

Sequence

```

1  tggaaaggct aattcactcc caaagaagac aagatatoct  tgatctgtgg atctaccaca
cacaaggcta cttccctgat tagcagaact acacaccagg
101 gccaggggct agatatccac tgaccttgg atggtgctac  aagctagtag cagttgagcc
agataaggtg gaagaggcca ataaaggaga gaacaccagc
201 ttgttacacc ctgtgagcct gcatgggatg gatgaccagg  agagagaagt gttagagtgg
aggtttgaca gccgcctagc atttcatcac gtggcccag
301 agctgcatcc ggagtacttc aagaactgct gatatcgagc  ttgctacaag ggactttccg
ctggggactt tccagggagg cgtggcctgg gcgggactgg
401 ggagtggcga gccctcagat cctgcatata agcagctgct  ttttgctgtg actgggtctc
tctggttaga ccagatctga gcctgggagc tctctggcta
501 actaggggaa ccactgctta agcctcaata aagcttgctt  tgagtgtctc aagtagtgtg
tgcccgtctg ttgtgtgact ctggtaacta gagatccctc
601 agaccctttt agtcagtgtg gaaaatctct agcagtggcg  cccgaacagg gacttgaaag
cgaaagggaa accagaggag ctctctcgac gcaggactcg
701 gcttgctgaa gcgcgcacgg caagaggcga ggggcggcga  ctggtgagta cgccaaaaat
ttgactagc ggaggctaga aggagagaga tgggtgagc
801 agcgtcagta ttaagcgggg gagaattaga tcgcgatggg  aaaaaattcg gttaaggcca
gggggaaaga aaaaatataa attaaaacat atagtatggg
901 caagcaggga gctagaacga ttgcagttta atcctggcct  gttagaacaa tcagaaggct
gtagacaaat actgggacag ctacaacat cccttcagac
1001 aggatcagaa gaacttagat cattatataa tacagtagca  accctctatt gtgtgcatca
aaggatagag ataaaagaca ccaaggaagc tttagacaag
1101 atagaggaag agcaaaacaa aagtaagacc accgcacagc  aagcggccgg ccgctgatct
tcagacctgg aggaggagat atgagggaca attggagaag
1201 tgaattatat aatatataag tagtaaaaat tgaaccatta  ggagtagcac ccaccaaggc
aaagagaaga gtggtgcaga gagaaaaaag agcagtggga
1301 ataggagctt tgttccttgg gttcttggga gcagcaggaa  gcactatggg cgcagcgtca
atgacgctga cggtagcagg cagacaatta ttgtctggta
1401 tagtgcagca gcagaacaat ttgctgaggg ctattgaggg  gcaacagcat ctggtgcaac
tcacagtctg gggcatcaag cagctccagg caagaatcct
1501 ggctgtggaa agatacctaa aggatcaaca gctcctgggg  atttgggggt gctctggaaa
actcatttgc accactgctg tgccttggaa tgctagtgtg
1601 agtaataaat ctctggaaca gatttgggat cacacgacct  ggatggagtg ggacagagaa
attaacaatt acacaagctt aatacactcc ttaattgaag
1701 aatcgcaaaa ccagcaagaa aagaatgaac agaattatt  ggaattagat aaatgggcaa
gtttgtgaaa ttggtttaac ataacaat ggctgtggta
1801 tataaaatta ttcataatga tagtaggagg cttggttaggt  ttaagaatag tttttgctgt
actttctata gtgaatagag ttaggcaggg atattcacca
1901 ttatcgttcc agaccacct cccaaccccg aggggaccog  acaggcccca aggaatagaa
gaagaagggt gagagagaga cagagacaga tccattcgat
2001 tagtgaacgg atctcgacgg tatcgccgaa ttcacaaatg  gcagtattca tccacaattt
taaaagaaaa ggggggattg ggggtacag tgcaggggaa
2101 agaatagtag acataatagc aacagacata caaactaaag  aattacaaaa acaaatata
aaaattcaaa attttcgggt ttattacagg gacagcagag
2201 atccagtttg gactagtgga gttccgcggt acataactta  cggtaaatgg cccgcctggc
tgaccgcca acgacccccc cccattgacg tcaataatga
2301 cgtatgttcc catagtaacg ccaatagggg ctttccattg  acgtcaatgg gtggagtatt
tacggtaaac tgcccacttg gcagtacac aagtgtatca
2401 tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa  tggcccgcct ggcattatgc
ccagtacatg accttatggg actttcctac ttggcagtac
2501 atctacgtat tagtcatcgc tattaccatg gtgatcggg  tttggcagta catcaatggg
cgtggatagc ggtttgactc acggggattt ccaagtctcc
2601 acccatttga cgtcaatggg agtttgtttt ggcacaaaa  tcaacgggac tttccaaaat
gtcgtaacaa ctccgcccc tgacgcaaa tgggcggtag
2701 cggtgtacgg tgggaggtct atataagcag agctcgttta  gtgaaccgtc agatgcctg
gagacgcat ccacgtgtt ttgacctcca tagaagacac

```

2801 cggcggccgc atgccgtctt ctgtctcgtg gggcaccctc ctgctggcag gcctgtgctg
cctggtcctt gtctccctgg ctgaggatcc ccagggagat
2901 gctgcccaga agacagatac atcccacat gatcaggatc acccaacctt caacaagatc
accccccaacc tggctgagtt cgccttcagc ctataccgcc
3001 agctggcaca ccagtcaca agcaccaata tctctctc cccagtgagc atcgctacag
cctttgcaat gctctccctg gggaccaagg ctgacactca
3101 cgatgaaatc ctggagggcc tgaatttcaa cctcacggag attccggagg ctcagatcca
tgaaggcttc caggaactcc tccgtaccct caaccagcca
3201 gacagccagc tccagctgac caccggcaat ggctgttcc tcagcgaggg cctgaagcta
gtggataagt ttttggagga tgtaaaaag ttgtaccact
3301 cagaagcctt cactgtcaac ttcggggaca ccgaagaggc caagaaacag atcaacgatt
acgtggagaa gggacttcaa gggaaaattg tggatttggg
3401 caaggagctt gacagagaca cagtttttgc tctggtgaat tacatcttct ttaaaggcaa
atgggagaga ccttttgaag tcaaggacac cgaggaagag
3501 gacttccacg tggaccagg gaccaccgtg aagggtccta tgatgaagcg tttaggcatg
tttaacatcc agcactgtaa gaagctgtcc agctgggtgc
3601 tgctgatgaa atacctggg aatgccaccg ccatcttctt cctgcctgat gaggggaaac
tacagcacct gaaaaatgaa ctcaccacg atatcatcac
3701 caagtctctg gaaaatgaag acagaaggtc tgccagctta catttaccba aactgtccat
tactggaacc tatgatctga agagcgtcct gggtaactg
3801 ggcactacta aggtcttcag caatggggct gacctctccg gggtcacaga ggaggcacc
ctgaagctct ccaaggccgt gcataaggct gtgctgacca
3901 tcgacgagaa agggactgaa gctgctgggg ccatgttttt agaggccata cccatgtcta
tccccccga ggtcaagttc acaaaacct ttgtcttctt
4001 aatgattgaa caaatacca agtctccctt ctcatggga aaagtggta atcccaccca
aaaataaaga tctaatcaa cctctggatt acaaaatttg
4101 tgaagattg actggtattc ttaactatgt tgctctttt acgctatgtg gatacgtg
tttaatgctt ttgtatcatg ctattgcttc ccgtatggct
4201 ttcattttct cctccttgta taaatcctgg ttgctgtctc tttatgagga gttgtggccc
gttgtcaggc aacgtggcgt ggtgtgact gtgtttgctg
4301 acgcaacccc cactggttg ggcattgcca ccactgtca gctcctttcc gggactttcg
ctttccctt cctattgcc acggcgaac tcatcgccg
4401 ctgccttgc cgtgctgga caggggctcg gctgttgggc actgacaatt ccgtgggtgtt
gtcggggaaa tcatcgtcct ttccttggct gctcgcctgt
4501 gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc cctcaatcca
gcgaccttc cttcccggc cctgctgccc gctctgccc
4601 ctcttcggc tcttcgctt cgcctcaga cgagtcggat ctccctttgg gccgctccc
cgcctgagat cctttaagac caatgactta caaggcagct
4701 gtagatctta gccacttttt aaaagaaaag gggggactgg aagggcta tcaactccaa
cgaagacaag atctgctttt tgcctgtact gggctctctt
4801 ggttagacca gatctgagcc tgggagctct ctggctaact agggaaacca ctgcttaagc
ctcaataaag cttgccttga gtgctcaag tagtgtgtg
4901 cgtctgttg tgtgactctg gtaactagag atccctcaga cccttttagt cagtgtggaa
aatctctagc agtagtagtt catgcatct tattattcag
5001 tattataaac ttgcaaagaa atgaatatca gagagtgaga ggccccgggtt aattaaggaa
agggctagat cattctttaa gacgaaagg cctcgtgata
5101 cgcctatctt tataggtaa tgcctgata ataatggttt cttagacgct aggtggcact
tttcggggaa atgtgcgccc aaccctatt tgtttatctt
5201 tctaaataca tcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat
aatattgaaa aaggaagagt atgagtattc aacatttccg
5301 tgtcgcctt attccctttt ttgcggcatt ttgccttct gtttttgcct acccagaaac
gctgggtgaaa gtaaaagatg ctgaagatca gttgggtgca
5401 cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccctgagag ttttcgcccc
gaagaacgct tccaatgat gagcactttt aaagtctctg
5501 tatgtggcgc ggtattatcc cgtgttgacg ccgggcaaga gcaactcggc cgcgcatac
actattctca gaatgacttg gttgagtact caccagtcac
5601 agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgtct ccataacat
gagtataac actgcggcca acttacttct gacaacgatc
5701 ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt
gatcgttggg aaccggagct gaatgaagcc ataccaaacg
5801 acgagcgtga caccagatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg
gcgaactact tactctagct tcccggcaac aattaataga
5901 ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg
gtttattgct gataaatctg gagccggtga cgtgggtct

6001 cgcggtatca ttgcagcaact ggggccagat ggtaagccct cccgtatcgt agttatctac
 acgacgggga gtcaggcaac tatggatgaa cgaaatagac
 6101 agatcgtga gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact
 catatatact ttagattgat ttaaaacttc atttttaatt
 6201 taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc cttaacgtga
 gttttcgttc cactgagcgt cagaccccg agaaaagatc
 6301 aaaggatcct cttgagatcc ttttttctg cgcgtaatct gctgcttga aacaaaaaaa
 ccaccgctac cagcgggtgt ttgtttgccg gatcaagagc
 6401 taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc
 ttctagtgtg gccgtagtta ggccaccact tcaagaactc
 6501 tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg
 cgataagtcg tgtcttaccg ggttgactc aagacgatag
 6601 ttaccggata aggcgcagc gtcgggctga acggggggtt cgtgcacaca gccagcttg
 gagcgaacga cctacaccga actgagatac ctacagcgtg
 6701 agctatgaga aagcgcacg cttcccgaag ggagaaaggc ggacaggtat ccggtaaagc
 gcagggctcg aacaggagag cgcagaggg agcttccagg
 6801 gggaaacgcc tggatcttt atagctctgt cgggtttcgc cacctctgac ttgagcgtcg
 atttttgtga tgctcgtcag gggggcggag cctatggaaa
 6901 aacgccagca acgcggcctt ttacgggtc ctggcctttt gctggccttt tgctcacatg
 ttctttctcg cgttatcccc tgattctgtg gataaccgta
 7001 ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt
 cagtgagcga ggaagcggaa gagcgcctca tacgcaaacc
 7101 gcctctcccc gcgcgttggc cgattcatta atgcagcaag ctcatggctg actaatTTTT
 tttatTTatg cagagggcga ggccgcctcg gcctctgagc
 7201 tattccagaa gtagtgagga ggcttttttg gaggcctagg cttttgcaaa aagctccccg
 tggcacgaca ggtttcccga ctggaagcg ggcagtgagc
 7301 gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta cactttatgc
 ttccggctcg tatgttgtgt ggaattgtga gcggataaca
 7401 atttcacaca ggaacagct atgacatgat tacgaatttc acaataaag catttttttc
 actgcattct agttgtgggt tgcctaaaact catcaatgta
 7501 tcttatcatg tctggatcaa ctggataact caagctaacc aaaatcatcc caaacttccc
 accccatacc ctattaccac tgccaattac ctgtggtttc
 7601 atttactcta aacctgtgat tcctctgaat tttttcatt ttaaagaat tgtatTTgtt

Component Fragments