



General Description

DNA pHAGE-CMV-eGFP-W
Entire molecule length: 7149 bp

Standard Fields

Original Source Database: GenBank
Division: SYN
Modification Date in the Original DB: 31-JUL-2003
Original Accession Number(s): Your; mama'
Sequence Source: Unknown.
Organism(s): Unknown(Unknown)
Taxonomy: Unclassified
Associated Genes: LTR; psi; rre; cppt; CMVp; eGFP; WPRE; dU3 LTR; 3'flank; AMPr; ORI;

User Fields

Author(s)

Author

Original author

Darrell Kotton
EMAIL: dkotton@bu.edu

References

REFERENCE 1 (bases 1 to 8541)
AUTHORS Self
JOURNAL Unpublished.

Comments

HRST Ligated into PUC
WPRE fixed
Changed cPPT to include Spel site for cloning promoters
cloned in clean CMV promoter without extraneous sequence
9/19/03 Completely sequenced with 4x coverage to confirm perfect sequence
11/13/03 Confirmed shortened cloning to add SmaI/PacI and tag
11/20/03 Tested Virus production and expression from viral integration in parallel with all Intermediate

Feature Map

CDS (2 total)

eGFP

Start: 2812 End: 3531
Enhanced Green Fluorescent Protein (Mammalian Codon Optimized)
Original Location Description:
2837..3556
Qualifiers:
/gene="eGFP"

AMPr

Start: 4763 End: 5602
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: 3538 End: 4129
Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element
Original Location Description:
3563..4154
Qualifiers:
/gene="WPRE"

HIV nef

Start: 4135 End: 4223
HIV nef

HIV PPT

Start: 4185 End: 4199
Polypurine Tract (Not degraded by RNaseH so it can serve as a primer for + strand DNA)

dU3 LTR

Start: 4200 End: 4435
Deleted U3 Long-Terminal Repeat
Original Location Description:
4225..4460
Qualifiers:
/gene="dU3 LTR"

HIV R

Start: 4237 End: 4352
HIV Repeat Region

HIV U5

Start: 4353 End: 4435
HIV U5

3' Flank

Start: 4436 End: 4506
Chromosomal 3' Flanking Sequence carried over from original HIV integration site

SV40 ori

Start: 6601 End: 6724
Original Location Description:
7991..8114
Qualifiers:
/gene="SV40 ori"

Modified Base (1 total)

Difference from 5' LTR

Start: 4248 End: 4248
This T is a C in the 5' LTR

Primer Binding Site (31 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

pHAGE-S3001-3500
Start: 2901 End: 2918

pHAGE-A3050-2550
Start: 3144 End: 3164 (Complementary)

pHAGE-S3501-4000
Start: 3399 End: 3420

pHAGE 3' CDS
Start: 3566 End: 3600 (Complementary)

pHAGE-A3550-3050
Start: 3644 End: 3667 (Complementary)

pHAGE-S4001-4500
Start: 3797 End: 3817

pHAGE-A4050-3550
Start: 4125 End: 4147 (Complementary)

pHAGE-S4501-5000
Start: 4433 End: 4464

pHAGE-S6001-6500
Start: 4539 End: 4561

pHAGE-A6050-5550
Start: 4812 End: 4835 (Complementary)

pHAGE-S6501-7000
Start: 5012 End: 5036

pHAGE-A6550-6050
Start: 5281 End: 5304 (Complementary)

pHAGE-S7001-7500
Start: 5532 End: 5556

pHAGE-A7050-6550
Start: 5792 End: 5813 (Complementary)

pHAGE-S7501-8000

Start: 6036 End: 6060

pHAGE-A7550-7050

Start: 6303 End: 6324 (Complementary)

pHAGE-S8001-8500

Start: 6539 End: 6558

pHAGE-A8050-7550

Start: 6779 End: 6805 (Complementary)

pHAGE-S1-500

Start: 7044 End: 7072

Promoter Prokaryotic (1 total)

P-Bla

Start: 4710 End: 4744

Beta Lactamase Promoter

Replication Origin (1 total)

ORI

Start: 5831 End: 6364

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
AflIII	1	6421
ApaLI	3	3749 4861 6107
AvaI	3	296 1928 4508
BamHI	1	3533
BglII	4	474 4168 4234 4275
BsrGI	1	3522
BssHII	1	712
DraI	6	2062 4185 4955 5647 5666 7107
EagI	3	1145 1149 2805
EcoRI	1	2029
EcoRV	3	36 115 334

Fsel	1	1151
Fspl	1	5310
HindIII	4	532 1088 1676 4333
NaeI	2	1149 4054
NarI	1	639
NcoI	2	2527 2811
NdeI	1	2401
NotI	1	2805
PacI	1	4518
PmlI	1	291
PvuI	1	5164
PvuII	2	436 4163
SacI	4	492 683 2735 4293
SacII	1	4043
SfiI	1	6653
SmaI	1	4510
SpeI	1	2213
StuI	1	6699
XmaI	1	4508

No cuts: Acc65I, Afel, Agel, Ascl, BclI, ClaI, HpaI, KpnI, MluI, MscI, NheI, PmeI, PstI, Sall, SbfI, SrfI,

Sequence

```

1  tgggaagggt aattcactcc caaagaagac aagatatacct tgatctgtgg atctaccaca
cacaaggcta cttccctgat tagcagaact acacaccagg
101 gccaggggtc agatataccac tgaccttgg atggtgctac aagctagtag cagttgagcc
agataaggta gaagaggcca ataaaggaga gaacaccagc
201 ttgttacacc ctgtgagcct gcatgggatg gatgaccgag agagagaagt gttagagtgg
aggtttgaca gccgcctagc atttcatcac gtggcccag
301 agctgcatcc ggagtacttc aagaactgct gatatacgagc ttgctacaag ggactttccg
ctggggactt tccagggagg cgtggcctgg gcgggactgg
401 ggagtggcga gccctcagat cctgcatata agcagctgct ttttgccctg actgggtctc
tctgggttaga ccagatctga gcttgggagc tctctggcta
501 actagggaac cactgctta agcctcaata aagcttgctt tgagtgcttc aagtagtgtg
tgcccgtctg ttgtgtgact ctgtaacta gagatccctc
601 agaccctttt agtcagtgtg gaaaatctct agcagtggcg cccgaacagg gacttgaaag
cgaaagggaa accagaggag ctctctcgac gcaggactcg
701 gcttgctgaa gcgcgcacgg caagaggcga gggcgggcga ctggtgagta cgccaaaaat
tttgactagc ggaggctaga aggagagaga tgggtgagc
801 agcgtcagta ttaagcgggg gagaattaga tcgcatggg aaaaaattcg gttaaaggcca
gggggaaaga aaaaatataa attaaaacat atagtatggg
901 caagcagggg gctagaacga ttgcgagtta atcctggcct gttagaaca tcagaaggct
gtagacaaat actgggacag ctacaacat cccttcagac
1001 aggatcagaa gaacttagat cattatataa tacagtagca accctctatt gtgtgcatca
aaggatagag ataaaagaca ccaaggaagc tttagacaag
1101 atagaggaag agcaaaaca aagtaagacc accgcacagc aagcggcccg ccgctgatct
tcagacctgg aggaggagat atgagggaca attggagaag
1201 tgaattatat aatatataag tagtaaaaaat 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 gctctgggg atttgggggt gctctggaaa
actcatttgc accactgctg tgccctggaa tgctagtgtg

```

1601 agtaataaat ctctggaaca gatttggaaat cacacgacct ggatggagtg ggacagagaa
attaacaatt acacaagctt aatacactcc ttaattgaag
1701 aatcgcaaaa ccagcaagaa aagaatgaac agaattatt ggaattagat aaatgggcaa
gtttgtggaa ttggtttaac ataacaaatt ggctgtggta
1801 tataaaatta ttcataatga tagtaggagg cttggtaggt ttaagaatag tttttgctgt
actttctata gtgaatagag ttaggcaggg atattcacca
1901 ttatcgtttc agaccacct cccaacccc aggggacccg acaggcccga aggaatagaa
gaagaagggtg gagagagaga cagagacaga tccattcgat
2001 tagtgaacgg atctcgacgg tatcgccgaa ttcacaaatg gcagtattca tccacaattt
taaaagaaaa ggggggattg ggggtacag tgcaggggaa
2101 agaatagtag acataatagc aacagacata caaactaaag aattacaaaa acaaattaca
aaaattcaaa attttcgggt ttattacagg gacagcagag
2201 atccagtttg gactagtga gttccggtt acataactta cggtaaattg cccgcctggc
tgaccgcca acgaccccc cccattgacg tcaataatga
2301 cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt
tacggtaaac tgcccacttg gcagtacatc aagtgtatca
2401 tatgcaagt acgcccccta ttgacgtcaa tgacggtaaa tggccccctt ggcattatgc
ccagtacatg accttatggg actttcctac ttggcagtag
2501 atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg
cgtggatagc ggtttgactc acggggattt ccaagtctcc
2601 accccattga cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tttccaaaat
gtcgtacaaa ctccgcccc ttgacgcaaa tgggcggtag
2701 gcgtgtacgg tgggaggtct atataagcag agctcgttta gtgaaccgtc agatcgcctg
gagacgccat ccacgctgtt ttgacctca tagaagacac
2801 cggcgccgc catggtgagc aaggcgagg agctgttcc cggggtggtg cccatcctgg
tcgagctgga cggcgacgta aacggccaca agttcagcgt
2901 gtccgcgag ggcgagggcg atgccacta cggcaagctg accctgaagt tcatctgcac
caccggcaag ctgcccgtgc cctggcccac cctcgtgacc
3001 accctgacct acggcgtgca gtgcttcagc cgctacccc accacatgaa gcagcacgac
ttcttcaagt ccgccatgcc cgaaggctac gtccaggagc
3101 gcaccatctt cttcaaggac gacggcaact acaagacccg cgccgaggtg aagttcgagg
gcgacacctt ggtgaaccgc atcgagctga agggcatcga
3201 cttcaaggag gacggcaaca tctggggca caagctggag tacaactaca acagccacaa
cgtctatata atggccgaca agcagaagaa cggcatcaag
3301 gtgaacttca agatccgcca caacatcgag gacggcagcg tgcagctcgc cgaccactac
cagcagaaca cccccatcgg cgacggcccc gtgctgctgc
3401 ccgacaacca ctacctgagc accagtcgg cctgagcaa agaccccaac gagaagcggc
atcacatggt cctgctggag ttcgtgaccg ccgcccggat
3501 cactctcggc atggacgagc tgtacaagta aggatcctaa tcaacctctg gattacaaaa
tttgtgaaag attgactggt attcttaact atgttctcc
3601 ttttacgcta tgtggatacg ctgctttaat gcctttgtat catgctattg cttcccgtat
ggctttcatt ttctctcct tgtataaatc ctggttgcg
3701 tctctttatg aggagttgtg gccggttgc aggcaacgtg gcgtggtgtg cactgtgttt
gctgacgcaa cccccactgg ttgggcatt gccaccact
3801 gtcagctcct ttcgggact ttcgcttcc cctccctat tgccacggcg gaactcatcg
ccgctgctt tgcccgtgc tggacagggg ctggctgtt
3901 gggcactgac aattccgtg tgttgcggg gaaatcatcg tcttttctt ggctgctcgc
ctgtgttgc acctggatc tgcgcgggac gtccttctgc
4001 tacgtccctt cggccctcaa tccagcggac cttccttccc gggcctgct gccggctctg
cggcctctc cgcgtcttcg ccttcgccc cagacgagtc
4101 ggatctccc ttgggcccgc tcccgcctg agatccttta agaccaatga cttacaaggc
agctgtagat cttagccact ttttaaaaga aaagggggga
4201 ctggaagggc taattcactc ccaacgaaga caagatctgc tttttgcttg tactgggtct
ctctggttag accagatctg agcctgggag ctctctggct
4301 aactagggaa cccactgctt aagcctcaat aaagcttgc ttgagtgctt caagtagtgt
gtgcccgtct gttgtgtgac tctgtaact agagatcct
4401 cagaccctt tagtcagtgt ggaaaatctc tagcagtagt agttcatgct atcttattat
tcagtattta taacttgcaa agaaatgaat atcagagagt
4501 gagagcccg ggttaattaa ggaaagggt agatcattct tgaagacgaa agggcctcgt
gatacgccta tttttatagg ttaatgcat gataataatg
4601 gtttcttaga cgtcaggtg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta
tttttctaaa tacattcaaa tatgtatccg ctcatgagac
4701 aaaaacctg ataatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt
tccgtgctgc ccttatccc tttttgagg cattttgctt

4801 tctgtttttt gctcaccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg
 tgcacgagtg ggttacatcg aactggatct caacagcggg
 4901 aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt
 ctgctatgtg gcgcggtatt atcccgtgtt gacgccgggc
 5001 aagagcaact cggtcgccc atacactatt ctcagaatga cttggttgag tactcaccag
 tcacagaaaa gcactttacg gatggcatga cagtaagaga
 5101 attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac ttctgacaac
 gatcggagga ccgaaggagc taaccgcttt ttgacacaac
 5201 atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca
 aacgacgagc gtgacaccac gatgcctgta gcaatggcaa
 5301 caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg caacaattaa
 tagactggat ggaggcggat aaagttgcag gaccacttct
 5401 gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg
 gtctcgggt atcattgcag cactggggcc agatggtaag
 5501 cctcccgtg taagtattat ctacacgacg gggagtcagg caactatgga tgaacgaaat
 agacagatcg ctgagatagg tgcctcactg attaagcatt
 5601 ggtaactgct agaccaagtt tactcatata tacttttagat tgatttaaaa cttcattttt
 aatttaaaag gatctaggtg aagatccttt ttgataatct
 5701 catgacaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa
 gatcaaaagga tcttcttgag atcctttttt tctgcgctg
 5801 atctgctgct tgcaaaaaaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa
 gagctaccaa ctctttttcc gaagtaact ggcttcagca
 5901 gagcgcagat accaaatact gttcttctag ttagcgcgta gttaggccac cacttcaaga
 actctgtagc accgcctaca tacctcgtc tgctaactct
 6001 gttaccagtg gctgctgcca gtggcgataa gctggtctt accgggttgg actcaagacg
 atagttaccg gataaggcgc agcgtcggg ctgaacgggg
 6101 ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag
 cgtgagctat gagaaagcgc cacgctccc gaagggagaa
 6201 agcggacag gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc
 cagggggaaa cgctgggtat ctttatagtc ctgctgggtt
 6301 tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg
 gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc
 6401 ttttctggc cttttgctca catgttcttt cctgcgttat ccctgattc tgtggataac
 cgtattaccg cttttgagtg agctgatacc gctcgcgcga
 6501 gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgc
 aaccgcctct ccccgcgctg tggccgattc attaatgcag
 6601 caagctcatg gctgactaat tttttttatt tatgcagagg ccgagggcgc ctccgctct
 gagctattcc agaagtagtg aggagcttt tttggaggcc
 6701 taggcttttg caaaaagctc cccgtggcac gacaggtttc ccgactggaa agcgggcagt
 gagcgcaacg caattaatgt gagttagctc actcattagg
 6801 caccocaggc tttacacttt atgcttccgg ctgctatggt gtgtggaatt gtgagcggat
 aacaatttca cacaggaaac agctatgaca tgattacgaa
 6901 tttcacaaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca aactcatcaa
 tgtatcttat catgctgga tcaactggat aactcaagct
 7001 aacccaaatc atcccaact tcccaccca taccctatta ccaactgcaa ttacctgtgg
 tttcatttac tctaaacctg tgattcctct gaattatfff
 7101 cattttaag aaattgtatt tgttaaataat gtactacaaa cttagtagt

Component Fragments