

HA-mPB

General Description

DNA 'HA-mPB'

Expression vector containing the mPB transposase coding sequence in a pcDNA3-Kz-HA backbone

Currently local object. Original author: s

Created: 07/11/07 12:33PM

Last Modified: 07/11/07 12:38PM

length: 7253 bp

storage type: Basic

form: Circular

Standard Fields

Original Source Database: GenBank

Modification Date in the Original DB: 11-JAN-2007

Original Author

Comments

Annotations

Feature Map

CDS (3 total)

mPB

Start: 983 End: 2770

Optimized piggyBac transposase CDS

Original Location Description:

983..2770

Neo(R)

Start: 3958 End: 4752

Original Location Description:

3958..4752

Amp(R)

Start: 6257 End: 7117 (Complementary)

Original Location Description:

complement(6257..7117)

Misc. Feature (2 total)

Kozak seq

Start: 905 End: 917

Original Location Description:

905..917

HA tag

Start: 929 End: 949

Original Location Description:

929..949

PolyA Signal (2 total)

bGH polyA

Start: 2825 End: 3056

Original Location Description:

2825..3056

SV40 pA

Start: 4926 End: 5056
Original Location Description:
4926..5056

Promoter Eukaryotic (2 total)

CMV promoter

Start: 236 End: 823
Original Location Description:
236..823

SV40 early promoter

Start: 3553 End: 3922
Original Location Description:
3553..3922

Replication Origin (2 total)

f1 origin

Start: 3119 End: 3547
Original Location Description:
3119..3547

pUC origin

Start: 5439 End: 6112 (Complementary)
Original Location Description:
complement(5439..6112)

Restriction/Methylation Map

BamHI: 2 sites

GATC
CCTAG

N1: 918
N2: 1225

BglII: 3 sites

AGATCT
TCTAGA

N1: 960
N2: 1507
N3: 1720

Clal: 1 site

ATCGAT
TAGCTA

N1: 16

EcoRI: 1 site

GAATTC
CTTAAG

N1: 969

EcoRV: 0 sites

GATATC
CTATAG

HindIII: 1 site

AGCTT
TTCGAA

N1: 894

KpnI: 0 sites

GGTACC
CATGG

NotI: 1 site

GCGCCG
CGCCGC

N1: 2773

SacI: 2 sites

GAGCTC
CTCGAG

N1: 823
N2: 2833

HA-mPB

SacII: 0 sites

ccgcgg
ggcgcc

Sall: 3 sites

ctcgac
cagctg

N1: 37

N2: 5064

N3: 7252

XbaI: 1 site

ctctaga
agatct

N1: 2791

XhoI: 2 sites

ctcgag
gagctc

N1: 900








N2: 2779

Restriction Fragments

5374: HA-mPB: XhoI(2779) - XhoI(900)

1879: HA-mPB: XhoI(900) - XhoI(2779)

HA-mPB

																																		
1	GACGGATCGG	GAGATCGATC	TCCCGATCCC	CTATGGTCTGA	CTCTCAGTAC	AATCTGCTCT	GATGCCGCAT	AGTTAAGCCA	GTATCTGCTC	CCTGCTTGTG																								
	CTGCCTAGCC	CTCTAGCTAG	AGGGCTAGGG	GATACCAGCT	GAGAGTCATG	TTAGACGAGA	CTACGGCGTA	TCAATTCGGT	CATAGACGAG	GGACGAACAC																								
101	TGTTGGAGGT	CGCTGAGTAG	TGCGCGAGCA	AAATTTAAGC	TACAACAAGG	CAAGGCTTGA	CCGACAATTG	CATGAAGAAT	CTGCTTAGGG	TTAGGCGTTT																								
	ACAACCTCCA	GCGACTCATC	ACGCGCTCGT	TTTAAATTCG	ATGTTGTTCC	GTTCCGAAC	GGCTGTTAAC	GTACTTCTTA	GACGAATCCC	AATCCGCAAA																								
201	TGCGCTGCTT	CGCGATGTAC	GGGCCAGATA	TACGCGTTGA	CATTGATTAT	TGACTAGTTA	TTAATAGTAA	TCAATTACGG	GGTCATTAGT	TCATAGCCCCA																								
	ACGCGACGAA	GCGCTACATG	CCCGGTCTAT	ATGCGCAACT	GTAACATAA	ACTGATCAAT	AATTATCATT	AGTTAATGCC	CCAGTAATCA	AGTATCGGGT																								
301	TATATGGAGT	TCCGCGTTAC	ATAACTTACG	GTAAATGGCC	CGCCTGGCTG	ACCGCCCAAC	GACCCCGCC	CATTGACGTC	AATAATGACG	TATGTTCCCA																								
	ATATACCTCA	AGGCGCAATG	TATTGAATGC	CATTTACCGG	GCGGACCGAC	TGGCGGGTTG	CTGGGGGCGG	GTAACGTCAG	TTATTACTGC	ATACAAGGGT																								
401	TAGTAACGCC	AATAGGGACT	TTCCATTGAC	GTCAATGGGT	GGACTATTTA	CGGTAAACTG	CCCACTTGGC	AGTACATCAA	GTGTATCATA	TGCCAAGTAC																								
	ATCATTGCGG	TTATCCCTGA	AAGGTAACG	CAGTTACCCA	CCTGATAAAT	GCCATTTGAC	GGGTGAACCG	TCATGTAGTT	CACATAGTAT	ACGGTTCATG																								
501	GCCCCCTATT	GACGTCAATG	ACGGTAAATG	GCCCGCCTGG	CATTATGCCC	AGTACATGAC	CTTATGGGAC	TTTCCTACTT	GGCAGTACAT	CTACGTATTA																								
	CGGGGGATAA	CTGCAGTTAC	TGCCATTTAC	CGGGCGGACC	GTAATACGGG	TCATGTACTG	GAATACCCTG	AAAGGATGAA	CCGTCATGTA	GATGCATAAT																								
601	GTCATCGCTA	TTACCATGGT	GATGCGGTTT	TGGCAGTACA	TCAATGGGCG	TGGATAGCGG	TTTGACTCAC	GGGGATTTC	AAGTCTCCAC	CCCATTGACG																								
	CAGTAGCGAT	AATGGTACCA	CTACGCCAAA	ACCGTCATGT	AGTTACCCGC	ACCTATCGCC	AAACTGAGTG	CCCCTAAAGG	TTCAGAGGTG	GGGTAACTGC																								
701	TCAATGGGAG	TTTGTTTTGG	CACCAAAATC	AACGGGACTT	TCCAAAATGT	CGTAACAAC	CCGCCCCATT	GACGCAAATG	GGCGGTAGGC	GTGTACGGTG																								
	AGTTACCCTC	AAACAAAACC	GTGGTTTTAG	TTGCCCTGAA	AGGTTTTTACA	GCATTGTTGA	GGCGGGGTAA	CTGCGTTTAC	CCGCCATCCG	CACATGCCAC																								
801	GGAGGTCTAT	ATAAGCAGAG	CTCTCTGGCT	AACTAGAGAA	CCCACTGCTT	ACTGGCTTAT	CGAAATTAAT	ACGACTCACT	ATAGGGAGAC	CCAAGCTTCT																								
	CCTCCAGATA	TATTCGTCTC	GAGAGACCGA	TTGATCTCTT	GGGTGACGAA	TGACCGAATA	GCTTTAATTA	TGCTGAGTGA	TATCCCTCTG	GGTTCGAAGA																								
+2																																		
		Met	Gly	Ser	Thr	Met	Tyr	Asp	Val	Pro	Asp	Tyr	Ala	Ser	Leu	Pro	Arg	Ser	Pro	Glu	Phe	Ala	Ala	Thr	Met	Gly	Ser	Ser	Leu	Asp				
901	CGAGGCCGCC	ACCATGGGAT	CCACCATGTA	CGACGTTTCT	GATTACGCTA	GCCTCCCGAG	ATCTCCTGAA	TTGCGCGCCA	CCATGGGCAG	CAGCCTGGAC																								
	GCTCCGCGCG	TGGTACCCTA	GGTGGTACAT	GCTGCAAGGA	CTAATGCGAT	CGGAGGGCTC	TAGAGGACTT	AAGCGGCGGT	GGTACCCGTC	GTCGGACCTG																								
+2	Asp	Glu	His	Ile	Leu	Ser	Ala	Leu	Leu	Gln	Ser	Asp	Asp	Glu	Leu	Val	Gly	Glu	Asp	Ser	Asp	Ser	Glu	Ile	Ser	Asp	His	Val	Ser	Glu	Asp	Asp	Val	Gln
1001	GACGAGCACA	TCCTGAGCGC	CCTGCTGCAG	AGCGACGACG	AGCTGGTCTG	CGAGGACAGC	GACAGCGAGA	TCAGCGACCA	CGTGAGCGAG	GACGACGTGC																								
	CTGCTCTGTG	AGGACTCGCG	GGACGACGTC	TCGCTGCTGC	TCGACCAGCC	GCTCCTGTGC	CTGTCTGCTCT	AGTCGCTGGT	GCACTCGCTC	CTGCTGCACG																								
+2	·Glr	Ser	Asp	Thr	Glu	Glu	Ala	Phe	Ile	Asp	Glu	Val	His	Glu	Val	Gln	Pro	Thr	Ser	Ser	Gly	Ser	Glu	Ile	Leu	Asp	Glu	Gln	Asn	Val	Ile	Glu	Gln	Pro
1101	AGTCCGACAC	CGAGGAGGCC	TTCATCGACG	AGGTGCACGA	GGTGCAGCCT	ACCAGCAGCG	GCTCCGAGAT	CCTGGACGAG	CAGAACGTGA	TCGAGCAGCC																								
	TCAGGCTGTG	GCTCCTCCGG	AAGTAGCTGC	TCCACGTGCT	CCACGTCGGA	TGGTCGTCGC	CGAGGCTCTA	GGACCTGCTC	GTCTTGCACT	AGCTCGTCGG																								

HA-mPB

	+2	Pro Gly Ser Ser Leu Ala Ser Asn Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg
1201		CGGCAGCTCC CTGGCCAGCA ACAGGATCCT GACCCTGCCC CAGAGGACCA TCAGGGGCAA GAACAAGCAC TGCTGGTCCA CCTCCAAGAG CACCAGGCGG GCCGTCGAGG GACCGGTCGT TGTCTAGGA CTGGGACGGG GTCTCCTGGT AGTCCCCGTT CTTGTTCGTG ACGACCAGGT GGAGGTTCTC GTGGTCCGCC
	+2	Ser Arg Val Ser Ala Leu Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr
1301		AGCAGGGTGT CCGCCCTGAA CATCGTGAGA AGCCAGAGGG GCCCCACCAG GATGTGCAGG AACATCTACG ACCCCCTGCT GTGCTTCAAG CTGTTCTTCA TCGTCCCACA GGCGGGACTT GTAGCACTCT TCGGTCTCCC CGGGGTGGTC CTACACGTCC TTGTAGATGC TGGGGGACGA CACGAAGTTC GACAAGAAGT
	+2	Thr Asp Glu Ile Ile Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg Glu Ser Met Thr Gly Ala Thr Phe Arg Asp Thr Asn Glu
1401		CCGACGAGAT CATCAGCGAG ATCGTGAAGT GGACCAACGC CGAGATCAGC CTGAAGAGGC GGGAGAGCAT GACCGGCGCC ACCTTCAGGG ACACCAACGA GGCTGCTCTA GTAGTCGCTC TAGCACTTCA CCTGGTTGCG GCTCTAGTCG GACTTCTCCG CCCTCTCGTA CTGGCCGCGG TGGAAGTCCC TGTGGTTGCT
	+2	Glu Asp Glu Ile Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met
1501		GGACGAGATC TACGCCTTCT TCGGCATCCT GGTGATGACC GCCGTGAGGA AGGACAACCA CATGAGCACC GACGACCTGT TCGACAGATC CCTGAGCATG CCTGCTCTAG ATGCGGAAGA AGCCGTAGGA CCACTACTGG CGGCACTCCT TCCTGTTGGT GTACTCGTGG CTGCTGGACA AGCTGTCTAG GGACTCGTAC
	+2	Val Tyr Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
1601		GTGTACGTGA GCGTGATGAG CAGGGACAGA TTCGACTTCC TGATCAGATG CCTGAGGATG GACGACAAGA GCATCAGGCC CACCCTGCGG GAGAACGACG CACATGCACT CGCACTACTC GTCCCTGTCT AAGCTGAAGG ACTAGTCTAC GGACTCCTAC CTGCTGTTCT CGTAGTCCGG GTGGGACGCC CTCTTGCTGC
	+2	Val Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu Gly
1701		TGTTACACCC CGTGAGAAAG ATCTGGGACC TGTTTCATCCA CCAGTGCATC CAGAACTACA CCCCTGGCGC CCACCTGACC ATCGACGAGC AGCTGCTGGG ACAAGTGGGG GCACTCTTTC TAGACCCTGG ACAAGTAGGT GGTCACGTAG GTCTTGATGT GGGGACCGCG GGTGGACTGG TAGCTGCTCG TCGACGACCC
	+2	Gly Phe Arg Gly Arg Cys Pro Phe Arg Met Tyr Ile Pro Asn Lys Pro Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys Tyr Met
1801		CTTCAGGGGC AGGTGCCCCCT TCAGGATGTA TATCCCCAAC AAGCCCAGCA AGTACGGCAT CAAGATCCTG ATGATGTGCG ACAGCGGCAC CAAGTACATG GAAGTCCCCG TCCACGGGGA AGTCCTACAT ATAGGGGTTG TTCGGGTCGT TCATGCCGTA GTTCTAGGAC TACTACACGC TGTCGCCGTG GTTCATGTAC
	+2	Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val His Gly Ser Cys
1901		ATCAACGGCA TGCCCTACCT GGGCAGGGGC ACCCAGACCA ACGGCGTGCC CCTGGGCGAG TACTACGTGA AGGAGCTGTC CAAGCCCCTC CACGGCAGCT TAGTTGCCGT ACGGGATGGA CCCGTCCCCG TGGGTCTGGT TGCCGCACGG GGACCCGCTC ATGATGCACT TCCTCGACAG GTTCGGGCAG GTGCCGTCTGA
	+2	Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val Gly Thr Val Arg Ser
2001		GCAGAAACAT CACCTGCGAC AACTGGTTCA CCAGCATCCC CCTGGCCAAG AACCTGCTGC AGGAGCCCTA CAAGCTGACC ATCGTGGGCA CCGTGAGAAG CGTCTTTGTA GTGGACGCTG TTGACCAAGT GGTTCGTAGGG GGACCGGTTT TTGGACGACG TCCTCGGGAT GTTCGACTGG TAGCACCCGT GGCACCTCTC

HA-mPB

	+2	-Ser	Asn	Lys	Arg	Glu	Ile	Pro	Glu	Val	Leu	Lys	Asn	Ser	Arg	Ser	Arg	Pro	Val	Gly	Thr	Ser	Met	Phe	Cys	Phe	Asp	Gly	Pro	Leu	Thr	Leu	Val	Ser	Tyr	
2101		CAACAAGAGA	GAGATCCCCG	AGGTCCTGAA	GAACAGCAGG	TCCAGGCCCG	TGGGCACCAG	CATGTTCTGC	TTCGACGGCC	CCCTGACCCT	GGTGTCTTAC	GTGTTCTCT	CTCTAGGGG	TCCAGGACTT	CTTGTCGTCC	AGGTCCGGG	ACCCGTGGT	GTACAAGAC	AAGCTGCCG	GGGACTGGGA	CCACAGGAT															
	+2	Lys	Pro	Lys	Pro	Ala	Lys	Met	Val	Tyr	Leu	Leu	Ser	Ser	Cys	Asp	Glu	Asp	Ala	Ser	Ile	Asn	Glu	Ser	Thr	Gly	Lys	Pro	Gln	Met	Val	Met	Tyr	Tyr	Asn	
2201		AAGCCCAAGC	CCGCCAAGAT	GGTGTACCTG	CTGTCCAGCT	GCGACGAGGA	CGCCAGCATC	AACGAGAGCA	CCGGCAAGCC	CCAGATGGTG	ATGTACTACA	TTCGGGTTTC	GGCGGTTCTA	CCACATGGAC	GACAGGTCGA	CGCTGCTCCT	GCGGTCTGTAG	TTGCTCTCTG	GGCCGTTTCG	GGTCTACCAC	TACATGATGT															
	+2	-Asn	Gln	Thr	Lys	Gly	Gly	Val	Asp	Thr	Leu	Asp	Gln	Met	Cys	Ser	Val	Met	Thr	Cys	Ser	Arg	Lys	Thr	Asn	Arg	Trp	Pro	Met	Ala	Leu	Leu	Tyr	Gly	Met	
2301		ACCAGACCAA	GGGCGGCGTG	GACACCCTGG	ACCAGATGTG	CAGCGTGATG	ACCTGCAGCA	GAAAGACCAA	CAGGTGGCCC	ATGGCCCTGC	TGTACGGCAT	TGGTCTGGTT	CCCGCCGCAC	CTGTGGGACC	TGGTCTACAC	GTCGCACTAC	TGGACGTCGT	CTTTCTGGTT	GTCCACCGGG	TACCGGGACG	ACATGCCGTA															
	+2	-Met	Ile	Asn	Ile	Ala	Cys	Ile	Asn	Ser	Phe	Ile	Ile	Tyr	Ser	His	Asn	Val	Ser	Ser	Lys	Gly	Glu	Lys	Val	Gln	Ser	Arg	Lys	Lys	Phe	Met	Arg	Asn	Leu	
2401		GATCAACATC	GCCTGCATCA	ACAGCTTCAT	CATCTACAGC	CACAACGTGA	GCAGCAAGGG	CGAGAAGGTG	CAGAGCCGGA	AAAAGTTCAT	GCGGAACCTG	CTAGTTGTAG	CGGACGTAGT	TGTCGAAGTA	GTAGATGTCT	GTGTTGCACT	CGTCGTTCCC	GCTCTTCCAC	GTCTCGGCCT	TTTTCAAGTA	CGCCTTGGAC															
	+2	Tyr	Met	Ser	Leu	Thr	Ser	Ser	Phe	Met	Arg	Lys	Arg	Leu	Glu	Ala	Pro	Thr	Leu	Lys	Arg	Tyr	Leu	Arg	Asp	Asn	Ile	Ser	Asn	Ile	Leu	Pro	Asn	Glu	Val	
2501		TACATGAGCC	TGACCTCCAG	CTTCATGAGG	AAGAGGCTGG	AGGCCCCCAC	CCTGAAGAGA	TACCTGAGGG	ACAACATCAG	CAACATCCTG	CCCAACGAGG	ATGTAATCGG	ACTGGAGGTC	GAAGTACTCC	TTCTCCGACC	TCCGGGGGTG	GGACTTCTCT	ATGGACTCCC	TGTTGTAGTC	GTTGTAGGAC	GGGTTGCTCC															
	+2	-Val	Pro	Gly	Thr	Ser	Asp	Asp	Ser	Thr	Glu	Glu	Pro	Val	Met	Lys	Lys	Arg	Thr	Tyr	Cys	Thr	Tyr	Cys	Pro	Ser	Lys	Ile	Arg	Arg	Lys	Ala	Asn	Ala	Ser	
2601		TGCCCCGCAC	CAGCGACGAC	AGCACCGAGG	AGCCCGTGAT	GAAGAAGAGG	ACCTACTGCA	CCTACTGTCC	CAGCAAGATC	AGAAGAAAGG	CCAACGCCAG	ACGGGCCGTG	GTCGCTGCTG	TCGTGGCTCC	TCGGGCACTA	CTTCTTCTCC	TGGATGACGT	GGATGACAGG	GTCGTTCTAG	TCTTCTTTCC	GGTTGCGGTC															
	+2	-Ser	Cys	Lys	Lys	Cys	Lys	Lys	Val	Ile	Cys	Arg	Glu	His	Asn	Ile	Asp	Met	Cys	Gln	Ser	Cys	Phe	***	***	<div><div>XhoI</div><div>NotI</div><div>XbaI</div></div>										
2701		CTGCAAGAAG	TGTAAGAAGG	TCATCTGCCG	GGAGCACAAC	ATCGACATGT	GCCAGAGCTG	TTTCTGATGA	GCGGCCGCTC	GAGCATGCAT	CTAGAGGGCC	GACGTTCTTC	ACATTCTTCC	AGTAGACGGC	CCTCGTGTTG	TAGCTGTACA	CGGTCTCGAC	AAAGACTACT	CGCCGGCGAG	CTCGTACGTA	GATCTCCCGG															
2801		CTATTCTATA	GTGTCACCTA	AATGCTAGAG	CTCGCTGATC	AGCCTCGACT	GTGCCTTCTA	GTTGCCAGCC	ATCTGTTGTT	TGCCCCCTCC	CCGTGCCTTC	GATAAGATAT	CACAGTGGAT	TTACGATCTC	GAGCGACTAG	TCGGAGCTGA	CACGGAAGAT	CAACGGTCGG	TAGACAACAA	ACGGGGAGGG	GGCACGGAAG															
2901		CTTGACCCTG	GAAGGTGCCA	CTCCCACTGT	CCTTTCCTAA	TAAAATGAGG	AAATTGCATC	GCATTGTCTG	AGTAGGTGTC	ATTCTATTCT	GGGGGGTGGG	GAAGTGGGAC	CTTCCACGGT	GAGGGTGACA	GGAAAGGATT	ATTTTACTCC	TTTAACGTAG	CGTAACAGAC	TCATCCACAG	TAAGATAAGA	CCCCCACC															
3001		GTGGGGCAGG	ACAGCAAGGG	GGAGGATTGG	GAAGACAATA	GCAGGCATGC	TGGGGATGCG	GTGGGGCTCTA	TGGCTTCTGA	GGCGGAAAGA	ACCAGCTGGG	CACCCCGTCC	TGTCGTTCCC	CCTCCTAACC	CTTCTGTTAT	CGTCCGTACG	ACCCCTACGC	CACCCGAGAT	ACCGAAGACT	CCGCCTTTCT	TGGTCGACCC															
3101		GCTCTAGGGG	GTATCCCCAC	GCGCCCTGTA	GCGGCGCATT	AAGCGCGGCG	GGTGTGGTGG	TTACGCGCAG	CGTGACCGCT	ACACTTGCCA	GCGCCCTAGC	CGAGATCCCC	CATAGGGGTG	CGCGGGACAT	CGCCGCGTAA	TTCGCGCCGC	CCACACCACC	AATGCGCGTC	GCACTGGCGA	TGTGAACGGT	CGCGGGATCG															

HA-mPB

3201	GCCCGCTCCT	TTCGCTTTCT	TCCCTTCCTT	TCTCGCCACG	TTCGCCGGCT	TTCCCCGTCA	AGCTCTAAAT	CGGGGCATCC	CTTTAGGGTT	CCGATTTAGT
	CGGGCGAGGA	AAGCGAAAGA	AGGGAAGGAA	AGAGCGGTGC	AAGCGGCCGA	AAGGGGCAGT	TCGAGATTTA	GCCCCGTAGG	GAAATCCCAA	GGCTAAATCA
3301	GCTTTACGGC	ACCTCGACCC	CAAAAACTT	GATTAGGGTG	ATGGTTTCACG	TAGTGGGGCCA	TCGCCCTGAT	AGACGGTTTTT	TCGCCCTTTG	ACGTTGGAGT
	CGAAATGCCG	TGGAGCTGGG	GTTTTTTGAA	CTAATCCCAC	TACCAAGTGC	ATCACCCGGT	AGCGGGACTA	TCTGCCAAAA	AGCGGGAAAC	TGCAACCTCA
3401	CCACGTTCTT	TAATAGTGGA	CTCTTGTTCC	AAACTGGAAC	AACACTCAAC	CCTATCTCGG	TCTATTCTTT	TGATTTATAA	GGGATTTTGG	GGATTTCCGGC
	GGTGCAAGAA	ATTATCACCT	GAGAACAAGG	TTTGACCTTG	TTGTGAGTTG	GGATAGAGCC	AGATAAGAAA	ACTAAATATT	CCCTAAAAACC	CCTAAAGCCG
3501	CTATTGGTTA	AAAAATGAGC	TGATTTAACA	AAAATTTAAC	GCGAATTAAT	TCTGTGGAAT	GTGTGTCAGT	TAGGGTGTGG	AAAGTCCCCA	GGCTCCCCAG
	GATAACCAAT	TTTTTACTCG	ACTAAATTGT	TTTTAAATTG	CGCTTAATTA	AGACACCTTA	CACACAGTCA	ATCCCACACC	TTTCAGGGGT	CCGAGGGGTC
3601	GCAGGCAGAA	GTATGCAAAG	CATGCATCTC	AATTAGTCAG	CAACCAGGTG	TGGAAAGTCC	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC
	CGTCCGTCTT	CATACGTTTC	GTACGTAGAG	TTAATCAGTC	GTTGGTCCAC	ACCTTTCAGG	GGTCCGAGGG	GTCGTCCGTC	TTCATACGTT	TCGTACGTAG
3701	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC	TAACTCCGCC	CATCCCGCC	CTAACTCCGC	CCAGTTCGCG	CCATTCTCCG	CCCCATGGCT	GACTAATTTT
	AGTTAATCAG	TCGTTGGTAT	CAGGGCGGGG	ATTGAGGCGG	GTAGGGCGGG	GATTGAGGCG	GGTCAAGGCG	GGTAAGAGGC	GGGGTACCGA	CTGATTAAAA
3801	TTTTATTTAT	GCAGAGGCCG	AGGCCGCCTC	TGCCTCTGAG	CTATTCCAGA	AGTAGTGAGG	AGGCTTTTTT	GGAGGCCTAG	GCTTTTGCAA	AAAGCTCCCG
	AAAATAAATA	CGTCTCCGGC	TCCGGCGGAG	ACGGAGACTC	GATAAGGTCT	TCATCACTCC	TCCGAAAAAA	CCTCCGGATC	CGAAAACGTT	TTTCGAGGGC
3901	GGAGCTTGTA	TATCCATTTT	CGGATCTGAT	CAAGAGACAG	GATGAGGATC	GTTTCGCATG	ATTGAACAAG	ATGGATTGCA	CGCAGGTTCT	CCGGCCGCTT
	CCTCGAACAT	ATAGGTAAAA	GCCTAGACTA	GTTCTCTGTC	CTACTCCTAG	CAAAGCGTAC	TAACTTGTTT	TACCTAACGT	GCGTCCAAGA	GGCCGGCGGAA
4001	GGGTGGAGAG	GCTATTCCGGC	TATGACTGGG	CACAACAGAC	AATCGGCTGC	TCTGATGCCG	CCGTGTTCCG	GCTGTCAGCG	CAGGGGCGCC	CGGTTCTTTT
	CCCACCTCTC	CGATAAGCCG	ATACTGACCC	GTGTTGTCTG	TTAGCCGACG	AGACTACGGC	GGCACAAGGC	CGACAGTCGC	GTCCCCGCGG	GCCAAGAAAA
4101	TGTCAAGACC	GACCTGTCCG	GTGCCCTGAA	TGAACTGCAG	GACGAGGCAG	CGCGGCTATC	GTGGCTGGCC	ACGACGGGCG	TTCTTGCGC	AGCTGTGCTC
	ACAGTTCTGG	CTGGACAGGC	CACGGGACTT	ACTTGACGTC	CTGCTCCGTC	GCGCCGATAG	CACCGACCGG	TGCTGCCCGC	AAGGAACGCG	TCGACACGAG
4201	GACGTTGTCA	CTGAAGCGGG	AAGGGACTGG	CTGCTATTGG	GCGAAGTGCC	GGGGCAGGAT	CTCCTGTCAT	CTCACCTTGC	TCCTGCCGAG	AAAGTATCCA
	CTGCAACAGT	GACTTCGCCC	TTCCCTGACC	GACGATAACC	CGCTTCACGG	CCCCGTCCTA	GAGGACAGTA	GAGTGGAACG	AGGACGGCTC	TTTCATAGGT
4301	TCATGGCTGA	TGCAATGCGG	CGGCTGCATA	CGCTTGATCC	GGCTACCTGC	CCATTCGACC	ACCAAGCGAA	ACATCGCATC	GAGCGAGCAC	GTACTCGGAT
	AGTACCGACT	ACGTTACGCC	GCCGACGTAT	GCGAACTAGG	CCGATGGACG	GGTAAGCTGG	TGGTTTCGCTT	TGTAGCGTAG	CTCGCTCGTG	CATGAGCCTA
4401	GGAAGCCGGT	CTTGTCGATC	AGGATGATCT	GGACGAAGAG	CATCAGGGGC	TCGCGCCAGC	CGAACTGTTT	GCCAGGCTCA	AGGCGCGCAT	GCCCGACGGC
	CCTTCGGCCA	GAACAGCTAG	TCCTACTAGA	CCTGCTTCTC	GTAGTCCCCG	AGCGCGGTCG	GCTTGACAAG	CGGTCCGAGT	TCCGCGCGTA	CGGGCTGCCG
4501	GAGGATCTCG	TCGTGACCCA	TGGCGATGCC	TGCTTGCCGA	ATATCATGGT	GGAAAATGGC	CGCTTTTCTG	GATTCATCGA	CTGTGGCCCG	CTGGGTGTGG
	CTCCTAGAGC	AGCACTGGGT	ACCGCTACGG	ACGAACGGCT	TATAGTACCA	CCTTTTACCG	GCGAAAAGAC	CTAAGTAGCT	GACACCGGCC	GACCCACACC
4601	CGGACCGCTA	TCAGGACATA	GCGTTGGCTA	CCCGTGATAT	TGCTGAAGAG	CTTGGCGGCG	AATGGGCTGA	CCGCTTCCTC	GTGCTTTACG	GTATCGCCGC
	GCCTGGCGAT	AGTCCTGTAT	CGCAACCGAT	GGGCACTATA	ACGACTTCTC	GAACCGCCGC	TTACCCGACT	GGCGAAGGAG	CACGAAATGC	CATAGCGGGC

HA-mPB

4701	TCCCGATTCTG	CAGCGCATCG	CCTTCTATCG	CCTTCTTGAC	GAGTTCTTCT	GAGCGGGACT	CTGGGGTTCTG	AAATGACCGA	CCAAGCGACG	CCCAACCTGC
	AGGGCTAAGC	GTCGCGTAGC	GGAAGATAGC	GGAAGAACTG	CTCAAGAAGA	CTCGCCCTGA	GACCCCAAGC	TTTACTGGCT	GGTTCGCTGC	GGGTTGGACG
4801	CATCACGAGA	TTTCGATTCC	ACCGCCGCCT	TCTATGAAAG	GTTGGGCTTC	GGAATCGTTT	TCCGGGACGC	CGGCTGGATG	ATCCTCCAGC	GCGGGGATCT
	GTAGTGCTCT	AAAGCTAAGG	TGGCGGCGGA	AGATACTTTC	CAACCCGAAG	CCTTAGCAAA	AGGCCCTGCG	GCCGACCTAC	TAGGAGGTCTG	CGCCCCTAGA
4901	CATGCTGGAG	TTCTTCGCCC	ACCCCAACTT	GTTTATTGCA	GCTTATAATG	GTTACAAATA	AAGCAATAGC	ATCACAAATT	TCACAAATAA	AGCATTTTTT
	GTACGACCTC	AAGAAGCGGG	TGGGGTTGAA	CAAATAACGT	CGAATATTAC	CAATGTTTAT	TTCGTTATCG	TAGTGTTTAA	AGTGTTTATT	TCGTAAAAAA
5001	TCACTGCATT	CTAGTTGTGG	TTTGTCCAAA	CTCATCAATG	TATCTTATCA	TGTCTGTATA	CCGTCGACCT	CTAGCTAGAG	CTTGGCGTAA	TCATGGTCAT
	AGTGACGTAA	GATCAACACC	AAACAGGTTT	GAGTAGTTAC	ATAGAATAGT	ACAGACATAT	GGCAGCTGGA	GATCGATCTC	GAACCGCATT	AGTACCAGTA
5101	AGCTGTTTCC	TGTGTGAAAT	TGTTATCCGC	TCACAATTCC	ACACAACATA	CGAGCCGGAA	GCATAAAGTG	TAAAGCCTGG	GGTGCCTAAT	GAGTGAGCTA
	TCGACAAAGG	ACACACTTTA	ACAATAGGCG	AGTGTTAAGG	TGTGTTGTAT	GCTCGGCCTT	CGTATTTTAC	ATTCGGACC	CCACGGATTA	CTCACTCGAT
5201	ACTCACATTA	ATTGCGTTGC	GCTCACTGCC	CGCTTTCAG	TCGGGAAAAC	TGTCGTGCCA	GCTGCATTAA	TGAATCGGCC	AACGCGCGGG	GAGAGGCGGT
	TGAGTGTAAT	TAACGCAACG	CGAGTGACGG	GCGAAAAGGTC	AGCCCTTTGG	ACAGCACGGT	CGACGTAATT	ACTTAGCCGG	TTGCGCGCCC	CTCTCCGCCA
5301	TTGCGTATTG	GGCGCTCTTC	CGCTTCCTCG	CTCACTGACT	CGCTGCGCTC	GGTCGTTTCGG	CTGCGGCGAG	CGGTATCAGC	TCACTCAAAG	GCGGTAATAC
	AACGCATAAC	CCGCGAGAAG	GCGAAGGAGC	GAGTGACTGA	GCGACGCGAG	CCAGCAAGCC	GACGCCGCTC	GCCATAGTCG	AGTGAGTTTC	CGCCATTATG
5401	GGTTATCCAC	AGAATCAGGG	GATAACGCAG	GAAAGAACAT	GTGAGCAAAA	GGCCAGCAAA	AGGCCAGGAA	CCGTAAAAAG	GCCGCGTTGC	TGGCGTTTTT
	CCAATAGGTG	TCTTAGTCCC	CTATTGCGTC	CTTCTTTGTA	CACTCGTTTT	CCGGTCGTTT	TCCGGTCCTT	GGCATTTTTT	CGGCGCAACG	ACCGCAAAAA
5501	CCATAGGCTC	CGCCCCCTTG	ACGAGCATCA	CAAAAAATCGA	CGCTCAAGTC	AGAGGTGGCG	AAACCCGACA	GGACTATAAA	GATACCAGGC	GTTTCCCCCT
	GGTATCCGAG	GCGGGGGGAC	TGCTCGTAGT	GTTTTTAGCT	GCGAGTTTCA	TCTCCACCGC	TTTGGGCTGT	CCTGATATTT	CTATGGTCCG	CAAAGGGGGA
5601	GGAAGCTCCC	TCGTGCGCTC	TCCTGTTCCG	ACCCTGCCGC	TTACCGGATA	CCTGTCCGCC	TTTCTCCCTT	CGGGAAGCGT	GGCGCTTTCT	CAATGCTCAC
	CCTTCGAGGG	AGCACGCGAG	AGGACAAGGC	TGGGACGGCG	AATGGCCTAT	GGACAGGCGG	AAAGAGGGAA	GCCCTTCGCA	CCGCGAAAGA	GTTACGAGTG
5701	GCTGTAGGTA	TCTCAGTTCTG	GTGTAGGTCG	TTCGCTCCAA	GCTGGGCTGT	GTGCACGAAC	CCCCCGTTCA	GCCCGACCGC	TGCGCCTTAT	CCGGTAACTA
	CGACATCCAT	AGAGTCAAGC	CACATCCAGC	AAGCGAGGTT	CGACCCGACA	CACGTGCTTG	GGGGGCAAGT	CGGGCTGGCG	ACGCGGAATA	GGCCATTGAT
5801	TCGTCTTGAG	TCCAACCCGG	TAAGACACGA	CTTATCGCCA	CTGGCAGCAG	CCACTGGTAA	CAGGATTAGC	AGAGCGAGGT	ATGTAGGCGG	TGCTACAGAG
	AGCAGAACTC	AGGTTGGGGC	ATTCTGTGCT	GAATAGCGGT	GACCGTCGTC	GGTGACCATT	GTCCTAATCG	TCTCGCTCCA	TACATCCGCC	ACGATGTCTC
5901	TTCTTGAAAGT	GGTGGCCTAA	CTACGGCTAC	ACTAGAAGGA	CAGTATTTGG	TATCTGCGCT	CTGCTGAAGC	CAGTTACCTT	CGGAAAAAGA	GTTGGTAGCT
	AAGAACTTCA	CCACCGGATT	GATGCCGATG	TGATCTTCTT	GTCATAAACC	ATAGACGCGA	GACGACTTCG	GTCAATGGAA	GCCTTTTTCT	CAACCATCGA
6001	CTTGATCCGG	CAAACAAACC	ACCGCTGGTA	GCGGTGGTTT	TTTTGTTTGC	AAGCAGCAGA	TTACGCGCAG	AAAAAAAGGA	TCTCAAGAAG	ATCCTTTGAT
	GAAC TAGGCC	GTTTGT TTGG	TGGCGACCAT	CGCCACCAAA	AAAACAAACG	TTCTGTCGCT	AATGCGCGTC	TTTTTTTCTT	AGAGTTCTTC	TAGGAAACTA
6101	CTTTTCTACG	GGGCTGACG	CTCAGTGGA	CGAAAACTCA	CGTTAAGGGA	TTTTGGTCAT	GAGATTATCA	AAAAGGATCT	TCACCTAGAT	CCTTTTAAAT
	GAAAAGATGC	CCCAGACTGC	GAGTCACCTT	GCTTTTGAGT	GCAATTCCCT	AAAACCAGTA	CTCTAATAGT	TTTTCTCTAG	AGTGGAATCT	GGAAAAATTA

HA-mPB

6201	TAAAAATGAA	GTTTTAAATC	AATCTAAAGT	ATATATGAGT	AAACTTGGTC	TGACAGTTAC	CAATGCTTAA	TCAGTGAGGC	ACCTATCTCA	GCGATCTGTC
	ATTTTTACTT	CAAAATTTAG	TTAGATTTCA	TATATACTCA	TTTGAACCAG	ACTGTCAATG	GTTACGAATT	AGTCACTCCG	TGGATAGAGT	CGCTAGACAG
6301	TATTTTCGTT	ATCCATAGTT	GCCTGACTCC	CCGTCTGTGA	GATAACTACG	ATACGGGAGG	GCTTACCATC	TGGCCCCAGT	GCTGCAATGA	TACCGCGAGA
	ATAAAGCAAG	TAGGTATCAA	CGGACTGAGG	GGCAGCACAT	CTATTGATGC	TATGCCCTCC	CGAATGGTAG	ACCGGGGTCA	CGACGTTACT	ATGGCGCTCT
6401	CCCACGCTCA	CCGGCTCCAG	ATTTATCAGC	AATAAACCCAG	CCAGCCGGAA	GGGCCGAGCG	CAGAAGTGGT	CCTGCAACTT	TATCCGCCTC	CATCCAGTCT
	GGGTGCGAGT	GGCCGAGGTC	TAAATAGTCG	TTATTTGGTC	GGTCGGCCTT	CCCGGCTCGC	GTCTTCACCA	GGACGTTGAA	ATAGGCGGAG	GTAGGTCAGA
6501	ATTAATTGTT	GCCGGGAAGC	TAGAGTAAGT	AGTTCGCCAG	TTAATAGTTT	GCGCAACGTT	GTTGCCATTG	CTACAGGCAT	CGTGGTGTCA	CGCTCGTCGT
	TAATTAACAA	CGGCCCTTCG	ATCTCATTCA	TCAAGCGGTC	AATTATCAAA	CGCGTTGCAA	CAACGGTAAC	GATGTCCGTA	GCACCACAGT	GCGAGCAGCA
6601	TTGGTATGCG	TTCATTTCAGC	TCCGGTTCCC	AACGATCAAG	GCGAGTTACA	TGATCCCCCA	TGTTGTGCAA	AAAAGCGGTT	AGCTCCTTCG	GTCTCCGAT
	AACCATACCG	AAGTAAGTCG	AGGCCAAGGG	TTGCTAGTTC	CGCTCAATGT	ACTAGGGGGT	ACAACACGTT	TTTTTCGCCAA	TCGAGGAAGC	CAGGAGGCTA
6701	CGTTGTCAGA	AGTAAGTTGG	CCGCAGTGTT	ATCACTCATG	GTTATGGCAG	CACTGCATAA	TTCTCTTACT	GTCATGCCAT	CCGTAAGATG	CTTTTCTGTG
	GCAACAGTCT	TCATTCAACC	GGCGTCACAA	TAGTGAGTAC	CAATACCGTC	GTGACGTATT	AAGAGAATGA	CAGTACGGTA	GGCATTCTAC	GAAAAGACAC
6801	ACTGGTGAGT	ACTCAACCAA	GTCAATCTGA	GAATAGTGTA	TGCGGCGACC	GAGTTGCTCT	TGCCCGGCGT	CAATACGGGA	TAATACCGCG	CCACATAGCA
	TGACCACTCA	TGAGTTGGTT	CAGTAAGACT	CTTATCACAT	ACGCCGCTGG	CTCAACGAGA	ACGGGCCGCA	GTTATGCCCT	ATTATGGCGC	GGTGTATCGT
6901	GAACTTTAAA	AGTGCTCATC	ATTGGA AAAAC	GTTCTTCGGG	GCGAAAACCTC	TCAAGGATCT	TACCGCTGTT	GAGATCCAGT	TCGATGTAAC	CCACTCGTGC
	CTTGAAATTT	TCACGAGTAG	TAACCTTTTG	CAAGAAGCCC	CGCTTTTGAG	AGTTCCTAGA	ATGGCGACAA	CTCTAGGTCA	AGCTACATTG	GGTGAGCACG
7001	ACCCAACCTGA	TCTTCAGCAT	CTTTTACTTT	CACCAGCGTT	TCTGGGTGAG	CAAAAAACAGG	AAGGC AAAAAT	GCCGCAAAAA	AGGGAATAAG	GGCGACACGG
	TGGGTTGACT	AGAAGTCGTA	GAAAAATGAAA	GTGGTCGCAA	AGACCCACTC	GTTTTTGTCC	TTCCGTTTTA	CGGCGTTTTT	TCCCTTATTC	CCGCTGTGCC
7101	AAATGTTGAA	TACTCATACT	CTTCCTTTTT	CAATATTATT	GAAGCATTTA	TCAGGGTTAT	TGTCTCATGA	GCGGATACAT	ATTTGAATGT	ATTTAGAAAA
	TTTACAACCT	ATGAGTATGA	GAAGGAAAAA	GTTATAATAA	CTTCGTAAAT	AGTCCCAATA	ACAGAGTACT	CGCCTATGTA	TAAACTTACA	TAAATCTTTT
7201	ATAAACAAAT	AGGGGTTCCG	CGCACATTTT	CCCGAAAAGT	GCCACCTGAC	GTC				
	TATTTGTTTA	TCCCAAGGC	GCGTGTAAG	GGGCTTTTCA	CGGTGGACTG	CAG				

Sall
~~~~~