

HA-iPB

General Description

DNA 'HA-iPB'
Local object
Created: 10/01/07 11:35AM
Last Modified: 10/01/07 11:35AM
length: 7253 bp
storage type: Basic
form: Circular

Standard Fields

Original Source Database: GenBank
Modification Date in the Original DB: 12-JAN-2007

Comments

Annotations

Feature Map

CDS (3 total)

iPB

Start: 983 End: 2770
Native PB transposase CDS
Original Location Description:
983..2770

Neo(R)

Start: 3958 End: 4752
Original Location Description:
2151..2945

Amp(R)

Start: 6257 End: 7117 (Complementary)
Original Location Description:
6257..7117

Misc. Feature (3 total)

Kozak seq

Start: 905 End: 917

HA tag

Start: 929 End: 949

bGH polyA

Start: 2825 End: 3056

PolyA Signal (1 total)

SV40 pA

Start: 4926 End: 5056
Original Location Description:
3119..3249

Promoter Eukaryotic (2 total)

CMV promoter

Start: 236 End: 823

HA-iPB

Original Location Description:
232..819

SV40 early promoter

Start: 3553 End: 3922
Original Location Description:
1746..2115

Replication Origin (2 total)


f1 origin

Start: 3119 End: 3547
Original Location Description:
1312..1740

pUC origin

Start: 5439 End: 6112 (Complementary)
Original Location Description:
complement(3632..4305)

Restriction/Methylation Map

EcoRI: 1 site 

N1: 969

NcoI: 5 sites 

N1: 615
N2: 913
N3: 982
N4: 3784
N5: 4519

NotI: 1 site 

N1: 2773

Restriction Fragments

3349: pcDNA3 Kz HA Native PB transposase: NcoI(4519) - NcoI(615)
2802: pcDNA3 Kz HA Native PB transposase: NcoI(982) - NcoI(3784)
735: pcDNA3 Kz HA Native PB transposase: NcoI(3784) - NcoI(4519)
298: pcDNA3 Kz HA Native PB transposase: NcoI(615) - NcoI(913)
69: pcDNA3 Kz HA Native PB transposase: NcoI(913) - NcoI(982)

HA-iPB

1	GACGGATCGG	GAGATCGATC	TCCCGATCCC	CTATGGTCGA	CTCTCAGTAC	AATCTGCTCT	GATGCCGCAT	AGTTAAGCCA	GTATCTGCTC	CCTGCTTGTG	
	CTGCCTAGCC	CTCTAGCTAG	AGGGCTAGGG	GATACCAGCT	GAGAGTCATG	TTAGACGAGA	CTACGGCGTA	TCAATTTCGGT	CATAGACGAG	GGACGAACAC	
101	TGTTGGAGGT	CGCTGAGTAG	TGCGCGAGCA	AAATTTAAGC	TACAACAAGG	CAAGGCTTGA	CCGACAATTG	CATGAAGAAT	CTGCTTAGGG	TTAGGCGTTT	
	ACAACCTCCA	GCGACTCATC	ACGCGCTCGT	TTTAAATTCTG	ATGTTGTTCC	GTTCCGAAC	GGCTGTTAAC	GTACTTCTTA	GACGAATCCC	AATCCGCAAA	
201	TGCGCTGCTT	CGCGATGTAC	GGGCCAGATA	TACGCGTTGA	CATTGATTAT	TGACTAGTTA	TTAATAGTAA	TCAATTACGG	GGTCATTAGT	TCATAGCCCA	
	ACGCGACGAA	GCGCTACATG	CCCGGTCTAT	ATGCGCAACT	GTAACATAA	ACTGATCAAT	AATTATCATT	AGTTAATGCC	CCAGTAATCA	AGTATCGGGT	
301	TATATGGAGT	TCCGCGTTAC	ATAACTTACG	GTAATGGCC	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	GCCCCGCTGG	CATTATGCCC	AGTACATGAC	CTTATGGGAC	TTTCCTACTT	GGCAGTACAT	CTACGTATTA	
	CGGGGGATAA	CTGCGATTAC	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	TTTGTTTTTG	CACCAAAATC	AACGGGACTT	TCCAAAATGT	CGTAACAAC	CCGCCCCATT	GACGCAAATG	GGCGGTAGGC	GTGTACGGTG	
	AGTTACCCTC	AAACAAAACC	GTGGTTTTAG	TTGCCCTGAA	AGGTTTTACA	GCATTGTTGA	GGCGGGGTAA	CTGCGTTTAC	CCGCCATCCG	CACATGCCAC	
801	GGAGGTCTAT	ATAAGCAGAG	CTCTCTGGCT	AACTAGAGAA	CCCACTGCTT	ACTGGCTTAT	CGAAATTAAT	ACGACTCACT	ATAGGGAGAC	CCAAGCTTCT	
	CCTCCAGATA	TATTCTGCTC	GAGAGACCGA	TTGATCTCTT	GGGTGACGAA	TGACCGAATA	GCTTTAATTA	TGCTGAGTGA	TATCCCTCTG	GGTTTCGAAGA	
+2											
901	CGAGGCCGCC	ACCATGGGAT	CCACCATGTA	CGACGTTCCCT	GATTACGCTA	GCCTCCCGAG	ATCTCCTGAA	TTGCGCGCCA	CCATGGGTAG	TTCTTTTAGAC	
	GCTCCGGCGG	TGGTACCCTA	GGTGGTACAT	GCTGCAAGGA	CTAATGCGAT	CGGAGGGCTC	TAGAGGACTT	AAGCGGCGGT	GGTACCCATC	AAGAAATCTG	
+2											
1001	GATGAGCATA	TCCTCTCTGC	TCTTCTGCAA	AGCGATGACG	AGCTTGTGTTG	TGAGGATTCT	GACAGTGAAA	TATCAGATCA	CGTAAGTGAA	GATGACGTCC	
	CTACTCGTAT	AGGAGAGACG	AGAAGACGTT	TCGCTACTGC	TCGAACAACC	ACTCCTAAGA	CTGTCACTTT	ATAGTCTAGT	GCATTCACTT	CTACTGCAGG	
+2											
1101	AGAGCGATAC	AGAAGAAGCG	TTTATAGATG	AGGTACATGA	AGTGCAGCCA	ACGTCAAGCG	GTAGTGAAAT	ATTAGACGAA	CAAAATGTTA	TTGAACAACC	
	TCTCGCTATG	TCTTCTTCGC	AAATATCTAC	TCCATGTACT	TCACGTCGGT	TGCAGTTCGC	CATCACTTTA	TAATCTGCTT	GTTTTACAAT	AACTTGTTGG	

HA-iPB

	+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		AGGTTCTTCA TTGGCTTCTA ACAGAATCTT GACCTTGCCA CAGAGGACTA TTAGAGGTAA GAATAAACAT TGTTGGTCAA CTTCAAAGTC CACGAGGCGT TCCAAGAAGT AACCGAAGAT TGTCTTAGAA CTGGAACGGT GTCTCCTGAT AATCTCCATT CTTATTTGTA ACAACCAGTT GAAGTTTCAG GTGCTCCGCA
	+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		AGCCGAGTCT CTGCACTGAA CATTGTCAGA TCTCAAAGAG GTCCGACGCG TATGTGCCGC AATATATATG ACCCACTTTT ATGCTTCAAA CTATTTTTTA TCGGCTCAGA GACGTGACTT GTAACAGTCT AGAGTTTCTC CAGGCTGCGC ATACACGGCG TTATATATAC TGGGTGAAAA TACGAAGTTT GATAAAAAAT
	+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		CTGATGAGAT AATTTTCGGAA ATTGTAAAT GGACAAATGC TGAGATATCA TTGAAACGTC GGGAATCTAT GACAGGTGCT ACATTTTCGTG ACACGAATGA GACTACTCTA TTAAAGCCTT TAACATTTTA CCTGTTTACG ACTCTATAGT AACTTTGCAG CCCTTAGATA CTGTCCACGA TGTAAGCAC TGTGCTTACT
	+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		AGATGAAATC TATGCTTTCT TTGGTATTCT GGTAATGACA GCAGTGAGAA AAGATAACCA CATGTCCACA GATGACCTCT TTGATCGATC TTTGTCAATG TCTACTTTAG ATACGAAAGA AACCATAAGA CCATTACTGT CGTCACTCTT TTCTATTGGT GTACAGGTGT CTACTGGAGA AACTAGCTAG AAACAGTTAC
	+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		GTGTACGTCT CTGTAATGAG TCGTGATCGT TTTGATTTTT TGATACGATG TCTTAGAATG GATGACAAAA GTATACGGCC CACTTTCGA GAAAACGATG CACATGCAGA GACATTACTC AGCACTAGCA AAATAAAAA ACTATGCTAC AGAATCTTAC CTACTGTTTT CATATGCCGG GTGTGAAGCT CTTTGTCTAC
	+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		TATTTACTCC TGTTAGAAAA ATATGGGATC TCTTTATCCA TCAGTGCATA CAAAATTACA CTCCAGGGGC TCATTTGACC ATAGATGAAC AGTTACTTGG ATAAATGAGG ACAATCTTTT TATACCCTAG AGAAATAGGT AGTCACGTAT GTTTTAATGT GAGGTCCCCG AGTAAACTGG TATCTACTTG TCAATGAACC
	+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		TTTTAGAGGA CGGTGTCCGT TTAGGATGTA TATCCCAAAC AAGCCAAGTA AGTATGGAAT AAAAATCCTC ATGATGTGTG ACAGTGGTAC GAAGTATATG AAAATCTCCT GCCACAGGCA AATCCTACAT ATAGGGTTTG TTCGGTTCAT TCATACCTTA TTTTATAGGAG TACTACACAC TGTCACCATG CTTTCATATAC
	+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		ATAAATGGAA TGCCTTATTT GGGAAGAGGA ACACAGACCA ACGGAGTACC ACTCGGTGAA TACTACGTGA AGGAGTTATC AAAGCCTGTG CACGGTAGTT TATTTACCTT ACGGAATAAA CCCTTCTCCT TGTGTCTGGT TGCCTCATGG TGAGCCACTT ATGATGCACT TCCTCAATAG TTTCGGACAC GTGCCATCAA
	+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		GTCGTAATAT TACGTGTGAC AATTGGTTCA CCTCAATCCC TTTGGCAAAA AACTTACTAC AAGAACCGTA TAAGTTAACC ATTGTGGGAA CCGTGCGATC CAGCATTATA ATGCACACTG TTAACCAAGT GGAGTTAGGG AAACCGTTTT TTGAATGATG TTCTTGGCAT ATTCAATTGG TAACACCCCTT GGCACGCTAG
	+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		AAACAAACGC GAGATACCGG AAGTACTGAA AAACAGTCGC TCCAGGCCAG TGGGAACATC GATGTTTTGT TTTGACGGAC CCCTTACTCT CGTCTCATAT TTTGTTTGCG CTCTATGGCC TTCATGACTT TTTGTCAGCG AGGTCCGGTC ACCCTTGTAG CTACAAAACA AAAGTGCCTG GGGAATGAGA GCAGAGTATA

HA-iPB

+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	AAACCGAAGC	CAGCTAAGAT	GGTATACTTA	TTATCATCTT	GTGATGAGGA	TGCTTCTATC	AACGAAAGTA	CCGGTAAACC	GCAAATGGTT	ATGTATTATA	TTTGGCTTCG	GTCGATTCTA	CCATATGAAT	AATAGTAGAA	CACTACTCCT	ACGAAGATAG	TTGCTTTCAT	GGCCATTTGG	CGTTTACCAA	TACATAATAT														
+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	ATCAAACATA	AGGCGGAGTG	GACACGCTAG	ACCAAATGTG	TTCTGTGATG	ACCTGCAGTA	GGAAGACGAA	TAGGTGGCCT	ATGGCATTAT	TGTACGGAAT	TAGTTTGATT	TCCGCCTCAC	CTGTGCGATC	TGGTTTACAC	AAGACACTAC	TGGACGTCAT	CCTTCTGCTT	ATCCACCGGA	TACCGTAATA	ACATGCCTTA														
+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	GATAAACATT	GCCTGCATAA	ATTCTTTTAT	TATATACAGC	CATAATGTCA	GTAGCAAGGG	AGAAAAGGTT	CAAAGTCGCA	AAAAATTTAT	GAGAAACCTT	CTATTTGTAA	CGGACGTATT	TAAGAAAATA	ATATATGTCG	GTATTACAGT	CATCGTTCCC	TCTTTTCCAA	GTTTCAGCGT	TTTTTAAATA	CTCTTTGGAA														
+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	TGACGTCATC	GTTTATGCGT	AAGCGTTTAG	AAGCTCCTAC	TTTGAAGAGA	TATTTGCGCG	ATAATATCTC	TAATATTTTG	CCAAATGAAG	ATGTACTCGG	ACTGCAGTAG	CAAATACGCA	TTCGCAAATC	TTCGAGGATG	AAACTTCTCT	ATAAACCGCG	TATTATAGAG	ATTATAAAAC	GGTTTACTTC														
+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	TGCCTGGTAC	ATCAGATGAC	AGTACTGAAG	AGCCAGTAAT	GAAAAAACGT	ACTTACTGTA	CTTACTGCCC	CTCTAAAATA	AGGCGAAAGG	CAAATGCATC	ACGGACCATG	TAGTCTACTG	TCATGACTTC	TCGGTCATTA	CTTTTTTTGCA	TGAATGACAT	GAATGACGGG	GAGATTTTAT	TCCGCTTTCC	GTTTACGTAG														
+2	Ser	Cys	Lys	Lys	Cys	Lys	Lys	Val	Ile	Cys	Arg	Glu	His	Asn	Ile	Asp	Met	Cys	Gln	Ser	Cys	Phe	***	***	NotI	Ala	Ala							
2701	GTGCAAAAAA	TGCAAAAAAG	TTATTTGTCT	AGAGCATAAT	ATTGATATGT	GCCAAAGTTG	TTTCTGATGA	GCGGCCGCTC	GAGCATGCAT	CTAGAGGGCC	CACGTTTTTTT	ACGTTTTTTTC	AATAAACAGC	TCTCGTATTA	TAACTATACA	CGGTTTCAAC	AAAGACTACT	CGCCGGCGAG	CTCGTACGTA	GATCTCCCGG														
2801	CTATTCTATA	GTGTCACCTA	AATGCTAGAG	CTCGCTGATC	AGCCTCGACT	GTGCCTTCTA	GTTGCCAGCC	ATCTGTTGTT	TGCCCCTCCC	CCGTGCCTTC	GATAAGATAT	CACAGTGGAT	TTACGATCTC	GAGCGACTAG	TCGGAGCTGA	CACGGAAGAT	CAACGGTCGG	TAGACAACAA	ACGGGGAGGG	GGCACGGAAG														
2901	CTTGACCCTG	GAAGGTGCCA	CTCCCACTGT	CCTTTCCTAA	TAAAATGAGG	AAATTGCATC	GCATTGTCTG	AGTAGGTGTC	ATTCTATTCT	GGGGGGTGGG	GAACCTGGGAC	CTTCCACGGT	GAGGGTGACA	GGAAAGGATT	ATTTTACTCC	TTTAACGTAG	CGTAACAGAC	TCATCCACAG	TAAGATAAGA	CCCCCACACC														
3001	GTGGGGCAGG	ACAGCAAGGG	GGAGGATTGG	GAAGACAATA	GCAGGCATGC	TGGGGATGCG	GTGGGCTCTA	TGGCTTCTGA	GGCGGAAAAG	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														
3201	GCCCGCTCCT	TTCGCTTTCT	TCCCTTCCTT	TCTCGCCACG	TTCGCGGGCT	TTCCCCGTCA	AGCTCTAAAT	CGGGGCATCC	CTTTAGGGTT	CCGATTTAGT	CGGGCGAGGA	AAGCGAAAAG	AGGGAAGGAA	AGAGCGGTGC	AAGCGGCCGA	AAGGGGCAGT	TCGAGATTTA	GCCCCGTAGG	GAAATCCCCA	GGCTAAATCA														
3301	GCTTTACGGC	ACCTCGACCC	CAAAAACTT	GATTAGGGTG	ATGGTTCACG	TAGTGGGGCA	TCGCCCTGAT	AGACGGTTTTT	TCGCCCTTTG	ACGTTGGAGT	CGAAATGCCG	TGGAGCTGGG	GTTTTTTTGAA	CTAATCCCAC	TACCAAGTGC	ATCACCCGGT	AGCGGGACTA	TCTGCCAAAA	AGCGGGAAAC	TGCAACCTCA														

HA-iPB

3401	CCACGTTCTT	TAATAGTGGA	CTCTTGTTCC	AAACTGGAAC	AACACTCAAC	CCTATCTCGG	TCTATTCTTT	TGATTTATAA	GGGATTTTGG	GGATTTCTGGC
	GGTGCAAGAA	ATTATCACCT	GAGAACAAGG	TTTGACCTTG	TTGTGAGTTG	GGATAGAGCC	AGATAAGAAA	ACTAAATATT	CCCTAAAACC	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	TGGAAAAGTCC	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC
	CGTCCGTCTT	CATACGTTTC	GTACGTAGAG	TTAATCAGTC	GTTGGTCCAC	ACCTTTCAGG	GGTCCGAGGG	GTCGTCCGTC	TTCATACGTT	TCGTACGTAG
3701	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC	TAACTCCGCC	CATCCCGCCC	CTAACTCCGC	CCAGTTCCGC	CCATTCTCCG	CCCCATGGCT	ACTAATTTT
	AGTTAATCAG	TCGTTGGTAT	CAGGGCGGGG	ATTGAGGCGG	GTAGGGCGGG	GATTGAGGCG	GGTCAAGGCG	GGTAAGAGGC	GGGGTACCGA	CTGATTAAAA
3801	TTTTATTTAT	GCAGAGGCCG	AGGCCGCCTC	TGCCTCTGAG	CTATTCCAGA	AGTAGTGAGG	AGGCTTTTTT	GGAGGCCTAG	GCTTTTGCAA	AAAGCTCCCC
	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	GGCCGGCGAA
4001	GGGTGGAGAG	GCTATTCTGC	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	TTCTTTGCGC	AGCTGTGCTC
	ACAGTTCTGG	CTGGACAGGC	CACGGGACTT	ACTTGACGTC	CTGCTCCGTC	GCGCCGATAG	CACCGACCGG	TGCTGCCCCG	AAGGAACGCG	TCGACACGAG
4201	GACGTTGTCA	CTGAAGCGGG	AAGGGACTGG	CTGCTATTGG	GCGAAGTGCC	GGGGCAGGAT	CTCCTGTGTC	CTCACCTTGC	TCCTGCCGAG	AAAGTATCCA
	CTGCAACAGT	GACTTCGCCC	TTCCCTGACC	GACGATAACC	CGCTTCACGG	CCCCGTCCCTA	GAGGACAGTA	GAGTGGAACG	AGGACGGCTC	TTTCATAGGT
4301	TCATGGCTGA	TGCAATGCGG	CGGCTGCATA	CGCTTGATCC	GGCTACCTGC	CCATTGACCC	ACCAAGCGAA	ACATCGCATC	GAGCGAGCAC	GTACTCGGAT
	AGTACCGACT	ACGTTACGCC	GCCGACGTAT	GCGAACTAGG	CCGATGGACG	GGTAAGCTGG	TGGTTCGCTT	TGTAGCGTAG	CTCGCTCGTG	CATGAGCCTA
4401	GGAAGCCGGT	CTTGTCGATC	AGGATGATCT	GGACGAAGAG	CATCAGGGGC	TCGCGCCAGC	CGAACTGTTC	GCCAGGCTCA	AGGCGCGCAT	GCCCGACGGC
	CCTTCGGCCA	GAACAGCTAG	TCCTACTAGA	CCTGCTTCTC	GTAGTCCCCG	AGCGCGGTCG	GCTTGACAAG	CGGTCCGAGT	TCCGCGCGTA	CGGGCTGCCC
4501	GAGGATCTCG	TCGTGACCCA	TGGCGATGCC	TGCTTGCCGA	ATATCATGGT	GGAAAATGGC	CGCTTTTCTG	GATTCATCGA	CTGTGGCCGG	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	CATAGCGGCG
4701	TCCCGATTCT	CAGCGCATCG	CCTTCTATCG	CCTTCTTGAC	GAGTTCTTCT	GAGCGGGACT	CTGGGGTTCT	AAATGACCGA	CCAAGCGACG	CCCAACCTGC
	AGGGCTAAGC	GTCGCGTAGC	GGAAGATAGC	GGAAGAACTG	CTCAAGAAGA	CTCGCCCTGA	GACCCCAAGC	TTTACTGGCT	GGTTCGCTGC	GGGTGAGACG

HA-iPB

4801	CATCACGAGA	TTTCGATTCC	ACCGCCGCCT	TCTATGAAAG	GTTGGGCTTC	GGAATCGTTT	TCCGGGACGC	CGGCTGGATG	ATCCTCCAGC	GCGGGGATCT
	GTAGTGCTCT	AAAGCTAAGG	TGGCGGCGGA	AGATACTTTC	CAACCCGAAG	CCTTAGCAAA	AGGCCCTGCG	GCCGACCTAC	TAGGAGGTCG	CGCCCCTAGA
4901	CATGCTGGAG	TTCTTCGCCC	ACCCCAACTT	GTTTATTGCA	GCTTATAATG	GTTACAAATA	AAGCAATAGC	ATCACAAATT	TCACAAATAA	AGCATTTTTTT
	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	ATTTCTGGACC	CCACGGATTA	CTCACTCGAT
5201	ACTCACATTA	ATTGCGTTGC	GCTCACTGCC	CGCTTTCAG	TCGGGAAACC	TGTCGTGCCA	GCTGCATTAA	TGAATCGGCC	AACGCGCGGG	GAGAGGCGGT
	TGAGTGTAAT	TAACGCAACG	CGAGTGACGG	GCGAAAGGTC	AGCCCTTTGG	ACAGCACGGT	CGACGTAATT	ACTTAGCCGG	TTGCGCGCCC	CTCTCCGCCA
5301	TTGCGTATTG	GGCGCTCTTC	CGCTTCCTCG	CTCACTGACT	CGCTGCGCTC	GGTCGTTCCG	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	CTTTCTTGTA	CACTCGTTTT	CCGGTCGTTT	TCCGGTCCTT	GGCATTTTTT	CGGCGCAACG	ACCGCAAAAA
5501	CCATAGGCTC	CGCCCCCTG	ACGAGCATCA	CAAAAATCGA	CGCTCAAGTC	AGAGGTGGCG	AAACCCGACA	GGACTATAAA	GATACCAGGC	GTTTCCCCCT
	GGTATCCGAG	GCGGGGGGAC	TGCTCGTAGT	GTTTTTAGCT	GCGAGTTCAG	TCTCCACCGC	TTTGGGCTGT	CCTGATATTT	CTATGGTCCG	CAAAGGGGGA
5601	GGAAGCTCCC	TCGTGCGCTC	TCCTGTTCCTG	ACCCTGCCGC	TTACCGGATA	CCTGTCCGCC	TTTCTCCCTT	CGGGAAGCGT	GGCGCTTTCT	CAATGCTCAC
	CCTTCGAGGG	AGCACGCGAG	AGGACAAGGC	TGGGACGGCG	AATGGCCTAT	GGACAGGCGG	AAAGAGGGAA	GCCCTTCGCA	CCGCGAAAAG	GTTACGAGTG
5701	GCTGTAGGTA	TCTCAGTTCG	GTGTAGGTCG	TTGCTCCAA	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	AGGTTGGGCC	ATTCTGTGCT	GAATAGCGGT	GACCGTCGTC	GGTGACCATT	GTCTAATTCG	TCTCGCTCCA	TACATCCGCC	ACGATGTCTC
5901	TTCTTGAAGT	GGTGGCCTAA	CTACGGCTAC	ACTAGAAGGA	CAGTATTTGG	TATCTGCGCT	CTGCTGAAGC	CAGTTACCTT	CGGAAAAAGA	GTTGGTAGCT
	AAGAACTTCA	CCACCGGATT	GATGCCGATG	TGATCTTCCT	GTCATAAAAC	ATAGACGCGA	GACGACTTCG	GTCAATGGAA	GCCTTTTTCT	CAACCATCGA
6001	CTTGATCCGG	CAAACAAACC	ACCGCTGGTA	GCGGTGGTTT	TTTTGTTTGC	AAGCAGCAGA	TTACGCGCAG	AAAAAAAGGA	TCTCAAGAAG	ATCCTTTGAT
	GAAC TAGGCC	GTTTGTTTGG	TGGCGACCAT	CGCCACCAAA	AAAAACAAAC	TTGTCGTCCT	AATGCGCGTC	TTTTTTTCCT	AGAGTTCTTC	TAGGAAACTA
6101	CTTTTCTACG	GGGTCTGACG	CTCAGTGGA	CGAAAACTCA	CGTTAAGGGA	TTTTGGTCAT	GAGATTATCA	AAAAGGATCT	TCACCTAGAT	CCTTTTAAAT
	GAAAAGATGC	CCCAGACTGC	GAGTCACCTT	GCTTTTGAGT	GCAATTCCCT	AAAACCAGTA	CTCTAATAGT	TTTTCTTAGA	AGTGGATCTA	GGAAAATTTA
6201	TAAAAATGAA	GTTTTAAATC	AATCTAAAAGT	ATATATGAGT	AAACTTGGTC	TGACAGTTAC	CAATGCTTAA	TCAGTGAGGC	ACCTATCTCA	GCGATCTGTC
	ATTTTTACTT	CAAAATTTAG	TTAGATTTCA	TATATACTCA	TTTGAACCAG	ACTGTCAATG	GTTACGAATT	AGTCACTCCG	TGGATAGAGT	CGCTAGACAG

HA-iPB

6301	TATTTTCGTTT	ATCCATAGTT	GCCTGACTCC	CCGTCGTGTA	GATAACTACG	ATACGGGAGG	GCTTACCATC	TGGCCCCAGT	GCTGCAATGA	TACCGCGAGA
	ATAAAGCAAG	TAGGTATCAA	CGGACTGAGG	GGCAGCACAT	CTATTGATGC	TATGCCCTCC	CGAATGGTAG	ACCGGGGTCA	CGACGTTACT	ATGGCGCTCT
6401	CCCACGCTCA	CCGGCTCCAG	ATTTATCAGC	AATAAACCAG	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	ATCTCATTTCA	TCAAGCGGTC	AATTATCAAA	CGCGTTGCAA	CAACGGTAAC	GATGTCCGTA	GCACCACAGT	GCGAGCAGCA
6601	TTGGTATGGC	TTCATTTCAGC	TCCGGTTCCC	AACGATCAAG	GCGAGTTACA	TGATCCCCCA	TGTTGTGCAA	AAAAGCGGTT	AGCTCCTTCG	GTCTCCGAT
	AACCATAACCG	AAGTAAGTCG	AGGCCAAGGG	TTGCTAGTTC	CGCTCAATGT	ACTAGGGGGT	ACAACACGTT	TTTTCGCCAA	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	GTCATTCTGA	GAATAGTGTA	TGCGGCGACC	GAGTTGCTCT	TGCCCCGGCGT	CAATACGGGA	TAATACCGCG	CCACATAGCA
	TGACCACTCA	TGAGTTGGTT	CAGTAAGACT	CTTATCACAT	ACGCCGCTGG	CTCAACGAGA	ACGGGCCGCA	GTTATGCCCT	ATTATGGCGC	GGTGTATCGT
6901	GAACTTTAAA	AGTGCTCATC	ATTGGAAAAAC	GTTCTTCGGG	GCGAAAACTC	TCAAGGATCT	TACCGCTGTT	GAGATCCAGT	TCGATGTAAC	CCACTCGTGC
	CTTGAAATTT	TCACGAGTAG	TAACCTTTTG	CAAGAAGCCC	CGCTTTTGAG	AGTTCCTAGA	ATGGCGACAA	CTCTAGGTCA	AGCTACATTG	GGTGAGCACG
7001	ACCCAACCTGA	TCTTCAGCAT	CTTTTACTTT	CACCAGCGTT	TCTGGGTGAG	CAAAAACAGG	AAGGCAAAAT	GCCGCAAAAA	AGGGAATAAG	GGCGACACGG
	TGGGTTGACT	AGAAGTCGTA	GAAAATGAAA	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	AGGGGTTCG	CGCACATTC	CCCGAAAAGT	GCCACCTGAC	GTC				
	TATTTGTTTA	TCCCCAAGGC	GCGTGTAAG	GGGCTTTTCA	CGGTGGACTG	CAG				