

pHZ5 (NLS) attR\_2

		MAL61/62 promoter				
1	GTCTTGGATG	TAATTCTTAT	TGTTATACTG	AATACGCTAA	AACCACTCAC	
	CAGAACCTAC	ATTAAGAATA	ACAATATGAC	TTATGCGATT	TTGGTGAGTG	
MAL61/62 promoter						
51	AACAAGTATG	GAGTATATTG	TGTCTCTTTA	TATGCTGAGT	ACTTATGCAA	
	TTGTTTCATAC	CTCATATAAC	ACAGAGAAAT	ATACGACTCA	TGAATACGTT	
MAL61/62 promoter						
101	TATGCGCTCA	CTCAGGATGA	AATGTACACA	GCCGAAAGTA	TATTGAAAGC	
	ATACGCGAGT	GAGTCCTACT	TTACATGTGT	CGGCTTTCAT	ATAACTTTCG	
MAL61/62 promoter						
151	TGCCTCTGTG	GAAACTTCTA	TCTAATCTTG	TCTCCAGATG	TAGACTATGA	
	ACGGAGACAC	CTTTGAAGAT	AGATTAGAAC	AGAGGTCTAC	ATCTGATACT	
MAL61/62 promoter						
201	GGCCTGAAGA	AGTCTTTAGG	CACCTGTTGG	AGAGTATAAG	GAGACTGCTA	
	CCGGACTTCT	TCAGAAATCC	GTGGACAACC	TCTCATATTC	CTCTGACGAT	
MAL61/62 promoter						
251	CAACAACGTC	TTCCCACAAA	AAATTATCTG	GAGGCCGGTA	TGATACCTGC	
	GTTGTTGCAG	AAGGGTGTTT	TTTAATAGAC	CTCCGGCCAT	ACTATGGACG	
MAL61/62 promoter						
301	ACAAAGGTTA	AGTTACACAT	GAAAAAGAAA	CTGACATAAC	TTTGATCTCT	
	TGTTTCCAAT	TCAATGTGTA	CTTTTTCTTT	GACTGTATTG	AAACTAGAGA	
MAL61/62 promoter						
351	GAAAATATCT	TTTCCCCTGA	GTAGCTTCAC	TGCTTGGATA	CCAATAGGAA	
	CTTTTATAGA	AAAGGGGACT	CATCGAAGTG	ACGAACCTAT	GGTTATCCTT	
MAL61/62 promoter						
401	TAGACCTTGG	CTATAGTAAG	TTGCATCTGT	ACCGTAGAGA	TTCTTGCAAC	
	ATCTGGAACC	GATATCATTG	AACGTAGACA	TGGCATCTCT	AAGAACGTTG	
MAL61/62 promoter						
451	TCGCTTAAAC	TCTCGCTTTT	AGATAATATT	TCTCCTTATT	GCGCGCTTCG	
	AGCGAATTTG	AGAGCGAAAA	TCTATTATAA	AGAGGAATAA	CGCGCGAAGC	
MAL61/62 promoter						
501	TTGAAAATTT	CGCTAAACAC	GGGGTTTAAG	TTAAAGTTTA	CAGGATTTAT	
	AACTTTTAAA	GCGATTTGTG	CCCCAAATTC	AATTTCAAAT	GTCCTAAATA	
MAL61/62 promoter						
551	CCGGAAATTT	TCGCGGACCC	CACACAATTA	AGAATTGGCT	CGAAGAGTGA	
	GGCCTTTAAA	AGCGCCTGGG	GTGTGTTAAT	TCTTAACCGA	GCTTCTCACT	
MAL61/62 promoter						
601	TAACGCATAC	TTTTCTTTTC	TTTTTTTAGT	TCCTAGCGTA	CCTAACGTAG	
	ATTGCGTATG	AAAAGAAAAG	AAAAAAATCA	AGGATCGCAT	GGATTGCATC	

pHZ5 (NLS) attR\_2

MAL61/62 promoter

651 GTAACATGAT TTGGATCGTG GGATGATACA AACACGTAAGATGACTACT  
 CATTGTACTA AACCTAGCAC CCTACTATGT TTGTTGCATT CTACTGATGA

MAL61/62 promoter

701 TCCTTCCTCA ATTCTTCTTG CAGCATCATT TTCTTGAGGC GCTCTGGGCA  
 AGGAAGGAGT TAAGAAGAAC GTCGTAGTAA AAGAACTCCG CGAGACCCGT

MAL61/62 promoter

751 AGGTATAAAA AGTTCCATTA ATACGTCTCT AAAAAATTAA ATCATCCATC  
 TCCATATTTT TCAAGGTAAT TATGCAGAGA TTTTTTAATT TAGTAGGTAG

MAL61/62 promoter

801 TCTTAAGCAG TTTTTTTGAT AATCTCAAAT GTACATCAGT CAAGCGTAAC  
 AGAATTCGTC AAAAAAATAA TTAGAGTTTA CATGTAGTCA GTTCGCATTG

MAL61/62 promoter

+3 TAAATTACAT AAATGTACAT CAGTCAAGCG TAACTAAATT ACATAAATGA  
 ATTTAATGTA TTTACATGTA GTCAGTTCGC ATTGATTAA TGTATTTACT

Agel  
M T

Agel

+3 T G K A L T A R Q Q E V F D L I R

901 CCGGTAAAGC GTTAACGGCC AGGCAACAAG AGGTGTTTGA TCTCATCCGT  
 GGCCATTTTCG CAATTGCCGG TCCGTTGTTC TCCACAAACT AGAGTAGGCA

Agel

+3 D H I S Q T G M P P T R A E I A Q

951 GATCACATCA GCCAGACAGG TATGCCGCCG ACGCGTGCGG AAATCGCGCA  
 CTAGTGTAGT CGGTCTGTCC ATACGGCGGC TGCGCACGCC TTTAGCGCGT

Agel

+3 Q R L G F R S P N A A E E H L K A L

1001 GCGTTTGGGG TTCCGTTCCC CAAACGCGGC TGAAGAACAT CTGAAGGCGC  
 CGCAAACCCC AAGGCAAGGG GTTTGCGCCG ACTTCTTGTA GACTTCCGCG

Agel

+3 L A R K G V I E I V S G A S R G I

1051 TGGCACGCAA AGGCGTTATT GAAATTGTTT CCGGCGCATC ACGCGGGATT  
 ACCGTGCGTT TCCGCAATAA CTTTAACAAA GGCCGCGTAG TGCGCCCTAA

Agel

+3 R L L Q E E E E G L P L V G R V A

1101 CGTCTGTTGC AGGAAGAGGA AGAAGGGTTG CCGCTGGTAG GTCGTGTGGC  
 GCAGACAACG TCCTTCTCCT TCTTCCAAC GGCGACCATC CAGCACACCG

Agel

+3 A A G E P L L A Q Q H I E G H Y Q V

1151 TGCCGGTGAA CCACTTCTGG CGCAACAGCA TATTGAAGGT CATTATCAGG  
 ACGGCCACTT GGTGAAGACC GCGTTGTCGT ATAACCTCCA GTAATAGTCC

Agel

+3 V D P S L F K P N A D F L L R V S

1201 TCGATCCTTC CTTATTCAAG CCGAATGCTG ATTTCTGCT GCGCGTCAGC  
 AGCTAGGAAG GAATAAGTTC GGCTTACGAC TAAAGGACGA CGCGCAGTCG

Agel

+3 G M S M K D I G I M D G D L L A V

1251 GGGATGTCGA TGAAAGATAT CGGCATTATG GATGGTGACT TGCTGGCAGT  
 CCCTACAGCT ACTTTCTATA GCCGTAATAC CTACCACTGA ACGACCGTCA

pHZ5(NLS)attR\_2

```

+3 · V H K T Q D V R N G Q V V V A R I D ·
1301 GCATAAAACT CAGGATGTAC GTAACGGTCA GGTCGTTGTC GCACGTATTG
    CGTATTTTGA GTCCTACATG CATTGCCAGT CCAGCAACAG CGTGCATAAC
.....
+3 · D D E V T V K R L K K Q G N K V E
1351 ATGACGAAGT TACCGTTAAG CGCCTGAAAA AACAGGGCAA TAAAGTCGAA
    TACTGCTTCA ATGGCAATTC GCGGACTTTT TTGTCCCATT ATTTTCAGCTT
.....
+3 · L L P E N S E F K P I V V D L R Q ·
1401 CTGTTGCCAG AAAATAGCGA GTTTAAACCA ATTGTCGTAG ATCTTCGTCA
    GACAACGGTC TTTTATCGCT CAAATTTGGT TAACAGCATC TAGAAGCAGT
.....
+3 · Q Q S F T I E G L A V G V I R N G D ·
1451 GCAGAGCTTC ACCATTGAAG GGCTGGCGGT TGGGGTTATT CGCAACGGCG
    CGTCTCGAAG TGGTAACTTC CCGACCGCCA ACCCCAATAA GCGTTGCCGC
.....
+3 · D W L E L A P K K K R K V E F
1501 ACTGGCTGGA ATTGGCCCC AAGAAAAAGA GAAAGGTGGA ATTCCCAGGG
    TGACCGACCT TAACCGGGGG TTCTTTTCT CTTTCCACCT TAAGGGCCCC
.....
1551 ATCCGTCGAC CATGGCGGCC GCTCGAGTCG ACCTGCAGCC AAGCTAATTC
    TAGGCAGCTG GTACCGCCGG CGAGCTCAGC TGGACGTCGG TTCGATTAAG
.....
1601 CGGGACAAGT TTGTACAAAA AAGCTGAACG AGAAACGTAA AATGATATAA
    GCCCTGTTCA AACATGTTTT TTCGACTTGC TCTTTGCATT TTAATAATTT
.....
1651 ATATCAATAT ATTAAATTAG ATTTTGCATA AAAAACAGAC TACATAATAC
    TATAGTTATA TAATTTAATC TAAAACGTAT TTTTGTCTG ATGTATTATG
.....
1701 TGTA AACAC AACATATCCA GTCACTATGG CGGCCGCATT AGGCACCCCA
    ACATTTTGTG TTGTATAGGT CAGTGATACC GCCGGCGTAA TCCGTGGGGT
.....
1751 GGCTTTACAC TTTATGCTTC CGGCTCGTAT AATGTGTGGA TTTTGAGTTA
    CCGAAATGTG AAATACGAAG GCCGAGCATA TTACACACCT AAAACTCAAT
.....
1801 GGATCCGTCG AGATTTTCAG GAGCTAAGGA AGCTAAAATG GAGAAAAAAA
    CCTAGGCAGC TCTAAAAGTC CTCGATTCCCT TCGATTTTAC CTCTTTTTTT
.....
1851 TCACTGGATA TACCACCGTT GATATATCCC AATGGCATCG TAAAGAACAT
    AGTGACCTAT ATGGTGGCAA CTATATAGGG TTACCGTAGC ATTTCTTGTA
.....
1901 TTTGAGGCAT TTCAGTCAGT TGCTCAATGT ACCTATAACC AGACCGTTCA
    AACTCCGTA AAGTCAGTCA ACGAGTTACA TGGATATTGG TCTGGCAAGT
.....
1951 GCTGGATATT ACGGCCTTTT TAAAGACCGT AAAGAAAAAT AAGCACAAGT
    CGACCTATAA TGCCGAAAA ATTTCTGGCA TTTCTTTTTA TTCGTGTTCA
.....
2001 TTTATCCGGC CTTTATTAC ATTCTTGCCC GCCTGATGAA TGCTCATCCG
    AAATAGGCCG GAAATAAGTG TAAGAACGGG CGGACTACTT ACGAGTAGGC
.....

```

XmaI  
EcoRI  
BamHI

XhoI

BamHI

NotI

NotI

BamHI

pHZ5 (NLS) attR\_2

EcoRI					
2051	GAATTCCGTA	TGGCAATGAA	AGACGGTGAG	CTGGTGATAT	GGGATAGTGT
	CTTAAGGCAT	ACCGTTACTT	TCTGCCACTC	GACCACTATA	CCCTATCACA
2101	TCACCCTTGT	TACACCGTTT	TCCATGAGCA	AACTGAAACG	TTTTCATCGC
	AGTGGGAACA	ATGTGGCAA	AGGTACTCGT	TTGACTTTGC	AAAAGTAGCG
2151	TCTGGAGTGA	ATACCACGAC	GATTTCCGGC	AGTTTCTACA	CATATATTCTG
	AGACCTCACT	TATGGTGCTG	CTAAAGGCCG	TCAAAGATGT	GTATATAAGC
2201	CAAGATGTGG	CGTGTTACGG	TGAAAACCTG	GCCTATTTCC	CTAAAGGGTT
	GTTCTACACC	GCACAATGCC	ACTTTTGGAC	CGGATAAAGG	GATTTCCCAA
2251	TATTGAGAAT	ATGTTTTTCG	TCTCAGCCAA	TCCCTGGGTG	AGTTTCACCA
	ATAACTCTTA	TACAAAAGC	AGAGTCGGTT	AGGGACCCAC	TCAAAGTGGT
2301	GTTTTGATTT	AAACGTGGCC	AATATGGACA	ACTTCTTCGC	CCCCGTTTTT
	CAAACATAAA	TTTGCACCGG	TTATACCTGT	TGAAGAAGCG	GGGGCAAAAG
2351	ACCATGGGCA	AATATTATAC	GCAAGGCGAC	AAGGTGCTGA	TGCCGCTGGC
	TGGTACCCGT	TTATAATATG	CGTTCGCTG	TCCACGACT	ACGGCGACCG
2401	GATTCAGGTT	CATCATGCCG	TCTGTGATGG	CTTCCATGTC	GGCAGAATGC
	CTAAGTCCAA	GTAGTACGGC	AGACACTACC	GAAGGTACAG	CCGTCTTACG
2451	TTAATGAATT	ACAACAGTAC	TGCGATGAGT	GGCAGGGCGG	GGCGTAAAGA
	AATTACTTAA	TGTTGTCATG	ACGCTACTCA	CCGTCCC GCC	CCGCATTTCT
BamHI					
2501	TCTGGATCCG	GCTTACTAAA	AGCCAGATAA	CAGTATGCGT	ATTTGCGCGC
	AGACCTAGGC	CGAATGATTT	TCGGTCTATT	GTCATACGCA	TAAACGCGCG
2551	TGATTTTTGC	GGTATAAGAA	TATATACTGA	TATGTATACC	CGAAGTATGT
	ACTAAAACG	CCATATTCTT	ATATATGACT	ATACATATGG	GCTTCATACA
2601	CAAAAAGAGG	TGTGCTATGA	AGCAGCGTAT	TACAGTGACA	GTTGACAGCG
	GTTTTTCTCC	ACACGATACT	TCGTGCGATA	ATGTCACTGT	CAACTGTCGC
2651	ACAGCTATCA	GTTGCTCAAG	GCATATATGA	TGTCAATATC	TCCGGTCTGG
	TGTCGATAGT	CAACGAGTTC	CGTATATACT	ACAGTTATAG	AGGCCAGACC
2701	TAAGCACAAC	CATGCAGAAT	GAAGCCCGTC	GTCTGCGTGC	CGAACGCTGG
	ATTCGTGTTG	GTACGTCTTA	CTTCGGGCAG	CAGACGCACG	GCTTGCAGCC
2751	AAAGCGGAAA	ATCAGGAAGG	GATGGCTGAG	GTCGCCCGGT	TTATTGAAAT
	TTTCGCCTTT	TAGTCCTTCC	CTACCGACTC	CAGCGGGCCA	AATAACTTTA
2801	GAACGGCTCT	TTTGCTGACG	AGAACAGGGA	CTGGTGAAAT	GCAGTTTAAG
	CTTGCCGAGA	AAACGACTGC	TCTTGTCCTT	GACCACTTTA	CGTCAAATTC
2851	GTTTACACCT	ATAAAAAGAGA	GAGCCGTTAT	CGTCTGTTTG	TGGATGTACA
	CAAATGTGGA	TATTTTCTCT	CTCGGCAATA	GCAGACAAAC	ACCTACATGT
XmaI					
2901	GAGTGATATT	ATTGACACGC	CCGGGCGACG	GATGGTGATC	CCCCTGGCCA
	CTCACTATAA	TAACGTGTCG	GGCCCGCTGC	CTACCACTAG	GGGGACCGGT
2951	GTGCACGTCT	GCTGTCAGAT	AAAGTCTCCC	GTGAACTTTA	CCCAGTGGTG
	CACGTGCAGA	CGACAGTCTA	TTTCAGAGGG	CACTTGAAAT	GGGCCACCAC

pHZ5 (NLS) attR\_2

```

3001 CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA TGGCCAGTGT
      GTATAGCCCC TACTTTCGAC CGCGTACTAC TGGTGGCTAT ACCGGTCACA
.....
3051 GCCGGTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA
      CGGCCAGAGG CAATAGCCCC TTCTTCACCG ACTAGAGTCG GTGGCGCTTT
.....
3101 ATGACATCAA AAACGCCATT AACCTGATGT TCTGGGGAAT ATAAATGTCA
      TACTGTAGTT TTTGCGGTAA TTGGACTACA AGACCCCTTA TATTTACAGT
.....
3151 GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATAGT GACTGGATAT
      CCGAGGGAAT ATGTGTCGGT CAGACGTCCA GCTGGTATCA CTGACCTATA
.....
3201 GTTGTGTTTT ACAGTATTAT GTAGTCTGTT TTTTATGCAA AATCTAATTT
      CAACACAAA TGTCATAATA CATCAGACAA AAAATACGTT TTAGATTAAA
.....
3251 AATATATTGA TATTTATATC ATTTTACGTT TCTCGTTCAG CTTTCTTGTA
      TTATATAACT ATAAATATAG TAAAATGCAA AGAGCAAGTC GAAAGAACAT
.....
3301 CAAAGTGGTG TTATAAAAAA AATAAGTGTA TACAAATTTT AAAGTGACTC
      GTTTCACCAC AATATTTTTT TTATTCACAT ATGTTTAAAA TTTCACTGAG
.....
3351 TTAGGTTTTA AAACGAAAAT TCTTGTTCTT GAGTAACTCT TTCCTGTAGG
      AATCCAAAAT TTTGCTTTTA AGAACAAAGAA CTCATTGAGA AAGGACATCC
.....
3401 TCAGGTTGCT TTCTCAGGTA TAGCATGAGG TCGCTCTTAT TGACCACACC
      AGTCCAACGA AAGAGTCCAT ATCGTACTCC AGCGAGAATA ACTGGTGTGG
.....
3451 TCTACCGGCA TGCCGAGCAA ATGCCTGCAA ATCGCTCCCC ATTTCACCCA
      AGATGGCCGT ACGGCTCGTT TACGGACGTT TAGCGAGGGG TAAAGTGGGT
.....
3501 ATTGTAGATA TGCTAACTCC AGCAATGAGT TGATGAATCT CGGTGTGTAT
      TAACATCTAT ACGATTGAGG TCGTTACTCA ACTACTTAGA GCCACACATA
.....
3551 TTTATGTCCCT CAGAGGACAA CACCTGTTGT AATCGTTCTT CCACACGGAT
      AAATACAGGA GTCTCCTGTT GTGGACAACA TTAGCAAGAA GGTGTGCCTA
.....
3601 CGATCCACAG GACGGGTGTG GTCGCCATGA TCGCGTAGTC GATAGTGGCT
      GCTAGGTGTC CTGCCACAC CAGCGGTACT AGCGCATCAG CTATCACCGA
.....
3651 CCAAGTAGCG AAGCGAGCAG GACTGGGCGG CGGCCAAAGC GGTCCGACAG
      GGTTCATCGC TTCGCTCGTC CTGACCCGCC GCCGGTTTCG CCAGCCTGTC
.....
3701 TGCTCCGAGA ACGGGTGCGC ATAGAAATTG CATCAACGCA TATAGCGCTA
      ACGAGGCTCT TGCCCACGCG TATCTTTAAC GTAGTTGCGT ATATCGCGAT
.....
3751 GCAGCACGCC ATAGTGACTG GCGATGCTGT CGGAATGGAC GATATCCCGC
      CGTCGTGCGG TATCACTGAC CGCTACGACA GCCTTACCTG CTATAGGGCG
.....
3801 AAGAGGCCCG GCAGTACCGG CATAACCAAG CCTATGCCTA CAGCATCCAG
      TTCTCCGGGC CGTCATGGCC GTATTGGTTC GGATACGGAT GTCGTAGGTC
.....
3851 GGTGACGGTG CCGAGGATGA CGATGAGCGC ATTGTTAGAT TTCATACACG
      CCACTGCCAC GGCTCCTACT GCTACTCGCG TAACAATCTA AAGTATGTGC
.....
3901 GTGCCTGACT GCGTTAGCAA TTTAACTGTG ATAACTACC GCATTAAAGC
      CACGGACTGA CGCAATCGTT AAATTGACAC TATTTGATGG CGTAATTTTCG
.....
3951 TAGCTTTGAA GAAAAATGCG CCTTATTCAA TCTTTGCTAT AAAAAATGGC
      ATCGAAACTT CTTTTTACGC GGAATAAGTT AGAAACGATA TTTTTTACCG
.....
4001 CCAAAATCTC ACATTGGAAG ACATTTGATG ACCTCATTTT TTTCAATGAA
      GGTTTTAGAG TGTAACCTTC TGTAACACTAC TGGAGTAAAG AAAGTTACTT
.....

```

pHZ5 (NLS) attR\_2

4051	GGGCCTAACG	GAGTTGACTA	ATGTTGTGGG	AAATTGGAGC	GATAAGCGTG
	CCCGGATTGC	CTCAACTGAT	TACAACACCC	TTTAACCTCG	CTATTTCGCAC
4101	CTTCTGCCGT	GGCCAGGACA	ACGTATACTC	ATCAGATAAC	AGCAATACCT
	GAAGACGGCA	CCGGTCCTGT	TGCATATGAG	TAGTCTATTG	TCGTTATGGA
		Spel ~~~~~			
4151	GATCACTACT	TCGCACTAGT	TTCTCGGTAC	TATGCATATG	ATCCAATATC
	CTAGTGATGA	AGCGTGATCA	AAGAGCCATG	ATACGTATAC	TAGGTTATAG
4201	AAAGGAAATG	ATAGCATTGA	AGGATGAGAC	TAATCCAATT	GAGGAGTGGC
	TTTCCTTTAC	TATCGTAACT	TCCTACTCTG	ATTAGGTTAA	CTCCTCACCG
4251	AGCATATAGA	ACAGCTAAAG	GGTAGTGCTG	AAGGAAGCAT	ACGATACCCC
	TCGTATATCT	TGTCGATTTC	CCATCACGAC	TTCCTTCGTA	TGCTATGGGG
4301	GCATGGAATG	GGATAATATC	ACAGGAGGTA	CTAGACTACC	TTTCATCCTA
	CGTACCTTAC	CCTATTATAG	TGTCCTCCAT	GATCTGATGG	AAAGTAGGAT
4351	CATAAATAGA	CGCATATAAG	TACGCATTTA	AGCATAAACA	CGCACTATGC
	GTATTTATCT	GCGTATATTC	ATGCGTAAAT	TCGTATTTGT	GCGTGATACG
4401	CGTTCTTCTC	ATGTATATAT	ATATACAGGC	AACACGCAGA	TATAGGTGCG
	GCAAGAAGAG	TACATATATA	TATATGTCCG	TTGTGCGTCT	ATATCCACGC
4451	ACGTGAACAG	TGAGCTGTAT	GTGCGCAGCT	CGCGTTGCAT	TTTCGGAAGC
	TGCACTTGTC	ACTCGACATA	CACGCGTCGA	GCGCAACGTA	AAAGCCTTCG
				XbaI ~~~~~	
4501	GCTCGTTTTT	GGAAACGCTT	TGAAGTTCCT	ATTCCGAAGT	TCCTATTCTC
	CGAGCAAAAG	CCTTTGCGAA	ACTTCAAGGA	TAAGGCTTCA	AGGATAAGAG
	XbaI ~~~~~				
4551	TAGAAAGTAT	AGGAACTTCA	GAGCGCTTTT	GAAAACCAAA	AGCGCTCTGA
	ATCTTTCATA	TCCTTGAAGT	CTCGCGAAAA	CTTTTGTTTT	TCGCGAGACT
4601	AGACGCACTT	TCAAAAAACC	AAAAACGCAC	CGGACTGTAA	CGAGCTACTA
	TCTGCGTGAA	AGTTTTTTGG	TTTTTGGCGTG	GCCTGACATT	GCTCGATGAT
4651	AAATATTGCG	AATACCGCTT	CCACAAACAT	TGCTCAAAAG	TATCTCTTTG
	TTTATAACGC	TTATGGCGAA	GGTGTTTGTA	ACGAGTTTTT	ATAGAGAAAC
4701	CTATATATCT	CTGTGCTATA	TCCCTATATA	ACCTACCCAT	CCACCTTTCG
	GATATATAGA	GACACGATAT	AGGGATATAT	TGGATGGGTA	GGTGGAAAGC
4751	CTCCTTGAAC	TTGCATCTAA	ACTCGACCTC	TACATTTTTT	ATGTTTATCT
	GAGGAACTTG	AACGTAGATT	TGAGCTGGAG	ATGTAAAAAA	TACAAATAGA
4801	CTAGTATTAC	TCTTTAGACA	AAAAAATTGT	AGTAAGAACT	ATTCATAGAG
	GATCATAATG	AGAAATCTGT	TTTTTTAACA	TCATTCTTGA	TAAGTATCTC
4851	TGAATCGAAA	ACAATACGAA	AATGTAAACA	TTTCCTATAC	GTAGTATATA
	ACTTAGCTTT	TGTTATGCTT	TTACATTTGT	AAAGGATATG	CATCATATAT
4901	GAGACAAAAT	AGAAGAAACC	GTCATAATT	TTCTGACCAA	TGAAGAATCA
	CTCTGTTTTA	TCTTCTTTGG	CAAGTATTAA	AAGACTGGTT	ACTTCTTAGT
4951	TCAACGCTAT	CACTTTCTGT	TCACAAAGTA	TGCGCAATCC	ACATCGGTAT
	AGTTGCGATA	GTGAAAGACA	AGTGTTTCAT	ACGCGTTAGG	TGTAGCCATA

pHZ5 (NLS) attR\_2

```

5001  AGAATATAAT  CGGGGATGCC  TTTATCTTGA  AAAAATGCAC  CCGCAGCTTC
      TCTTATATTA  GCCCCTACGG  AAATAGAACT  TTTTACGTG  GCGTTCGAAG
.....
5051  GCTAGTAATC  AGTAAACGCG  GGAAGTGGAG  TCAGGCTTTT  TTTATGGAAG
      CGATCATTAG  TCATTTGCGC  CCTTCACCTC  AGTCCGAAAA  AAATACCTTC
.....
5101  AGAAAAATAGA  CACCAAAGTA  GCCTTCTTCT  AACCTTAACG  GACCTACAGT
      TCTTTTATCT  GTGGTTTCAT  CGGAAGAAGA  TTGGAATTGC  CTGGATGTCA
.....
5151  GCAAAAAGTT  ATCAAGAGAC  TGCATTATAG  AGCGCACAAA  GGAGAAAAAA
      CGTTTTTCAA  TAGTTCTCTG  ACGTAATATC  TCGCGTGTTT  CCTCTTTTTT
.....
5201  AGTAATCTAA  GATGCTTTGT  TAGAAAAATA  GCGCTCTCGG  GATGCATTTT
      TCATTAGATT  CTACGAAACA  ATCTTTTTAT  CGCGAGAGCC  CTACGTAAAA
.....
5251  TGTAGAACAA  AAAAGAAGTA  TAGATTCTTT  GTTGGTAAAA  TAGCGCTCTC
      ACATCTTGTT  TTTTCTTCAT  ATCTAAGAAA  CAACCATTTT  ATCGCGAGAG
.....
5301  GCGTTGCATT  TCTGTTCTGT  AAAAATGCAG  CTCAGATTCT  TTGTTTGAAA
      CGCAACGTAA  AGACAAGACA  TTTTACGTG  GAGTCTAAGA  AACAACTTT
.....
5351  AATTAGCGCT  CTCGCGTTGC  ATTTTGTGTT  TACAAAAATG  AAGCACAGAT
      TTAATCGCGA  GAGCGCAACG  TAAAAACAAA  ATGTTTTTAC  TTCGTGTCTA
.....
5401  TCTTCGTTGG  TAAAATAGCG  CTTTCGCGTT  GCATTTCTGT  TCTGTAAAAA
      AGAAGCAACC  ATTTTATCGC  GAAAGCGCAA  CGTAAAGACA  AGACATTTTT
.....
5451  TGCAGCTCAG  ATTCTTTGTT  TGAAAAATTA  GCGCTCTCGC  GTTGCATTTT
      ACGTCGAGTC  TAAGAAACAA  ACTTTTTAAT  CGCGAGAGCG  CAACGTAAAA
.....
5501  TGTTCTACAA  AATGAAGCAC  AGATGCTTCG  TTAACAAAGA  TATGCTATTG
      ACAAGATGTT  TTACTTCGTG  TCTACGAAGC  AATTGTTTCT  ATACGATAAC
.....
5551  AAGTGCAAGA  TGGAAACGCA  GAAAATGAAC  CGGGGATGCG  ACGTGCAAGA
      TTCACGTTCT  ACCTTTGCGT  CTTTTACTTG  GCCCCTACGC  TGCACGTTCT
.....
5601  TTACCTATGC  AATAGATGCA  ATAGTTTCTC  CAGGAACCGA  AATACATACA
      AATGGATACG  TTATCTACGT  TATCAAAGAG  GTCCTTGGCT  TTATGTATGT
.....
5651  TTGTCTTCCG  TAAAGCGCTA  GACTATATAT  TATTATACAG  GTTCAAATAT
      AACAGAAGGC  ATTTGCGGAT  CTGATATATA  ATAATATGTC  CAAGTTTATA
.....
5701  ACTATCTGTT  TCAGGGAAAA  CTCCCAGGTT  CGGATGTTCA  AAATTCAATG
      TGATAGACAA  AGTCCCTTTT  GAGGGTCCAA  GCCTACAAGT  TTTAAGTTAC
.....
5751  ATGGGTAACA  AGTACGATCG  TAAATCTGTA  AAACAGTTTG  TCGGATATTA
      TACCCATTGT  TCATGCTAGC  ATTTAGACAT  TTTGTCAAAC  AGCCTATAAT
.....
5801  GGCTGTATCT  CCTCAAAGCG  TATTCGAATA  TCATTGAGAA  GCTGCAGCAG
      CCGACATAGA  GGAGTTTCGC  ATAAGCTTAT  AGTAACTCTT  CGACGTCGTC
.....
5851  GCGTGAAGTT  AGACGACAAC  TTCTCTCTGG  AAACGCATAC  CGATATTCAG
      CGCACTTCAA  TCTGCTGTTG  AAGAGAGACC  TTTGCGTATG  GCTATAAGTC
.....
5901  GCTGCTGCAA  AGGCACAGGC  TAGTGCCCGT  GCGAGTGCAT  CCGGTACCAC
      CGACGACGTT  TCCGTGTCCG  ATCACGGGCA  CGCTCACGTA  GGCCATGGTG
.....
5951  CCCAGATGCT  GTAGTAGCTT  CTGGTAGCAC  TGCAATGAGC  CATGCTTATC
      GGGTCTACGA  CATCATCGAA  GACCATCGTG  ACGTTACTCG  GTACGAATAG
.....
6001  AAGAAAACAC  AGGTTTTGGT  ACTCGTCCCA  TATATCTTGA  CATGCAAGCC
      TTCTTTTGTG  TCCAAAACCA  TGAGCAGGGT  ATATAGAACT  GTACGTTCCG
.....

```

pHZ5 (NLS) attR\_2

```

6051 ACTACACCAA CAGACCCTAG GGTTTTGGAT ACGATGTTGA AGTTTTTATAC
      TGATGTGGTT GTCTGGGATC CAAAACCTA TGCTACAAC TCAAAATATG
.....
6101 GGGACTTTAT GGTAATCCTC ATTCCAACAC TCACTCTTAC GGTGGGAAA
      CCCTGAAATA CCATTAGGAG TAAGGTGTG AGTGAGAATG CCAACCCTT
.....
6151 CAAATACTGC TGTGGAAAAT GCTAGAGCTC ACGTAGCAA GATGATCAAT
      GTTTATGACG ACACCTTTTA CGATCTCGAG TGCATCGTT CTACTAGTTA
.....
6201 GCCGACCCCA AGGAAATAAT ATTCACTTCG GGAGCGACCG AATCTAATAA
      CGGCTGGGGT TCCTTTATTA TAAGTGAAG CCTCGCTGGC TTAGATTATT
.....
6251 TATGGTTCTT AAGGGTGTCC CAAGATTTTA TAAGAAGACT AAGAAACACA
      ATACCAAGAA TTCCACAGG GTTCTAAAAT ATTCTTCTGA TTCTTTGTGT
.....
6301 TCATCACCAC TAGAACGGAA CACAAGTGTG TCTTGGAAAGC CGCACGGGCC
      AGTAGTGGTG ATCTTGCCTT GTGTTACAC AGAACCTTCG GCGTGCCCGG
.....
6351 ATGATGAAGG AGGGATTTGA AGTCACTTTC CTAAATGTGG ACGATCAAGG
      TACTACTTCC TCCCTAAACT TCAGTGAAAG GATTTACACC TGCTAGTTCC
.....
6401 TCTTATCGAT TTGAAGGAAT TGGAAAGATGC CATTAGACCA GATACCTGTC
      AGAATAGCTA AACTTCCTTA ACCTTCTACG GTAATCTGGT CTATGGACAG
.....
6451 TCGTCTCTGT GATGGCTGTC AATAATGAAA TCGGTGTCAT TCAACCTATT
      AGCAGAGACA CTACCGACAG TTATTACTTT AGCCACAGTA AGTTGGATAA
.....
6501 AAAGAAATTG GAGCAATTTG TAGAAAGAAT AAGATCCTCG GGGACACCAA
      TTTCTTTAAC CTCGTTAAAC ATCTTTCTTA TTCTAGGAGC CCCTGTGGTT
.....
6551 ATATGGCGAT CTCGGCCTTT TCGTTTCTTG GAGCTGGGAC ATGTTTGCCA
      TATACCGCTA GAGCCGAAA AGCAAAGAAC CTCGACCCTG TACAAACGGT
.....
6601 TCGATCCATC TACCACCAGA ACGGCCGTTA GATCTGCTGC CACCGTTGTT
      AGCTAGGTAG ATGGTGGTCT TGCCGGCAAT CTAGACGACG GTGGCAACAA
.....
6651 TCCACCGAAG AAACCACCGT TGCCGTAACC ACCACGACGG TTGTTGCTAA
      AGGTGGCTTC TTTGGTGGCA ACGGCATTGG TGGTGCTGCC AACACGATT
.....
6701 AGAAGCTGCC ACCGCCACGG CCACCGTTGT AGCCGCCGTT GTTGTATTG
      TCTTCGACGG TGGCGGTGCC GGTGGCAACA TCGGCGGCAA CAACAATAAC
.....
6751 TAGTTGCTAC TGTTATTTCT GGCACCTCTT GGTTTTCTC TTAAGTGAGG
      ATCAACGATG ACAATAAAGA CCGTGAAGAA CAAAAGGAG AATTCACTCC
.....
6801 AGGAACATAA CCATTCTCGT TGTTGTCGTT GATGCTTAAA TTTTGCCTT
      TCCTTGATTT GGTAAGAGCA ACAACAGCAA CTACGAATTT AAAACGTGAA
.....
6851 GTTCGCTCAG TTCAGCCATA ATATGAAATG CTTTCTTGT TGTTCTTACG
      CAAGCGAGTC AAGTCGGTAT TATACTTTAC GAAAAGAACA ACAAGAATGC
.....
6901 GAATACCACT TGCCACCTAT CACCACAAC TACTTTTCC CGTTCCTCCA
      CTTATGGTGA ACGGTGGATA GTGGTGTGTA TTGAAAAGG GCAAGGAGGT
.....
6951 TCTCTTTTAT ATTTTTTTTC TCGATCGAGT TCAAGAGAAA AAAAAAGAAA
      AGAGAAAATA TAAAAAAAAG AGCTAGCTCA AGTTCTCTTT TTTTTCTTT
.....
7001 AAGCAAAAAG AAAAAAGGAA AGCGCGCCTC GTTCAGAATG ACACGTATAG
      TTCGTTTTTC TTTTTTCTT TCGCGCGGAG CAAGTCTTAC TGTGCATATC
.....
7051 AATGATGCAT TACCTTGTC TCTTCAGTAT CATACTGTT GTATACATAC
      TTACTIONGTA ATGGAACAGT AGAAGTCATA GTATGACAAG CATATGTATG
.....

```



pHZ5 (NLS) attR\_2

7101	TTACTGACAT	TCATAGGTAT	ACATATATAC	ACATGTATAT	ATATCGTATG
	AATGACTGTA	AGTATCCATA	TGTATATATG	TGTACATATA	TATAGCATA
7151	CTGCAGCTTT	AAATAATCGG	TGTCACTACA	TAAGAACACC	TTTGGTGGAG
	GACGTCGAAA	TTTATTAGCC	ACAGTGATGT	ATTCTTGTGG	AAACCACCTC
7201	GGAACATCGT	TGGTACCATT	GGGCGAGGTG	GCTTCTCTTA	TGGCAACCGC
	CCTTGTAGCA	ACCATGGTAA	CCCGCTCCAC	CGAAGAGAAT	ACCGTTGGCG
7251	AAGAGCCTTG	AACGCACTCT	CACTACGGTG	ATGATCATT	TTGCCTCGCA
	TTCTCGGAAC	TTGCGTGAGA	GTGATGCCAC	TACTAGTAAG	AACGGAGCGT
				HindIII	
7301	GACAATCAAC	GTGGAGGGTA	ATTCTGCTAG	CCTCTGCAA	GCTTTCAAGA
	CTGTTAGTTG	CACCTCCCAT	TAAGACGATC	GGAGACGTTT	CGAAAGTTCT
7351	AAATGCGGGA	TCATCTCGCA	AGAGAGATCT	CCTACTTTCT	CCCTTTGCAA
	TTTACGCCCT	AGTAGAGCGT	TCTCTCTAGA	GGATGAAAGA	GGGAAACGTT
7401	ACCAAGTTCG	ACAACCTGCGT	ACGGCCTGTT	CGAAAGATCT	ACCACCGCTC
	TGGTTCAAGC	TGTTGACGCA	TGCCGGACAA	GCTTTCTAGA	TGGTGGCGAG
7451	TGGAAAGTGC	CTCATCCAAA	GGCGCAAATC	CTGATCCAAA	CCTTTTTACT
	ACCTTTCACG	GAGTAGGTTT	CCGCGTTTAG	GACTAGGTTT	GGAAAAATGA
			HindIII		
7501	CCACGCGCCA	GTAGGGCCTC	TTTAAAAGCT	TGACCGAGAG	CAATCCCGCA
	GGTGCGCGGT	CATCCCGGAG	AAATTTTCGA	ACTGGCTCTC	GTTAGGGCGT
7551	GTCTTCAGTG	GTGTGATGGT	CGTCTATGTG	TAAGTCACCA	ATGCACTCAA
	CAGAAGTCAC	CACACTACCA	GCAGATACAC	ATTCAGTGGT	TACGTGAGTT
7601	CGATTAGCGA	CCAGCCGGAA	TGCTTGGCCA	GAGCATGTAT	CATATGGTCC
	GCTAATCGCT	GGTCGGCCTT	ACGAACCGGT	CTCGTACATA	GTATACCAGG
7651	AGAAACCCTA	TACCTGTGTG	GACGTTAATC	ACTTGCGATT	GTGTGGCCTG
	TCTTTGGGAT	ATGGACACAC	CTGCAATTAG	TGAACGCTAA	CACACCGGAC
7701	TTCTGCTACT	GCTTCTGCCT	CTTTTTCTGG	GAAGATCGAG	TGCTCTATCG
	AAGACGATGA	CGAAGACGGA	GAAAAAGACC	CTTCTAGCTC	ACGAGATAGC
7751	CTAGGGGACC	ACCCTTTAAA	GAGATCGCAA	TCTGAATCTT	GGTTTCATTT
	GATCCCCTGG	TGGGAAATTT	CTCTAGCGTT	AGACTTAGAA	CCAAAGTAAA
7801	GTAATACGCT	TTACTAGGGC	TTTCTGCTCT	GTCATCTTTG	CCTTCGTTTA
	CATTATGCGA	AATGATCCCG	AAAGACGAGA	CAGTAGAAAC	GGAAGCAAAT
7851	TCTTGCCTGC	TCATTTTTTA	GTATATTCTT	CGAAGAAATC	ACATTACTTT
	AGAACGGACG	AGTAAAAAAT	CATATAAGAA	GCTTCTTTAG	TGTAATGAAA
7901	ATATAATGTA	TAATTCATTA	TGTGATAATG	CCAATCGCTA	AGAAAAAATA
	TATATTACAT	ATTAAGTAAT	ACACTATTAC	GGTTAGCGAT	TCTTTTTTTT
7951	AGAGTCATCC	GCTAGGTGGA	AAAAAAAAAA	TGAAAATCAT	TACCGAGGCA
	TCTCAGTAGG	CGATCCACCT	TTTTTTTTTT	ACTTTTAGTA	ATGGCTCCGT
8001	TAAAAAATA	TAGAGTGATC	TAGAGGAGGC	CAAGAGTAAT	AGAAAAAGAA
	ATTTTTTTAT	ATCTCACATG	ATCTCCTCCG	GTTCTCATT	TCTTTTTCTT
8051	AATTGCGGGA	AAGGACTGTG	TTATGACTTC	CCTGACTAAT	GCCGTGTTCA
	TTAACGCCCT	TTCCTGACAC	AATACTGAAG	GGACTGATTA	CGGCACAAGT

pHZ5 (NLS) attR\_2

8101	AACGATACCT TTGCTATGGA	GGCAGTGA CCGTCAC	GGACTG GGATCG	CGCTC CGCGAG	ACCAAGCTCT TGGTTCGAGA	TAAAACGAGA ATTTTGCTCT
8151	ATTAAGAAAA TAATTCTTTT	AGTCGTCATC TCAGCAGTAG	TTTCGATAAG AAAGCTATTC	TTTTTCCCAC AAAAAGGGTG	AGCAAAGCAA TCGTTTCGTT	
8201	TAGTAGAAAA ATCATCTTTT	ACAATGGGAA TGTTACCCTT	ACGTTGAATG TGCAACTTAC	AAGACAAAGC TTCTGTTTCG	GTCGTGGTTT CAGCACCAAA	
8251	AAAAGGAAAT TTTTCCTTTA	ACGCTCACGT TGCGAGTGCA	ACATGCTAGG TGTACGATCC	GAACAGGACC CTTGTCCTGG	GTGCAGCGGA CACGTCGCTT	
8301	TCTAATGAAT AGATTACTTA	CCATTTGTTA GGTAAACAAT	GTTAATAGTT CAATTATCAA	TAAATGTTTT ATTTACAAAA	TATCGGAAGA ATAGCCTTCT	
8351	GGTTTTGTCA CCAAAACAGT	TCACATCAGC AGTGTAGTCG	AATGTTCTTC TTACAAGAAG	TTGGTCTCGA AACCAGAGCT	TGTAGTATAC ACATCATATG	
8401	GTATAAATTA CATATTTAAT	TTACCTGATA AATGGACTAT	CTTCATCTCT GAAGTAGAGA	AAGTCTCATT TTCAGAGTAA	GCCTTTGTGC CGGAAACACG	
8451	CAAAAAATCT GTTTTTTTGA	GTTTCTAAAT CAAAGATTTA	TTCTCTTCAT AAGAGAAGTA	TTGTAGACTT AACATCTGAA	AATTATACTG TTAATATGAC	
8501	ATCGTTGATC TAGCAACTAG	TACTATCAGT ATGATAGTCA	AAGTAAGCCT TTCATTCGGA	TTAAAAAAA AATTTTTTTT	AAAAAAAAAA TTTTTTTTTT	
8551	AAAAAAAAAAA TTTTTTTTTTT	AAACCTGTAA TTTGGACATT	CAATAGCAAT GTTATCGTTA	ACCCCAAATA TGGGGTTTAT	CCTAATGTAG GGATTACATC	
8601	TTCCAGCAAG AAGGTCGTT	CAAGCTAAAA GTTTCGATTTT	AGTAAAGCAA TCATTTTCGTT	CAACATAACT GTTGTATTGA	CACCCCTGCA GTGGGGACGT	
		Xmal ~~~~~				
8651	TCTGCAGCTT AGACGTCGAA	TTGCCCGGGC AACGGGCCCG	AGCCTGCTCT TCGGACGAGA	GCCTGTGTTT CGGACACAAA	TCTTTAATTG AGAAATTAAC	
8701	AGCAGTAGAC TCGTCATCTG	CATTTAGCAG GTAAATCGTC	TTGCATGAAT AACGTACTTA	AGCTGCAGCG TCGACGTCGC	TCACATCGGA AGTGTAGCCT	
8751	TAATAATGAT ATTATTACTA	GGCAGCCATT CCGTCGGTAA	GTAGAAGTGC CATCTTCACG	CTTTTGCATT GAAAACGTAA	TCTAGTCTCT AGATCAGAGA	
8801	TTCTCGGTCT AAGAGCCAGA	AGCTAGTTTT TCGATCAAAA	ACTACATCGC TGATGTAGCG	GAAGATAGAA CTTCTATCTT	TCTTAGATCA AGAATCTAGT	
8851	CACTGCCTTT GTGACGGAAA	GCTGAGCTGG CGACTCGACC	ATCAATAGAG TAGTTATCTC	TAACAAAAGA ATTGTTTTCT	GTGGTAAGGC CACCATTCCG	
8901	CTCGTTAAAG GAGCAATTTT	GACAAGGACC CTGTTCTCTG	TGAGCGGAAG ACTCGCCTTC	TGTATCGTAC ACATAGCATG	AGTAGACGGA TCATCTGCCT	
		SpeI ~~~~~				
8951	GTATACTAGT CATATGATCA	ATAGTCTATA TATCAGATAT	GTCGGTGGAA CAGGCACCTT	TTAATTCTTG AATTAAGAAC	AAGACGAAAG TTCTGCTTTC	
9001	GGCCTCGTGA CCGGAGCACT	TACGCCTATT ATGCGGATAA	TTTATAGGTT AAATATCCAA	AATGTCATGA TTACAGTACT	TAATAATGGT ATTATTACCA	
9051	TTCTTAGACG AAGAATCTGC	TCAGGTGGCA AGTCCACCGT	CTTTTCGGGG GAAAAGCCCC	AAATGTGCGC TTTACACGCG	GGAACCCCTA CCTTGGGGAT	

pHZ5 (NLS) attR\_2

```

9101  TTTGTTTATT  TTTCTAAATA  CATTCAAATA  TGTATCCGCT  CATGAGACAA
      AAACAAATAA  AAAGATTTAT  GTAAGTTTAT  ACATAGGCGA  GTACTCTGTT
.....
9151  TAACCCTGAT  AAATGCTTCA  ATAATATTGA  AAAAGGAAGA  GTATGAGTAT
      ATTGGGACTA  TTTACGAAGT  TATTATAACT  TTTTCCTTCT  CATACTCATA
.....
9201  TCAACATTTT  CGTGTCGCCC  TTATTCCCTT  TTTTGCGGCA  TTTTGCCTTC
      AGTTGTAAAG  GCACAGCGGG  AATAAGGGAA  AAAACGCCGT  AAAACGGAAG
.....
9251  CTGTTTTTTG  TCACCCAGAA  ACGCTGGTGA  AAGTAAAAGA  TGCTGAAGAT
      GACAAAAACG  AGTGGGTCTT  TGCGACCACT  TTCATTTTCT  ACGACTTCTA
.....
9301  CAGTTGGGTG  CACGAGTGGG  TTACATCGAA  CTGGATCTCA  ACAGCGGTAA
      GTC AACCCAC  GTGCTCACCC  AATGTAGCTT  GACCTAGAGT  TGTCGCCATT
.....
9351  GATCCTTGAG  AGTTTTTCGCC  CCGAAGAACG  TTTTCCAATG  ATGAGCACTT
      CTAGGAACTC  TCAAAGCGG  GGCTTCTTGC  AAAAGGTTAC  TACTCGTGAA
.....
9401  TTAAAGTTCT  GCTATGTGGC  GCGGTATTAT  CCCGTGTTGA  CGCCGGGCAA
      AATTTCAAGA  CGATACACCG  CGCCATAATA  GGGCACAACT  GCGGCCCGTT
.....
9451  GAGCAACTCG  GTCGCCGCAT  ACACTATTCT  CAGAATGACT  TGGTTGAGTA
      CTCGTTGAGC  CAGCGGCGTA  TGTGATAAGA  GTCCTACTGA  ACCAACTCAT
.....
9501  CTCACCAGTC  ACAGAAAAGC  ATCTTACGGA  TGGCATGACA  GTAAGAGAAT
      GAGTGGTCAG  TGTCTTTTCG  TAGAATGCCT  ACCGTACTGT  CATTCTCTTA
.....
9551  TATGCAGTGC  TGCCATAACC  ATGAGTGATA  ACACTGCGGC  CAACTTACTT
      ATACGTCACG  ACGGTATTGG  TACTCACTAT  TGTGACGCCG  GTTGAATGAA
.....
9601  CTGACAACGA  TCGGAGGACC  GAAGGAGCTA  ACCGCTTTTT  TGCACAACAT
      GACTGTTGCT  AGCCTCCTGG  CTTCTTCGAT  TGGCGAAAAA  ACGTGTGTA
.....
9651  GGGGGATCAT  GTAACTCGCC  TTGATCGTTG  GGAACCGGAG  CTGAATGAAG
      CCCCTAGTA  CATTGAGCGG  AACTAGCAAC  CCTTGGCCTC  GACTTACTTC
.....
9701  CCATACCAA  CGACGAGCGT  GACACCACGA  TGCCTGCAGC  AATGGCAACA
      GGTATGGTTT  GCTGCTCGCA  CTGTGGTGCT  ACGGACGTCG  TTACCGTTGT
.....
9751  ACGTTGCGCA  AACTATTAAC  TGGCGAACTA  CTTACTCTAG  CTTCCCGGCA
      TGCAACGCGT  TTGATAATTG  ACCGCTTGAT  GAATGAGATC  GAAGGGCCGT
.....
9801  ACAATTAATA  GACTGGATGG  AGGCGGATAA  AGTTGCAGGA  CCACTTCTGC
      TGTTAATTAT  CTGACCTACC  TCCGCCTATT  TCAACGTCTT  GGTGAAGACG
.....
9851  GCTCGGCCCT  TCCGGCTGGC  TGGTTTATTG  CTGATAAATC  TGGAGCCGGT
      CGAGCCGGGA  AGGCCGACCG  ACCAAATAAC  GACTATTTAG  ACCTCGGCCA
.....
9901  GAGCGTGGGT  CTCGCGGTAT  CATTGCAGCA  CTGGGGCCAG  ATGGTAAGCC
      CTCGCACCCA  GAGCGCCATA  GTAACGTCGT  GACCCCGGTC  TACCATTCCG
.....
9951  CTCCCGTATC  GTAGTTATCT  ACACGACGGG  GAGTCAGGCA  ACTATGGATG
      GAGGGCATAG  CATCAATAGA  TGTGCTGCC  CTCAGTCCGT  TGATACCTAC
.....
10001 AACGAAATAG  ACAGATCGCT  GAGATAGGTG  CCTCACTGAT  TAAGCATTGG
      TTGCTTTATC  TGTCTAGCGA  CTCTATCCAC  GGAGTACTA  ATTCGTAACC
.....
10051 TAACTGTCAG  ACCAAGTTTA  CTCATATATA  CTTTAGATTG  ATTTAAAAC
      ATTGACAGTC  TGGTTCAAAT  GAGTATATAT  GAAATCTAAC  TAAATTTTGA
.....
10101 TCATTTTTTAA  TTTAAAAGGA  TCTAGGTGAA  GATCCTTTTT  GATAATCTCA
      AGTAAAAATT  AAATTTTCTT  AGATCCACTT  CTAGGAAAAA  CTATTAGAGT
.....

```

pHZ5 (NLS) attR\_2

10151	TGACCAAAAT	CCCTTAACGT	GAGTTTTCGT	TCCACTGAGC	GTCAGACCCC
	ACTGGTTTTA	GGGAATTGCA	CTCAAAGCA	AGGTGACTCG	CAGTCTGGGG
10201	GTAGAAAAGA	TCAAAGGATC	TTCTTGAGAT	CCTTTTTTTC	TGCGCGTAAT
	CATCTTTTCT	AGTTTCCTAG	AAGAACTCTA	GGAAAAAAG	ACGCGCATT
10251	CTGCTGCTTG	CAAACAAAAA	AACCACCGCT	ACCAGCGGTG	GTTTGTTTGC
	GACGACGAAC	GTTTGTTTTT	TTGGTGCCGA	TGGTCGCCAC	CAAACAAACG
10301	CGGATCAAGA	GCTACCAACT	CTTTTTCCGA	AGGTAAGTGG	CTTCAGCAGA
	GCCTAGTTCT	CGATGGTTGA	GAAAAAGGCT	TCCATTGACC	GAAGTCGTCT
10351	GCGCAGATAC	CAAATACTGT	CCTTCTAGTG	TAGCCGTAGT	TAGGCCACCA
	CGCGTCTATG	GTTTATGACA	GGAAGATCAC	ATCGGCATCA	ATCCGGTGGT
10401	CTTCAAGAAC	TCTGTAGCAC	CGCCTACATA	CCTCGCTCTG	CTAATCCTGT
	GAAGTTCTTG	AGACATCGTG	GCGGATGTAT	GGAGCGAGAC	GATTAGGACA
10451	TACCAGTGGC	TGCTGCCAGT	GCGGATAAGT	CGTGTCTTAC	CGGGTTGGAC
	ATGGTCACCG	ACGACGGTCA	CCGCTATTCA	GCACAGAATG	GCCCAACCTG
10501	TCAAGACGAT	AGTTACCGGA	TAAGGCGCAG	CGGTCCGGCT	GAACGGGGGG
	AGTTCTGCTA	TCAATGGCCT	ATTCCGCGTC	GCCAGCCCGA	CTTGCCCCC
10551	TTCGTGCACA	CAGCCCAGCT	TGGAGCGAAC	GACCTACACC	GAAGTGAAGT
	AAGCACGTGT	GTCGGGTCGA	ACCTCGCTTG	CTGGATGTGG	CTTGACTCTA
10601	ACCTACAGCG	TGAGCTATGA	GAAAGCGCCA	CGCTTCCCGA	AGGGAGAAAG
	TGGATGTCGC	ACTCGATACT	CTTTCGCGGT	GCGAAGGGCT	TCCCTCTTTC
10651	GCGGACAGGT	ATCCGGTAAG	CGGCAGGGTC	GGAACAGGAG	AGCGCACGAG
	CGCCTGTCCA	TAGGCCATTC	GCCGTCCAG	CCTTGTCCTC	TCGCGTGCTC
10701	GGAGCTTCCA	GGGGGAAACG	CCTGGTATCT	TTATAGTCCT	GTCGGGTTTC
	CCTCGAAGGT	CCCCCTTTGC	GGACCATAGA	AATATCAGGA	CAGCCCAAAG
10751	GCCACCTCTG	ACTTGAGCGT	CGATTTTTGT	GATGCTCGTC	AGGGGGGCGG
	CGGTGGAGAC	TGAACTCGCA	GCTAAAAACA	CTACGAGCAG	TCCCCCGCC
10801	AGCCTATGGA	AAAACGCCAG	CAACGCGGCC	TTTTTACGGT	TCCTGGCCTT
	TCGGATACCT	TTTTGCGGTC	GTTGCGCCGG	AAAAATGCCA	AGGACCGGAA
10851	TTGCTGGCCT	TTTGCTCACA	TGTTCTTTCC	TGCGTTATCC	CCTGATTCTG
	AACGACCGGA	AAACGAGTGT	ACAAGAAAGG	ACGCAATAGG	GGACTAAGAC
10901	TGGATAACCG	TATTACCGCC	TTTGAGTGAG	CTGATACCGC	TCGCCGCAGC
	ACCTATTGGC	ATAATGGCGG	AAACTCACTC	GACTATGGCG	AGCGGCGTCG
10951	CGAACGACCG	AGCGCAGCGA	GTCAGTGAGC	GAGGAAGCGG	AAGAGCGCCT
	GCTTGCTGGC	TCGCGTCGCT	CAGTCACTCG	CTCCTTCGCC	TTCTCGCGGA
11001	GATGCGGTAT	TTTCTCCTTA	CGCATCTGTG	CGGTATTTCA	CACCGCATAT
	CTACGCCATA	AAAGAGGAAT	GCGTAGACAC	GCCATAAAGT	GTGGCGTATA
11051	GGTGCACTCT	CAGTACAATC	TGCTCTGATG	CCGCATAGTT	AAGCCAGTAT
	CCACGTGAGA	GTCATGTTAG	ACGAGACTAC	GGCGTATCAA	TTCGGTCATA
11101	ACACTCCGCT	ATCGCTACGT	GACTGGGTCA	TGGCTGCGCC	CCGACACCCG
	TGTGAGGCGA	TAGCGATGCA	CTGACCCAGT	ACCGACGCGG	GGCTGTGGGC
11151	CCAACACCCG	CTGACGCGCC	CTGACGGGCT	TGTCTGCTCC	CGGCATCCGC
	GGTTGTGGGC	GACTGCGCGG	GACTGCCCGA	ACAGACGAGG	GCCGTAGGCG

pHZ5(NLS)attR\_2

11201	TTACAGACAA	GCTGTGACCG	TCTCCGGGAG	CTGCATGTGT	CAGAGGTTTT
	AATGTCTGTT	CGACACTGGC	AGAGGCCCTC	GACGTACACA	GTCTCCAAAA
11251	CACCGTCATC	ACCGAAACGC	GCGAGGCAGG	ATGATCCGGG	ATCGAAGAAA
	GTGGCAGTAG	TGGCTTTGCG	CGCTCCGTCC	TACTAGGCC	TAGCTTCTTT
11301	TGATGGTAAA	TGAAATAGGA	AATCAAGGAG	CATGAAGGCA	AAAGACAAAT
	ACTACCATTT	ACTTTATCCT	TTAGTTCCTC	GTACTTCCGT	TTTCTGTTTA
11351	ATAAGGGTCG	AACGAAAAAT	AAAGTGAAAA	GTGTTGATAT	GATGTATTTG
	TATTCCCAGC	TTGCTTTTTA	TTTCACTTTT	CACAACATA	CTACATAAAC
		PacI			
11401	GCTTTGCGGC	GATTAATTAA	TA		
	CGAAACGCCG	CTAATTAATT	AT		