

pHZ5

MAL61/62 promoter

1 GTCTTGGATG TAATTCTTAT TGTTATACTG AATACGCTAA AACCACTCAC
CAGAACCTAC ATTAAGAATA ACAATATGAC TTATGCGATT TTGGTGAGTG

MAL61/62 promoter

51 AACAAAGTATG GAGTATATTG TGTCTCTTTA TATGCTGAGT ACTTATGCAA
TTGTTTCATAC CTCATATAAC ACAGAGAAAT ATACGACTCA TGAATACGTT

MAL61/62 promoter

101 TATGCGCTCA CTCAGGATGA AATGTACACA GCCGAAAGTA TATTGAAAGC
ATACGCGAGT GAGTCCTACT TTACATGTGT CGGCTTTCAT ATAAC TTTCG

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 AAAC TAGAGA

MAL61/62 promoter

351 GAAAATATCT TTTCCCCTGA GTAGCTTCAC TGCTTGGATA CCAATAGGAA
CTTTTATAGA AAAGGGGACT CATCGAAGTG ACGAACCTAT GGTTATCCTT

MAL61/62 promoter

401 TAGACCTTGG CTATAGTAAG TTGCATCTGT ACCGTAGAGA TTCTTGCAAC
ATCTGGAACC GATATCATT C 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

MAL61/62 promoter

651 GTAACATGAT TTGGATCGTG GGATGATACA AACACGTA GATGACTACT
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 AAAAAAATA TTAGAGTTTA CATGTAGTCA GTTCGCATTG

MAL61/62 promoter

Agel
M T

+3

851 TAAATTACAT AAATGTACAT CAGTCAAGCG TAACTAAATT ACATAAATGA
ATTTAATGTA TTTACATGTA GTCAGTTCGC ATTGATTTAA TGTATTTACT

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

+3

901 CCGGTAAAGC GTTAACGGCC AGGCAACAAG AGGTGTTTGA TCTCATCCGT
GGCCATTTTCG CAATTGCCGG TCCGTTGTTT TCCACAAACT AGAGTAGGCA

+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

+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

+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

+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 TCTTCCCAAC GGCGACCATC CAGCACACCG

+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

+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

+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

```

+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 TTGTCCCCTT 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
.....
                                          XmaI
                                          EcoRI
                                          BamHI
+3 · D W L E L A P K K K R K V E F
1501 ACTGGCTGGA ATTGGCCCCC AAGAAAAAGA GAAAGGTGGA ATTCCCAGGG
      TGACCGACCT TAACCGGGGG TTCTTTTTCT CTTTCCACCT TAAGGGCCCC
.....
                                          XhoI
BamHI
1551 ATCCGTCGAC CATGGCGGCC GCTCGAGTCG ACCTGCAGCC AAGCTAATTC
      TAGGCAGCTG GTACCGCCGG CGAGCTCAGC TGGACGTCGG TTCGATTAAG
.....
NotI
1601 CGGGCGAATT TCTTATGATT TATGATTTTT ATTATTA AAT AAGTTATAAA
      GCCCGCTTAA AGAATACTAA ATACTAAAA TAATAATTTA TTCAATATTT
.....
1651 AAAAATAAGT GTATACAAAT TTTAAAGTGA CTCTTAGGTT TTTAAAACGAA
      TTTTTATTCA CATATGTTTA AAATTTCACT GAGAATCCAA AATTTTGCTT
.....
1701 AATTCTTGTT CTTGAGTAAC TCTTTCCTGT AGGTCAGGTT GCTTTCTCAG
      TTAAGAACAA GAACTCATTG AGAAAGGACA TCCAGTCCAA CGAAAGAGTC
.....
1751 GTATAGCATG AGGTCGCTCT TATTGACCAC ACCTCTACCG GCATGCCGAG
      CATATCGTAC TCCAGCGAGA ATAACTGGTG TGGAGATGGC CGTACGGCTC
.....
1801 CAAATGCCTG CAAATCGCTC CCCATTTTAC CCAATTGTAG ATATGCTAAC
      GTTTACGGAC GTTTAGCGAG GGGTAAAGTG GGTAAACATC TATACGATTG
.....
1851 TCCAGCAATG AGTTGATGAA TCTCGGTGTG TATTTTATGT CCTCAGAGGA
      AGGTCGTTAC TCAACTACTT AGAGCCACAC ATAAAATACA GGAGTCTCCT
.....
1901 CAACACCTGT TGTAATCGTT CTTCACACG GATCGATCCA CAGGACGGGT
      GTTGTGGACA ACATTAGCAA GAAGGTGTGC CTAGCTAGGT GTCCTGCCCA
.....
1951 GTGGTCGCCA TGATCGCGTA GTCGATAGTG GCTCCAAGTA GCGAAGCGAG
      CACCAGCGGT ACTAGCGCAT CAGCTATCAC CGAGGTTTCA CGCTTCGCTC
.....
2001 CAGGACTGGG CGGCGGCCAA AGCGGTTCGA CAGTGCTCCG AGAACGGGTG
      GTCCTGACCC GCCGCCGGTT TCGCCAGCCT GTCACGAGGC TCTTGCCCA
.....
2051 CGCATAGAAA TTGCATCAAC GCATATAGCG CTAGCAGCAC GCCATAGTGA
      GCGTATCTTT AACGTAGTTG CGTATATCGC GATCGTCGTG CGGTATCACT
.....

```

pHZ5

2101	CTGGCGATGC	TGTCGGAATG	GACGATATCC	CGCAAGAGGC	CCGGCAGTAC
	GACCGCTACG	ACAGCCTTAC	CTGCTATAGG	GCGTTCTCCG	GGCCGTCATG
2151	CGGCATAACC	AAGCCTATGC	CTACAGCATC	CAGGGTGACG	GTGCCGAGGA
	GCCGTATTGG	TTCGGATACG	GATGTCGTAG	GTCCCACTGC	CACGGCTCCT
2201	TGACGATGAG	CGCATTGTTA	GATTTCATAc	ACGGTGCCTG	ACTGCGTTAG
	ACTGCTACTC	GCGTAACAAT	CTAAAGTATG	TGCCACGGAC	TGACGCAATC
2251	CAATTTAACT	GTGATAAACT	ACCGCATTAA	AGCTAGCTTT	GAAGAAAAAT
	GTTAAATTGA	CACTATTTGA	TGGCGTAATT	TCGATCGAAA	CTTCTTTTTTA
2301	GCGCCTTATT	CAATCTTTGC	TATAAAAAAT	GGCCCAAAT	CTCACATTGG
	CGCGGAATAA	GTTAGAAACG	ATATTTTTTA	CCGGGTTTTA	GAGTGTAACC
2351	AAGACATTTG	ATGACCTCAT	TTCTTTCAAT	GAAGGGCCTA	ACGGAGTTGA
	TTCTGTAAAC	TACTGGAGTA	AAGAAAGTTA	CTTCCCGGAT	TGCCTCAACT
2401	CTAATGTTGT	GGGAAATTGG	AGCGATAAGC	GTGCTTCTGC	CGTGGCCAGG
	GATTACAACA	CCCTTTAACC	TCGCTATTcG	CACGAAGACG	GCACCGGTCC
2451	ACAACGTATA	CTCATCAGAT	AACAGCAATA	CCTGATCACT	ACTTCGCACT
	TGTTGCATAT	GAGTAGTCTA	TTGTCGTTAT	GGACTAGTGA	TGAAGCGTGA
2501	AGTTTCTCGG	TACTATGCAT	ATGATCCAAT	ATCAAAGGAA	ATGATAGCAT
	TCAAAGAGCC	ATGATACGTA	TACTAGGTTA	TAGTTTCCTT	TACTATCGTA
2551	TGAAGGATGA	GACTAATCCA	ATTGAGGAGT	GGCAGCATAT	AGAACAGCTA
	ACTTCTACT	CTGATTAGGT	TAACTCCTCA	CCGTCGTATA	TCTTGTCGAT
2601	AAGGGTAGTG	CTGAAGGAAG	CATACGATAC	CCCGCATGGA	ATGGGATAAT
	TTCCCATCAC	GACTTCCTTC	GTATGCTATG	GGGCGTACCT	TACCCTATTA
2651	ATCACAGGAG	GTACTAGACT	ACCTTTCATC	CTACATAAAT	AGACGCATAT
	TAGTGTCCTC	CATGATCTGA	TGGAAAGTAG	GATGTATTTA	TCTGCGTATA
2701	AAGTACGCAT	TTAAGCATAA	ACACGCACTA	TGCCGTTCTT	CTCATGTATA
	TTCATGCGTA	AATTTCGTATT	TGTGCGTGAT	ACGGCAAGAA	GAGTACATAT
2751	TATATATACA	GGCAACACGC	AGATATAGGT	GCGACGTGAA	CAGTGAGCTG
	ATATATATGT	CCGTTGTGCG	TCTATATCCA	CGCTGCACTT	GTCACTCGAC
2801	TATGTGCGCA	GCTCGCGTTG	CATTTTCGGA	AGCGCTCGTT	TTCGGAAACG
	ATACACGCGT	CGAGCGCAAC	GTAAGAGCCT	TCGCGAGCAA	AAGCCTTTGC
2851	CTTTGAAGTT	CCTATTCCGA	AGTTCCTATT	CTCTAGAAAG	TATAGGAACT
	GAAACTTCAA	GGATAAGGCT	TCAAGGATAA	GAGATCTTTC	ATATCCTTGA
2901	TCAGAGCGCT	TTTGAAAACC	AAAAGCGCTC	TGAAGACGCA	CTTTCAAAAA
	AGTCTCGCGA	AAACTTTTGG	TTTTCGCGAG	ACTTCTGCGT	GAAAGTTTTT
2951	ACCAAAAACG	CACCGGACTG	TAACGAGCTA	CTAAAATATT	GCGAATACCG
	TGGTTTTTGC	GTGGCCTGAC	ATTGCTCGAT	GATTTTATAA	CGCTTATGGC
3001	CTTCCACAAA	CATTGCTCAA	AAGTATCTCT	TTGCTATATA	TCTCTGTGCT
	GAAGGTGTTT	GTAACGAGTT	TTCATAGAGA	AACGATATAT	AGAGACACGA

Spel

~~~~~

Spel

~~~~~

XbaI

~~~~~

## pHZ5

|      |             |             |             |            |             |
|------|-------------|-------------|-------------|------------|-------------|
| 3051 | ATATCCCTAT  | ATAACCTACC  | CATCCACCTT  | TCGCTCCTTG | AACTTGCATC  |
|      | TATAGGGATA  | TATTGGATGG  | GTAGGTGGAA  | AGCGAGGAAC | TTGAACGTAG  |
| 3101 | TAAACTCGAC  | CTCTACATTT  | TTTATGTTTA  | TCTCTAGTAT | TACTCTTTAG  |
|      | ATTTGAGCTG  | GAGATGTAAA  | AAATACAAAT  | AGAGATCATA | ATGAGAAATC  |
| 3151 | ACAAAAAAT   | TGTAGTAAGA  | ACTATTCATA  | GAGTGAATCG | AAAACAATAC  |
|      | TGTTTTTTTA  | ACATCATTCT  | TGATAAGTAT  | CTCACTTAGC | TTTTGTTATG  |
| 3201 | GAAAATGTAA  | ACATTTCCCTA | TACGTAGTAT  | ATAGAGACAA | AATAGAAGAA  |
|      | CTTTTACATT  | TGTAAAGGAT  | ATGCATCATA  | TATCTCTGTT | TTATCTTCTT  |
| 3251 | ACCGTTCATA  | ATTTTCTGAC  | CAATGAAGAA  | TCATCAACGC | TATCACTTTC  |
|      | TGGCAAGTAT  | TAAAAGACTG  | GTTACTTCTT  | AGTAGTTGCG | ATAGTGAAAG  |
| 3301 | TGTTCAACAA  | GTATGCGCAA  | TCCACATCGG  | TATAGAATAT | AATCGGGGAT  |
|      | ACAAGTGTTT  | CATACGCGTT  | AGGTGTAGCC  | ATATCTTATA | TTAGCCCCTA  |
| 3351 | GCCTTTATCT  | TGAAAAAATG  | CACCCGCAGC  | TTCGCTAGTA | ATCAGTAAAC  |
|      | CGGAAATAGA  | ACTTTTTTAC  | GTGGGCGTCG  | AAGCGATCAT | TAGTCATTTG  |
| 3401 | GCGGGAAGTG  | GAGTCAGGCT  | TTTTTTTATGG | AAGAGAAAAT | AGACACCAAA  |
|      | CGCCCTTCAC  | CTCAGTCCGA  | AAAAAATACC  | TTCTCTTTTA | TCTGTGGTTT  |
| 3451 | GTAGCCTTCT  | TCTAACCTTA  | ACGGACCTAC  | AGTGCAAAAA | GTTATCAAGA  |
|      | CATCGGAAGA  | AGATTGGAAT  | TGCCTGGATG  | TCACGTTTTT | CAATAGTTCT  |
| 3501 | GACTGCATTA  | TAGAGCGCAC  | AAAGGAGAAA  | AAAAGTAATC | TAAGATGCTT  |
|      | CTGACGTAAT  | ATCTCGCGTG  | TTTCCTCTTT  | TTTTCATTAG | ATTCTACGAA  |
| 3551 | TGTTAGAAAA  | ATAGCGCTCT  | CGGGATGCAT  | TTTTGTAGAA | CAAAAAAGAA  |
|      | ACAATCTTTT  | TATCGCGAGA  | GCCCTACGTA  | AAAACATCTT | GTTTTTTCTT  |
| 3601 | GTATAGATTC  | TTTGTTGGTA  | AAATAGCGCT  | CTCGCGTTGC | ATTTCTGTTC  |
|      | CATATCTAAG  | AAACAACCAT  | TTTATCGCGA  | GAGCGCAACG | TAAAGACAAG  |
| 3651 | TGTAAAAATG  | CAGCTCAGAT  | TCTTTGTTTG  | AAAAATTAGC | GCTCTCGCGT  |
|      | ACATTTTTTAC | GTCGAGTCTA  | AGAAACAAAC  | TTTTTAATCG | CGAGAGCGCA  |
| 3701 | TGCATTTTTG  | TTTTACAAAA  | ATGAAGCACA  | GATTCTTCGT | TGGTAAAATA  |
|      | ACGTAAAAAC  | AAAATGTTTT  | TACTTCGTGT  | CTAAGAAGCA | ACCATTTTTAT |
| 3751 | GCGCTTTCGC  | GTTGCATTTT  | TGTTCTGTAA  | AAATGCAGCT | CAGATTCTTT  |
|      | CGCGAAAGCG  | CAACGTAAAG  | ACAAGACATT  | TTTACGTCGA | GTCTAAGAAA  |
| 3801 | GTTTGAAAAA  | TTAGCGCTCT  | CGCGTTGCAT  | TTTTGTTCTA | CAAATGAAG   |
|      | CAAACTTTTT  | AATCGCGAGA  | GCGCAACGTA  | AAAACAAGAT | GTTTTACTTC  |
| 3851 | CACAGATGCT  | TCGTTAACAA  | AGATATGCTA  | TTGAAGTGCA | AGATGGAAAC  |
|      | GTGTCTACGA  | AGCAATTGTT  | TCTATACGAT  | AACTTCACGT | TCTACCTTTG  |
| 3901 | GCAGAAAATG  | AACCGGGGAT  | GCGACGTGCA  | AGATTACCTA | TGCAATAGAT  |
|      | CGTCTTTTTAC | TTGGCCCCTA  | CGCTGCACGT  | TCTAATGGAT | ACGTTATCTA  |
| 3951 | GCAATAGTTT  | CTCCAGGAAC  | CGAAATACAT  | ACATTGTCTT | CCGTAAAGCG  |
|      | CGTTATCAAA  | GAGGTCCTTG  | GCTTTATGTA  | TGTAACAGAA | GGCATTTTCGC |
| 4001 | CTAGACTATA  | TATTATTATA  | CAGGTTCAAA  | TATACTATCT | GTTTCAGGGA  |
|      | GATCTGATAT  | ATAATAATAT  | GTCCAAGTTT  | ATATGATAGA | CAAAGTCCCT  |
| 4051 | AAACTCCCAG  | GTTTCGGATGT | TCAAAATTCA  | ATGATGGGTA | ACAAGTACGA  |
|      | TTTGAGGGTC  | CAAGCCTACA  | AGTTTTAAGT  | TACTACCCAT | TGTTTCATGCT |

## pHZ5

|      |                          |                          |                          |                           |                          |
|------|--------------------------|--------------------------|--------------------------|---------------------------|--------------------------|
| 4101 | TCGTAAATCT<br>AGCATTTAGA | GTA AACAGT<br>CATTTTGTCA | TTGTCGGATA<br>AACAGCCTAT | TTAGGCTGTA<br>AATCCGACAT  | TCTCCTCAAA<br>AGAGGAGTTT |
| 4151 | GCGTATTCGA<br>CGCATAAGCT | ATATCATTGA<br>TATAGTAACT | GAAGCTGCAG<br>CTTCGACGTC | CAGGCGTGAA<br>GTCCGCACTT  | GTTAGACGAC<br>CAATCTGCTG |
| 4201 | AACTTCTCTC<br>TTGAAGAGAG | TGGAAACGCA<br>ACCTTTGCGT | TACCGATATT<br>ATGGCTATAA | CAGGCTGCTG<br>GTCCGACGAC  | CAAAGGCACA<br>GTTTCCGTGT |
| 4251 | GGCTAGTGCC<br>CCGATCACGG | CGTGCGAGTG<br>GCACGCTCAC | CATCCGGTAC<br>GTAGGCCATG | CACCCAGAT<br>GTGGGGTCTA   | GCTGTAGTAG<br>CGACATCATC |
| 4301 | CTTCTGGTAG<br>GAAGACCATC | CACTGCAATG<br>GTGACGTTAC | AGCCATGCTT<br>TCGGTACGAA | ATCAAGAAAA<br>TAGTTCCTTTT | CACAGGTTTT<br>GTGTCCAAAA |
| 4351 | GGTACTCGTC<br>CCATGAGCAG | CCATATATCT<br>GGTATATAGA | TGACATGCAA<br>ACTGTACGTT | GCCACTACAC<br>CGGTGATGTG  | CAACAGACCC<br>GTTGTCTGGG |
| 4401 | TAGGGTTTTG<br>ATCCCAAAAC | GATACGATGT<br>CTATGCTACA | TGAAGTTTTA<br>ACTTCAAAAT | TACGGGACTT<br>ATGCCCTGAA  | TATGGTAATC<br>ATACCATTAG |
| 4451 | CTCATTCCAA<br>GAGTAAGGTT | CACTCACTCT<br>GTGAGTGAGA | TACGGTTGGG<br>ATGCCAACC  | AAACAAATAC<br>TTTGTTTATG  | TGCTGTGGAA<br>ACGACACCTT |
| 4501 | AATGCTAGAG<br>TTACGATCTC | CTCACGTAGC<br>GAGTGCATCG | AAAGATGATC<br>TTTCTACTAG | AATGCCGACC<br>TTACGGCTGG  | CCAAGGAAAT<br>GGTTCCTTTA |
| 4551 | AATATTCACT<br>TTATAAGTGA | TCGGGAGCGA<br>AGCCCTCGCT | CCGAATCTAA<br>GGCTTAGATT | TAATATGGTT<br>ATTATACCAA  | CTTAAGGGTG<br>GAATTCCCAC |
| 4601 | TCCAAGATT<br>AGGGTTCTAA  | TTATAAGAAG<br>AATATTCTTC | ACTAAGAAAC<br>TGATTCTTTG | ACATCATCAC<br>TGTAGTAGTG  | CACTAGAACG<br>GTGATCTTGC |
| 4651 | GAACACAAGT<br>CTTGTGTTCA | GTGTCTTGGA<br>CACAGAACCT | AGCCGCACGG<br>TCGGCGTGCC | GCCATGATGA<br>CGGTACTACT  | AGGAGGGATT<br>TCCTCCCTAA |
| 4701 | TGAAGTCACT<br>ACTTCAGTGA | TTCCTAAATG<br>AAGGATTTAC | TGGACGATCA<br>ACCTGCTAGT | AGGTCTTATC<br>TCCAGAATAG  | GATTTGAAGG<br>CTAAACTTCC |
| 4751 | AATTGGAAGA<br>TTAACCTTCT | TGCCATTAGA<br>ACGGTAATCT | CCAGATACCT<br>GGTCTATGGA | GTCTCGTCTC<br>CAGAGCAGAG  | TGTGATGGCT<br>ACACTACCGA |
| 4801 | GTCAATAATG<br>CAGTTATTAC | AAATCGGTGT<br>TTTAGCCACA | CATTCAACCT<br>GTAAGTTGGA | ATTAAAGAAA<br>TAATTTCTTT  | TTGGAGCAAT<br>AACCTCGTTA |
| 4851 | TTGTAGAAAG<br>AACATCTTTC | AATAAGATCC<br>TTATTCTAGG | TCGGGGACAC<br>AGCCCCTGTG | CAAATATGGC<br>GTTTATAACCG | GATCTCGGCC<br>CTAGAGCCGG |
| 4901 | TTTTCGTTTC<br>AAAAGCAAAG | TTGGAGCTGG<br>AACCTCGACC | GACATGTTTG<br>CTGTACAAAC | CCATCGATCC<br>GGTAGCTAGG  | ATCTACCACC<br>TAGATGGTGG |
| 4951 | AGAACGGCCG<br>TCTTGCCGGC | TTAGATCTGC<br>AATCTAGACG | TGCCACCGTT<br>ACGGTGGCAA | GTTTCCACCG<br>CAAAGGTGGC  | AAGAAACCAC<br>TTCTTTGGTG |
| 5001 | CGTTGCCGTA<br>GCAACGGCAT | ACCACCACGA<br>TGGTGGTGCT | CGGTTGTTGC<br>GCCAACAACG | TAAAGAAGCT<br>ATTTCTTCGA  | GCCACCGCCA<br>CGGTGGCGGT |
| 5051 | CGGCCACCGT<br>GCCGGTGGCA | TGTAGCCGCC<br>ACATCGGCGG | GTTGTTGTTA<br>CAACAACAAT | TTGTAGTTGC<br>AACATCAACG  | TACTGTTATT<br>ATGACAATAA |
| 5101 | TCTGGCACTT<br>AGACCGTGAA | CTTGGTTTTT<br>GAACCAAAAG | CTCTTAAGTG<br>GAGAATTCAC | AGGAGGAACA<br>TCCTCCTTGT  | TAACCATTCT<br>ATTGGTAAGA |

## pHZ5

|      |            |             |             |            |             |
|------|------------|-------------|-------------|------------|-------------|
| 5151 | CGTTGTTGTC | GTTGATGCTT  | AAATTTTGCA  | CTTGTTGCT  | CAGTTCAGCC  |
|      | GCAACAACAG | CAACTACGAA  | TTTAAAACGT  | GAACAAGCGA | GTCAAGTCGG  |
| 5201 | ATAATATGAA | ATGCTTTTCT  | TGTTGTTCTT  | ACGGAATACC | ACTTGCCACC  |
|      | TATTATACTT | TACGAAAAGA  | ACAACAAGAA  | TGCCTTATGG | TGAACGGTGG  |
| 5251 | TATCACCACA | ACTAACTTTT  | TCCCGTTCCT  | CCATCTCTTT | TATATTTTTT  |
|      | ATAGTGGTGT | TGATTGAAAA  | AGGGCAAGGA  | GGTAGAGAAA | ATATAAAAAA  |
| 5301 | TTCTCGATCG | AGTTCAAGAG  | AAAAAAAAAAG | AAAAAGCAA  | AAGAAAAAAG  |
|      | AAGAGCTAGC | TCAAGTTCTC  | TTTTTTTTTTC | TTTTTCGTTT | TTCTTTTTTTC |
| 5351 | GAAAGCGCGC | CTCGTTCAGA  | ATGACACGTA  | TAGAATGATG | CATTACCTTG  |
|      | CTTTCGCGCG | GAGCAAGTCT  | TACTGTGCAT  | ATCTTACTAC | GTAATGGAAC  |
| 5401 | TCATCTTCAG | TATCATACTG  | TTCGTATACA  | TACTTACTGA | CATTCATAGG  |
|      | AGTAGAAGTC | ATAGTATGAC  | AAGCATATGT  | ATGAATGACT | GTAAGTATCC  |
| 5451 | TATACATATA | TACACATGTA  | TATATATCGT  | ATGCTGCAGC | TTTAAATAAT  |
|      | ATATGTATAT | ATGTGTACAT  | ATATATAGCA  | TACGACGTCG | AAATTTATTA  |
| 5501 | CGGTGTCACT | ACATAAGAAC  | ACCTTTGGTG  | GAGGGAACAT | CGTTGGTACC  |
|      | GCCACAGTGA | TGTATTCTTG  | TGGAAACCAC  | CTCCCTTGTA | GCAACCATGG  |
| 5551 | ATTGGGCGAG | GTGGCTTCTC  | TTATGGCAAC  | CGCAAGAGCC | TTGAACGCAC  |
|      | TAACCCGCTC | CACCGAAGAG  | AATACCGTTG  | GCGTTCTCGG | AACTTGCGTG  |
| 5601 | TCTCACTACG | GTGATGATCA  | TTCTTGCCTC  | GCAGACAATC | AACGTGGAGG  |
|      | AGAGTGATGC | CACTACTAGT  | AAGAACGGAG  | CGTCTGTTAG | TTGCACCTCC  |
|      |            |             | HindIII     |            |             |
| 5651 | GTAATTCTGC | TAGCCTCTGC  | AAAGCTTTCA  | AGAAAATGCG | GGATCATCTC  |
|      | CATTAAGACG | ATCGGAGACG  | TTTCGAAAGT  | TCTTTTACGC | CCTAGTAGAG  |
| 5701 | GCAAGAGAGA | TCTCCTACTT  | TCTCCCTTTG  | CAAACCAAGT | TCGACAACCTG |
|      | CGTTCTCTCT | AGAGGATGAA  | AGAGGGAAAC  | GTTTGGTTCA | AGCTGTTGAC  |
| 5751 | CGTACGGCCT | GTTTCGAAAGA | TCTACCACCG  | CTCTGGAAAG | TGCCTCATCC  |
|      | GCATGCCGGA | CAAGCTTTCT  | AGATGGTGGC  | GAGACCTTTC | ACGGAGTAGG  |
| 5801 | AAAGGCGCAA | ATCCTGATCC  | AAACCTTTTT  | ACTCCACGCG | CCAGTAGGGC  |
|      | TTTCCGCGTT | TAGGACTAGG  | TTTGGAAAAA  | TGAGGTGCGC | GGTCATCCCG  |
|      |            | HindIII     |             |            |             |
| 5851 | CTCTTTAAAA | GCTTGACCGA  | GAGCAATCCC  | GCAGTCTTCA | GTGGTGTGAT  |
|      | GAGAAATTTT | CGAACTGGCT  | CTCGTTAGGG  | CGTCAGAAGT | CACCACACTA  |
| 5901 | GGTCGTCTAT | GTGTAAGTCA  | CCAATGCACT  | CAACGATTAG | CGACCAGCCG  |
|      | CCAGCAGATA | CACATTCAGT  | GGTTACGTGA  | GTTGCTAATC | GCTGGTCGGC  |
| 5951 | GAATGCTTGG | CCAGAGCATG  | TATCATATGG  | TCCAGAAACC | CTATACCTGT  |
|      | CTTACGAACC | GGTCTCGTAC  | ATAGTATACC  | AGGTCTTTGG | GATATGGACA  |
| 6001 | GTGGACGTTA | ATCACTTGCG  | ATTGTGTGGC  | CTGTTCTGCT | ACTGCTTCTG  |
|      | CACCTGCAAT | TAGTGAACGC  | TAACACACCG  | GACAAGACGA | TGACGAAGAC  |
| 6051 | CCTCTTTTTT | TGGGAAGATC  | GAGTGCTCTA  | TCGCTAGGGG | ACCACCCTTT  |
|      | GGAGAAAAAG | ACCCTTCTAG  | CTCACGAGAT  | AGCGATCCCC | TGGTGGGAAA  |
| 6101 | AAAGAGATCG | CAATCTGAAT  | CTTGGTTTCA  | TTTGTAATAC | GCTTTACTAG  |
|      | TTTCTCTAGC | GTTAGACTTA  | GAACCAAAGT  | AAACATTATG | CGAAATGATC  |

## pHZ5

|      |             |            |            |            |             |
|------|-------------|------------|------------|------------|-------------|
| 6151 | GGCTTTCTGC  | TCTGTCATCT | TTGCCTTCGT | TTATCTTGCC | TGCTCATTTT  |
|      | CCGAAAGACG  | AGACAGTAGA | AACGGAAGCA | AATAGAACGG | ACGAGTAAAA  |
| 6201 | TTAGTATATT  | CTTCGAAGAA | ATCACATTAC | TTTATATAAT | GTATAATTCA  |
|      | AATCATATAA  | GAAGCTTCTT | TAGTGTAATG | AAATATATTA | CATATTAAGT  |
| 6251 | TTATGTGATA  | ATGCCAATCG | CTAAGAAAAA | AAAAGAGTCA | TCCGCTAGGT  |
|      | AATACACTAT  | TACGGTTAGC | GATTCTTTTT | TTTTCTCAGT | AGGCGATCCA  |
| 6301 | GGAAAAAAA   | AAATGAAAAT | CATTACCGAG | GCATAAAAAA | ATATAGAGTG  |
|      | CCTTTTTTTTT | TTTACTTTTA | GTAATGGCTC | CGTATTTTTT | TATATCTCAC  |
| 6351 | TACTAGAGGA  | GGCCAAGAGT | AATAGAAAAA | GAAAATTGCG | GGAAAGGACT  |
|      | ATGATCTCCT  | CCGGTTCTCA | TTATCTTTTT | CTTTTAACGC | CCTTTCCTGA  |
| 6401 | GTGTTATGAC  | TTCCCTGACT | AATGCCGTGT | TCAAACGATA | CCTGGCAGTG  |
|      | CACAATACTG  | AAGGGACTGA | TTACGGCACA | AGTTTGCTAT | GGACCGTCAC  |
| 6451 | ACTCCTAGCG  | CTCACCAAGC | TCTTAAAACG | AGAATTAAGA | AAAAGTCGTC  |
|      | TGAGGATCGC  | GAGTGGTTCG | AGAATTTTGC | TCTTAATTCT | TTTTCAGCAG  |
| 6501 | ATCTTTCGAT  | AAGTTTTTCC | CACAGCAAAG | CAATAGTAGA | AAAACAATGG  |
|      | TAGAAAGCTA  | TTCAAAAAGG | GTGTCGTTTC | GTTATCATCT | TTTTGTTACC  |
| 6551 | GAAACGTTGA  | ATGAAGACAA | AGCGTCGTGG | TTTAAAAGGA | AATACGCTCA  |
|      | CTTTGCAACT  | TACTTCTGTT | TCGCAGCACC | AAATTTTCCT | TTATGCGAGT  |
| 6601 | CGTACATGCT  | AGGGAACAGG | ACCGTGCAGC | GGATCTAATG | AATCCATTTG  |
|      | GCATGTACGA  | TCCCTTGTC  | TGGCACGTCG | CCTAGATTAC | TTAGGTAAAC  |
| 6651 | TTAGTTAATA  | GTTTAAATGT | TTTTATCGGA | AGAGGTTTTG | TCATCACATC  |
|      | AATCAATTAT  | CAAATTTACA | AAAATAGCCT | TCTCCAAAAC | AGTAGTGTAG  |
| 6701 | AGCAATGTTT  | TTCTTGGTCT | CGATGTAGTA | TACGTATAAA | TTATTACCTG  |
|      | TCGTTACAAG  | AAGAACCAGA | GCTACATCAT | ATGCATATTT | AATAATGGAC  |
| 6751 | ATACTTCATC  | TCTAAGTCTC | ATTGCCTTTG | TGCCAAAAAA | TCTGTTTCTA  |
|      | TATGAAGTAG  | AGATTCAGAG | TAACGGAAAC | ACGGTTTTTT | AGACAAAGAT  |
| 6801 | AATTTCTCTT  | CATTTGTAGA | CTTAATTATA | CTGATCGTTG | ATCTACTATC  |
|      | TTAAAGAGAA  | GTAAACATCT | GAATTAATAT | GACTAGCAAC | TAGATGATAG  |
| 6851 | AGTAAGTAAG  | CCTTTAAAAA | AAAAAAAAAA | AAAAAAAAAA | AAAAAACCTG  |
|      | TCATTCATTC  | GGAAATTTTT | TTTTTTTTTT | TTTTTTTTTT | TTTTTTGGAC  |
| 6901 | TAACAATAGC  | AATACCCCAA | ATACCTAATG | TAGTTCACAG | AAGCAAGCTA  |
|      | ATTGTTATCG  | TTATGGGGTT | TATGGATTAC | ATCAAGGTCG | TTCGTTTCGAT |
| 6951 | AAAAGTAAAG  | CAACAACATA | ACTCACCCCT | GCATCTGCAG | CTTTTGCCCCG |
|      | TTTTCATTTT  | GTTGTTGTAT | TGAGTGGGGA | CGTAGACGTC | GAAAACGGGC  |
|      | Xmal        |            |            |            | Xmal        |
| 7001 | GGCAGCCTGC  | TCTGCCTGTG | TTTTCTTTAA | TTGAGCAGTA | GACCATTTAG  |
|      | CCGTCGGACG  | AGACGGACAC | AAAAGAAATT | AACTCGTCAT | CTGGTAAATC  |
| 7051 | CAGTTGCATG  | AATAGCTGCA | GCGTCACATC | GGATAATAAT | GATGGCAGCC  |
|      | GTCAACGTAC  | TTATCGACGT | CGCAGTGTAG | CCTATTATTA | CTACCGTCGG  |
| 7101 | ATTGTAGAAG  | TGCCTTTTGC | ATTTCTAGTC | TCTTTCTCGG | TCTAGCTAGT  |
|      | TAACATCTTC  | ACGGAAAACG | TAAAGATCAG | AGAAAGAGCC | AGATCGATCA  |



## pHZ5

|      |             |            |            |            |             |
|------|-------------|------------|------------|------------|-------------|
| 7151 | TTTACTACAT  | CGCGAAGATA | GAATCTTAGA | TCACACTGCC | TTTGCTGAGC  |
|      | AAATGATGTA  | GCGCTTCTAT | CTTAGAATCT | AGTGTGACGG | AAACGACTCG  |
| 7201 | TGGATCAATA  | GAGTAACAAA | AGAGTGGTAA | GGCCTCGTTA | AAGGACAAGG  |
|      | ACCTAGTTAT  | CTCATTGTTT | TCTCACCATT | CCGGAGCAAT | TTCCTGTTCC  |
| 7251 | ACCTGAGCGG  | AAGTGTATCG | TACAGTAGAC | GGAGTATACT | AGTATAGTCT  |
|      | TGGACTCGCC  | TTCACATAGC | ATGTCATCTG | CCTCATATGA | TCATATCAGA  |
| 7301 | ATAGTCCGTG  | GAATTAATTC | TTGAAGACGA | AAGGGCCTCG | TGATACGCC   |
|      | TATCAGGCAC  | CTTAATTAAG | AACTTCTGCT | TTCCCGGAGC | ACTATGCGGA  |
| 7351 | ATTTTTATAG  | GTTAATGTCA | TGATAATAAT | GGTTTCTTAG | ACGTCAGGTG  |
|      | TAAAAATATC  | CAATTACAGT | ACTATTATTA | CCAAAGAATC | TGCAGTCCAC  |
| 7401 | GCACTTTTCG  | GGGAAATGTG | CGCGGAACCC | CTATTTGTTT | ATTTTTCTAA  |
|      | CGTGAAAAGC  | CCCTTTACAC | GCGCCTTGGG | GATAAACAAA | TAAAAAGATT  |
| 7451 | ATACATTCAA  | ATATGTATCC | GCTCATGAGA | CAATAACCCT | GATAAATGCT  |
|      | TATGTAAGTT  | TATACATAGG | CGAGTACTCT | GTTATTGGGA | CTATTTACGA  |
| 7501 | TCAATAATAT  | TGAAAAAGGA | AGAGTATGAG | TATTCAACAT | TTCCGTGTCG  |
|      | AGTTATTATA  | ACTTTTTCTT | TCTCATACTC | ATAAGTTGTA | AAGGCACAGC  |
| 7551 | CCCTTATTCC  | CTTTTTTGCG | GCATTTTGCC | TTCCTGTTTT | TGCTCACCCA  |
|      | GGGAATAAGG  | GAAAAACGC  | CGTAAAACGG | AAGGACAAA  | ACGAGTGGGT  |
| 7601 | GAAACGCTGG  | TGAAAGTAAA | AGATGCTGAA | GATCAGTTGG | GTGCACGAGT  |
|      | CTTTGCGACC  | ACTTTCATTT | TCTACGACTT | CTAGTCAACC | CACGTGCTCA  |
| 7651 | GGGTTACATC  | GAAGTGGATC | TCAACAGCGG | TAAGATCCTT | GAGAGTTTTT  |
|      | CCCAATGTAG  | CTTGACCTAG | AGTTGTGCGC | ATTCTAGGAA | CTCTCAAAAG  |
| 7701 | GCCCCGAAGA  | ACGTTTTCCA | ATGATGAGCA | CTTTTAAAGT | TCTGCTATGT  |
|      | CGGGGCTTCT  | TGCAAAAGGT | TACTACTCGT | GAAAATTTCA | AGACGATACA  |
| 7751 | GGCGCGGTAT  | TATCCCGTGT | TGACGCCGGG | CAAGAGCAAC | TCGGTCGCCG  |
|      | CCGCGCCATA  | ATAGGGCACA | ACTGCGGCC  | GTTCTCGTTG | AGCCAGCGGC  |
| 7801 | CATACACTAT  | TCTCAGAATG | ACTTGGTTGA | GTACTCACCA | GTCACAGAAA  |
|      | GTATGTGATA  | AGAGTCTTAC | TGAACCAACT | CATGAGTGGT | CAGTGTCTTT  |
| 7851 | AGCATCTTAC  | GGATGGCATG | ACAGTAAGAG | AATTATGCAG | TGCTGCCATA  |
|      | TCGTAGAATG  | CCTACCGTAC | TGTCATTCTC | TTAATACGTC | ACGACGGTAT  |
| 7901 | ACCATGAGTG  | ATAACACTGC | GGCCAACCTA | CTTCTGACAA | CGATCGGAGG  |
|      | TGGTACTCAC  | TATTGTGACG | CCGGTTGAAT | GAAGACTGTT | GCTAGCCTCC  |
| 7951 | ACCGAAGGAG  | CTAACCGCTT | TTTTGCACAA | CATGGGGGAT | CATGTAACTC  |
|      | TGGCTTCCTC  | GATTGGCGAA | AAAACGTGTT | GTACCCCTA  | GTACATTGAG  |
| 8001 | GCCTTGATCG  | TTGGGAACCG | GAGCTGAATG | AAGCCATACC | AAACGACGAG  |
|      | CGGAAC TAGC | AACCCTTGGC | CTCGACTTAC | TTCGGTATGG | TTTGCTGCTC  |
| 8051 | CGTGACACCA  | CGATGCCTGC | AGCAATGGCA | ACAACGTTGC | GCAAAC TATT |
|      | GCACTGTGGT  | GCTACGGACG | TCGTTACCGT | TGTTGCAACG | CGTTTGATAA  |
| 8101 | AACTGGCGAA  | CTACTTACTC | TAGCTTCCCG | GCAACAATTA | ATAGACTGGA  |
|      | TTGACCGCTT  | GATGAATGAG | ATCGAAGGGC | CGTTGTTAAT | TATCTGACCT  |

Spel

## pHZ5

|      |            |             |            |            |            |
|------|------------|-------------|------------|------------|------------|
| 8151 | TGGAGGCGGA | TAAAGTTGCA  | GGACCACTTC | TGCGCTCGGC | CCTTCCGGCT |
|      | ACCTCCGCCT | ATTTCAACGT  | CCTGGTGAAG | ACGCGAGCCG | GGAAGGCCGA |
| 8201 | GGCTGGTTTA | TTGCTGATAA  | ATCTGGAGCC | GGTGAGCGTG | GGTCTCGCGG |
|      | CCGACCAAAT | AACGACTATT  | TAGACCTCGG | CCACTCGCAC | CCAGAGCGCC |
| 8251 | TATCATTGCA | GCACTGGGGC  | CAGATGGTAA | GCCCTCCCGT | ATCGTAGTTA |
|      | ATAGTAACGT | CGTGACCCCG  | GTCTACCATT | CGGGAGGGCA | TAGCATCAAT |
| 8301 | TCTACACGAC | GGGGAGTCAG  | GCAACTATGG | ATGAACGAAA | TAGACAGATC |
|      | AGATGTGCTG | CCCCTCAGTC  | CGTTGATACC | TACTTGCTTT | ATCTGTCTAG |
| 8351 | GCTGAGATAG | GTGCCTCACT  | GATTAAGCAT | TGGTAACTGT | CAGACCAAGT |
|      | CGACTCTATC | CACGGAGTGA  | CTAATTCGTA | ACCATTGACA | GTCTGGTTCA |
| 8401 | TTACTCATAT | ATACTTTAGA  | TTGATTTAAA | ACTTCATTTT | TAATTTAAAA |
|      | AATGAGTATA | TATGAAATCT  | AACTAAATTT | TGAAGTAAAA | ATTAAATTTT |
| 8451 | GGATCTAGGT | GAAGATCCTT  | TTTGATAATC | TCATGACCAA | AATCCCTTAA |
|      | CCTAGATCCA | CTTCTAGGAA  | AAACTATTAG | AGTACTGGTT | TTAGGGAATT |
| 8501 | CGTGAGTTTT | CGTTCCACTG  | AGCGTCAGAC | CCCGTAGAAA | AGATCAAAGG |
|      | GCACTCAAAA | GCAAGGTGAC  | TCGCAGTCTG | GGGCATCTTT | TCTAGTTTCC |
| 8551 | ATCTTCTTGA | GATCCTTTTT  | TTCTGCGCGT | AATCTGCTGC | TTGCAAACAA |
|      | TAGAAGAACT | CTAGGAAAAA  | AAGACGCGCA | TTAGACGACG | AACGTTTGTT |
| 8601 | AAAAACCACC | GCTACCAGCG  | GTGGTTTGTT | TGCCGGATCA | AGAGCTACCA |
|      | TTTTTGGTGG | CGATGGTCGC  | CACCAAACAA | ACGGCCTAGT | TCTCGATGGT |
| 8651 | ACTCTTTTTT | CGAAGGTAAC  | TGGCTTCAGC | AGAGCGCAGA | TACCAAATAC |
|      | TGAGAAAAAG | GCTTCCATTG  | ACCGAAGTCG | TCTCGCGTCT | ATGGTTTATG |
| 8701 | TGTCCTTCTA | GTGTAGCCGT  | AGTTAGGCCA | CCACTTCAAG | AACTCTGTAG |
|      | ACAGGAAGAT | CACATCGGCA  | TCAATCCGGT | GGTGAAGTTC | TTGAGACATC |
| 8751 | CACCGCCTAC | ATACCTCGCT  | CTGCTAATCC | TGTTACCAGT | GGCTGCTGCC |
|      | GTGGCGGATG | TATGGAGCGA  | GACGATTAGG | ACAATGGTCA | CCGACGACGG |
| 8801 | AGTGCGGATA | AGTCGTGTCT  | TACCGGGTTG | GACTCAAGAC | GATAGTTACC |
|      | TCACCGCTAT | TCAGCACAGA  | ATGGCCCAAC | CTGAGTTCTG | CTATCAATGG |
| 8851 | GGATAAGGCG | CAGCGGTCCG  | GCTGAACGGG | GGGTTCGTGC | ACACAGCCCA |
|      | CCTATTCCGC | GTCGCCAGCC  | CGACTTGCCC | CCCAAGCACG | TGTGTCGGGT |
| 8901 | GCTTGGAGCG | AACGACCTAC  | ACCGAACTGA | GATACCTACA | GCGTGAGCTA |
|      | CGAACCTCGC | TTGCTGGATG  | TGGCTTGACT | CTATGGATGT | CGCACTCGAT |
| 8951 | TGAGAAAGCG | CCACGCTTCC  | CGAAGGGAGA | AAGGCGGACA | GGTATCCGGT |
|      | ACTCTTTCGC | GGTGCGAAGG  | GCTTCCCTCT | TTCCGCCTGT | CCATAGGCCA |
| 9001 | AAGCGGCAGG | GTCGGAACAG  | GAGAGCGCAC | GAGGGAGCTT | CCAGGGGGAA |
|      | TTCGCCGTCC | CAGCCTTGTC  | CTCTCGCGTG | CTCCCTCGAA | GGTCCCCCTT |
| 9051 | ACGCCTGGTA | TCTTTATAGT  | CCTGTCCGGT | TTCGCCACCT | CTGACTTGAG |
|      | TGCGGACCAT | AGAAATATCA  | GGACAGCCCA | AAGCGGTGGA | GACTGAACTC |
| 9101 | CGTCGATTTT | TGTGATGCTC  | GTCAGGGGGG | CGGAGCCTAT | GGAAAAACGC |
|      | GCAGCTAAAA | AACTACGAG   | CAGTCCCCCC | GCCTCGGATA | CCTTTTTGCG |
| 9151 | CAGCAACGCG | GCCTTTTTTAC | GGTTCCTGGC | CTTTTGCTGG | CCTTTTGCTC |
|      | GTCGTTGCGC | CGGAAAAATG  | CCAAGGACCG | GAAAACGACC | GGAAAACGAG |

## pHZ5

|      |            |             |             |            |             |
|------|------------|-------------|-------------|------------|-------------|
| 9201 | ACATGTTCTT | TCCTGCGTTA  | TCCCCTGATT  | CTGTGGATAA | CCGTATTACC  |
|      | TGTACAAGAA | AGGACGCAAT  | AGGGGACTAA  | GACACCTATT | GGCATAATGG  |
| 9251 | GCCTTTGAGT | GAGCTGATAC  | CGCTCGCCGC  | AGCCGAACGA | CCGAGCGCAG  |
|      | CGGAAACTCA | CTCGACTATG  | GCGAGCGGCG  | TCGGCTTGCT | GGCTCGCGTC  |
| 9301 | CGAGTCAGTG | AGCGAGGAAG  | CGGAAGAGCG  | CCTGATGCGG | TATTTTCTCC  |
|      | GCTCAGTCAC | TCGCTCCTTC  | GCCTTCTCGC  | GGACTACGCC | ATAAAAGAGG  |
| 9351 | TTACGCATCT | GTGCGGTATT  | TCACACCGCA  | TATGGTGCAC | TCTCAGTACA  |
|      | AATGCGTAGA | CACGCCATAA  | AGTGTGGCGT  | ATACCACGTG | AGAGTCATGT  |
| 9401 | ATCTGCTCTG | ATGCCGCATA  | GTTAAGCCAG  | TATACACTCC | GCTATCGCTA  |
|      | TAGACGAGAC | TACGGCGTAT  | CAATTTCGGTC | ATATGTGAGG | CGATAGCGAT  |
| 9451 | CGTGACTGGG | TCATGGCTGC  | GCCCCGACAC  | CCGCCAACAC | CCGCTGACGC  |
|      | GCACTGACCC | AGTACCGACG  | CGGGGCTGTG  | GGCGGTTGTG | GGCGACTGCG  |
| 9501 | GCCCTGACGG | GCTTGTCTGC  | TCCCGGCATC  | CGCTTACAGA | CAAGCTGTGA  |
|      | CGGGACTGCC | CGAACAGACG  | AGGGCCGTAG  | GCGAATGTCT | GTTTCGACACT |
| 9551 | CCGTCTCCGG | GAGCTGCATG  | TGTCAGAGGT  | TTTCACCGTC | ATCACCGAAA  |
|      | GGCAGAGGCC | CTCGACGTAC  | ACAGTCTCCA  | AAAGTGGCAG | TAGTGGCTTT  |
| 9601 | CGCGCGAGGC | AGGATGATCC  | GGGATCGAAG  | AAATGATGGT | AAATGAAATA  |
|      | GCGCGCTCCG | TCCTACTAGG  | CCCTAGCTTC  | TTTACTACCA | TTTACTTTAT  |
| 9651 | GGAAATCAAG | GAGCATGAAG  | GCAAAAGACA  | AATATAAGGG | TCGAACGAAA  |
|      | CCTTTAGTTC | CTCGTACTTC  | CGTTTTCTGT  | TTATATTCCC | AGCTTGCTTT  |
| 9701 | AATAAAGTGA | AAAGTGTTGA  | TATGATGTAT  | TTGGCTTTGC | GGCGATTAAT  |
|      | TTATTTCACT | TTTCACAACCT | ATACTACATA  | AACCGAAACG | CCGCTAATTA  |
|      | PacI       |             |             |            |             |
| 9751 | TAATA      |             |             |            |             |
|      | ATTAT      |             |             |            |             |