



NK-TGCK

1 TAGTTATTAC TAGCGCTACC GGA<sup>AfeI</sup>CTCAGAT CTCGAGCTCC ACCGCGGTGG XhoI NotI

51 CGGCCGCTCT AGAACTAGTC CAGTCTGGGT CCTAATGCGG GTGGCGTCTC SpeI

101 TCTTGACAGG CAGCGTTTGG GGACAACAGC GGGGAAGGGA GATAAGATGA

151 CATACCAGAG CAGATTTGGT GTGCGCGCTG ATACTCCTGG CCCGACAGGA

201 AACTCGGAGC TATTTAAAAA GGCCCTATCG ATTACTTTAT CTTCCCCGGA

251 GGAAACTTC TTGCCGAGAG ACAAAGATG TCCCCCTACC TAAAGATACA

301 AGGCCACACA GCAGAGGGTT TGTACAGCG ACGCGGAGT AGATCCCTGG BsrGI

351 GAGCGCAGAG ACGCCTTCT ACCGGCAGAG ACTGAAGTTT GACTGGAGCG

401 AGGCGGGGCG GAGCCACCTC CCGGTCCCCG TGC<sup>BspEI</sup>ACTCCGG AATTGTGAAC

451 GCTCCCGCAA AGTCCCCGCG AGTGTGTGTG TCACACTAAT TGGTGTGCCT

501 CTGAGTTCCA TCCGTACTCC CCCCCCCCCA AGTTTAAATG CTCCTTTTAA

551 GGGCTTGAGT GTCTGCAGCC GTCATGTGCA CCTTGAAAGC TTGGCGTGTG

601 ACATTGTGTT CCTCTAGCGT CTCCATCCCG CTTTTACCA ATTCGTGCCC

651 TGGCCGAATC CCGTGGGAGC AGAGGTGGCC CCTCGAAGTC CCGGGCCACC XmaI SmaI

701 CAGGGAGGCT CCGGGTACAC CCAGATTCCC AGCAGTCCGT AGTCTGGGCT

751 TTCCCCGTGT GACACTTTGC CGCCTACCGG TTCACAATCC TCTCGGCCCG

801 CGTGAGAGAA GCGGCATGT TCCGTGGGAC ACAGTTTATT TTATGCACAT

851 TAAGGGCTGT AAAGCACGGT CTGAAAGAAA AAGATGGTGC GCGGGGAGG

901 AGCGGGAGAA GGGGCAGCGA TGATCCTGGG ACTGTGCAGC GACCACTTAC

951 AACCTGATTC ATGCCACCG GAGACCCAG CGTCTGGA GGGAGCGGA

1001 GGCTCCGAGG CTTCCGGTTG GTTTCAGCC ACCAGGTTCC GGCAGTGA<sup>BmgBI</sup>CT

1051 CTTGGACTCT GACCACGTCC CGCGGCACAC TCGGAGACCC GGCTCGTCCG Tth111I

1101 CGAAGCTAGG AGCGAATGGC CCGCAGCTGC GCGTCGCTCA ATCTGAGCTC

1151 TCAAGCCCAA AGCGTCTGCA GAGCCGCTTC CCAGAGCTCC AAGCCACCGC NheI BmtI

1201 TAGCAGAGCT CTGAAGGCAC CTTGAAGCCT CGACCCACTT CAATGTCACC

1251 TGATTTTTTCA AATTATCACC TTCTGAGACG CGATCCTTCG GAACCGGCAG

AfeI

1301 AGCCCCCACC TTTGAGCGCC GCCGTTTATG CTGTAGCGCT GTAGCGGTTT

1351 GCAGCAGTCA GGGCAGGTGG AGGACCTTGT TACTGGCACT GCGGAGCCCC

1401 GAGCTGAGGG GAAAGAAAAC GGAGATAAGG AAAAGGGCTG CAGAGGTCGG

1451 AAAGCAGCTC CCATCGGATC CATCAAATCA ATCACAGCCC CAAGTGACCA

1501 AAGAGAAATF TTGCGTTTTG GAAATATCCC CACCACCAA AGCACGAGGA

1551 AAGTGTGAGC AGCAAGAAAC TAACAATAAA TTTTAGGTGC AAAAAATATA

1601 AGAAGCACAC AGAATTTGAA AAAAAAACC CGCACATTCT GCTATTTGAG

1651 TTTTCCATAA ACATACAAAT AAAACCTGCG GAAGATGTAC TTCTTCAAGT

1701 CTGTGACAGA CCCTGGACTG ACTAACGGGG ACGCGCTCAT TTAAAATGTC

1751 AGCGGTTCCCT ATATTTACC GAACTTTCAA CGAGCTTTTC CACATGTCCA

1801 GACTCATGGG ACAGTATGCT TATGACCCTG TTAATTGTAG TTCAGACTTA

NdeI

1851 CTATGTCTCA ATAAAGTGAC TCAGCACTTG ACCAGCATAT GAGGCTCCCC

1901 GAGTCCAGTC CCCAGAAACA AGTAGAATAT ACATATATAT ATAATTCAT

1951 ATATCTATCT ATATATTATT CTATATATAT ATTCTACTTG TATATATATG

2001 TATTACATAA TATATATACA AGTAGAATAT ATATAATATA ATTTATATTA

2051 TATATAGAAT ATATATAGAA ACAAGTAGAT TATATATATA TATATATATA

2101 GTATATATAT ATATATATAT ATATATATAT ATAGTATTGT ATTACATACA

BstZ17I

AccI

2151 TGTGGGTATA CATCATACAC ATGGCTTTCA ATCCATCCTC ACCAGCCCAT

2201 TTAGTGATAC AGGACACAGT GCAAAGAATA TTTATTCAAA TCCTCGACCC

2251 ACCCCCAACC CTGCGTTTAG ACTCAGCATA ACAGAATCAG GTATTTTCTT

2301 TGAGTGTGTT GGCTGGGATT TTCAGGCTAA CGAGGAGCAA TCTCCTTGCA

2351 GGAGTGGGAG ACTATCCCAA ATGTGTCGCG GCCCCTTTAA AAAAAGCTGA

2401 GCAGGATTTT TTTTCCCAT CCCTATGTCT ATATTAGGTG ACGCAGAACT

2451 GCCCGTCGCT CCTGTCATCC AGGCCCTGG CCCAATGGCA GGCTGAATCC

2501 CCCCTACTCC AGCCTGCTCC CGCCTCTTCT GCCCCTGGTG CTCCGCGCTA

2551 CCTGCTGCCG CGCACCACAT CCAGGACAGA GAGGCGGGTG CGCGGGCGGG

2601 CGGCGGCACC ATGCGGGGAG GCTGTCCCA GGGGTGGCAG CACCACTCTC

KpnI      AjuI  
Acc65I    AjuI

2651 TGCTACCCAC CTGGCCCAAG CTGGGTACCT GCCAAGCTCT AGTTCTAGAA  
\* F      Frame 4

ScaI

2701 TTCGCTGTCT GCGAGGGCCA GCTGTTGGGG TGAGTACTCC CTCTCAAAAG  
E S D A L A L Q Q P S Y E R E F      Frame 4

2751 CGGGCATGAC TTCTGCGCTA AGATTGTCAG TTTCCAAAAA CGAGGAGGAT  
A P M V E A S L N D T E L F S S S      Frame 4

2801 TTGATATTCA CCTGGCCCGC GGTGATGCCT TTGAGGGTGG CCGCGTCCAT  
K I N V Q G A T I G K L T A A D M      Frame 4

2851 CTGGTCAGAA AAGACAATCT TTTTGTGTC AAGCTTGAGG TGTGGCAGGC  
Q D S F V I K K N D L K L H P L      Frame 4

2901 TTGAGATCTG GCCATACACT TGAGTGACAA TGACATCCAC TTTGCCTTTC  
M T S T L P F      Frame 2  
S S I Q G Y V Q T V I V D V K G K      Frame 4

2951 TCTCCACAGG TGTCCACTCC CAGGTCCAAC TGCAGCCCAA GCGGATCCGA  
S P Q V S T P R S N C S P S G S E      Frame 2  
E G C T D V G L D L Q L G L P D S      Frame 4

3001 ATTCACCATG TCCAGACTGG ACAAGAGCAA AGTCATCAAC TCTGCCTTGG  
F T M S R L D K S K V I N S A L E      Frame 2  
N V M D L S S L L L T M L E A K      Frame 4

3051 AGCTCGTGAA TGAAGTTGGC ATTGAGGGCT TGACCACCAG GAAGCTGGCC  
L V N E V G I E G L T T R K L A      Frame 2  
S S T F S T P M S P K V V L F S A      Frame 4

3101 CAGAAGCTGG GTGTGGAGCA GCCTACCCTG TACTGGCATG TGAAGAACAA  
Q K L G V E Q P T L Y W H V K N K      Frame 2  
W F S P T S C G V R Y Q C T F F L      Frame 4

3151 GAGGGCTCTG CTTGATGCCC TGGCCATTGA GATGTTGGAC AGGCACCACA  
R A L L D A L A I E M L D R H H T      Frame 2  
L A R S S A R A M      Frame 4

EcoNI

3201 CCCACTTCTG CCCTCTGGAA GGGGAGTCCT GGCAGGACTT CCTGAGGAAC  
H F C P L E G E S W Q D F L R N      Frame 2

3251 AATGCCAAGA GCTTCAGATG TGCCTTGCTC TCCCACCGGG ATGGTGCCAA  
N A K S F R C A L L S H R D G A K      Frame 2

3301 AGTTCACTTG GGCACCAGGC CTACAGAGAA GCAGTATGAG ACCCTGGAGA  
V H L G T R P T E K Q Y E T L E N      Frame 2

3351 ACCAGCTGGC ATTCCTGTGC CAACAAGGCT TCTCCCTGGA AAATGCCTTG  
Q L A F L C Q Q G F S L E N A L      Frame 2

3401 TATGCCCTCT CTGCTGTGGG CCACTTCACC TTGGGCTGTG TGCTGGAGGA  
Y A L S A V G H F T L G C V L E D Frame 2

3451 CCAGGAGCAC CAAGTTGCCA AGGAGGAGAG GGAGACCCCC ACCACTGACT  
Q E H Q V A K E E R E T P T T D S Frame 2

3501 CCATGCCACC ACTGCTGCGG CAAGCTATTG AGTTGTTTGA CCACCAAGGG  
M P P L L R Q A I E L F D H Q G Frame 2

BclI

3551 GCTGAGCCTG CATTCTTTTT TGGCCTGGAA CTGATCATCT GTGGCCTGGA  
A E P A F L F G L E L I I C G L E Frame 2  
\* R H G P Frame 6

3601 AAAGCAGCTG AAATGTGAGT CTGGCTCTCA CGTGCCCAA AAGAGAAAGC  
K Q L K C E S G S H V P K K R K H Frame 2  
F A A S I H T Q S E R A W F S F A Frame 6

3651 ACGTGCCTGC TGATGCCTTG GATGATTTTG ACCTGGACAT GCTGCCTGCT  
V P A D A L D D F D L D M L P A Frame 2  
R A Q Q H R P H N Q G P C A A Q Frame 6

3701 GATGCCCTGG ATGACTTTGA TTTGGACATG CTCCTGCTG ATGCACTTGA  
D A L D D F D L D M L P A D A L D Frame 2  
Q H G P H S Q N P C A G Q Q H V Q Frame 6

3751 TGATTTTGAC CTGGATATGC TGTGATGGAT CCAGACATGA TAAGATACAT  
D F D L D M L \* Frame 2  
H N Q G P Y A T I S G S M I L Y M Frame 6

3801 TGATGAGTTT GGACAAACCA CAACTAGAAT GCAGTAAAA AAATGCTTTA  
S S N P C V V V L I C H F F H K Frame 6

3851 TTTGTGAAAT TTGTGATGCT ATTGCTTTAT TTGTAACCAT TATAAGCTGC  
I Q S I Q S A I A K N T V M I L Q Frame 6

HpaI MfeI

3901 ATAAACAAGT TAACAACAAC AATTGCATTC ATTTTATGTT TCAGGTTTCAG  
M F L N V V V I A N M K H K L N L Frame 6

3951 GGGGAGGTGT GGGAGGTTTT TAAAAGCAAG TAAAACCTCT ACAAATGTGG  
P L H P L N K L A L L V E V F T Frame 6

4001 TATGGCTGAT TATGATCCTG CAAGCCTCGT CGTCCTGGCC GGACCACGCT  
T H S I I I R C A E D D Q G S W A Frame 6

4051 ATCTGTGCAA GGTCCCCGGC CCCGGACGCG CGCTCCATGA GCAGAGCGCC  
I Q A L D G A G S A R E M Frame 6

SfoI

NarI

KasI

BbsI

BbeI

4101 CGCCGCCGAG GCGAAGACTC GGGCGGCGCC CTGCCCGTCC CACCAGGTCA

4151 ACAGGCGGTA ACGTAACCGG CCTCTTCATC GGAATGCGC GCGACCTTCA

4201 GCATCGCCGG CATGTCCCC TGGCGGACGG GAAGTATCCA GCTCGACCAA

XmaI

SmaI

4251 GCTTCTCGAC GATAGATCTC CCGGGTGGCA TCCCTGTGAC CCCTCCCCAG

CspCI

CspCI

4301 TGCCTCTCCT GGCCCTGGAA GTTGCCACTC CAGTGCCCAC CAGCCTTGTC

AhdI

4351 CTAATAAAAT TAAGTTGCAT CTTTTGTCT GACTAGGTGT CCTTCTATAA

BbsI

4401 TATTATGGGG TGGAGGGGG TGGTATGGAG CAAGGGGCAA GTTGGGAAGA

4451 CAACCTGTAG GGCCTGCGGG GTCTATTGGG AACCAAGCTG GAGTGCAGTG

4501 GCACAATCTT GGCTCACTGC AATCTCCGCC TCCTGGGTTT AAGCGATTCT

4551 CCTGCCTCAG CCTCCGAGT TGTTGGGATT CCAGGCATGC ATGACCAGGC

4601 TCAGCTAATT TTTGTTTTTT TGGTAGAGAC GGGGTTTCAC CATATTGGCC

4651 AGGCTGGTCT CCAACTCCTA ATCTCAGGTG ATCTACCCAC CTTGGCCTCC

AloI

AloI

4701 CAAATTGCTG GGATTACAGG CGTGAACCAC TGCTCCCTTC CCTGTCCTTC

ZraI

AatII

4751 TGATTTTAAA ATAACTATAC CAGCAGAGGA CGTCCAGACA CAGCATAGCT

4801 ACCTGGCCAT GCCCAACCGG TGGGACATTT GAGTTGCTTG CTTGGCACTG

KpnI

Acc65I

4851 TCCTCTCATG CGTTGGGTCC ACTCAGTAGA TGCCTGTTGA TTCGGTACCG

4901 AGCTCGACTT TCACTTTTCT CTATCACTGA TAGGGAGTGT AAACCTCGACT

4951 TTCACTTTTC TCTATCACTG ATAGGAGTGT AAACCTCGACT TTCACTTTTC

5001 TCTTATCTGA TAGGGAGTGG TAAACTCCAC TTTAAATTC TACTCCGCGG

XhoI

5051 ACCCGACTTC CTGCAATTAA AGCTTCTCGA GTCTAGATCT GCACCGGTCC

5101 CCACCATGGT GAGCAAGGGC GAGGAGCTGT TCACCGGGGT GGTGCCCATC

M V S K G E E L F T G V V P I

Frame 3

5151 CTGGTTCGAGC TGGACGGCGA CGTAAACGGC CACAAGTTCA GCGTGTCCGG

L V E L D G D V N G H K F S V S G

Frame 3

5201 CGAGGGCGAG GGCGATGCCA CCTACGGCAA GCTGACCCTG AAGTTCATCT

E G E G D A T Y G K L T L K F I C

Frame 3

5251 GCACCACCGG CAAGCTGCCC GTGCCCTGGC CCACCCTCGT GACCACCCTG  
T T G K L P V P W P T L V T T L Frame 3

5301 ACCTACGGCG TGCAGTGCTT CAGCCGCTAC CCGACCACA TGAAGCAGCA  
T Y G V Q C F S R Y P D H M K Q H Frame 3

5351 CGACTTCTTC AAGTCCGCCA TGCCCGAAGG CTACGTCCAG GAGCGCACCA  
D F F K S A M P E G Y V Q E R T I Frame 3

5401 TCTTCTTCAA GGACGACGGC AACTACAAGA CCCGCGCCGA GGTGAAGTTC  
F F K D D G N Y K T R A E V K F Frame 3

5451 GAGGGCGACA CCCTGGTGAA CCGCATCGAG CTGAAGGGCA TCGACTTCAA  
E G D T L V N R I E L K G I D F K Frame 3

5501 GGAGGACGGC AACATCCTGG GGCACAAGCT GGAGTACAAC TACAACAGCC  
E D G N I L G H K L E Y N Y N S H Frame 3

5551 ACAACGTCTA TATCATGGCC GACAAGCAGA AGAACGGCAT CAAGGTGAAC  
N V Y I M A D K Q K N G I K V N Frame 3

5601 TTCAAGATCC GCCACAACAT CGAGGACGGC AGCGTGCAGC TCGCCGACCA  
F K I R H N I E D G S V Q L A D H Frame 3

5651 CTACCAGCAG AACACCCCCA TCGGCGACGG CCCCCTGCTG CTGCCCGACA  
Y Q Q N T P I G D G P V L L P D N Frame 3

5701 ACCACTACCT GAGCACCCAG TCCGCCCTGA GCAAAGACCC CAACGAGAAG  
H Y L S T Q S A L S K D P N E K Frame 3

5751 CGCGATCACA TGGTCTGCT GGAGTTCGTG ACCGCCGCCG GGATCACTCT  
R D H M V L L E F V T A A G I T L Frame 3

BsrGI

5801 CGGCATGGAC GAGCTGTACA AGTCCGGTAC AGCTCTCGAC GGAGAAAGCT  
G M D E L Y K S G T A L D G E S S Frame 3

5851 CAGGCTCTGG CTCAGAGTCT GACTCCATGG CCAATTTACT GACCGTACAC  
G S G S E S D S M A N L L T V H Frame 3

5901 CAAAATTTGC CTGCATTACC GGTCGATGCA ACCAGTGATG AGGTTGCAA  
Q N L P A L P V D A T S D E V R K Frame 3

5951 GAACCTGATG GACATGTTCA GGGATCGCCA GCGTTTTTCT GAGCATACT  
N L M D M F R D R Q A F S E H T W Frame 3

6001 GGAAAATGCT TCTGTCCGTT TGCCGGTCGT GGGCGGCATG GTGCAAGTTG  
K M L L S V C R S W A A W C K L Frame 3

XmnI      NruI

6051 AATAACCGGA AATGGTTTCC CGCAGAACCT GAAGATGTTT GCGATTATCT  
N N R K W F P A E P E D V R D Y L Frame 3

6101 TCTATATCTT CAGGCGCGCG GTCTGGCAGT AAAAATATC CAGCAACATT  
L Y L Q A R G L A V K T I Q Q H L Frame 3

RsrII

6151 TGGGCCAGCT AACATGCTT CATCGTCGGT CCGGGCTGCC ACGACCAAGT

G Q L N M L H R R S G L P R P S Frame 3

BsrDI AclI  
6201 GACAGCAATG CTGTTTCACT GGTTATGCGG CGGATCCGAA AAGAAAACGT  
D S N A V S L V M R R I R K E N V Frame 3

BstBI  
6251 TGATGCCGGT GAACGTGCAA AACAGGCTCT AGCGTTCGAA CGCACTGATT  
D A G E R A K Q A L A F E R T D F Frame 3

PvuI AsiSI SnaBI  
6301 TCGACCAGGT TCGTTCACCTC ATGGAAAATA GCGATCGCTG CCAGGATATA  
D Q V R S L M E N S D R C Q D I Frame 3

SnaBI  
6351 CGTAATCTGG CATTCTGGG GATTGCTTAT AACACCCTGT TACGTATAGC  
R N L A F L G I A Y N T L L R I A Frame 3

EcoRV  
6401 CGAAATTGCC AGGATCAGGG TTAAAGATAT CTCACGTACT GACGGTGGGA  
E I A R I R V K D I S R T D G G R Frame 3

6451 GAATGTAAAT CCATATTGGC AGAACGAAAA CGCTGGTTAG CACCGCAGGT  
M L I H I G R T K T L V S T A G Frame 3

6501 GTAGAGAAGG CACTTAGCCT GGGGGTAACT AAAGTGGTCG AGCGATGGAT  
V E K A L S L G V T K L V E R W I Frame 3

6551 TTCCGTCTCT GGTGTAGCTG ATGATCCGAA TAACTACCTG TTTTGCCGGG  
S V S G V A D D P N N Y L F C R V Frame 3

6601 TCAGAAAAAA TGGTGTGGCC GCGCCATCTG CCACCAGCCA GCTATCAACT  
R K N G V A A P S A T S Q L S T Frame 3

6651 CGCGCCCTGG AAAGGATTTT TGAAGCAACT CATCGATTGA TTTACGGCGC  
R A L E R I F E A T H R L I Y G A Frame 3

PshAI  
6701 TAAGGATGAC TCTGGTCAGA GATACCTGGC CTGGTCTGGA CACAGTGCCC  
K D D S G Q R Y L A W S G H S A R Frame 3

6751 GTGTCCGAGC CGCGCGAGAT ATGGCCGCGC TGGAGTTCAT ACCGGAGATC  
V G A A R D M A A L E F I P E I Frame 3

6801 ATGCAGCTGG TGCTGGACCA TGTAATATGT CATGACTATA TCGTACTGAT  
M Q L V L D H V I C H D Y I V L I Frame 3

MluI  
6851 AGTGAACAGG CATGTGCGCT GCTGAAGATG GCGATTAGCC ATTAACGCGT  
V N R H V R C \* Frame 3

6901 AAATGATTGC TCTAGAGTGA GGGTCCCCAC CTGGGACCCT TGAGAGTATC

6951 AGGTCTCCCA CGTGGGAGAC AAGAAATCCC TGTTTAATAT TTAAACAGCA

7001 GTGTTCCCCA TCTGGGTCCT TGCACCCCTC ACTCTGGCCT CAGCCGACTG



7051 CACAGCGGCC CCTGCATCCC CTTGGCTGTG AGGCCCTGG ACAAGCAGAG

7101 GTGGCCAGAG CTGGGAGGCA TGGCCCTGGG GTCCCACGAA TTTGCTGGGG

AflII

7151 AATCTCGTTT TTCTTCTTAA GACTTTTGGG ACATGGTTTG ACTCCCGAAC

BmgBI

7201 ATCACCGACG TGTCTCCTGT TTTTCTGGGT GGCCTCGGGA CACCTGCCCT

7251 GCCCCACGA GGGTCAGGAC TGTGACTCTT TTTAGGGCCA GGCAGGTGCC

7301 TGGACATTTG CCTTGCTGGA CGGGGACTGG GGATGTGGGA GGGAGCAGAC

7351 AGGAGGAATC ATGTCAGGCC TGTGTGTGAA AGGAAGCTCC ACTGTCACCC

BarI

BarI

7401 TCCACCTCTT CACCCCCAC TCACCAGTGT CCCCTCCACT GTCACATTGT

7451 AACTGAACTT CAGGATAATA AAGTGTTC CTCCAAAAA AAAAAAAAAA

NotI

7501 ACCGAATTCA GCTTGAAGTT CGCGGCCCGG ACTCTAGATC ATAATCAGCC

7551 ATACCACATT TGTAGAGGTT TTAAGTTC TAAAAACCT CCCACACCTC

MfeI

HpaI

7601 CCCCTGAACC TGAAACATAA AATGAATGCA ATTGTTGTTG TTAACCTGTT

7651 TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA

7701 CAAATAAAGC ATTTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC

AflII

7751 ATCAATGTAT CTTAAGGCGT AAATTGTAAG CGTTAATATT TTGTTAAAAT

7801 TCGCGTTAAA TTTTGTGTTAA ATCAGCTCAT TTTTAAACCA ATAGGCCGAA

7851 ATCGGCAAAA TCCCTTATAA ATCAAAGAA TAGACCGAGA TAGGGTTGAG

AloI

AloI

7901 TGTTGTTCCA GTTTGAACA AGAGTCCACT ATTAAAGAAC GTGGACTCCA

DraIII

7951 ACGTCAAAGG GCGAAAACC GTCTATCAGG GCGATGGCCC ACTACGTGAA

8001 CCATCACCTT AATCAAGTTT TTTGGGTCG AGGTGCCGTA AAGCACTAAA

8051 TCGGAACCCT AAAGGGAGCC CCCGATTTAG AGCTTGACGG GGAAAGCCGG

8101 CGAACGTGGC GAGAAAGGAA GGGAAGAAAG CGAAAGGAGC GGGCGCTAGG

8151 GCGCTGGCAA GTGTAGCGGT CACGCTGCGC GTAACCACCA CACCCGCCGC

8201 GCTTAATGCG CCGCTACAGG GCGCGTCAGG TGGCACTTTT CGGGGAAATG

8251 TCGCGGGAAC CCCTATTTGT TTATTTTTCT AAATACATTC AAATATGTAT  
8301 CCGCTCATGA GACAATAACC CTGATAAATG CTTCAATAAT ATTGAAAAAG  
8351 GAAGAGTCCT GAGGCGGAAA GAACCAGCTG TGGAAATGTGT GTCAGTTAGG  
8401 GTGTGGAAAG TCCCCAGGCT CCCAGCAGG CAGAAGTATG CAAAGCATGC  
8451 ATCTCAATTA GTCAGCAACC AGGTGTGGAA AGTCCCCAGG CTCCCCAGCA  
8501 GGCAGAAGTA TGCAAAGCAT GCATCTCAAT TAGTCAGCAA CCATAGTCCC  
8551 GCCCCTAACT CCGCCCATCC CGCCCCTAAC TCCGCCAGT TCCGCCATT

SfiI

8601 CTCCGCCCA TGGCTGACTA ATTTTTTTTA TTTATGCAGA GGCCGAGGCC  
8651 GCCTCGGCCT CTGAGCTATT CCAGAAGTAG TGAGGAGGCT TTTTGGAGG

AvrII

8701 CCTAGGCTTT TGCAAAGATC GATCAAGAGA CAGGATGAGG ATCGTTTCGC  
8751 ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG CTGGGGTGGAA  
M I E Q D G L H A G S P A A W V E

Frame 3

8801 GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC TGCTCTGATG  
R L F G Y D W A Q Q T I G C S D A

Frame 3

SfoI

NarI

KasI

EbeI

8851 CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT TTTTGTCAAG  
A V F R L S A Q G R P V L F V K

Frame 3

8901 ACCGACCTGT CCGGTGCCCT GAATGAACTG CAAGACGAGG CAGCGCGGCT  
T D L S G A L N E L Q D E A A R L

Frame 3

FspI

Tth111I

8951 ATCGTGGCTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG CTCGACGTTG  
S W L A T T G V P C A A V L D V V

Frame 3

9001 TCACTGAAGC GGAAGGGAC TGGCTGCTAT TGGGCGAAGT GCCGGGGCAG  
T E A G R D W L L L G E V P G Q

Frame 3

9051 GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT CCATCATGGC  
D L L S S H L A P A E K V S I M A  
\* R V K S R G L F Y G D H

Frame 3

Frame 6

BsrDI

9101 TGATGCAATG CCGCGGCTGC ATACGCTTGA TCCGGCTACC TGCCCATTCC  
D A M R R L H T L D P A T C P F D  
S I C H P P Q M R K I R S G A W E

Frame 3

Frame 6

9151 ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTAICTG GATGGAAGCC  
H Q A K H R I E R A R T R M E A  
V V L R F M A D L S C T S P H F G

Frame 3

Frame 6

## SapI

9201 GGTCTTGTGCG ATCAGGATGA TCTGGACGAA GAGCATCAGG GGCTCGCGCC  
 G L V D Q D D L D E E H Q G L A P Frame 3  
 T K D I L I I Q V F L M L P E R Frame 6

9251 AGCCGAAC TG TCGCCAGGC TCAAGGCGAG CATGCCCGAC GGCGAGGATC  
 A E L F A R L K A S M P D G E D L Frame 3  
 W G F Q E G P E L R A H G V A L I Frame 6

9301 TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT GGTGGAAAAT  
 V V T H G D A C L P N I M V E N Frame 3  
 E D H G M A I G A Q R I D H H F I Frame 6

## RsrII

9351 GGCCGCTTTT CTGGATTCAT CGACTGTGGC CGGCTGGGTG TGGCGGACCG  
 G R F S G F I D C G R L G V A D R Frame 3  
 A A K R S E D V T A P Q T H R V Frame 6

## SapI

9401 CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA GAGCTTGGCG  
 Y Q D I A L A T R D I A E E L G G Frame 3  
 A I L V Y R Q S G T I N S F L K A Frame 6

9451 GCGAATGGGC TGACCGCTTC CTCGTGCTTT ACGGTATCGC CGCTCCCGAT  
 E W A D R F L V L Y G I A A P D Frame 3  
 A F P S V A E E H K V T D G S G I Frame 6

9501 TCGCAGCGCA TGCCTTCTA TGCCTTCTT GACGAGTTCT TCTGAGCGGG  
 S Q R I A F Y R L L D E F F \* Frame 3  
 R L A D G E I A K K V L E E S R Frame 6

## BstBI

9551 ACTCTGGGGT TCGAAATGAC CGACCAAGCG ACGCCCAACC TGCCATCAGG  
 M T D Q A T P N L P S R Frame 2  
 S E P T R F S R G L S A W G A M Frame 6

9601 AGATTTCGAT TCCACCGCCG CCTTCTATGA AAGGTTGGGC TTCGGAATCG  
 D F D S T A A F Y E R L G F G I V Frame 2

9651 TTTTCCGGGA CGCCGGCTGG ATGATCCTCC AGCGCGGGGA TCTCATGCTG  
 F R D A G W M I L Q R G D L M L Frame 2

AvrII

9701 GAGTTCTTCG CCCACCCTAG GGGGAGGCTA ACTGAAACAC GGAAGGAGAC  
 E F F A H P R G R L T E T R K E T Frame 2

9751 AATACCGGAA GGAACCCGCG CTATGACGGC AATAAAAAGA CAGAATAAAA  
 I P E G T R A M T A I K R Q N K T Frame 2

9801 CGCACGGTGT TGGGTCGTTT GTTCATAAAC GCGGGGTTTC GTCCCAGGGC  
 H G V G S F V H K R G V R S Q G Frame 2

9851 TGGCACTCTG TCGATACCCC ACCGAGACCC CATTGGGGCC AATACGCCCC  
 W H S V D T P P R P H W G Q Y A R Frame 2

9901 CGTTTCTTCC TTTTCCCCAC CCCACCCCC AAGTTCGGGT GAAGGCCCCAG  
 V S S F S P P H P P S S G E G P G Frame 2

9951 GGCTCGCAGC CAACGTCGGG GCGGCAGGCC CTGCCATAGC CTCAGGTTAC  
L A A N V G A A G P A I A S G Y  
10001 TCATATATAC TTTAGATTGA TTTAAAACCTT CATTTTTAAT TTAAAAGGAT  
S Y I L \* Frame 2  
10051 CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG  
10101 AGTTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT  
10151 TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA  
10201 ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC  
10251 TTTTCCGAA GGTAAGTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC  
10301 CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC  
10351 GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG  
10401 GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT  
10451 AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGACAC AGCCCAGCTT  
10501 GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG  
10551 AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC  
10601 GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC  
10651 CTGGTATCTT TATAGTCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC  
10701 GATTTTTGTG ATGCTCGTCA GGGGGCGGA GCCTATGGAA AAACGCCAGC  
10751 AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT  
10801 GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCA  
10851 TGCAT

Frame 2

pNK-TGCK, 10855 bp. (CIRCULAR)  
Restriction analysis 8/6/2009  
pDRAW32 revision 1.1.101  
Rebase containing 280 enzymes.  
56 enzymes match enzyme selection criteria.

\*\*\*\*\*

**AatII** (G\_ACGT'C) [ZraI]  
Cuts 1 time.  
Cuts at position 4783.

**AccI** (GT'mk\_AC) [FblI,XmiI]  
Cuts 1 time.  
Cuts at position 2157.

**Acc65I** (G'GTAC\_C) [KpnI,Asp718I]  
[dcm methylated]  
Cuts 2 times.  
Cuts at positions 2674, 4894.  
Fragment sizes 2220, 8635.

**AcII** (AA'CG\_TT) [Psp1406I]  
Cuts 1 time.  
Cuts at position 6247.

**AfeI** (AGC'GCT) [Eco47III,Aor51HI]  
Cuts 2 times.  
Cuts at positions 14, 1337.  
Fragment sizes 1323, 9532.

**AflIII** (C'TTAA\_G) [BfrI,BspTI,Bst98I,BstAFI,MspCI,Vha464I]  
Cuts 2 times.  
Cuts at positions 7166, 7761.  
Fragment sizes 595, 10260.

**AhdI** (GACnn\_n'nnGTC) [Eam1105I,AspEI,BmeRI,DriI]  
Cuts 1 time.  
Cuts at position 4386.

**AjuI** (GAAnnnnnnnTTGnnnnnn\_nnnnn')  
Cuts 1 time.  
Cuts at position 2675.

**AjuI** (CCAAnnnnnnnTTCnnnnnn\_nnnnn')  
Cuts 1 time.  
Cuts at position 2707.

**AloI** (GAACnnnnnnTCCnnnnnn\_nnnnn')  
Cuts 2 times.  
Cuts at positions 4748, 7961.  
Fragment sizes 3213, 7642.

**AloI** (GGAnnnnnnGTTcnnnnnnn\_nnnnn')

Cuts 2 times.

Cuts at positions 4716, 7929.

Fragment sizes 3213, 7642.

**AsiSI** (GCG\_AT'CGC) [SgfI,RgaI,SfaAI]

Cuts 1 time.

Cuts at position 6335.

**AvrII** (C'CTAG\_G) [AspA2I,BlnI,XmaJI]

Cuts 2 times.

Cuts at positions 8701, 9716.

Fragment sizes 1015, 9840.

**BarI** (GAAGnnnnnnTACnnnnnnn\_nnnnn')

Cuts 1 time.

Cuts at position 7441.

**BarI** (GTAnnnnnnCTTcnnnnnnn\_nnnnn')

Cuts 1 time.

Cuts at position 7473.

**BbeI** (G\_GCGC'C) [NarI,DinI,EgeI,EheI,KasI,Mly113I,SfoI]

Cuts 2 times.

Cuts at positions 4129, 8882.

Fragment sizes 4753, 6102.

**BbsI** (GAAGACnn'nnnn\_) [BpiI,BpuAI,BstV2I]

Cuts 2 times.

Cuts at positions 4120, 4453.

Fragment sizes 333, 10522.

**BclI** (T'GATC\_A) [FbaI,Ksp22I]

[dam methylated]

Cuts 1 time.

Cuts at position [3582].

**BmgBI** (CAC'GTC) [BtrI,AjiI]

Cuts 2 times.

Cuts at positions 1066, 7209.

Fragment sizes 6143, 4712.

**BmtI** (G\_CTAG'C) [NheI,AsuNHI,BspOI]

Cuts 1 time.

Cuts at position 1203.

**BspEI** (T'CCGG\_A) [AccIII,Aor13HI,BseAI,Bsp13I,Kpn2I,MroI]

[dam methylated]

Cuts 1 time.

Cuts at position 436.

**BsrDI** (GCAATG\_nn') [Bse3DI,BseMI]

Cuts 2 times.

Cuts at positions 6212, 9112.

Fragment sizes 2900, 7955.

**BsrGI** (T'GTAC\_A) [Bsp1407I,BstAUI]

Cuts 2 times.

Cuts at positions 321, 5815.

Fragment sizes 5494, 5361.

**BstBI** (TT'CG\_AA) [AsuII,Bpu14I,Bsp119I,BspT104I,Csp45I,NspV,SfuI]

Cuts 2 times.

Cuts at positions 6286, 9561.

Fragment sizes 3275, 7580.

**BstZ17I** (GTA'TAC) [BssNAI,Bst1107I]

Cuts 1 time.

Cuts at position 2158.

**CspCI** (CAAnnnnnGTGGnnnnnnnnnn\_nn')

Cuts 1 time.

Cuts at position 4326.

**CspCI** (CCACnnnnnTTGnnnnnnnnnn\_nn')

Cuts 1 time.

Cuts at position 4361.

**DraIII** (CAC\_nnn'GTG) [AdeI]

Cuts 1 time.

Cuts at position 7995.

**EcoNI** (CCTnn'n\_nnAGG) [BstENI,XagI]

Cuts 1 time.

Cuts at position 3216.

**EcoRV** (GAT'ATC) [Eco32I]

Cuts 1 time.

Cuts at position 6428.

**FspI** (TGC'GCA) [Acc16I,AviII,NsbI]

Cuts 1 time.

Cuts at position 8981.

**HpaI** (GTT'AAC) [KspAI]

Cuts 2 times.

Cuts at positions 3911, 7642.

Fragment sizes 3731, 7124.

**KasI** (G'GCGC\_C) [NarI,BbeI,DinI,EgeI,EheI,Mly113I,SfoI]

Cuts 2 times.

Cuts at positions 4125, 8878.  
Fragment sizes 4753, 6102.

**KpnI** (G\_GTAC'C) [Acc65I,Asp718I]

Cuts 2 times.  
Cuts at positions 2678, 4898.  
Fragment sizes 2220, 8635.

**MfeI** (C'AATT\_G) [MunI]

Cuts 2 times.  
Cuts at positions 3920, 7629.  
Fragment sizes 3709, 7146.

**MluI** (A'CGCG\_T)

Cuts 1 time.  
Cuts at position 6895.

**NarI** (GG'CG\_CC) [BbeI,DinI,EgeI,EheI,KasI,Mly113I,SfoI]

Cuts 2 times.  
Cuts at positions 4126, 8879.  
Fragment sizes 4753, 6102.

**NdeI** (CA'TA\_TG) [FauNDI]

Cuts 1 time.  
Cuts at position 1887.

**NheI** (G'CTAG\_C) [AsuNHI,BmtI,BspOI]

Cuts 1 time.  
Cuts at position 1199.

**NotI** (GC'GGCC\_GC) [CciNI]

Cuts 2 times.  
Cuts at positions 51, 7523.  
Fragment sizes 7472, 3383.

**NruI** (TCG'CGA) [Bsp68I,BtuMI]

[dam methylated]  
Cuts 1 time.  
Cuts at position 6091.

**PshAI** (GACnn'nnGTC) [BoxI,BstPAI]

Cuts 1 time.  
Cuts at position 6712.

**PvuI** (CG\_AT'CG) [BpvUI,MvrI,Ple19I]

Cuts 1 time.  
Cuts at position 6335.

**RsrII** (CG'GwC CG) [CpoI,CspI,Rsr2I]

Cuts 2 times.



Cuts at positions 6178, 9395.  
Fragment sizes 3217, 7638.

**SapI** (GCTCTTCn'nnn\_) [BspQI,LguI,PciSI]

Cuts 2 times.

Cuts at positions 9223, 9433.

Fragment sizes 210, 10645.

**ScaI** (AGT'ACT) [AssI,BmcAI,ZrmI]

Cuts 1 time.

Cuts at position 2735.

**SfiI** (GGCCn\_nnn'nGGCC)

[dcm methylated]

Cuts 1 time.

Cuts at position 8654.

**SfoI** (GGC'GCC) [NarI,BbeI,DinI,EgeI,EheI,KasI,Mly113I]

Cuts 2 times.

Cuts at positions 4127, 8880.

Fragment sizes 4753, 6102.

**SmaI** (CCC'GGG) [Cfr9I,TspMI,XmaI,XmaCI]

Cuts 2 times.

Cuts at positions 692, 4272.

Fragment sizes 3580, 7275.

**SnaBI** (TAC'GTA) [BstSNI,Eco105I]

Cuts 2 times.

Cuts at positions 6351, 6393.

Fragment sizes 42, 10813.

**SpeI** (A'CTAG\_T) [AhlI,BcuI]

Cuts 1 time.

Cuts at position 64.

**Tth111I** (GACn'n\_nGTC) [AspI,PflFI,PsyI]

Cuts 2 times.

Cuts at positions 1064, 8997.

Fragment sizes 7933, 2922.

**XhoI** (C'TCGA\_G) [PaeR7I,Sfr274I,SlaI,StrI,TliI]

Cuts 2 times.

Cuts at positions 31, 5076.

Fragment sizes 5045, 5810.

**XmaI** (C'CCGG\_G) [SmaI,Cfr9I,TspMI,XmaCI]

Cuts 2 times.

Cuts at positions 690, 4270.

Fragment sizes 3580, 7275.

**XmnI** (GAAnn'nnTTC) [Asp700I, MroXI, PdmI]  
Cuts 1 time.  
Cuts at position 6085.

**ZraI** (GAC'GTC) [AatII]  
Cuts 1 time.  
Cuts at position 4781.