

NheI  
~~~~~

1 TGAAAGACCC CACCTGTAGG TTTGGCAAGC TAGCTTAAGT AACGCCATTT TGCAAGGCAT GGAAAATACA TAACTGAGAA TAGAGAAGTT CAGATCAAGG  
ACTTTCTGGG GTGGACATCC AAACCGTTCG ATCGAATTCA TTGCGGTAAT ACGTTCCGTA CCTTTTATGT ATTGACTCTT ATCTCTTCAA GTCTAGTTCC

EcoRV  
~~~~~

101 TTAGGAACAG AGAGACAGCA GAATATGGGC CAAACAGGAT ATCTGTGGTA AGCAGTTCCT GCCCCGGCTC AGGGCCAAGA ACAGATGGTC CCCAGATGCG  
AATCCTTGTC TCTCTGTCGT CTTATAACCCG GTTGTGTCCTA TAGACACCAT TCGTCAAGGA CGGGGCCGAG TCCCGGTTCT TGTCTACCAG GGGTCTACGC

201 GTCCCGCCCT CAGCAGTTTC TAGAGAACCA TCAGATGTTT CCAGGGTGCC CCAAGGACCT GAAAATGACC CTGTGCCTTA TTTGAACTAA CCAATCAGTT  
CAGGGCGGGA GTCGTCAAAG ATCTCTTGGT AGTCTACAAA GGTCCCACGG GGTTCCTGGA CTTTTACTGG GACACGGAAT AAACCTTGATT GGTTAGTCAA

SacI  
~~~~~

301 CGCTTCTCGC TTCTGTTTCG CCGCTTCTGC TCCCCGAGCT CAATAAAAGA GCCACAACC CCTCACTCGG CGCGCCAGTC CTCCGATAGA CTGCGTCCGC  
GCGAAGAGCG AAGACAAGCG CCGAAGACG AGGGGCTCGA GTTATTTTCT CGGGTGTTGG GGAGTGAGCC GCGCGGTCAG GAGGCTATCT GACGCAGCGG

KpnI  
~~~~~

401 CGGGTACCCG TATTCCCAAT AAAGCCTCTT GCTGTTTGCA TCCGAATCGT GGACTCGCTG ATCCTTGGGA GGGTCTCCTC AGATTGATTG ACTGCCACC  
GCCCATGGGC ATAAGGGTTA TTTCCGGAGAA CGACAAACGT AGGCTTAGCA CCTGAGCGAC TAGGAACCCCT CCCAGAGGAG TCTAACAAAC TGACGGGTGG

501 TCGGGGGTCT TTCATTTGGA GGTTCACCG AGATTTGGAG ACCCCTGCCT AGGGACCACC GACCCCCCG CCGGGAGGTA AGCTGGCCAG CGGTCTGTTT  
AGCCCCAGA AAGTAAACCT CCAAGGTGGC TCTAAACCTC TGGGGACGGA TCCCTGGTGG CTGGGGGGGC GGCCCTCCAT TCGACCGGTC GCCAGCAAAG

601 GTGTCTGTCT CTGTCTTTGT GCGTGTGTTGT GCCGGCATCT AATGTTTGGC CCTGCGTCTG TACTAGTTAG CTAAGTAGCT CTGTATCTGG CGGACCCGTG  
CACAGACAGA GACAGAAACA CGCACAAACA CGCCCGTAGA TTACAAACGC GGACGCAGAC ATGATCAATC GATTGATCGA GACATAGACC GCCTGGGCAC

701 GTGGAAGTGA CGAGTCTGA ACACCCGGCC GCAACCCTGG GAGACGTCCC AGGGAAGTTG GGGGCCGTTT TTGTGGCCCG ACCTGAGGAA GGGAGTTCGAT  
CACCTTGACT GCTCAAGACT TGTGGGCCGG CGTFTGGACC CTCTGCAGG TCCCTGAAAC CCCCGGCAA AACACCGGGC TGGACTCCTT CCCTCAGCTA

801 GTGGAATCCG ACCCCGTCAG GATATGTGGT TCTGGTAGGA GACGAGAACC TAAAACAGTT CCCGCTCCG TCTGAATTTT TGCTTTCCGT TTGGAACCGA  
CACCTTAGGC TGGGGCAGTC CTATACACCA AGACCATCCT CTGCTCTTGG ATTTTGTCAA GGGCGGAGGC AGACTTAAAA ACGAAAGCCA AACCTTGCTT

901 AGCCGCGCGT CTTGTCTGCT GCAGCGCTGC AGCATCGTTC TGTGTTGTCT CTGTCTGACT GTGTTTCTGT ATTTGTCTGA AAATTAGGGC CAGACTGTTA  
TCGGCGCGCA GAACAGACGA CGTCGCGACG TCGTAGCAAG ACACAACAGA GACAGACTGA CACAAAGACA TAAACAGACT TTTAATCCCG GTCTGACAAT

1001 CCACTCCCTT AAGTTTGACC TTAGGTCACT GGAAAGATGT CGAGCGGATC GCTCACAACC AGTCGGTAGA TGTCAAGAAG AGACGTTGGG TTACCTTCTG  
GGTGAGGGAA TTCAAACTGG AATCCAGTGA CCTTTCTACA GCTCGCCTAG CGAGTGTTGG TCAGCCATCT ACAGTTCTTC TCTGCAACCC AATGGAAGAC

1101 CTCTGCAGAA TGGCCAACCT TTAACGTCGG ATGGCCGCGA GACGGCACCT TTAACCGAGA CCTCATCACC CAGGTTAAGA TCAAGGTCTT TTCACCTGGC  
GAGACGTCTT ACCGGTTGGA AATTCGAGCC TACCGCGCCT CTGCCGTGGA AATTGGCTCT GGAGTAGTGG GTCCAATTCT AGTTCCAGAA AAGTGGACCG

1201 CCGCATGGAC ACCCAGACCA GGTCGCCCTAC ATCGTGACCT GGGAAAGCCT GGCTTTTGAC CCCCTCCCT GGGTCAAGCC CTTTGTACAC CCTAAGCCTC  
GGCGTACCTG TGGGTCTGGT CCAGGGGATG TAGCACTGGA CCCTTCGGAA CCGAAAACCTG GGGGGAGGGA CCCAGTTCGG GAAACATGTG GGATTCGGAG

1301 CGCCTCCTCT TCCTCCATCC GCCCGTCTC TCCCCCTTGA ACCTCCTCGT TCGACCCCG CTGATCCTC CTTTTATCCA GCCCTCCTC CTTCTTAGG  
GCGGAGGAGA AGGAGGTAGG CGGGGCAGAG AGGGGGAAC TGGAGGAGCA AGCTGGGGCG GAGCTAGGAG GAAAATAGGT CCGGAGTGAG GAAGAGATCC

EcoRI  
~~~~~

1401 CGCCGGAATT CGGCCCCCC CCCCTTGCCG CATCCACGAA ACTTTGCCCA TAGCAGCGGG CGGGCACTTT GCACTGGAAC TTACAACACC CGAGCAAGGA  
GCGGCCTTAA GCCGGGGGG GGGGAACGGC GTAGGTGCTT TGAAACGGGT ATCGTCGCCC GCCCGTGAAA CGTGACCTTG AATGTTGTGG GCTCGTTCCT

1501 CGCGACTCTC CCGACGCGGG GAGGCTATTC TGCCATTTG GGGACACTTC CCCGCGCTG CCAGGACCCG CTTCTCTGAA AGGCTCTCCT TGCAGCTGCT  
GCGCTGAGAG GGCTGCGCCC CTCGATAAG ACGGGTAAAC CCTGTGTAAG GGGCGGCGAC GGTCTGGGC GAAGAGACTT TCCGAGAGGA ACGTCCGACG

1601 TAGACGCTGG ATTTTTTTTC GGTTAGTGAA AACCAGCCTC CCGCGACGAT GCCCTCAAC GTTAGCTTCA CCAACAGGAA CTATGACCTC GACTACGACT  
ATCTGCGACC TAAAAAAGC CCATCACCTT TTGGTCCGAG GCGCTGCTA CGGGGAGTTG CAATCGAAGT GGTGTCTCTT GATACTGGAG CTGATGCTGA

EcoRV

~~~~~

1701 CGGTGCAGCC GTATTTCTAC TGCACGAGG AGGAGAACTT CTACCAGCAG CAGCAGCAGA GCGAGCTGCA GCCCCCCGGCG CCCAGCGAGG ATATCTGGAA  
 GCCACGTCGG CATAAAGATG ACGCTGTCTC TCCTCTTGAA GATGGTCTGT GTCGTCTGCT CGTCTGACGT CGGGGGCCGC GGGTCTGCTC TATAGACCTT

1801 GAAATTCGAG TAGCTGCCCC CCCC GCCCT GTCCTCTGCT CGCCCTCTAC GCCCTCTCTC CACCCTTCTC CCTTCGGGTA  
 CTTTTAAGCTC CACGACGGGT GGGCGGGGA CAGGGGATCG GGGCGAGGC CCGAGACGAG CGGGAGGATG CAACGCCAGT GTGGGAAGAG GGAAGCCCTC

1901 GACAACGACG GCGGTGGCGG GAGCTTCTCC ACGGCCGACC AGCTGGAGAT GGTGACCGAG CTGCTGGGAG GAGACATGGT GAACCAGAGT TTCATCTGCG  
 CTGTTGCTGC CGCCACCGCC CTCGAAGAGG TGCCGGCTGG TCGACCTCTA CCACTGGCTC GACGACCTC CTCTGTACCA CTTGGTCTCA AAGTAGACGC

2001 ACCCGGACGA CGAGACCTTC ATCAAAAACA TCATCATCCA GGAAGTGTATG TGGAGCGGCT TCTCGGCCGC CGCCAAGCTC GTCTCAGAGA AGCTGGCCTC  
 TGGGCTGTCT GCTCTGGAAG TAGTTTTTGT AGTAGTAGGT CCTGACATAC ACCTCGCCGA AGAGCCGGCG GCGGTTTCGAG CAGAGTCTCT TCGACCGGAG

2101 CTACCAGGCT GCGCGCAAAG ACAGCGGCAG CCCGAACCCC GCCCGCGGG ACAGCGTCTG CTCCACCTCC AGCTTGTACC TGCAGGATCT GAGCGCCGCC  
 GATGGTCCGA GCGCGTTC TGTCGCCGTC GGGCTTGGGG CGGGCGCCGG TGTCGCAGAC GAGGTGGAGG TCGAACATGG ACGTCTTAGA CTCGCGGCGG

2201 GCCTCAGAGT GCATCGACCC CTCGGTGGTC TTTCCCTACC CTCTCAACGA CAGCAGCTCG CCCAAGTCTT GCGCCTCGCA AGACTCCAGC GCCTTCTCTC  
 CGGAGTCTCA CGTAGCTGGG GAGCCACCAG AAGGGGATGG GAGAGTTGCT GTCGTGAGC GGGTTCAGGA CGCGGAGCGT TCTGAGGTCG CGGAAGAGAG

2301 CGTCTCGGA TTCTCTGCTC TCCTCGACGG AGTCTCTCCC GCAGGGCAGC CCCGAGCCCC TGGTGTCTCA TGAGGAGACA CCGCCCACCA CCAGCAGCGA  
 GCAGGAGCCT AAGAGACGAG AGGAGCTGCC TCAGGAGGGG CGTCCCCTCG GGGCTCGGGG ACCACGAGGT ACTCCTCTGT GCGGGGTGTT GGTCGTCTGT

2401 CTCTGAGGAG GAACAAGAAG ATGAGGAAGA AATCGATGTT GTTCTGTGG AAAAGAGGCA GGCTCCTGGC AAAAGGTCAG AGTCTGGATC ACCTTCTGCT  
 GAGACTCCTC CTTGTCTTTC TACTCCTTCT TTAGCTACAA CAAAGACACC TTTTCTCCGT CCGAGGACCC TTTTCCAGT TCAGACCTAG TGGAAGACGA

2501 GGAGGCCACA GCAAACCTCC TCACAGCCCA CTGGTCTCTA AGAGGTGCCA CGTCTCCACA CATCAGCACA ACTACGCAGC ACTCGGAGG GCCTCGGAGG  
 CCTCCGGTGT CGTTTGGAGG AGTGTCCGGT GACCAGGAGT TCTCCACGGT GCAGAGGTGT GTAGTCGTGT TGATGCGTCG CGGAGGGAGG TGAGCCTTCC

2601 ACTATCCTGC TGCCAAGAGG GTCAAGTTGG ACAGTGTGAG AGTCTGAGA CAGATCAGCA ACAACCGAAA ATGCACCAGC CCCAGGTCCT CGGACACCGA  
 TGATAGGACG ACGGTCTCTC CAGTTC AAC TGTACAGTC TCAGGACTCT GTCTAGTCGT TGTTGGCTTT TACGTGGTCG GGGTCCAGGA GCCTGTGGCT

2701 GGAGAAATGTC AAGAGCGCAA CACACAACGT CTTGGAGCGC CAGAGGAGGA ACGAGCTAAA ACGGAGCTTT TTTGCCCTGC GTGACCAGAT CCCGGAGTTG  
 CCTCTTACAG TTCTCCGCTT GTGTGTTGCA GAACCTCGCG GTCTCCTCCT TGCTCGATTT TGCCTCGAAA AAACGGGACG CACTGGTCTA GGGCCTCAAC

2801 GAAAACAATG AAAAGGCCCC CAAGGTAGTT ATCCTTAAAA AAGCCACAGC ATACATCTG TCCGTCCAAG CAGAGGAGCA AAAGCTCAT TCTGAAGAGG  
 CTTTTGTAC TTTTCCGGGG GTTCCATCAA TAGGAATTTT TTCGGTGTG TATGTAGGAC AGGCAGGTTT GTCTCCTCGT TTTTCGAGTAA AGACTTCTCC

2901 ACTTGTGTCG GAAACGACGA GAACAGTTGA AACACAAACT TGAACAGCTA CGGAACTCTT GTGCGTAAGG AAAAGTAAGG AAAATGATTC CTTCTAACAG  
 TGAACAACGC CTTTGTCTGCT CTTGTCAACT TTGTGTTTGA ACTTGTGTCGAT GCCTTGAGAA CACGCATTCC TTTTCATTCC TTTTACTAAG GAAGATTGTC

3001 AAATGTCCTG AGCAATCACC TATGAACTTG TTTCAAATGC ATGATCAAAT GCAACCTCAC AACCTTGGCT GAGTCTTGAG ACTGAAAGAT TTAGCCATAA  
 TTTACAGGAC TCGTTAGTGG ATACTTGAAC AAAGTTTACG TACTAGTTTA CGTTGGAGTG TTGGAACCGA CTCAGAATC TGACTTTCTA AATCGGTATT

3101 TGTAAGTGC CTCAAATTGG ACTTTGGGCA TAAAAGAACT TTTTATGCT TACCATCTTT TTTTTTCTT TAACAGATTT GTATTTAAGA ATTTGTTTTA  
 ACATTTGACG GAGTTTAACC TGAACCCCGT ATTTCTTTGA AAAAATACGA ATGGTAGAAA AAAAAAGAAA ATTTGTCTAAA CATAAATTCT TAACAAAAAT

EcoRI XhoI  
 ~~~~~ ~~~~~

3201 AAAAATTTTA AGATTACAC AATGTTTCTC TGTAATATAT GCCATTAAAT GTAAATAACT TCCGAATTCG TTAACCTCGA GCGGGATCAA TTCCGCCCCC  
 TTTTTAAAAAT TCTAAATGTG TTACAAAGAG ACATTTATAA CGGTAATTTA CATTTATTGA AGGCTTAAGC AATTGGAGCT CGCCCTAGTT AAGGCGGGGG

3301 CCCCTAACGT TACTGGCCGA AGCCGCTTGG AATAAGGCCG GTGTGCGTTT GTCTATATGT TATTTTCCAC CATATTGCCG TCTTTTGGCA ATGTGAGGGC  
 GGGGATTGCA ATGACCGGT TCGGCGAACC TTATTTCCGG CACACGCAA CAGATATACA ATAAAAGGTG GTATAACGGC AGAAAACCGT TACACTCCCG

3401 CCGGAAACT GGCCCTGTCT TCTTGACGAG CATTCCTAGG GTTCTTTCCC CTCTCGCCAA AGGAATGCAA GGTCTGTTGA ATGTCGTGAA ATGAGCAGTT  
 GGCCTTTGGA CCGGGACAGA AGAAGTGTCT GTAAGGATCC CCAGAAAAGG GAGAGCGGTT TCCTTACGTT CCAGACAACT TACAGCACTT CCTTCGTCAA

HindIII  
 ~~~~~

3501 CCTCTGGAAG CTTCTGGAAG ACAACAACG TCTGTAGCGA CCCTTTGCAG GCAGCGGAAC CCCCCACCTG GCGACAGGTG CCTCTGCGGC CAAAAGCCAC  
 GGAGACCTTC GAAGAACTTC TGTTTGTGTC AGACATCGCT GGGAAACGTC CGTCGCCTTG GGGGGTGGAC CGCTGTCCAC GGAGACGCCG GTTTTCGGTG

3601 GTGTATAAGA TACACCTGCA AAGGCGGCAC AACCCAGTG CCACGTTGTG AGTTGGATAG TTGTGGAAAG AGTCAAATGG CTCTCTCAA GCGTATTCAA  
 CACATATTCT ATGTGGACGT TTCCGCCGTG TTGGGGTCAC GGTGCAACAC TCAACCTATC AACACCTTTC TCAGTTTACC GAGAGGAGTT CGCATAAGTT

KpnI  
~~~~~

3701 CAAGGGGCTG AAGGATGCC AGAAGGTACC CCATTGTATG GGATCTGATC TGGGGCCTCG GTGCACATGC TTTACATGTG TTTAGTCGAG GTTAAAAAAC  
GTTCCCCGAC TTCCTACGGG TCTTCCATGG GGTAACATA CCTAGACTAG ACCCCGGAGC CACGTGTACG AAATGTACAC AAATCAGCTC CAATTTTTTTG

NcoI  
~~~~~

3801 GTCTAGGCC CCCGAACCAC GGGGACGTGG TTTTCCTTTG AAAAAACACGA TAATACCATG GTGAGCAAGG GCGAGGAGCT GTTCACCGGG GTGGTGCCCA  
CAGATCCGGG GGGCTTGGTG CCCCTGCACC AAAAGGAAAC TTTTGTGTCT ATTATGGTAC CACTCGTTCC CGCTCCTCGA CAAGTGGCCC CACCACGGGT

3901 TCCTGGTCGA GCTGGACGGC GACGTAAACG GCCACAAGTT CAGCGTGTCC GGCGAGGGCG AGGGCGATGC CACCTACGGC AAGCTGACCC TGAAGTTCAT  
AGGACCAGCT CGACCTGCCG CTGCATTTGC CCGTGTTCAG GTCGCACAGG CCGCTCCC GC TCCCGCTACG GTGGATGCCG TTCGACTGGG ACTTCAAGTA

4001 CTGCACCACC GGCAAGCTGC CCGTGCCCTG GCCACCCTC GTGACCACC TGACCTACGG CGTGCAGTGC TTCAGCCGCT ACCCCGACCA CATGAAGCAG  
GACGTGGTGG CCGTTCGACG GGCACGGGAC CGGGTGGGAG CACTGGTGGG ACTGGATGCC GCACGTACG AAGTCGGCGA TGGGGCTGGT GTACTTCGTC

4101 CACGACTTCT TCAAGTCCGC CATGCCCGAA GGCTACGTCC AGGAGCGCAC CATCTTCTTC AAGGACGACG GCAACTACAA GACCCGCGCC GAGGTGAAGT  
GTGCTGAAGA AGTTCAGGCG GTACGGGCTT CCGATGCAGG TCCTCGCGTG GTAGAAGAAG TTCCTGCTGC CGTTGATGTT CTGGGCGCGG CTCCACTTCA

4201 TCGAGGGCGA CACCTGGTG AACCGCATCG AGCTGAAGGG CATCGACTTC AAGGAGGACG GCAACATCCT GGGGCACAAG CTGGAGTACA ACTACAACAG  
AGCTCCCCTG GTGGGACCAC TTGGCGTAGC TCGACTTCCC GTAGCTGAAG TTCCTCCTGC CGTTGTAGGA CCCCCTGTTC GACCTCATGT TGATGTTGTC

4301 CCACAACGTC TATATCATGG CCGACAAGCA GAAGAACGGC ATCAAGGTGA ACTTCAAGT CCGCCACAAC ATCGAGGACG GCAGCGTGCA GCTCGCCGAC  
GGTGTTCGAG ATATAGTACC GGCTGTTCGT CTCTTGCCG TAGTCCACT TGAAGTTCTA GCGGTGTGTG TAGCTCCTGC CGTCGCACGT CGAGCGGCTG

4401 CACTACCAGC AGAACACCCC CATCGGCAC GGGCCCGTGC TGCTGCCCCA CAACCACTAC CTGAGCACCC AGTCCGCCCT GAGCAAAGAC CCCAACGAGA  
GTGATGGTGG TCTTGTGGGG GTAGCCGCTG CCGGGGCACG ACGACGGGCT GTTGGTGATG GACTCGTGGG TCAGGCGGGA CTCGTTTTCTG GGGTTGCTCT

NotI  
~~~~~

4501 AGCGCGATCA CATGGTCCTG CTGGAGTTCTG TGACCGCCGC CGGGATCACT CTCGGCATGG ACGAGCTGTA CAAGTAAAGC GGCCGCGACT CTAGAGTCGA  
TCGCGCTAGT GTACCAGGAC GACCTCAAGC ACTGGCGGGC GCCCTAGTGA GAGCCGTACC TGCTCGACAT GTTCATTTCTG CCGGCGCTGA GATCTCAGCT

HindIII KpnI NheI EcoRV  
~~~~~

4601 CCTGCAGGCA TGCAAGCTTC AGGTAGCCGG CTAACGTTAA CAACCGGTAC CTCTAGAACT ATAGCTAGCA TGCGCAAATT TAAAGCGCTG ATATCGATAA  
GGACGTCCGT ACGTTCGAAG TCCATCGGCC GATTGCAATT GTTGGCCATG GAGATCTTGA TATCGATCGT ACGCGTTTAA ATTTCCGCGAC TATAGCTATT

NheI  
~~~~~

4701 AATAAAAGAT TTTATTTAGT CTCCAGAAAA AGGGGGGAAT GAAAGACCCC ACCTGTAGGT TTGGCAAGCT AGCTTAAGTA ACGCCATTTT GCAAGGCATG  
TTATTTTCTA AAATAAATCA GAGGTCTTTT TCCCCCTTA CTTTCTGGGG TGGACATCCA AACCGTTCGA TCGAATTCAT TGCGGTAAAA CGTTCCGTAC

EcoRV  
~~~~~

4801 GAAAATACAT AACTGAGAAT AGAGAAGTTC AGATCAAGGT TAGGAACAGA GAGACAGCAG AATATGGGCC AAACAGGATA TCTGTGGTAA GCAGTTCCTG  
CTTTTATGTA TTGACTCTTA TCTCTTCAAG TCTAGTTCCA ATCCTTGTCT CTCTGTGCTC TTATACCCGG TTTGTCCTAT AGACACCATT CGTCAAGGAC

4901 CCCC GGCTCA GGGCCAAGAA CAGATGGTCC CCAGATGCGG TCCCGCCCTC AGCAGTTTCT AGAGAACCAT CAGATGTTTC CAGGGTGCCC CAAGGACCTG  
GGGGCCGAGT CCCGTTCTT GTCTACCAGG GGTCTACGCC AGGGCGGGAG TCGTCAAAGA TCTCTTGGTA GTCTACAAAG GTCCCACGGG GTTCTTGAC

SacI  
~~~~~

5001 AAAATGACCC TGTGCCTTAT TTGAACATA CAATCAGTTC GCTTCTCGCT TCTGTTCGCG CGCTTCTGCT CCCCAGGCTC AATAAAAGAG CCCACAACCC  
TTTTACTGGG ACACGGAATA AACTTGATTG GTTAGTCAAG CGAAGAGCGA AGACAAGCGC GCGAAGACGA GGGGCTCGAG TTATTTTCTC GGGTGTGGG

KpnI  
~~~~~

5101 CTCACTCGGC GCGCCAGTCC TCCGATAGAC TGCGTCGCC GGTACCCGT GTATCCAATA AACCTCTTG CAGTTGCATC CGACTTGTGG TCTCGCTGTT

5201 GAGTGAGCCG CGCGGTCAGG AGGCTATCTG ACGCAGCGGG CCCATGGGCA CATAGGTTAT TTGGGAGAAC GTCAACGTAG GCTGAACACC AGAGCGACAA  
CCTTGGGAGG GTCTCCTCTG AGTGATTGAC TACCCGTCAG CCGGGGTCTT TCATGGGTAA CAGTTTCTTG AAGTTGGAGA ACAACATTCT GAGGGTAGGA  
5301 GGAACCCCTCC CAGAGGAGAC TCACTAAGTGC ATGGGCAGTC GCCCCAGAA AGTACCCATT GTCAAAGAAT TTCAACCTCT TGTTGTAAGA CTCCCATCCT  
GTCTGAATATT AAGTAATCCT GACTCAATTA GCCACTGTTT TGAATCCACA TACTCCAATA CTCCGTAATC AGTTCATTAT GGACAGCGCA GAAAGATCTG  
CAGCTTATAA TTCATTAGGA CTGAGTTAAT CGGTGACAAA AXTTAGGTAT ATGAGGTTAT GAGGCATTTA TCAAGTAATA CCTGTTCGGT GTTTCTCGAC  
5401 GGGAGAATTG TGAATTGTT ATCCGCTCAC AATTCACAC AACATACGAG CCGGAAGCAT AAAGTGATAA GCCTGGGGTG CCTAATGAGT GAGCTAACTC  
CCCTCTAAC ACTTTAACAA TAGGCGAGTG TTAAGGTGTG TTGTATGCTC GGCCTTCGTA TTTACATTT CGGACCCAC GGATTACTCA CTCGATTGAG  
5501 ACATTAATTG CGTTGCGCTC ACTGCCCCTG TCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA TCGGCCAACG CGCGGGGAGA GCGGGTTTG  
TGTAATTAAC GCAACGCGAG TGACGGGCGA AAGGTCAGCC CTTTGGACAG CACGGTTCGAC GTAATTACTT AGCCGGTTGC GCGCCCCTCT CCGCCAAACG  
5601 GTATTGGGGC CTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTG GTTCGGCTG GCGGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT  
CATAACCCGC GAGAAGGCGA AGGAGCGAGT GACTGAGCGA CGCGAGCCAG CAAGCCGACG CCGCTCGCCA TAGTTCGAGT AGTTTCCGCC ATTAGCCAA  
5701 ATCCACAGAA TCAGGGGATA ACCCAGGAAA GAACATGTGA GCAAAAAGGCC AGCAAAAAGGC CAGGAACCGT AAAAAGGCGC CGTTGCTGCG GTTTTTCCAT  
TAGGTGTCTT AGTCCCCTAT TGCGTCTTTT CTTGTACACT CGTTTTCCGG TCGTTTTCCG GTCTTTGGCA TTTTTCCGGC GCAACGACCG CAAAAGGTA  
5801 AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCCTGAA  
TCCGAGGCGG GGGGACTGCT CGTAGTGTTT TTAGCTGCGA GTTCAGTCTC CACCGCTTTG GGCTGTCTGT ATATTTCTAT GGTCCGCAA GGGGGACCTT  
5901 GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA GCTCACGCTG  
CGAGGGAGCA CGCGAGAGGA CAAGGCTGGG ACGGCGAATG GCCTATGGAC AGGCGGAAAG AGGGAAGCCC TTTCGACCCG GAAAGAGTAT CGAGTGGCAG  
6001 TAGGTATCTC AGTTCGGTGT AGGTCGTTCT GCTCAAGCTG CTCCAAGCTG ACGACCCCC CGTTCAGCCC GACCGCTGCG CTTATCCGG TAACATGCGT  
ATCCATAGAG TCAAGCCACA TCCAGCAAGC GAGGTTTCGAC CCGACACACG TGCTTGGGGG GCAAGTCCGG CTGGCGACGC GGAATAGGCC ATTAGTAGCA  
6101 CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT  
GAACTCAGGT TGGGCCATTG TGTGCTGAAT AGCGGTGACC GTCGTGCGTG ACCATTGTCC TAATCGTCTC GCTCCATACA TCCGCCACGA TGTCTCAAGA  
6201 TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTGT GTAGCTCTTG  
ACTTCACCAC CGGATTTGAT CCGATGTGAT CTTCTGTGTA TAAACCATAG ACGCGAGACG ACTTCCGGTCA ATGGAAGCCT TTTTCTCAAC CATCGAGAAC  
6301 ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT  
TAGGCCGTTT GTTTGGTGGC GACCATCGCC ACCAAAAAAA CAAACGTTCTG TCGTCTAATG CCGCTCTTTT TTTCCCTAGAG TTCTTCTAGG AAACATAGAAA  
6401 TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTA  
AGATGCCCA GACTGCGAGT CACCTTGCTT TTGAGTGCAA TTCCCTAAAA CCAGTACTCT AATAGTTTTT CCTAGAAGTG GATCTAGGAA AATTTAATTT  
6501 AATGAAGTTT TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT  
TTACTTCAAA ATTTAGTTAG ATTTTCATATA TACTCATTTG AACCAGACTG TCAATGGTTA CGAATTAGTC ACTCCGTGGA TAGAGTCCGT AGACAGATAA  
6601 TCGTTCATCC ATAGTTGCTT GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCACTGTCG CCCAGTGCCT CAATGATACC GCGAGACCCA  
AGCAAGTAGG TATCAACGGA CTGAGGGGCA GCACATGTAT TGATGCTATG CCCTCCCGAA TGGTAGACCG GGTACGCTAC GTTACTATGG CGTCTGGGT  
6701 CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG CCGGAAGGGC CGAGCGCAGA AGTGGTCCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA  
GCGAGTGGCC GAGGTCTAAA TAGTCGTTAT TTGGTTCGGT GGCCTTCCCG GCTCGCGTCT TCACCAGGAC GTTGAAATAG GCGGAGGTAG GTCAGATAAT  
6801 ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGCCT AACGTTGTTG CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTTGG  
TAACAACGGC CCTTCGATCT CATTCATCAA GCGGTCAATT ATCAAACGCG TTGCAACAAC GGTAACGATG TCCGTAGCAC CACAGTGCAG GCAGCAAACC  
6901 TATGGCTTCA TTCAGCTCCG GTTCCCAACG ATCAAGGCGA GTTACATGAT CCCCATGTT GTGCAAAAAA GCGGTTAGCT CTTTCGGTCC TCCGATCGTT  
ATACCGAAGT AAGTCGAGGC CAAGGGTTGC TAGTTCGGCT CAATGTACTA GGGGGTACAA CAGTTTTTTT CGCCAATCGA GGAAGCCAGG AGGCTAGCAA  
7001 GTCAGAAATA AGTTGGCCGC AGTGTATCA CTCTATGGTTA TGGCAGCACT GCATAATTCT TGTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG  
CAGTCTTCAT TCAACCGCGC TCACAATAGT GAGTACCAAT ACCGTCGTGA CGTATTAAGA GAATGACAGT ACGGTAGGCA TTCTACGAAA AGACACTGAC  
7101 GTGAGTACTC AACCAAGTCA TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTTGCC CGGCGTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC  
CACTCATGAG TTGGTTCAGT AAGACTCTTA TCACATACGC CGCTGGCTCA ACGAGAACGG GCCGCAGTTA TGCCCTATTA TGCGCGGGTG TATCGTCTTG  
7201 TTTAAAAGTG CTCATCATTG GAAAACGTTT TTCGGGGCGA AACTCTCAA GGATCTTACC GCTGTTGAGA TCCAGTTCGA TGTAACCCAC TCGTGCACCC  
AAATTTTCAC GAGTAGTAAC CTTTTGCAAG AAGCCCCGCT TTTGAGAGTT CCTAGAATGG CGACAACCTCT AGGTCAAGCT ACATTTGGGTG AGCACGTTGGG  
7301 AACTGATCTT CAGCATCTTT TACTTTCACC AGCGTTTCTG GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAGGG AATAAGGGCG ACACGGAAAT  
TTGACTAGAA GTCGTAGAAA ATGAAAGTGG TCGCAAAGAC CCACTCGTTT TTGTCTTCC GTTTTACGGC GTTTTTTCCC TTATTTCCCGC TGTGCTTTA

7401 GTTGAATACT CATACTCTTC CTTTTTCAAT ATTATTGAAG CATTTATCAG GGTATTGTGTC TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA  
CAACTTATGA GTATGAGAAG GAAAAAGTTA TAATAACTTC GTAAATAGTC CCAATAACAG AGTACTCGCC TATGTATAAA CTTACATAAA TCTTTTTATT  
7501 ACAAATAGGG GTTCCGCGCA CATTTCCCGG AAAAGTGCCA CCTGACGTCT AAGAAACCAT TATTATCATG ACATTAACCT ATAAAAATAG GCGTATCACG  
TGTTTATCCC CAAGGCGCGT GTAAAGGGGG TTTTCACGGT GGACTGCAGA TTCTTTGGTA ATAATAGTAG TGTAATTGGA TATTTTTATC CGCATAGTGC  
7601 AGGCCCTTTC GTCTCGCGCG TTTCCGGTGAT GACGGTGAAA ACCTCTGACA CATGCAGCTC CCGGAGACGG TCACAGCTTG TCTGTAAGCG GATGCCGGGA  
TCCGGGAAAAG CAGAGCGCGC AAAGCCACTA CTGCCACTTT TGGAGACTGT GTACGTCGAG GGCCTCTGCC AGTGTGCGAAC AGACATTCGC CTACGGCCCT  
7701 GCAGACAAGC CCGTCAGGGC GCGTCAGCGG GTGTTGGCGG GTGTCGGGGC TGGCTTAACT ATGCGGCATC AGAGCAGATT GTACTGAGAG TGCACCATAT  
CGTCTGTTCG GGCAGTCCCG CGCAGTCGCC CACAACCGCC CACAGCCCCG ACCGAATTGA TACGCCGTAG TCTCGTCTAA CATGACTCTC ACGTGGTATA  
7801 GCGGTGTGAA ATACCGCACA GATGCGTAA GAGAAAATAC CGCATCAGGC GCCATTTCGCC ATTCAGGCTG CGCAACTGTT GGGGAGGGCG ATCGGTGCGG  
CGCCACACTT TATGGCGTGT CTACGCATTC CTCTTTTATG GCGTAGTCCG CGGTAAGCGG TAAGTCCGAC GCGTTGACAA CCCTTCCCGC TAGCCACGCC  
7901 GCCTCTTCGC TATTACGCCA GCTGGCGAAA GGGGGATGTG CTGCAAGGCG ATTAAGTTGG GTAACGCCAG GGTTTTCCCA GTCACGACGT TGTA AAAACGA  
CGGAGAAGCG ATAATGCGGT CGACCGCTTT CCCCTACAC GACGTTCCGC TAATTCAACC CATTGCGGTC CCAAAAAGGGT CAGTGCTGCA ACATTTTGCT  
8001 CGGCCAGTGC CANNNNCGCT CTCCTTATG C GACTCCTGC ATTAGGAAGC AGCCCAGTAG TAGGTTGAGG CCGTTGAGCA CCGCCGCCG AAGGAATGGT  
GCCGTCACG GTNNNNCGCA GAGGGAATAC GCTGAGGACG TAATCCTTCG TCGGGTCATC ATCCAAC TCCG GCAACTCGT GCGCGCGGCG TTCCTTACCA  
8101 GCATGCAAGG AGATGGCGCC CAACAGTCCC CCGCCACGG GGCCTGCCAC CATACCCACG CCGAAAACAAG CGCTCATGAG CCCGAAGTGG CGAGCCCGAT  
CGTACGTTC TCTACCGCGG GTTGTCAGGG GGC CGGTGCC CCGGACGGTG GTATGGGTGC GGCTTTGTTC GCGAGTACTC GGGCTTCACC GCTCGGGCTA  
8201 CTCCCCATC GGTGATGTCG GCGATATAGG CGCCAGCAAC CGCACCTGTG GCGCCGGTGA TGCCGGCCAC GATGCGTCCG GCGTAGAGGC GATTAGTCCA  
GAAGGGGTAG CCACTACAGC CGCTATATCC GCGGTCGTTG GCGTGACAC CGCGGCCACT ACGGCCGGTG CTACGCAGGC CGCATCTCCG CTAATCAGGT  
EcoRV  
~~~~~  
8301 ATTTGTTAAA GACAGGATAT CAGTGGTCCA GGCTCTAGTT TTGACTCAAC AATATCACCA GCTGAAGCCT ATAGAGTACG AGCCATAGAT AAAATAAAAG  
TAAACAATTT CTGTCCTATA GTCACCAGGT CCGAGATCAA AACTGAGTTG TTATAGTGGT CGACTTCGGA TATCTCATGC TCGGTATCTA TTTTATTTTC  
8401 ATTTTATTTA GTCTCCAGAA AAAAAAAAAA A  
TAAAATAAAT CAGAGGTCTT TTTTTTTTTT T