```
       ScaI
   1   AGTACTTTGA TCCAACCCCT CCGCTGCTAT AGTGCAGTCG GCTTCTGACG TTCAGTGCAG CCGTCTTCTG AAAACGACAT GTCGCACAAG TCCTAAGTTA
       TCATGAAACT AGGTTGGGGA GGCGACGATA TCACGTCAGC CGAAGACTGC AAGTCACGTC GGCAGAAGAC TTTTGCTGTA CAGCGTGTTC AGGATTCAAT
 101   CGCGACAGGC TGCCGCCCTG CCCTTTTCCT GGCGTTTTCT TGTCGCGTGT TTTAGTCGCA TAAAGTAGAA TACTTGCGAC TAGAACCGGA GACATTACGC
       GCGCTGTCCG ACGGCGGGAC GGGAAAAGGA CCGCAAAAGA ACAGCGCACA AAATCAGCGT ATTTCATCTT ATGAACGCTG ATCTTGGCCT CTGTAATGCG
 201   CATGAACAAG AGCGCCGCCG CTGGCCTGCT GGGCTATGCC CGCGTCAGCA CCGACGACCA GGACTTGACC AACCAACGGG CCGAACTGCA CGCGGCCGGC
       GTACTTGTTC TCGCGGCGGC GACCGGACGA CCCGATACGG GCGCAGTCGT GGCTGCTGGT CCTGAACTGG TTGGTTGCCC GGCTTGACGT GCGCCGGCCG
 301   TGCACCAAGC TGTTTTCCGA GAAGATCACC GGCACCAGGC GCGACCGCCC GGAGCTGGCC AGGATGCTTG ACCACCTACG CCCTGGCGAC GTTGTGACAG
       ACGTGGTTCG ACAAAAGGCT CTTCTAGTGG CCGTGGTCCG CGCTGGCGGG CCTCGACCGG TCCTACGAAC TGGTGGATGC GGGACCGCTG CAACACTGTC
 401   TGACCAGGCT AGACCGCCTG GCCCGCAGCA CCCGCGACCT ACTGGACATT GCCGAGCGCA TCCAGGAGGC CGGCGCGGGC CTGCGTAGCC TGGCAGAGCC
       ACTGGTCCGA TCTGGCGGAC CGGGCGTCGT GGGCGCTGGA TGACCTGTAA CGGCTCGCGT AGGTCCTCCG GCCGCGCCCG GACGCATCGG ACCGTCTCGG
 501   GTGGGCCGAC ACCACCACGC CGGCCGGCCG CATGGTGTTG ACCGTGTTCG CCGGCATTGC CGAGTTCGAG CGTTCCCTAA TCATCGACCG CACCCGGAGC
       CACCCGGCTG TGGTGGTGCG GCCGGCCGGC GTACCACAAC TGGCACAAGC GGCCGTAACG GCTCAAGCTC GCAAGGGATT AGTAGCTGGC GTGGGCCTCG
 601   GGGCGCGAGG CCGCCAAGGC CCGAGGCGTG AAGTTTGGCC CCCGCCCTAC CCTCACCCCG GCACAGATCG CGCACGCCCG CGAGCTGATC GACCAGGAAG
       CCCGCGCTCC GGCGGTTCCG GGCTCCGCAC TTCAAACCGG GGGCGGGATG GGAGTGGGGC CGTGTCTAGC GCGTGCGGGC GCTCGACTAG CTGGTCCTTC
 701   GCCGCACCGT GAAAGAGGCG GCTGCACTGC TTGGCGTGCA TCGCTCGACC CTGTACCGCG CACTTGAGCG CAGCGAGGAA GTGACGCCCA CCGAGGCCAG
       CGGCGTGGCA CTTTCTCCGC CGACGTGACG AACCGCACGT AGCGAGCTGG GACATGGCGC GTGAACTCGC GTCGCTCCTT CACTGCGGGT GGCTCCGGTC
 801   GCGGCGCGGT GCCTTCCGTG AGGACGCATT GACCGAGGCC GACGCCCTGG CGGCCGCCGA GAATGAACGC CAAGAGGAAC AAGCATGAAA CCGCACCAGG
       CGCCGCGCCA CGGAAGGCAC TCCTGCGTAA CTGGCTCCGG CTGCGGGACC GCCGGCGGCT CTTACTTGCG GTTCTCCTTG TTCGTACTTT GGCGTGGTCC
 901   ACGGCCAGGA CGAACCGTTT TTCATTACCG AAGAGATCGA GGCGGAGATG ATCGCGGCCG GGTACGTGTT CGAGCCGCCC GCGCACGTCT CAACCGTGCG
       TGCCGGTCCT GCTTGGCAAA AAGTAATGGC TTCTCTAGCT CCGCCTCTAC TAGCGCCGGC CCATGCACAA GCTCGGCGGG CGCGTGCAGA GTTGGCACGC
1001   GCTGCATGAA ATCCTGGCCG GTTTGTCTGA TGCCAAGCTG GCGGCCTGGC CGGCCAGCTT GGCCGCTGAA GAAACCGAGC GCCGCCGTCT AAAAAGGTGA
       CGACGTACTT TAGGACCGGC CAAACAGACT ACGGTTCGAC CGCCGGACCG GCCGGTCGAA CCGGCGACTT CTTTGGCTCG CGGCGGCAGA TTTTTCCACT
1101   TGTGTATTTG AGTAAAACAG CTTGCGTCAT GCGGTCGCTG CGTATATGAT GCGATGAGTA AATAAACAAA TACGCAAGGG GAACGCATGA AGGTTATCGC
       ACACATAAAC TCATTTTGTC GAACGCAGTA CGCCAGCGAC GCATATACTA CGCTACTCAT TTATTTGTTT ATGCGTTCCC CTTGCGTACT TCCAATAGCG
1201   TGTACTTAAC CAGAAAGGCG GGTCAGGCAA GACGACCATC GCAACCCATC TAGCCCGCGC CCTGCAACTC GCCGGGGCCG ATGTTCTGTT AGTCGATTCC
       ACATGAATTG GTCTTTCCGC CCAGTCCGTT CTGCTGGTAG CGTTGGGTAG ATCGGGCGCG GGACGTTGAG CGGCCCCGGC TACAAGACAA TCAGCTAAGG
1301   GATCCCCAGG GCAGTGCCCG CGATTGGGCG GCCGTGCGGG AAGATCAACC GCTAACCGTT GTCGGCATCG ACCGCCCGAC GATTGACCGC GACGTGAAGG
       CTAGGGGTCC CGTCACGGGC GCTAACCCGC CGGCACGCCC TTCTAGTTGG CGATTGGCAA CAGCCGTAGC TGGCGGGCTG CTAACTGGCG CTGCACTTCC
1401   CCATCGGCCG GCGCGACTTC GTAGTGATCG ACGGAGCGCC CCAGGCGGCG GACTTGGCTG TGTCCGCGAT CAAGGCAGCC GACTTCGTGC TGATTCCGGT
       GGTAGCCGGC CGCGCTGAAG CATCACTAGC TGCCTCGCGG GGTCCGCCGC CTGAACCGAC ACAGGCGCTA GTTCCGTCGG CTGAAGCACG ACTAAGGCCA
                                 NdeI
1501   GCAGCCAAGC CCTTACGACA TATGGGCCAC CGCCGACCTG GTGGAGCTGG TTAAGCAGCG CATTGAGGTC ACGGATGGAA GGCTACAAGC GGCCTTTGTC
       CGTCGGTTCG GGAATGCTGT ATACCCGGTG GCGGCTGGAC CACCTCGACC AATTCGTCGC GTAACTCCAG TGCCTACCTT CCGATGTTCG CCGGAAACAG
1601   GTGTCGCGGG CGATCAAAGG CACGCGCATC GGCGGTGAGG TTGCCGAGGC GCTGGCCGGG TACGAGCTGC CCATTCTTGA GTCCCGTATC ACGCAGCGCG
       CACAGCGCCC GCTAGTTTCC GTGCGCGTAG CCGCCACTCC AACGGCTCCG CGACCGGCCC ATGCTCGACG GGTAAGAACT CAGGGCATAG TGCGTCGCGC
1701   TGAGCTACCC AGGCACTGCC GCCGCCGGCA CAACCGTTCT TGAATCAGAA CCCGAGGGCG ACGCTGCCCG CGAGGTCCAG GCGCTGGCCG CTGAAATTAA
       ACTCGATGGG TCCGTGACGG CGGCGGCCGT GTTGGCAAGA ACTTAGTCTT GGGCTCCCGC TGCGACGGGC GCTCCAGGTC CGCGACCGGC GACTTTAATT
1801   ATCAAAACTC ATTTGAGTTA ATGAGGTAAA GAGAAAATGA GCAAAAGCAC AAACACGCTA AGTGCCGGCC GTCCGAGCGC ACGCAGCAGC AAGGCTGCAA
       TAGTTTTGAG TAAACTCAAT TACTCCATTT CTCTTTTACT CGTTTTCGTG TTTGTGCGAT TCACGGCCGG CAGGCTCGCG TGCGTCGTCG TTCCGACGTT
1901   CGTTGGCCAG CCTGGCAGAC ACGCCAGCCA TGAAGCGGGT CAACTTTCAG TTGCCGGCGG AGGATCACAC CAAGCTGAAG ATGTACGCGG TACGCCAAGG
       GCAACCGGTC GGACCGTCTG TGCGGTCGGT ACTTCGCCCA GTTGAAAGTC AACGGCCGCC TCCTAGTGTG GTTCGACTTC TACATGCGCC ATGCGGTTCC
2001   CAAGACCATT ACCGAGCTGC TATCTGAATA CATCGCGCAG CTACCAGAGT AAATGAGCAA ATGAATAAAT GAGTAGATGA ATTTTAGCGG CTAAAGGAGG
       GTTCTGGTAA TGGCTCGACG ATAGACTTAT GTAGCGCGTC GATGGTCTCA TTTACTCGTT TACTTATTTA CTCATCTACT TAAAATCGCC GATTTCCTCC
2101   CGGCATGGAA AATCAAGAAC AACCAGGCAC CGACGCCGTG GAATGCCCCA TGTGTGGAGG AACGGGCGGT TGGCCAGGCG TAAGCGGCTG GGTTGTCTGC
       GCCGTACCTT TTAGTTCTTG TTGGTCCGTG GCTGCGGCAC CTTACGGGGT ACACACCTCC TTGCCCGCCA ACCGGTCCGC ATTCGCCGAC CCAACAGACG
2201   CGGCCCTGCA ATGGCACTGG AACCCCCAAG CCCGAGGAAT CGGCGTGACG GTCGCAAACC ATCCGGCCCG GTACAAATCG GCGCGGCGCT GGGTGATGAC
       GCCGGGACGT TACCGTGACC TTGGGGGTTC GGGCTCCTTA GCCGCACTGC CAGCGTTTGG TAGGCCGGGC CATGTTTAGC CGCGCCGCGA CCCACTACTG
```

```
2301  CTGGTGGAGA AGTTGAAGGC CGCGCAGGCC GCCCAGCGGC AACGCATCGA GGCAGAAGCA CGCCCCGGTG AATCGTGGCA AGCGGCCGCT GATCGAATCC
      GACCACCTCT TCAACTTCCG GCGCGTCCGG CGGGTCGCCG TTGCGTAGCT CCGTCTTCGT GCGGGGCCAC TTAGCACCGT TCGCCGGCGA CTAGCTTAGG
2401  GCAAAGAATC CCGGCAACCG CCGGCAGCCG GTGCGCCGTC GATTAGGAAG CCGCCCAAGG GCGACGAGCA ACCAGATTTT TTCGTTCCGA TGCTCTATGA
      CGTTTCTTAG GGCCGTTGGC GGCCGTCGGC CACGCGGCAG CTAATCCTTC GGCGGGTTCC CGCTGCTCGT TGGTCTAAAA AAGCAAGGCT ACGAGATACT
2501  CGTGGGCACC CGCGATAGTC GCAGCATCAT GGACGTGGCC GTTTTCCGTC TGTCGAAGCG TGACCGACGA GCTGGCGAGG TGATCCGCTA CGAGCTTCCA
      GCACCCGTGG GCGCTATCAG CGTCGTAGTA CCTGCACCGG CAAAAGGCAG ACAGCTTCGC ACTGGCTGCT CGACCGCTCC ACTAGGCGAT GCTCGAAGGT
2601  GACGGGCACG TAGAGGTTTC CGCAGGGCCG GCCGGCATGG CCAGTGTGTG GGATTACGAC CTGGTACTGA TGGCGGTTTC CCATCTAACC GAATCCATGA
      CTGCCCGTGC ATCTCCAAAG GCGTCCCGGC CGGCCGTACC GGTCACACAC CCTAATGCTG GACCATGACT ACCGCCAAAG GGTAGATTGG CTTAGGTACT
2701  ACCGATACCG GGAAGGGAAG GGAGACAAGC CCGGCCGCGT GTTCCGTCCA CACGTTGCGG ACGTACTCAA GTTCTGCCGG CGAGCCGATG GCGGAAAGCA
      TGGCTATGGC CCTTCCCTTC CCTCTGTTCG GGCCGGCGCA CAAGGCAGGT GTGCAACGCC TGCATGAGTT CAAGACGGCC GCTCGGCTAC CGCCTTTCGT
2801  GAAAGACGAC CTGGTAGAAA CCTGCATTCG GTTAAACACC ACGCACGTTG CCATGCAGCG TACGAAGAAG GCCAAGAACG GCCGCCTGGT GACGGTATCC
      CTTTCTGCTG GACCATCTTT GGACGTAAGC CAATTTGTGG TGCGTGCAAC GGTACGTCGC ATGCTTCTTC CGGTTCTTGC CGGCGGACCA CTGCCATAGG
2901  GAGGGTGAAG CCTTGATTAG CCGCTACAAG ATCGTAAAGA GCGAAACCGG GCGGCCGGAG TACATCGAGA TCGAGCTAGC TGATTGGATG TACCGCGAGA
      CTCCCACTTC GGAACTAATC GGCGATGTTC TAGCATTTCT CGCTTTGGCC CGCCGGCCTC ATGTAGCTCT AGCTCGATCG ACTAACCTAC ATGGCGCTCT
                                                                      ClaI
3001  TCACAGAAGG CAAGAACCCG GACGTGCTGA CGGTTCACCC CGATTACTTT TTGATCGATC CCGGCATCGG CCGTTTTCTC TACCGCCTGG CACGCCGCGC
      AGTGTCTTCC GTTCTTGGGC CTGCACGACT GCCAAGTGGG GCTAATGAAA AACTAGCTAG GGCCGTAGCC GGCAAAAGAG ATGGCGGACC GTGCGGCGCG
3101  CGCAGGCAAG GCAGAAGCCA GATGGTTGTT CAAGACGATC TACGAACGCA GTGGCAGCGC CGGAGAGTTC AAGAAGTTCT GTTTCACCGT GCGCAAGCTG
      GCGTCCGTTC CGTCTTCGGT CTACCAACAA GTTCTGCTAG ATGCTTGCGT CACCGTCGCG GCCTCTCAAG TTCTTCAAGA CAAAGTGGCA CGCGTTCGAC
3201  ATCGGGTCAA ATGACCTGCC GGAGTACGAT TTGAAGGAGG AGGCGGGGCA GGCTGGCCCG ATCCTAGTCA TGCGCTACCG CAACCTGATC GAGGGCGAAG
      TAGCCCAGTT TACTGGACGG CCTCATGCTA AACTTCCTCC TCCGCCCCGT CCGACCGGGC TAGGATCAGT ACGCGATGGC GTTGGACTAG CTCCCGCTTC
3301  CATCCGCCGG TTCCTAATGT ACGGAGCAGA TGCTAGGGCA AATTGCCCTA GCAGGGGAAA AAGGTCGAAA AGGTCTCTTT CCTGTGGATA GCACGTACAT
      GTAGGCGGCC AAGGATTACA TGCCTCGTCT ACGATCCCGT TTAACGGGAT CGTCCCCTTT TTCCAGCTTT TCCAGAGAAA GGACACCTAT CGTGCATGTA
3401  TGGGAACCCA AAGCCGTACA TTGGGAACCG GAACCCGTAC ATTGGGAACC CAAAGCCGTA CATTGGGAAC CGGTCACACA TGTAAGTGAC TGATATAAAA
      ACCCTTGGGT TTCGGCATGT AACCCTTGGC CTTGGGCATG TAACCCTTGG GTTTCGGCAT GTAACCCTTG GCCAGTGTGT ACATTCACTG ACTATATTTT
3501  GAGAAAAAAG GCGATTTTTC CGCCTAAAAC TCTTTAAAAC TTATTAAAAC TCTTAAAACC CGCCTGGCCT GTGCATAACT GTCTGGCCAG CGCACAGCCG
      CTCTTTTTTC CGCTAAAAAG GCGGATTTTG AGAAATTTTG AATAATTTTG AGAATTTTGG GCGGACCGGA CACGTATTGA CAGACCGGTC GCGTGTCGGC
3601  AAGAGCTGCA AAAAGCGCCT ACCCTTCGGT CGCTGCGCTC CCTACGCCCC GCCGCTTCGC GTCGGCCTAT CGCGGCCGCT GGCCGCTCAA AAATGGCTGG
      TTCTCGACGT TTTTCGCGGA TGGGAAGCCA GCGACGCGAG GGATGCGGGG CGGCGAAGCG CAGCCGGATA GCGCCGGCGA CCGGCGAGTT TTTACCGACC
3701  CCTACGGCCA GGCAATCTAC CAGGGCGCGG ACAAGCCGCG CCGTCGCCAC TCGACCGCCG GCGCCCACAT CAAGGCACCC TGCCTCGCGC GTTTCGGTGA
      GGATGCCGGT CCGTTAGATG GTCCCGCGCC TGTTCGGCGC GGCAGCGGTG AGCTGGCGGC CGCGGGTGTA GTTCCGTGGG ACGGAGCGCG CAAAGCCACT
3801  TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG
      ACTGCCACTT TTGGAGACTG TGTACGTCGA GGGCCTCTGC CAGTGTCGAA CAGACATTCG CCTACGGCCC TCGTCTGTTC GGGCAGTCCC GCGCAGTCGC
3901  GGTGTTGGCG GGTGTCGGGG CGCAGCCATG ACCCAGTCAC GTAGCGATAG CGGAGTGTAT ACTGGCTTAA CTATGCGGCA TCAGAGCAGA TTGTACTGAG
      CCACAACCGC CCACAGCCCC GCGTCGGTAC TGGGTCAGTG CATCGCTATC GCCTCACATA TGACCGAATT GATACGCCGT AGTCTCGTCT AACATGACTC
                 NdeI
4001  AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA AGGAGAAAAT ACCGCATCAG GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG
      TCACGTGGTA TACGCCACAC TTTATGGCGT GTCTACGCAT TCCTCTTTTA TGGCGTAGTC CGCGAGAAGG CGAAGGAGCG AGTGACTGAG CGACGCGAGC
4101  GTCGTTCGGC TGCGGCGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG TGAGCAAAAG
      CAGCAAGCCG ACGCCGCTCG CCATAGTCGA GTGAGTTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC TTTCTTGTAC ACTCGTTTTC
4201  GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCGTTTTTC CATAGGCTCC GCCCCCCTGA CGAGCATCAC AAAAATCGAC GCTCAAGTCA
      CGGTCGTTTT CCGGTCCTTG GCATTTTTCC GGCGCAACGA CCGCAAAAAG GTATCCGAGG CGGGGGGACT GCTCGTAGTG TTTTTAGCTG CGAGTTCAGT
4301  GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC
      CTCCACCGCT TTGGGCTGTC CTGATATTTC TATGGTCCGC AAAGGGGGAC CTTCGAGGGA GCACGCGAGA GGACAAGGCT GGGACGGCGA ATGGCCTATG
4401  CTGTCCGCCT TTCTCCCTTC GGGAAGCGTG GCGCTTTCTC ATAGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG
      GACAGGCGGA AAGAGGGAAG CCCTTCGCAC CGCGAAAGAG TATCGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGGTTC GACCCGACAC
4501  TGCACGAACC CCCCGTTCAG CCCGACCGCT GCGCCTTATC CGGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC TGGCAGCAGC
      ACGTGCTTGG GGGGCAAGTC GGGCTGGCGA CGCGGAATAG GCCATTGATA GCAGAACTCA GGTTGGGCCA TTCTGTGCTG AATAGCGGTG ACCGTCGTCG
```

```
4601  CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC AGTATTTGGT
      GTGACCATTG TCCTAATCGT CTCGCTCCAT ACATCCGCCA CGATGTCTCA AGAACTTCAC CACCGGATTG ATGCCGATGT GATCTTCCTG TCATAAACCA
4701  ATCTGCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT TTTGTTTGCA
      TAGACGCGAG ACGACTTCGG TCAATGGAAG CCTTTTTCTC AACCATCGAG AACTAGGCCG TTTGTTTGGT GGCGACCATC GCCACCAAAA AAACAAACGT
4801  AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT
      TCGTCGTCTA ATGCGCGTCT TTTTTTCCTA GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGAGTG CAATTCCCTA
4901  TTTGGTCATG CATGATATAT CTCCCAATTT GTGTAGGGCT TATTATGCAC GCTTAAAAAT AATAAAAGCA GACTTGACCT GATAGTTTGG CTGTGAGCAA
      AAACCAGTAC GTACTATATA GAGGGTTAAA CACATCCCGA ATAATACGTG CGAATTTTTA TTATTTTCGT CTGAACTGGA CTATCAAACC GACACTCGTT
                                                 HpaI
5001  TTATGTGCTT AGTGCATCTA ATCGCTTGAG TTAACGCCGG CGAAGCGGCG TCGGCTTGAA CGAATTTCTA GCTAGACATT ATTTGCCGAC TACCTTGGTG
      AATACACGAA TCACGTAGAT TAGCGAACTC AATTGCGGCC GCTTCGCCGC AGCCGAACTT GCTTAAAGAT CGATCTGTAA TAAACGGCTG ATGGAACCAC
5101  ATCTCGCCTT TCACGTAGTG GACAAATTCT TCCAACTGAT CTGCGCGCGA GGCCAAGCGA TCTTCTTCTT GTCCAAGATA AGCCTGTCTA GCTTCAAGTA
      TAGAGCGGAA AGTGCATCAC CTGTTTAAGA AGGTTGACTA GACGCGCGCT CCGGTTCGCT AGAAGAAGAA CAGGTTCTAT TCGGACAGAT CGAAGTTCAT
5201  TGACGGGCTG ATACTGGGCC GGCAGGCGCT CCATTGCCCA GTCGGCAGCG ACATCCTTCG GCGCGATTTT GCCGGTTACT GCGCTGTACC AAATGCGGGA
      ACTGCCCGAC TATGACCCGG CCGTCCGCGA GGTAACGGGT CAGCCGTCGC TGTAGGAAGC CGCGCTAAAA CGGCCAATGA CGCGACATGG TTTACGCCCT
5301  CAACGTAAGC ACTACATTTC GCTCATCGCC AGCCCAGTCG GGCGGCGAGT TCCATAGCGT TAAGGTTTCA TTTAGCGCCT CAAATAGATC CTGTTCAGGA
      GTTGCATTCG TGATGTAAAG CGAGTAGCGG TCGGGTCAGC CCGCCGCTCA AGGTATCGCA ATTCCAAAGT AAATCGCGGA GTTTATCTAG GACAAGTCCT
                                                                                                  PvuI
5401  ACCGGATCAA AGAGTTCCTC CGCCGCTGGA CCTACCAAGG CAACGCTATG TTCTCTTGCT TTTGTCAGCA AGATAGCCAG ATCAATGTCG ATCGTGGCTG
      TGGCCTAGTT TCTCAAGGAG GCGGCGACCT GGATGGTTCC GTTGCGATAC AAGAGAACGA AAACAGTCGT TCTATCGGTC TAGTTACAGC TAGCACCGAC
5501  GCTCGAAGAT ACCTGCAAGA ATGTCATTGC GCTGCCATTC TCCAAATTGC AGTTCGCGCT TAGCTGGATA ACGCCACGGA ATGATGTCGT CGTGCACAAC
      CGAGCTTCTA TGGACGTTCT TACAGTAACG CGACGGTAAG AGGTTTAACG TCAAGCGCGA ATCGACCTAT TGCGGTGCCT TACTACAGCA GCACGTGTTG
5601  AATGGTGACT TCTACAGCGC GGAGAATCTC GCTCTCTCCA GGGGAAGCCG AAGTTTCCAA AAGGTCGTTG ATCAAAGCTC GCCGCGTTGT TTCATCAAGC
      TTACCACTGA AGATGTCGCG CCTCTTAGAG CGAGAGAGGT CCCCTTCGGC TTCAAAGGTT TTCCAGCAAC TAGTTTCGAG CGGCGCAACA AAGTAGTTCG
5701  CTTACGGTCA CCGTAACCAG CAAATCAATA TCACTGTGTG GCTTCAGGCC GCCATCCACT GCGGAGCCGT ACAAATGTAC GGCCAGCAAC GTCGGTTCGA
      GAATGCCAGT GGCATTGGTC GTTTAGTTAT AGTGACACAC CGAAGTCCGG CGGTAGGTGA CGCCTCGGCA TGTTTACATG CCGGTCGTTG CAGCCAAGCT
5801  GATGGCGCTC GATGACGCCA ACTACCTCTG ATAGTTGAGT CGATACTTCG GCGATCACCG CTTCCCCCAT GATGTTTAAC TTTGTTTTAG GGCGACTGCC
      CTACCGCGAG CTACTGCGGT TGATGGAGAC TATCAACTCA GCTATGAAGC CGCTAGTGGC GAAGGGGGTA CTACAAATTG AAACAAAATC CCGCTGACGG
5901  CTGCTGCGTA ACATCGTTGC TGCTCCATAA CATCAAACAT CGACCCACGG CGTAACGCGC TTGCTGCTTG GATGCCCGAG GCATAGACTG TACCCCAAAA
      GACGACGCAT TGTAGCAACG ACGAGGTATT GTAGTTTGTA GCTGGGTGCC GCATTGCGCG AACGACGAAC CTACGGGCTC CGTATCTGAC ATGGGGTTTT
6001  AAACATGTCA TAACAAGAAG CCATGAAAAC CGCCACTGCG CCGTTACCAC CGCTGCGTTC GGTCAAGGTT CTGGACCAGT TGCGTGACGG CAGTTACGCT
      TTTGTACAGT ATTGTTCTTC GGTACTTTTG GCGGTGACGC GGCAATGGTG GCGACGCAAG CCAGTTCCAA GACCTGGTCA ACGCACTGCC GTCAATGCGA
6101  ACTTGCATTA CAGCTTACGA ACCGAACGAG GCTTATGTCC ACTGGGTTCG TGCCCGAATT GATCACAGGC AGCAACGCTC TGTCATCGTT ACAATCAACA
      TGAACGTAAT GTCGAATGCT TGGCTTGCTC CGAATACAGG TGACCCAAGC ACGGGCTTAA CTAGTGTCCG TCGTTGCGAG ACAGTAGCAA TGTTAGTTGT
6201  TGCTACCCTC CGCGAGATCA TCCGTGTTTC AAACCCGGCA GCTTAGTTGC CGTTCTTCCG AATAGCATCG GTAACATGAG CAAAGTCTGC CGCCTTACAA
      ACGATGGGAG GCGCTCTAGT AGGCACAAAG TTTGGGCCGT CGAATCAACG GCAAGAAGGC TTATCGTAGC CATTGTACTC GTTTCAGACG GCGGAATGTT
6301  CGGCTCTCCC GCTGACGCCG TCCCGGACTG ATGGGCTGCC TGTATCGAGT GGTGATTTTG TGCCGAGCTG CCGGTCGGGG AGCTGTTGGC TGGCTGGTGG
      GCCGAGAGGG CGACTGCGGC AGGGCCTGAC TACCCGACGG ACATAGCTCA CCACTAAAAC ACGGCTCGAC GGCCAGCCCC TCGACAACCG ACCGACCACC
                                                                                                                XhoI
                                                                                                       EcoRI
6401  CAGGATATAT TGTGGTGTAA ACAAATTGAC GCTTAGACAA CTTAATAACA CATTGCGGAC GTTTTTAATG TACTGAATTA ACGCCGAATT GAATTCCTCG
      GTCCTATATA ACACCACATT TGTTTAACTG CGAATCTGTT GAATTATTGT GTAACGCCTG CAAAAATTAC ATGACTTAAT TGCGGCTTAA CTTAAGGAGC
                                        XbaI          SacII
      XhoI        BamHI                         AscI
6501  AGTACGTAGG ATCCATTTAA ATTCTAGAGG CGCGCCGCGG CCGCTCCTGC TGAGCCTCGA CATGTTGTCG CAAAATTCGC CCTGGACCCG CCCAACGATT
      TCATGCATCC TAGGTAAATT TAAGATCTCC GCGCGGCGCC GGCGAGGACG ACTCGGAGCT GTACAACAGC GTTTTAAGCG GGACCTGGGC GGGTTGCTAA
6601  TGTCGTCACT GTCAAGGTTT GACCTGCACT TCATTTGGGG CCCACATACA CCAAAAAAAT GCTGCATAAT TCTCGGGGCA GCAAGTCGGT TACCCGGCCG
      ACAGCAGTGA CAGTTCCAAA CTGGACGTGA AGTAAACCCC GGGTGTATGT GGTTTTTTTA CGACGTATTA AGAGCCCCGT CGTTCAGCCA ATGGGCCGGC
```

```
6701  CCGTGCTGGA CCGGGTTGAA TGGTGCCCGT AACTTTCGGT AGAGCGGACG GCCAATACTC AACTTCAAGG AATCTCACCC ATGCGCGCCG GCGGGGAACC
      GGCACGACCT GGCCCAACTT ACCACGGGCA TTGAAAGCCA TCTCGCCTGC CGGTTATGAG TTGAAGTTCC TTAGAGTGGG TACGCGCGGC CGCCCCTTGG
6801  GGAGTTCCCT TCAGTGAGCG TTATTAGTTC GCCGCTCGGT GTGTCGTAGA TACTAGCCCC TGGGGCACTT TTGAAATTTG AATAAGATTT ATGTAATCAG
      CCTCAAGGGA AGTCACTCGC AATAATCAAG CGGCGAGCCA CACAGCATCT ATGATCGGGG ACCCCGTGAA AACTTTAAAC TTATTCTAAA TACATTAGTC
6901  TCTTTTAGGT TTGACCGGTT CTGCCGCTTT TTTTAAAATT GGATTTGTAA TAATAAAACG CAATTGTTTG TTATTGTGGC GCTCTATCAT AGATGTCGCT
      AGAAAATCCA AACTGGCCAA GACGGCGAAA AAAATTTTAA CCTAAACATT ATTATTTTGC GTTAACAAAC AATAACACCG CGAGATAGTA TCTACAGCGA
7001  ATAAACCTAT TCAGCACAAT ATATTGTTTT CATTTTAATA TTGTACATAT AAGTAGTAGG GTACAATCAG TAAATTGAAC GGAGAATATT ATTCATAAAA
      TATTTGGATA AGTCGTGTTA TATAACAAAA GTAAAATTAT AACATGTATA TTCATCATCC CATGTTAGTC ATTTAACTTG CCTCTTATAA TAAGTATTTT
7101  ATACGATAGT AACGGGTGAT ATATTCATTA GAATGAACCG AAACCGGCGG TAAGGATCTG AGCTACACAT GCTCAGGTTT TTTACAACGT GCACAACAGA
      TATGCTATCA TTGCCCACTA TATAAGTAAT CTTACTTGGC TTTGGCCGCC ATTCCTAGAC TCGATGTGTA CGAGTCCAAA AAATGTTGCA CGTGTTGTCT
                                                                                                  XbaI
                                                                                   SpeI                   BamHI
7201  ATTGAAAGCA AATATCATGC GATCATAGGC GTCTCGCATA TCTCATTAAA GCAGGACTCT AGGGACTAGT TCTAGACTAG GTGGATCCTC AGAAGAACTC
      TAACTTTCGT TTATAGTACG CTAGTATCCG CAGAGCGTAT AGAGTAATTT CGTCCTGAGA TCCCTGATCA AGATCTGATC CACCTAGGAG TCTTCTTGAG
7301  GTCAAGAAGG CGATAGAAGG CGATGCGCTG CGAATCGGGA GCGGCGATAC CGTAAAGCAC GAGGAAGCGG TCAGCCCATT CGCCGCCAAG CTCTTCAGCA
      CAGTTCTTCC GCTATCTTCC GCTACGCGAC GCTTAGCCCT CGCCGCTATG GCATTTCGTG CTCCTTCGCC AGTCGGGTAA GCGGCGGTTC GAGAAGTCGT
7401  ATATCACGGG TAGCCAACGC TATGTCCTGA TAGCGGTCCG CCACACCCAG CCGGCCACAG TCGATGAATC CAGAAAAGCG GCCATTTTCC ACCATGATAT
      TATAGTGCCC ATCGGTTGCG ATACAGGACT ATCGCCAGGC GGTGTGGGTC GGCCGGTGTC AGCTACTTAG GTCTTTTCGC CGGTAAAAGG TGGTACTATA
                                NcoI                                   SphI
7501  TCGGCAAGCA GGCATCGCCA TGGGTCACGA CGAGATCATC GCCGTCGGGC ATGCGCGCCT TGAGCCTGGC GAACAGTTCG GCTGGCGCGA GCCCCTGATG
      AGCCGTTCGT CCGTAGCGGT ACCCAGTGCT GCTCTAGTAG CGGCAGCCCG TACGCGCGGA ACTCGGACCG CTTGTCAAGC CGACCGCGCT CGGGGACTAC
7601  CTCTTCGTCC AGATCATCCT GATCGACAAG ACCGGCTTCC ATCCGAGTAC GTGCTCGCTC GATGCGATGT TTCGCTTGGT GGTCGAATGG GCAGGTAGCC
      GAGAAGCAGG TCTAGTAGGA CTAGCTGTTC TGGCCGAAGG TAGGCTCATG CACGAGCGAG CTACGCTACA AAGCGAACCA CCAGCTTACC CGTCCATCGG
7701  GGATCAAGCG TATGCAGCCG CCGCATTGCA TCAGCCATGA TGGATACTTT CTCGGCAGGA GCAAGGTGAG ATGACAGGAG ATCCTGCCCC GGCACTTCGC
      CCTAGTTCGC ATACGTCGGC GGCGTAACGT AGTCGGTACT ACCTATGAAA GAGCCGTCCT CGTTCCACTC TACTGTCCTC TAGGACGGGG CCGTGAAGCG
                                                                PvuII
7801  CCAATAGCAG CCAGTCCCTT CCCGCTTCAG TGACAACGTC GAGCACAGCT GCGCAAGGAA CGCCCGTCGT GGCCAGCCAC GATAGCCGCG CTGCCTCGTC
      GGTTATCGTC GGTCAGGGAA GGGCGAAGTC ACTGTTGCAG CTCGTGTCGA CGCGTTCCTT GCGGGCAGCA CCGGTCGGTG CTATCGGCGC GACGGAGCAG
7901  CTGCAGTTCA TTCAGGGCAC CGGACAGGTC GGTCTTGACA AAAAGAACCG GGCGCCCCTG CGCTGACAGC CGGAACACGG CGGCATCAGA GCAGCCGATT
      GACGTCAAGT AAGTCCCGTG GCCTGTCCAG CCAGAACTGT TTTTCTTGGC CCGCGGGGAC GCGACTGTCG GCCTTGTGCC GCCGTAGTCT CGTCGGCTAA
                                                                                                      SalI      EcoRI
8001  GTCTGTTGTG CCCAGTCATA GCCGAATAGC CTCTCCACCC AAGCGGCCGG AGAACCTGCG TGCAATCCAT CTTGTTCAAT CATGTCGACT GCAGAATTCG
      CAGACAACAC GGGTCAGTAT CGGCTTATCG GAGAGGTGGG TTCGCCGGCC TCTTGGACGC ACGTTAGGTA GAACAAGTTA GTACAGCTGA CGTCTTAAGC
                 SacI
                      XhoI
      HindIII            BglII      BspEI
8101  AAGCTTGAGC TCGAGATCTG AGTCCGGATT GGTAGATTGC AAATATAATG GTTTGGCGAT TATCCTTGAG GCCACACCTT TAAATAGATC AATGAATGGG
      TTCGAACTCG AGCTCTAGAC TCAGGCCTAA CCATCTAACG TTTATATTAC CAAACCGCTA ATAGGAACTC CGGTGTGGAA ATTTATCTAG TTACTTACCC
8201  CAAGTATTGA CTTGCGATAA TCCTTTAACT TTCTGCCTAC CATCAATGTG GATGAGTTGT CGGTAAAAAG GATCTTACAG GCCAAATTCG CTCTTAGCCG
      GTTCATAACT GAACGCTATT AGGAAATTGA AAGACGGATG GTAGTTACAC CTACTCAACA GCCATTTTTC CTAGAATGTC CGGTTTAAGC GAGAATCGGC
8301  TACAATATTA CTCACCGGTG CGATGCCCCC CATCGTAGGT GAAGGTGGAA ATTAATGATC CATCTTGAGA CCACAGGCCC ACAACAGCTA CCAGTTTCCT
      ATGTTATAAT GAGTGGCCAC GCTACGGGGG GTAGCATCCA CTTCCACCTT TAATTACTAG GTAGAACTCT GGTGTCCGGG TGTTGTCGAT GGTCAAAGGA
                                                                                   HpaI                      AscI
8401  CAAGGGTCCA CCAAAAACGT AAGCGCTTAC GTACATGGTC GATAAGAAAA GGCAATTTGT AGATGTTAAC ATCCAACGTC GCTTTCAGCC CGGTGGCGCG
      GTTCCCAGGT GGTTTTTGCA TTCGCGAATG CATGTACCAG CTATTCTTTT CCGTTAAACA TCTACAATTG TAGGTTGCAG CGAAAGTCGG GCCACCGCGC
```

```
                EcoRV
        AscI                            SacI
8501    CCGATATCCT CTCTTAAGGT AGCGAGCTCT TAATTAATAG GGATAACAGG GTAATTAACT ATAACGGTCC TAAGGTAGCG ATGGCAAACA GCTATTATGG
        GGCTATAGGA GAGAATTCCA TCGCTCGAGA ATTAATTATC CCTATTGTCC CATTAATTGA TATTGCCAGG ATTCCATCGC TACCGTTTGT CGATAATACC
                                                                                      XmaI            StuI
                                                                          PvuI    KpnI   SmaI    SalI          HindIII
8601    GTATTATGGG TGGTTCTTTA TGCGGACACT GACGGCTTTA TGCCTGCAGG TCGCGAGCGA TCGCGGTACC GCCCGGGCGT CGACAGGCCT AAGCTTAGCT
        CATAATACCC ACCAAGAAAT ACGCCTGTGA CTGCCGAAAT ACGGACGTCC AGCGCTCGCT AGCGCCATGG CGGGCCCGCA GCTGTCCGGA TTCGAATCGA
                                          PmeI
8701    TGAGCTTGGA TCAGATTGTC GTTTCCCGCC TTCAGTTTAA ACTATCAGTG TTTGACAGGA TATATTGGCG GGTAAACCTA AGAGAAAAGA GCGTTTATTA
        ACTCGAACCT AGTCTAACAG CAAAGGGCGG AAGTCAAATT TGATAGTCAC AAACTGTCCT ATATAACCGC CCATTTGGAT TCTCTTTTCT CGCAAATAAT
                                                                          SphI
8801    GAATAACGGA TATTTAAAAG GGCGTGAAAA GGTTTATCCG TTCGTCCATT TGTATGTGCA TGCCAACCAC AGGGTTCCCC TCGGGATCAA
        CTTATTGCCT ATAAATTTTC CCGCACTTTT CCAAATAGGC AAGCAGGTAA ACATACACGT ACGGTTGGTG TCCCAAGGGG AGCCCTAGTT
```