



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA **EcoNI (96)**

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**HindIII (245)** 201 **Psp1406I (203)** **PvuII (239)** **Bsu36I (291)**  
**EcoNI (287)** GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCCCTCTGAAGTGCCTCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 **NgoMIV (441)** GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

501 **KasI (535)** **AgeI (552)** TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCCCTACCTGAGATCACCGGTCATCATGATTCTCGCATCCGCTGGCTGTGGCCACGATCCC  
1 M I P R I R L A V A T I P

**NcoI (602)** **SmaI (638)** **NsiI (678)** 601 AGCCATGGCCTTCTCCTGCCTGAGATCTGAGAGCTGGGACCTTGCCTGACAGGTGGTTCCTAACATTAGTTACCAATGCATGGAGCTGAATTTCTAC  
13> A M A F L S C L R S E S W D P C V Q V V P N I S Y Q C M E L N F Y  
701 AAAATCCCTGACAACATCCACATCAGTCAAGATACTGGACCTGAGCTTAACTACCTGAGTCATTTAGACAGCAATAGCTTCTCCAGCTTTCCAGAAC  
47> K I P D N I P T S V K I L D L S F N Y L S H L D S N S F S S F P E

**NsiI (848)** **BsaBI (847)** 801 TGCAGGTGCTGGATTATCCAGATGTGAAATTGACACAATTGACGATGATGCATATGACGGCCTAAATTACCTTTCCACCTTGATACTGACGGAAATCC  
80> L Q V L D L S R C E I Q T I D D D A Y Q G L N Y L S T L I L T G N P  
901 TATCCAGAGTTTAGCCCTGGGAGCCTTTCTGGACTACCAAGTTTACAGAAGCTGGTTCGCGTTGAGACAACTAGCGTCTCTAGAGGACTTCCCATT  
113> I Q S L A L G A F S G L P S L Q K L V A V E T N L A S L E D F P I  
1001 GGACATCTAAAACCTTGAAGGAGCTTAATGTGGCTCAACATCATATCATTCTTTCAAGTTACCTGAATATTTTCTAACCTGCCAACCTGGAGCATT  
147> G H L K T L K E L N V A H N H I H S F K L P E Y F S N L P N L E H  
1101 TGGATCTTTCTAAGAACAAAATCGAAAATATTTATCATGAACACTTGCAGGTTCTACATCAAGTGCCCTACTCAATCTCTTTTAGACTTGTCCCTGAA  
180> L D L S K N K I E N I Y H E H L Q V L H Q V P L L N L S L D L S L N  
1201 CCCTTTAACTTTATTGAACCAGGTGCCTTTAAACAAAATTAGGCTCAATGGATTGACTTTGCGAAGTAATTTTAAATAGTTGCGATGTAATGAAAACCTGT  
213> P L N F I E P G A F N K I R L N G L L R S N F N S D V M K T C  
1301 ATTCAAGGTCTGGCTGGTTCAAAAATCAATCAGTTGGTTTTGGGAGAATTTAAAAATGAAAGGAACCTGAAAAGTTTTGACAAAATCTGTCTGGAGGAAAC  
247> I Q G L A G S K I N Q L V L G E F K N E R N L E S F D K S V L E E

**BspEI (1424)** **EcoRI (1444)** 1401 TGTGCAATTTGACCCTTGAACAGTTCCGGATAGCACACTTTGGTGAATTTCCAGATGATGTTTCTGACTTATTTAATTGTTGGCAAATGCCTCTGTGAT  
280> L C N L T L E Q F R I A H F G E F P D D V S D L F N C L A N A S V I

**XmnI (1531)** **HpaI (1573)** 1501 TTCTCTGTTGAGTCTGAATTTGCGTGGTCTAGAAGCCCTTCTAATGATTTCCAGTGGCAACACTTAGAAGTGGTTAACTGTAACCTTCAACAATTTCT  
313> S L L S L N L R G L E A L P N D F R W Q H L E V V N C K L Q Q F P

**NdeI (1648)** 1601 GCATTGAAGTTCAACTCTCTCAAAAAGTTGTTTTCAAAGACAATAAACATATGCACACTTTTACTGAAATTAACCTGCCAAACCTTTCAGTTTCTAGATC  
347> A L K F N S L K K F V F K D N K H M H T F T E I N L P N L Q F L D  
1701 TCAGTGAAATCACTTGAGCTTTAAAGGTTGCTGTTCTACAATGAATTTGGGACAACCAAAGTGAAGCATTAGATCTGAGCTTCAATGAAATCATTAC  
380> L S G N H L S F K G C C S H N E F G T T K L K H L D L S F N E I I T  
1801 CATGAAGTCAAACTTATGAGCACTAGAGCAACTAGAACTCCAGTATCCAGCTTCCAGTGGTGAACAGGCTAATGACTTTTCCATTTCTATCCTC  
413> M K S N F M G L E Q L E Y L D F Q H S S L K Q A N D F S I A F L S L  
1901 AGAAACCTCCATTACCTTGATATTTCTTATACTAACATCCAGTGTGCTTCCGTGGCATTGTTGCTGGCTTGGTCAGCCTCCAAACCTTGAATAAGGCTG  
447> R N L H Y L D I S Y T N I H V V F R G I F A G L V S L Q T L K M A  
2001 GCAATCTTTTTCAGAACAACTTGCCTCCTGATGCTTACAGATCTGACTAATTAATCCTCCTGGACCTCTCAAGTGTCAACTGGAACAGGTATCCCA  
480> G N S F Q N N L L P D V F T D L T N L I L L D L S K C Q L E Q V S Q

**AvrII (2119)** **Bsu36I (2196)** 2101 GAGGGCATTTCATCCCTCCCTAGGCTTCAGGTGCTAAATATGAGTCACAACAGACTCTTGTCTTGATACGCTTCTTATAAACCACTCCACTCCCTC  
513> R A F H S L P R L Q V L N M S H N R L L F L D T L P Y K P L H S L  
2201 AGGATTCTGGATTGCAGTTACAATCTTATCGTGGCCTCCAAGGAACAAGAACTACAGCATCTGCCAAGGAGCCTAGCTTTTTTAAATCTCACTAAGAATG  
547> R I L D C S Y N L I V A S K E Q E L Q H L P R S L A F L N L T K N

**BsrGI (2386)** 2301 ACTTTTCTGTGCGTGTGAACACCAGACTTTCTTGCAGTGGGTCAGGACCAGAAGCAGCTCCTTGTGGGAGCTGAGCAAATGGTGTGTACACAACTTT  
580> D F S C A C E H Q T F L Q W V K D Q K Q L L V G A E Q M V C T Q P L  
2401 AGAAATCGAGGACTGCCTGTGCTGAGTTTCAGGAACGCCACCTGTGAGTAAAGCAGGCGCTCATTAGTGCCTGAGTCTCACCTTCTCTCTGGTATC  
613> E M Q D L P V L S F R N A T C Q I S E A V I S A S V L T F L L V S  
2501 GTGGCAGGAATCTGGTCTATAAGTTCTATTTCCACTTGTGCTTTTTGTGGGCTGCAAAAAATATGGCAGAGGTGAAAGCACCTATGACGCCTTTGTTA  
647> V A G I L V Y K F Y F H L L L F V G C K K Y G R G E S T Y D A F V  
2601 TCTACTCCAGCCAGGACGAAGACTGGGTGAGGAATGAAGTGGTAAAGAAGTGGAGGAGGGGGTGCCTTTTTCATCTCTGCCTTCACTACAGAGACTT  
680> I Y S S Q D E D W V R N E L V K N L E E G V P P F H L C L H Y R D F  
2701 CATTCCCGGTGTGGCCATCGCTGCTAACATCATCCAGGAAGGTTTCCACAAAAGTTCGGAAGGTTATTGTCGTGGTGTCCAGCATTCATACAGAGCCGA  
713> I P G V A I A A N I I Q E G F H K S R K V I V V V S Q H F I Q S R

**NgoMIV (2854)** 2801 TGGTGATCTTTGAGTATGAGATTGCCAGACCTGGCAGTTTCTGAGGAGTATGCCGGCATCATCTTCATCGTCTGAGAAAGTTGGAGAAGTCCCTGC  
747> W C I F E Y E I A Q T W Q F L R S H A G I I F I V L Q K L E K S L

**Tth111I (2959)**

2901 TGC GGCAGCAGGTGGAGCTGTATCGCCTTCTCAGCAGGAATACCTACCTGGAGTGGGAGGACAGCGTCTGGGGCGGCACATCTTCTGGAGACGACTCAA  
780▶ L R Q Q V E L Y R L L S R N T Y L E W E D S V L G R H I F W R R L K

**NcoI (3022)**

3001 GAAAGCCTTGCTGGATGGTAAACCATGGAGTCCAGAAGGAACAGAGGATTGAGAAAGCAACCAGCATGATACAACAGCCTTCACTTAAGGAGGGAAAAC  
813▶ K A L L D G K P W S P E G T E D S E S N Q H D T T A F T •

**NheI (3119)**

3101 CCCAACGTGTCCCTTGGTCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTAAAAAATGCTTTATT

**HpaI (3257)**

3201 TGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAACAATTGCATTCAATTTATGTTTCAGGTTTCAGG

**EcoRI (3353)**

3301 GGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTTAACCTCCAATCAAGCC  
3401 TCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGT  
3501 TTAAGATATAGTGTATTTTCCCAAGGTTTGAAC TAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCA

**SwaI (3606)**

3601 GAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTT  
3701 GGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGGAGGGGATGAGTTCCTCAA  
141▶ • N R T Y K L P I L E E I

**SacI (3867)** **BstXI (3896)**

3801 TGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGTCACATGCCACAGGGGCTGACCC  
128▶ T T K V L K G N M E I L V F C D P A Y D S I L E R C M G C P S V V  
3901 CCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATG  
95▶ R I S R D V E D S Y P H R V A V I T D F D K Q G N S V A S G I A I

**StuI (4031)**

4001 GCTTCAGCACAGACAGTGCACCTGCCAATGTAGGCCTCAATGGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCT  
61▶ A E A C V T V R G I Y A E I H V A S I I E G T K T R I A A G V H H K

**XmnI (4173)**

4101 TGTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCCTATA  
28▶ N D E Y L M T I K E T A V E V L E L D Q Q S I N F T K M

**AseI (4239)** **SacI (4296)**

4201 GTGAGTCGTATTATACTATGCGGATATACTATGCGGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGC

**SpeI (4394)**

4301 TCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAG  
4400 TCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTA CTGCCAAAACCGCATC

**SnaBI (4522)**

4500 ATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTA CTGGCATAATGCCAGGCGGGCCATTTACCGT

**NdeI (4627)**

4600 CATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAA  
4700 AGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTGTTGGCGGTGACCCAGGCGGGCCATTTACCGTAAAGTTAT

**SdaI (4805)** **PacI (4813)** **BspLU11I (4823)**

4800 GTAACGCCTGCAGGTTAA TTAAGAACATGTGAGCAAAAGGCCAGAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTC  
4898 CGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC  
4998 TCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTA

**ApaLI (5137)**

5098 TCTCAGTTCGGTGTAGGTCGTTCTGCTCCAAGCTGGGCTGTGTGACGAAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG  
5198 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGT  
5298 GGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGG  
5398 CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG

**EagI (5573)**

5498 GGGTCTGACGCTCAGTGGAAACGAAAACCTACGTTAAGGGATTTGGTTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATT  
5598 TTCATTACATCTGTGTGTTGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTC  
5698 CCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA