



**PvuI (7)**  
**SgfI (6)**  
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAAGTGGCGGGGTAACCTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

**Psp1406I (203)** **PvuII (239)** **Bsu36I (291)**  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCGCTACCTGAGGCC  
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTCGCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

**KasI (535)** **AgeI (552)** **BspHI (568)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTAGGAGGGCCATCATGATGCCTCCCTGGCTCCTGGCTAGGACT  
1 M M P P W L L A R T  
601 CTGATCATGGCACTGTTCTTCTCCTGCCTGACACCAGGAAGCTTGAATCCTGCATAGAGGTAGTTTCTAATATTACCTACCAATGCATGGATCAGAAAC  
11 L I M A L F F S C L T P G S L N P C I E V V P N I T Y Q C M D Q K  
701 TCAGCAAAGTCCCTGATGACATTCCTTCAACCAAGAACATAGATCTGAGCTTCAACCCCTGAAGATCTTAAAAAGCTATAGCTTCTCCAATTTTT  
44 L S K V P D D I P S S T K N I D L S F N P L K I L K S Y S F S N F S  
801 AGAAGTTCAGTGGCTGGATTATCCAGGTGTGAAATGAAACAATTGAAGACAAGGCATGGCATGGCTTACACCACCTCTCAACTTGATACTGACAGGA  
77 E L Q W L D L S R C E I E T I E D K A W H G L H H L S N L I L T G  
901 AACCTATCCAGAGTTTTTCCCAAGGAAGTTTCTGGACTAACAAAGTTTAGAGAATCTGGTGGCTGTGGAGACAAAATTGGCCTCTAGAAAGCTTCC  
111 N P I Q S F S P G S F S G L T S L E N L V A V E T K L A S L E S F  
1001 CTATTGGACAGCTTATAACCTTAAAGAACTCAATGTGGCTCACAATTTTATACATTCCTGTAAGTTACTGCATATTTTCCAATCTGACGAACCTAGT  
144 P I G Q L I T L K K L N V A H N F I H S C K L P A Y F S N L T N L V  
1101 ACATGTGGATCTTTCTTATAACTATATTCAAATATTACTGTCAACGACTTACAGTTTCTACGTGAAAATCCACAAGTCAATCTCTTTAGACATATCT  
177 H V D L S Y N Y I Q T I T V N D L Q F L R E N P Q V N L S L D I S  
1201 TTGAACCAATTGACTTCAAGACCAAGCCTTTCAGGGAATTAAGCTCCTGAAGTACTTAAGAGTAATTTTAAAGCTCAAATATAATGAAAA  
211 L N P I D F I Q D Q A F Q G I K L H E L T L R G N F N S S N I M K  
1301 CTTGCCTTCAAACCTGGCTGGTTTACACATCCATCGTTGATCTTGGGAGAATTTAAAGATGAAAGGAATCTGGAAATTTTGAACCTCTATCATGGA  
244 T C L Q N L A G L H I H R L I L G E F K D E R N L E I F E P S I M E

**HpaI (1433)**  
1401 AGGACTATGTGATGTGACCATTGATGAGTTCAGGTTAACATATACAAATGATTTTTTTCAGATGATATTGTTAAGTTCATTGCTTGGCGAATGTTTCTGCA  
277 G L C D V T I D E F R L T Y T N D F S D D I V K F H C L A N V S A  
1501 ATGTCTCTGGCAGGTGTATCTATAAAATATCTAGAAGATGTTCCCTAAACATTTCAAATGGCAATCCTTATCAATCATTAGATGTCAACTTAAGCAGTTTC  
311 M S L A G V S I K Y L E D V P K H F K W Q S L S I I R C Q L K Q F  
1601 CAACTCTGGATCTACCCTTTCTTAAAGTTTGACTTTAACTATGAACAAAGGGTCTATCAGTTTTAAAAAGTGGCCCTACCAAGTCTCAGCTATCTAGA  
344 P T L D L P F L K S L T L T M N K G S I S F K K V A L P S L S Y L D  
1701 TCTTAGTAGAAATGCACTGAGCTTTAGTGGTTGCTGTTCTTATTCTGATTGGGAACAAACAGCCTGAGACACTGACCTCAGCTTCAATGGTGGCCATC  
377 L S R N A L S F S G C C S Y S D L G T N S L R H L D L S F N G A I

**BstXI (1810)** **SapI (1827)** **EcoRI (1878)**  
1801 ATTATGAGTGCCAAATTTTCAGTCTAGAAAGCTGCAGCACCTGGATTTTTTTCAGCACTCTACTTTAAAAAGGGTCCACAGAATTCCTCAGCGTTCTTATCCC  
411 I M S A N F M G L E E L Q H L D F Q H S T L K R V T E F S A F L S  
1901 TTGAAAAGCTACTTTACCTTGACATCTCTTATACTAACACAAAATTGACTTCGATGGTATATTTTCTGGCTTGACCAGTCTCAACACATTAATAATGGC  
444 L E K L L Y L D I S Y T N T K I D F D G I F L G L T S L N T L K M A  
2001 TGGCAATTTCTTCAAAGACAACCCCTTTCAAATGTCTTTGCAAAACACAACAACTTGCATTCCTGGATCTTTCTAAATGTCAATTGGAACAAATATCT  
477 G N S F K D N T L S N V F A N T T N L T F L D L S K C Q L E Q I S

**PvuII (2187)**  
2101 TGGGGGTATTTGACACCCTCCATAGACTTCAATTATTAATATGAGTCAACAACATCTATTGTTTTTGGATTATCCATTATAACCAGCTGTATTCCC  
511 W G V F D T L H R L Q L L N M S H N N L L F L D S S H Y N Q L Y S  
2201 TCAGCACTTTGATTGCACTTCAATCGCATAGAGACATCTAAAGGAATACTGCAACATTTTCCAAAGAGTCTAGCCTTCTTCAATCTTACTAATAATTC  
544 L S T L D C S F N R I E T S K G I L Q H F P K S L A F F N L T N S  
2301 TGTGCTGTATATGTGAACATCAGAAATTCCTGCACTGGGTCAAGGACAGAAGCAGTTCTTGGTGAATGTTGAACAAATGACATGTGCAACACCTGTGA  
577 V A C I C E H Q K F L Q W V K D Q K Q F L V N V E Q M T C A T P V

**BsrGI (2449)**  
2401 GAGATGAATACCTCCTTAGTGTGGATTTAATAATTCTACCTGTTATATGTACAAGACAATCATCAGTGTGTCAGTGGTCAAGTGTGATTGGTATCCA  
611 E M N T S L V L D F N N S T C Y M Y K T I I S V S V V S V I V V S

**NsiI (2587)**  
2501 CTGTAGCATTCTGATATACCCTTCTATTTTTCACCTGATACTTATTGCTGGCTGTA AAAAGTACAGCAGAGGAGAAAGCATCTATGATGCATTTGTGAT  
644 T V A F L I Y H F Y F H L I L I A G C K K Y S R G E S I Y D A F V I

**XhoI (2603)**  
2601 CTACTCGAGTCAGAATGAGGACTGGGTGAGAAATGAGCTGGTAAAGAATTTAGAAGAAGGAGTGCCCGCTTTTCACTCTGCCTTCACTACAGAGACTTT  
677 Y S S Q N E D W V R N E L V K N L E E G V P R F H L C L H Y R D F  
2701 ATTCCTGGTGTAGCCATTGCTGCCAATATCATCCAGGAAGGCTTCCACAAGAGCCGGAAGGTTATTGTGGTAGTGTCTAGACACTTTATTCAGAGCCGTT  
711 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 R H F I Q S R  
2801 GGTGATCTTTGATGATGAGATTGCTCAAACATGGCAGTTTCTGAGCAGCCACTCTGGCATCATCTTCTTCTGCTTGGAGAGGTTGAGAAAGTCCCTGCT  
744 W C I F E Y E I A Q E T W F L S S H S G I I F I V L E K V E K S L L  
2901 GAGGCAGCAGGTGGAATGTATCGCCTTCTTAGCAGAAACACCTACCTGGAATGGGAGGACAATCCTCTGGGAGGCACATCTTCTGGAGAAGACTTAAA  
777 R Q Q V E L Y R L L S R N T Y L E W E D N P L G R H I F W R R L K

3001 AATGCCCTATTGGATGGAAAAGCCTCGAATCCTGAGCAACAGCAGAGGAAGAACAAGAAACGGCAACTTGGACCTGAGGAGAACAAAACCTCTGGGGCCT  
811▶ N A L L D G K A S N P E Q T A E E E Q E T A T W T •

**MscI (3112)**  
**NheI (3106)**

3101 AAACCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTG

**HpaI (3244)**

3201 ATGCTATTGCTTTATTTGTAAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGA

**EcoRI (3340)**

3301 GGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCC  
3401 TTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGT

**SapI (3522)** **SspI (3579)** **Swal (3593)**

3501 TATTTTCCAAGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAA

**EcoO109I (3654)**

3601 TACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAAC

3701 AAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGAC  
141▶ • N R T Y K L P I L E E I T T K V

**SacI (3854)** **BstXI (3883)**

3801 CAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTG  
124▶ 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 R I S R  
3901 TCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGA  
90▶ 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 A E A C V

**StuI (4018)**

4001 CAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTGTCTCATA  
57▶ 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 N D E Y  
4101 GAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTA  
24▶ 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 (4226)** **SacI (4283)**

4201 TACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAACGAGCTCTGCTTATATAG

**SpeI (4381)**

4301 ACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCC  
4401 CATTGACGTCAATGGGGTGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGC

**SnaBI (4509)**

4501 GATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGTTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATA

**NdeI (4614)**

4601 GGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGCAGTTTACCCTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCG

**SdaI (4792)**

4701 TTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGGCGTCCAGCCAGCGGGCCATTTACCCTAAGTTATGTAACGCCTGCAGG

**PacI (4800)**

4801 TTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGC  
4901 ATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGT  
5001 TCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGATAGTATCTCAGTTCGGTGTAG

**ApaLI (5124)**

5101 GTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTGCAACCCGGTAAGAC  
5201 ACGACTTATCGCCACTGGCAGCAGCCTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGG  
5301 CTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTGCGCAACAAACCACCGCT  
5401 GGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGT

**EagI (5560)** **NotI (5559)**

5501 GGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTG  
5601 GTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGAAGTGCAG  
5701 GTGCCAGAACATTTCTCTATCGAA