



**PvuI (7)**  
**SgfI (6)**  
1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**HindIII (245)**  
**Psp1406I (203)** **PvuII (239)** **Bsu36I (291)**  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCGCCGCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGTCTTTGTTTCGTTT

**AgeI (552)** **NcoI (568)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCCACCATGGCTTTGCTTTGCTTGGCTATCGGATGC  
1 M A F V C L A I G C

**HpaI (663)** **Bsu36I (672)** **BamHI (692)**  
601 TTATATACCTTTCTGATAAGCACAACTTTGGCTGTACTTTCATCTTCAGACACCGAGATAAAAGTTAACCTCCTCAGGATTTGAGATAGTGGATCCCG  
11 L Y T F L I S T T F G C T S S S D T E I K V N P P Q D F E I V D P  
701 GATACTTAGGTTATCTCTATTTGCAATGGCAACCCCACTGTCTCTGGATCATTTAAGGAATGCACAGTGAATATGAACTAAAATACCGAAACATTGG  
44 G Y L G Y L Y L Q W Q P P L S L D H F K E C T V E Y E L K Y R N I G

**NcoI (897)**  
801 TAGTGAACATGGAAGACCATCATTACTAAGAATCTACATTACAAAGATGGGTTTGTCTTAAACAAGGCATTGAAGCGAAGATACACACGCTTTTACCA  
77 S E T W K T I I T K N L H Y K D G F D L N K G I E A K I H T L L P

**EcoRV (956)** **EcoRI (970)**  
901 TGGCAATGCACAAATGGATCAGAAGTTCAAAGTTCCTGGGCAGAACTACTTATTGGATATCACCACAAGGAATCCAGAACTAAAGTTCAGGATATGG  
11 W Q C T N G S E V Q S S W A E T T Y W I S P Q G I P E T K V Q D M  
1001 ATTGCGTATATTACAATTTGGCAATTTACTCTGTTCTTGAAACCTGGCATAGGTGTACTTCTTGATACCAATTACAACCTGTTTTACTGGTATGAGGG  
144 D C V Y Y N W Q Y L L C S W K P G I G V L L D T N Y N L F Y W Y E G

**NsiI (1108)**  
1101 CTTGGATCATGCATTACAGTGTGTTGATTACATCAAGGCTGATGGACAAAATATAGGATGCAGATTTCCCTATTTGGAGGCATCAGACTATAAAGATTT  
17 L D H A L Q C V D Y I K A D G Q N I G C R F P Y L E A S D Y K D F  
1201 TATATTTGTGTTAATGGATCATCAGAGAACAAGCCTATCAGATCCAGTTATTTCACTTTTCAGCTTCAAATATAGTTAAACCTTTGCCGCCAGTCTATC  
21 Y I C V N G S S E N K P I R S S Y F T F Q L Q N I V K P L P P V Y

**EcoO109I (1355)**  
1301 TTACTTTTACTCGGGAGAGTTTCTGTAATTAAGCTGAAATGGAGCATACTTTGGGACCTATTCCAGCAAGGTGTTTTGATTATGAAATTGAGATCAG  
24 L T F T R E S S C E I K L K W S I P L G P I P A R C F D Y E I E I R  
1401 AGAAGATGATACTACCTTGGTACTGCTACAGTTGAAAATGAAACATACACCTTGAAAACAACAATGAAACCCGACAATTATGCTTTGTAGTAAGAAGC  
27 E D D T T L V T A T V E N E T Y T L K T T N E T R Q L C F V V R S

**NheI (1599)**  
1501 AAAGTGAATATTTATTGCTCAGATGACGGAAATTTGGAGTGAGTGGAGTGATAAACAATGCTGGGAAGGTGAAGACCTATCGAAGAAAACCTTTGCTATGAG  
31 K V N I Y C S D D G I W S E W S D K Q C W E G E D L S K K T L L •

**MscI (1605)**  
1601 CTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTAT

**HpaI (1737)**  
1701 TGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTT

**EcoRI (1833)**  
1801 TAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTG  
1901 AGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTT

**SapI (2015)** **SwaI (2086)**  
2001 CCAAGGTTTGAAC TAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGTAATAATTCAGAAATAATTTAAATACATCA

**EcoO109I (2147)**  
2101 TTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAA  
2201 CCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTG  
141 • N R T Y K L P I L E E I T T K V L K

**SacI (2347)** **BstXI (2376)**  
2301 CCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCT  
121 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 D V E  
2401 CATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGAC  
88 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 T V

**StuI (2511)**  
2501 CCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCACTTGGTCTGATGGCCGCCCGACATGGTGTGTTGTCTCATAGAGCATG  
55 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 L M

**XmnI (2653)**  
2601 GTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAGTCGTATTACTATG  
21 T I K E T A V E V L E L D Q Q S I N F T K M

2701 **AseI (2719)** **SacI (2776)**  
 CCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCC

2801 **SpeI (2874)**  
 ACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTA**CTAGT**CAAAACAACTCCCATTGAC

2901  
 GTC AATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTA**CTG**CCAAAACCGCATCATCATGGTAATAGCGATGACT

3001 **SnaBI (3002)**  
 AATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTA**CTG**GGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTC AATAGGGGGCG

3101 **NdeI (3107)**  
 TACTTGGCATATGATACACTTGATGTA**CTG**CCAAGTGGCAGTTTACCGTAAACTCCACCCATTGACGTC AATGGAAAGTCCCTATTGGCGTTACTAT

3201 **PacI (3293)**  
 GGG AACATACGTCATTATTGACGTC AATGGGCGGGGTCGTTGGGCGGT**CAGCCAGGCGGGCCATTTACCGTAA**GGTTATGTA**ACCCTGCAGGTTAATTA**

3301 **BspLU11I (3303)**  
 AGAACATGTGAGCAAAAGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA

3401  
 AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCC**CTG**GAAAGCTCCCTCGTGGCTCTCCTGTTCCGACC

3501  
 CTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT**C**

3601 **ApaLI (3617)**  
 GCTCCAAGCTGGGCTGTGTGCACGAACCCCGTT**CAGCCGACCGTGC**GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTT

3701  
 ATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACT

3801  
 AGAAGAACAGTATTTGGTATCTGC**GCTCTGCTGAAGCCAGTTAC**CTTCGGA<sup>4033</sup>AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCG

3901  
 GTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGA

4001 **EagI (4053)**  
 AA**ACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC** **NotI (4052)** **AGCGGCCG**CAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTGTT

4101  
 TTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTCAAGTGCAGGTGCCAG

4201 AACATTTCTCTATCGAA