



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

101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**Psp1406I (203)** **PvuII (239)** **Bsu36I (291)**  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCCGCCGCCCTACCTGAGGCC  
EcoNI (287)

301 GCCATCCACGCCGGTTGAGTGCAGTCTGCCGCCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCGTTT

**AgeI (552)** **BspHI (568)**  
501 TCTGTTTGTCCGCCGTTACAGATCCAAGCTGTGACCGGCGCCTACCTGAGATCACCGGTAGGAGGCCATCATGATGTGGGGACCAAACCTTCTGCCGGTC  
1 M M W G T K L L P V  
**BspHI (689)**

601 CTGTTGCTGCAGCATGTCCTCTGCACCTCCTCCTGCTTCATGTGCCATCCCCTATGCAGAAGGACAGAAGAAAAGAAGAAATACACCTTCATGAATTTA  
11 L L L Q H V L L H L L L H V A I P Y A E G Q K K R R N T L H E F

**BbsI (729)**  
701 AAAAGTCAGCAAAAACCTACTCTTACCAAGGAAGACCCATTACTGAAGATTAACCAAAAAAGTGAACCTGCAGATGAGTGTGCCAACAGGTGTATCAG  
44 K K S A K T T L T K E D P L L K I K T K K V N S A D E C A N R C I R  
801 GAACAGGGGCTTACGTTCACTTGAAGGCCTTCGTTTTGATAAGTCAAGAAAACGATGCTACTGGTATCCTTTCAATAGTATGTCAAGTGGAGTGAAA  
77 N R G F T F T C K A F V F D K S R K R C Y W Y P F N S M S S G V K

**MscI (908)**  
901 AAAGGGTTTGGCCATGAATTTGACCTCTATGAAAACAAAGACTATATTAGAACTGCATCATTGGTAAAGGAGGCAGCTATAAAGGGACGGTATCCATCA  
111 K G F G H E F D L Y E N K D Y I R N C I I G K G G S Y K G T V S I  
1001 CTAAGAGTGGCATCAAATGCCAGCCTTGAATTCATGATCCCCATGAACACAGCTTTTTGCCTTCGAGCTATCGCGGTAAGACCTACAGGAAAACCTA  
144 T K S G I K C Q P W N S M I P H E H S F L P S S Y R G K D L Q E N Y

**XhoI (1111)** **SmaI (1127)**  
1101 CTGTGAAATCCTCGAGGGGAAGAAGGGGGACCCTGGTGTTCACAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAGTGTTCAGAAGTT  
177 C R N P R G E E G G P W C F T S N P E V R Y E V C D I P Q C S E V

**NsiI (1202)** **Eco4VII (1270)**  
1201 GAATGCATGACCTGCAATGGTGAAGCTACAGAGGTCCATGGATCACACAGAATCAGGCAAGACTTGTACGCGCTGGGACCAGACACCACACCGGC  
211 E C M T C N G E S Y R G P M D H T E S G K T C Q R W D Q Q T P H R

**EcoRV (1318)**  
1301 ACAAGTCTTCCAGAAAAGATATCCCGACAAGGGCTTTGATGATAATTATTGCCCAATCCTGATGGCAAGCCGAGGCCATGGTGTACTACTCTTGACC  
244 H K F L P E R Y P D K G F D D N Y C R N P D G K P R P W C Y T L D P

**NsiI (1484)**  
1401 TGACACCCCTGGGAGTATTGTGCAATTAACAGTGGCTCACAGTGTGTAATGAGACTGATGTCCCTATGGAAACAACCTGAATGCATTCAAGGCCAA  
277 D T P W E Y C A I K T C A H S A V N E T D V P M E T T E C I Q G Q  
**XmnI (1482)**

**EcoRV (1584)**  
1501 GGAGAAGTTACAGGGGAACAGCAATACCATTGGAATGGAATCCCTGTCAGCGTTGGGATTCGCAGTACCCTCACAAGCATGATATCACTCCCGAGA  
311 G E G Y R G T S N T I W N G I P C Q R W D S Q Y P H K H D I T P E  
1601 ACTTCAAATGCAAGGACCTTAGAGAAAATTATTGCCGAATCCAGATGGGGCTGAATCACCATGGTGTTTTACCACTGACCCAAACATCCGAGTTGGCTA  
344 N F K C K D L R E N Y C R N P D G A E S P W C F T T D P N I R V G Y

**BsaBI (1740)**  
1701 CTGCTCTCAAATCCCAAGTGTGACGTGTCAAGTGGACAAGATTGTTATCGTGCAATGGGAAAAATTACATGGGCAACTTATCCAAAACAAGTCTGGA  
377 C S Q I P K C D V S S G Q D C Y R G N G K N Y M G N L S K T R S G

**BspLUII (1803)** **NheI (1866)**  
1801 CTTACATGTTCCATGTGGGACAAGAATATGGAGGATTTACACCGTCATATCTTCTGGGAGCCAGATGCTAGCAAATGAATAAGAATTAAGTCCGGAATC  
411 L T C S M W D K N M E D L H R H I F W E P D A S K L N K N Y C R N  
1901 CTGATGATGATGCCATGGACCTTGGTGTACACGGGGAATCCTTATTCTTGGGATTATTGCCCTATTTCCCGTTGTGAAGGAGATACTACACCTAC  
444 P D D D A H G P W C Y T G N P L I P W D Y C P I S R C E G D T T P T  
2001 AATTGTCAATTTGGACCATCCTGTAATATCCTGTGCCAAAACAAAACACTGCGGGTTGTAATGGCATTCCAACACAACACAGTAGGGTGGATGGTT  
477 I V N L D H P V I S C A K T K Q L R V V N G I P T Q T T V G W M V  
2101 AGTTTGAATAACAGAAAATAACATATCTGTGGAGGATCATTGATAAAGAAAAGTTGGGTTCTTACTGCAAGACAATGTTTTCCAGCCAGAAAACAAGACT  
511 S L K Y R N K H I C G G S L I K E S W V L T A R Q C F P A R N K D

**BspHI (2237)** **BglII (2266)** **XcmI (2282)** **PvuII (2283)**  
2201 TGAAGACTATGAAGCTTGGCTTGGCATCCACGATGTTTCATGAGAGAGGGAGGAGAAGCGCAAGCAGATCTTAAACATTTCCAGCTGGTCTATGGTCC  
544 L K D Y E A W L G I H D V H E R G E E K R K Q I L N I S Q L V Y G P  
**XcmI (2282)**

**BsrGI (2388)**  
2301 TGAAGCTCAGACTTGGTTTTACTGAAGCTTGGCTGCACCTGCAATCCTGGATAACTTTGTGAGTACAATTGATTTACCTAGTTATGGTTGACAACTCCCT  
577 E G S D L V L L K L A R P A I L D N F V S T I D L P S Y G C T I P  
2401 GAAAAGACCACTTGCAGTATTTACGGCTGGGGTACACTGGATTGATCAACCGGATGGTTTATTACGAGTAGCTCATCTGTATATTATGGGGAATGAGA  
611 E K T T C S I Y G W G Y T G L I N A D G L L R V A H L Y I M G N E  
2501 AATGCAGTCAGCACCACCAAGGTAAGGCTGACTTGAATGAGTCTGAGTATGTGCTGGTGTGCTGAAAAGATTGGATCAGGACCATGTGAGGGAGATTATGG  
644 K C S Q H H Q G K V T L N E S E L C A G A E K I G S G P C E G D Y G  
2601 TGGCCACTCATTGTGAACAACAAAATGAGAATGGTTCTTGGTGTGATTGTTCTGGTGTGGATGTGCCATCCCAAATCGTCTGGTATTTTTGTT  
677 G P L I C E Q H K M R M V L G V I V P G R G C A I P N R P G I F V

2701 CGAGTAGCATATTATGCAAAATGGATACACAAAGTAATTTTGACATACAAGTTGTAATAGCCATAGAAGAGGCCCTCTAGACTAGCTGGCCAGACATGATA  
711▶ R V A Y Y A K W I H K V I L T Y K L •  
2801 AGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA

---

2901 TAAGCTGCAATAAACAAGTTAACAAACAACATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTA  
3001 CAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGG

---

3101 CATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTTAAAGATATAGTGATTTTTCCCAAGGTTTGAACACTAGCTCT  
3201 TCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTT  
3301 TTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACA  
3401 GCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGTTGACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTTCATCTCAATGAGCAC  
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I L V

---

3501 AAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTG  
115▶ 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 D S Y P H R  
3601 ACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAA  
81▶ 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 R G I Y A E I  
3701 TGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGCTTGTGCTCATAGAGCATGGTGATCTTCTCAGTGGCGAC  
48▶ 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 T I K E T A V

---

3801 CTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGA  
15▶ E V L E L D Q Q S I N F T K M  
3901 TTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCC

---

4001 ATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTG  
4101 GAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCA

---

4201 AGTAGGAAAGTCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCCTGATTGACGTCAATAGGGGGGCTACTTGGCATATGATACACT  
4301 TGATGTAAGTCCAAAGTGGGCGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTG

---

4401 ACGTCAATGGGCGGGGGTCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGAGGTTAAATAAGAACATGTGAGCAAAAGGC  
4501 CAGCAAAAGGCCAGGAACCGTAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA  
4601 GGTGGCAAAACCCGACAGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTGTCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCT

---

4701 GTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTG  
4801 CACGAACCCCGCTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCA  
4901 CTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT  
5001 CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCTCTTGTATCCGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAG  
5101 CAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTT

---

5201 TGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTGTGTAATCGTAACTA  
5301 ACATACGCTCTCCATCAAAAACAAACGAAAACAAACAACTAGCAAAATAGGCTGTCCCCAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA