



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA MfeI (82)

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGCTGTACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**HindIII (245)**  
**Psp1406I (203)** 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCCGCCGCCCTACCTGAGGGCC Bsu36I (291)

301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)** 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

**NcoI (560)**  
**BstEII (555)**  
**KasI (535)** 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCCACATGGACGATCTGTTCCCTCATCTTCCCTCAGAGCC AgeI (552)  
1 M D D L F P L I F P S E P

**SphI (653)** 601 AGCCAGGCTTCTGGCCTTATGTGGAGATCATCGAACAGCCGAAGCAACGGGCATGCGATTCCGCTATAAATGCGAGGGGCGCTCAGCGGGCAGTATT  
13 A Q A S G P Y V E I I E Q P K Q R G M R F R Y K C E G R S A G S I

**BstBI (776)** **BstEII (789)** **BamHI (798)** 701 CCTGGCGAGAGAAGCACAGATAACACCAAGACACACCCACCATCAAGATCAATGGCTACACAGGACCAGGAACAGTTCGAATCTCCCTGGTCCACCAAGG  
47 P G E R S T D T T K T H P T I K I N G Y T G P G T V R I S L V T K

801 ATCCACCTCACGGCCTCATCCACATGAAGTGTGGGGAAGGACTGCCGGATGGCTACTATGAGGCTGACCTCTGCCAGACCGCAGTATCCATAGCTT  
80 D P P H R P H P H E L V G K D C R D G Y Y E A D L C P D R S I H S F

**BamHI (911)** 901 CCAGAACCTGGGATCCAGTGTGTGAAGAAGCGAGACCTGGAGCAAGCCATTAGCCAGCAATCCAGACCAACAATAACCCCTTTCACGTTCCCTATAGAG  
113 Q N L G I Q C V K K R D L E Q A I S Q R I Q T N N N P F H V P I E

**SmaI (1058)** **PshAI (1086)** 1001 GAGCAGCGCGGGACTATGACTTGAATGCAGTGCCTCTGCTTCCAGGTGACAGTGCAGGACCCAGCAGGACGCGCCCTCCTCCTGACCCCTGCTCTCT  
147 E Q R G D Y D L N A V R L C F Q V T V R D P A G R P L L L T P V L

**Bsp120I (1121)** 1101 CACATCCGATTTTTGATAACCGGCCCAACACTGCCGAGCTCAAGATCTGCCGAGTAAACCGGAACCTGGGAGCTGCCTCGGTGGGGATGAGATCTT  
180 S H P I F D N R A P N T A E L K I C R V N R N S G S C L G G D E I F

**BsaBI (1288)** 1201 CTTGCTGTGCGACAAGGTGCAGAAAGAAGACATTGAGGTGATTTACGGGACCAGGCTGGGAGGCACGAGGCTCCTTTTCTCAAGTGTGATGCATCGG  
213 L L C D K V Q K E D I E V Y F T G P G W E A R G S F S Q A D V H R

**MscI (1304)** **BspEI (1316)** **BsiWI (1328)** **NruI (1392)** 1301 CAAGTGGCCATTGTGTTCCGGACTCCTCCGTAGCCGACCCAGCCTCCAGGCTCCTGTTGAGTCTCCATGCAGCTACGGCGCCTTCTGATCGCGAGC  
247 Q V A I V F R T P P Y A D P S L Q A P V R V S M Q L R R P S D R E

**NcoI (1409)** **ScaI (1421)** **BspHI (1496)** 1401 TCAGTGAGCCCATGGAGTCCAGTACTTCCAGACACAGATGATGCCACCGGATTGAAGAGAAGCGCAAAGGACCTATGAGACCTTCAAGAGTATCAT  
280 L S E P M E F Q Y L P D T D D R H R I E E K R K R T Y E T F K S I M

1501 GAAGAAGAGTCTTTCAATGGACCAACTGAACCCCGCCTCAACCCCGTATTGCTGTGCCATCCCGAAACTCAACTTCTGTCCCAAGCCAGCCCGC  
313 K K S P F N G P T E P R P R I A V P T R N S T S V P T R N S T S V P K P A P

1601 CAGCCCTACACCTTCCAGCATCCCTCAGCACCATCAACTTTGATGAGTTTTCCCCATGCTGTTACCATCAGGGCAGATCTCAAACAGGCCCTGGCCT  
347 Q P Y T F P A S L S T I N F D E F S P M L L P S G Q I S N Q A L A

**Acc65I (1755)**  
**NcoI (1733)** **XcmI (1751)**  
**PshAI (1731)** **NcoI (1751)** 1701 TAGCACCGTCTCTGCCAGTCCCTTGGCCAGACCATGGTCCCTTCTCAGCCATGGTACCTCTGGCTCAGCCCCAGCTCTGCCCCAGTTCTAACCC  
380 L A P S S A P V L A Q T M V P S S A M V P L A Q P P A P A P V L T P

**XmaI (1797)**  
**PstI (1882)** 1801 GGGTCTCCCGAGTCCCTGTCTGCACCTGTTCCAAAGAGCACCCAGGCTGGGGAAGGCACGCTGTCGGAAGCCCTGCTGCACCTGCAGTTTGTGCTGAT  
413 G P P Q S L S A P V P K S T Q A G E G T L S E A L L H L Q F D A D

1901 GAAGACTTGGGGCCTTGTGGCAACAGCACAGACCCAGGAGTGTTCACAGACCTGGCATCTGTGGACAACCTCAGAGTTTCCAGCAGCTCTGAACCCAG  
447 E D L G A L L G N S T D P G V F T D L A S V D N S E F Q Q L L N Q

**SmaI (2074)** **Tth111I (2070)** 2001 GTGTGTCATGTCTCACTCCACAGCTGAGCCCATGCTGATGGAGTACCCTGAAGCTATAACTCGCTGGTGACAGGGTCCCAGAGGCCCTGACCCAGC  
480 G V S M S H S T A E P M L M E Y P E A I T R L V T G S Q R P P D P A

**BspEI (2140)** **XcmI (2162)** 2101 TCCCACACCCCTGGGACCTCGGGCTTCCCAATGGTCTCCGGAGATGAAGACTTCTCCTCCATTGGGACATGGACTTCTGCTCTTTTGTGAGTCA  
513 P T P L G T S G L P N G L S G D E D F S S I A D M D F S A L L S Q

**NheI (2215)** 2201 ATCAGTCTTAAGTGTCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTG  
547 I S S •

**HpaI (2353)** **MfeI (2364)** 2301 AAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGA

2401 **EcoRI (2449)**  
GGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTA

2501 CTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGAGCCTCACCTTCTTTTCATGGAGTTTAA

---

2601 **SspI (2688)**  
GATATAGTGTATTTTCCAAGGTTTGAAGTCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATAGTAAATATTCAGAAA

2701 **SwaI (2702)**  
TAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGAC

2801 TTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGT

2901 GGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCCACCTG  
127 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 R

3001 ATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTT  
93 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 A E

3101 **StuI (3127)**  
CAGCACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTT  
60 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 N

3201 **XmnI (3269)** **BspHI (3277)**  
GTCCTCATAGAGCATGGTGTCTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGA  
27 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

3301 **AseI (3335)**  
GTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACATAACGAGCTCT

3400 **SpeI (3490)**  
GCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCA

3499 AAACAACTCCCATTTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATC

---

3599 **SnaBI (3618)**  
ATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAAGTCAATGTACTGGGCATAATGCCAGGGCGGCCATTTACCGTCAT

---

3699 **NdeI (3723)**  
TGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGT

3799 CCCTATTGGCGTTACTATGGAAACATACGTCATTATTGACGTCAATGGGGCGGGTCTTGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTA

---

3899 **PstI (3902)** **SdaI (3901)** **PacI (3909)** **BspLU11I (3919)**  
ACGCCCTGCAGGTTAA TTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGC

3997 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCG

4097 TGGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCT

---

4197 **ApaLI (4233)**  
CAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC

4297 AACCCGGTAAGACAGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGT

4397 GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA

4497 ACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG

---

4597 **EagI (4669)** **PacI (4649)** **SwaI (4658)** **NotI (4668)**  
TCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAATATCTTTATTTTC

4697 ATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACATACGCTCTCCATCAAAACAAAACGAAAACAAAACAACTAGCAAAATAGGCTGTCCCC

4797 AGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA