



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
MfeI (82) EcoNI (96)  
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

**Psp1406I (203)** 201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCGCCCGCCGCTACCTGAGGCC  
HindIII (245) Bsu36I (291)  
301 GCCATCCACGCGGTTGAGTGCAGTTCGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
EcoNI (287)  
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

---

501 TCTGTTCTGCGCGCTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCATCATGAGCTACGACTATCATCAGAGCTGGAGCCGCGATGG  
**KasI (535)** **AgeI (552)** **BspHI (560)** EcoO109I  
1 M S Y D Y H Q S W S R D G

---

**Bsp120I (600)** 601 GGGCCACGGGGTCCGGCCAGGGTCCAGCGCGCGCGCGGAGGGAGCCGGGGTCCGGCGCGCGGGAGGGGGCCGCGCGCGGGGCGGCATCCC  
**SfiI (673)** **EagI (681)** **NgoMIV (689)**  
13 G P R G S G Q G S S G G G G G S R G S G G G G G R G G R G R H P  
**NruI (714)** 701 GCACACCTTAAGGGTCCGAGATCGGCCTGTGGTACGCCAAGAAGCAGACGCAGAAGAACAAGGAGGCTGAGAGGCAGGAGAGCTGTAGTGCACATGG  
**ApaLI (790)**  
47 A H L K G R E I G L W Y A K K Q T Q K N K E A E R Q E R A V V H M

---

801 ATGAACGTCGAGAAGCAAATGTGCAGCTGCTGAATTCAGTCCAAGCTAAGACTGACAAAGATTGAGAAGCCGAGATATCTGGTTGCTCCTGAGGA  
**BstAPI (816)** **EcoRI (834)** **EcoRV (876)** **Bsu36I (892)**  
80 D E R R E E Q I V Q L L N S V Q A K T D K D S E A Q I S W F A P E D  
901 TCATGGGTATGGTACTGAAGTTTCTCAGAGAAAAAATAAACTCAGAGAAGAACTTACAACAGGAAAAGAAATTGCTAAACCAAGAAAAAAGACA  
113 H G Y G T E V S S E K K I N S E K K L D N Q E K K L L N Q E K K T

---

1001 TTTAGGATCACAGACAAATCATATATTGACCGAGATACTGAGTATTTATTGCAAGAAAATGAGCCAAACCTAAGCTTAGATCAACATTTACTGGAAGATT  
**HindIII (1071)**  
147 F R I T D K S Y I D R D T E Y L L Q E N E P N L S L D Q H L L E D

---

1101 TACAAAGGAAAAAACTGACCCTCGATATATAGAGATGCAGCGTTTCAGAAAAAGCTGCCTTCATATGGAATGCAGAAGGAGCTGGTAAATCTAATCAA  
**NdeI (1163)**  
180 L Q R K K T D P R Y I E M Q R F R K K L P S Y G M Q K E L V N L I N  
1201 TAACCATCAGGTGACAGTAATAAGTGGTAAACTGGTTGGCAAAACCACTCAGGTTACGCGATTCATCTTGGATAACTACATCGAAAGAGGAAAAAGG  
213 N H Q V T V I S G E T T G C G K T T Q V T Q F I L D N Y I E R G K G

---

1301 TCTGCCTGCAGAATAGTGTACTCAGCCAAGAAGAATTAGTGCCATTTTCAGTGCCTGAGAGAGTGGCCACAGAAAGGGCAGAGTCTTGTGGCAATGGTA  
**PstI (1305)**  
247 S A C R I V C T Q P R R I S A I S V A E R V A T E R A E S C G N G  
1401 ATAGTACTGGATACAGATTCGTCTTCAAAGTCGGTTGCCAAGGAAACAAGTTCTATCTTATACTGCACAACAGGAATCATTCTTCAGTGGCTCCAGTC  
280 N S T G Y Q I R L Q S R L P R K Q G S I L Y C T T G I I L Q W L Q S

---

1501 CGACTCACGTTTGTCCAGTGTAGTGCATATTGTAAGTTCATGAAAGGAATCTACAGTCCGATGTTTTAATGACTGTTATTAAGATCTTCTC  
**BspHI (1545)** **BglII (1590)**  
313 D S R L S S V S H I V L D E I H E R N L Q S D V L M T V I K D L L

---

1601 CATTTTCGATCTGATCTCAAAGTAATATTGATGAGTGAACCTTGAATGCTGAGAAATTTTGAATATTTTGGTAACTGTCCAATGATACATATACCCG  
**XmaI (1696)**  
347 H F R S D L K V I L M S A T L N A E K F S E Y F G N C P M I H I P

---

1701 GGTACTTTTCCAGTTGTGGAATATCTTTTGGAAAGATATCATTGAAAAAATAAGATATGTTCCAGACAAAAAGAACATAGATCCCAGTTCAAGAGGGG  
**EcoRV (1735)**  
380 G F T F P V V E Y L L E D I I E K I R Y V P D Q K E H R S Q F K R G  
1801 TTTTCATGCAGGGTATGTAATAGACAAGAAAAAGAAGAAAAAGAGGCCATCTATAAGGAACGCTGGCCAGCGTATATAAAGGAAGTCCGGACAGATAC  
413 F M Q G H V N R Q E K E E K E A I Y K E R W P A Y I K E L R T R Y  
1901 TCTGCAAGTACCGTAGATGTTTTGCAAATGATGGATGATGATAAAGTTGATCTGAATTTGATTGCTGCCCTTATTCGATACATTGTTTTGGAAGAAGAGG  
447 S A S T V D V L Q M M D D D K V D L N L I A A L I R Y I V L E E E  
2001 ATGGTGAATATTGGTCTTTCTACCAGGCTGGGACAATATCAGTACTTTGCATGATCTCTTGTGTCACAAGTATGTTAAATCAGATAAGTTTCTCAT  
480 D G A I L V F L P G W D N I S T L H D L L M S Q V M F K S D K F L I  
2101 TATACCTTACATTCAGTGCCTACCGTCAACCAGACAGGATTTAAAAAACTCCTCCGGTGTTCGAAAAATAGTAATTGCTACCAACATTGCA  
513 I P L H S L M P T V N Q T Q V F K K T P P G V R K I V I A T N I A

---

2201 GAGACTAGCATCACCATAGATGATGTGGTTTATGTAATAGATGGAGGAAAAATTAAGAAACACACTTTGACACACAGAACAACATCAGTACCATGTCTG  
**XcmI (2213)**  
547 E T S I T I D D V V Y V I D G G K I K E T H F D T Q N N I S T M S  
2301 CTGAGTGGGTTAGTAAAGCTAATGCCAAACAGAGGAAAGGGAGAGCTGGAAGAGTTGAGCCTGGTCTTATCATCTGTATAATGGTCTTAGAGCAAG  
580 A E W V S K A N A K Q R K G R A G R V Q P G H C Y H L Y N G L R A S

---

2401 TCTTCTAGATGACTACCAACTACCAGAAATTTTGAAGAACTCCTTTGGAAGAACTGTTTGAATAAAGATCTTGAGGCTGGTGGAAATTGCTTATTTT  
**XbaI (2403)** **BglIII (2468)**  
613 L L D D Y Q L P E I L R T P L E E L C L Q I K I L R L G G I A Y F

---

2501 CTGAGTAGATTAATGGATCCACCATCAATGAGGCAGTGGTGTCTCCATAAAACACCTAATGGAAGTGAAGTCTTTGGATAAGCAAGAAGAAATTAACAC  
**BamHI (2514)** **AseI (2508)**  
647 L S R L M D P P S N E A V V L S I K H L M E L S A L D K Q E E L T

2601 CTCTTGGTGTCCATTTAGCCCGACTGCCTGTTGAGCCACATATTGAAAAATGATTCTCTTTGGAGCTTTGTTCTGCTGCTTAGATCCAGTTCTACCAT  
680▶ P L G V H L A R L P V E P H I G K M I L F G A L F C C L D P V L T I  
2701 TGCTGCCAGTCTCAGCTTTAAAGATCCCTTTGTATTCCATTGGGAAAAGAAAAGATTGCAGATGCAAGAAGAAAAGAACTGGCAAAGGAAACTAGAAGT  
713▶ A A S L S F K D P F V I P L G K E K I A D A R R K E L A K E T R S

**Ppu10I (2819)**

**NsiI (2819)**

**XmnI (2884)**

2801 GATCACCTGACAGTTGTGAATGCATTTGAGGGCTGGGAAGAAGCCAAACGACGTGGTTTCAGGTATGAAAAGGACTACTGCTGGGAATATTTTCTGTCTT  
747▶ D H L T V V N A F E G W E E A K R R G F R Y E K D Y C W E Y F L S  
2901 CAAACACACTACAGATGCTGCATAACATGAAGGGACAGTTTGGTGGAGCTGGATTGTAAGCAGTAGAAGTCCCAAAGATCCAAAAGC  
780▶ S N T L Q M L H N M K G Q F A E H L L G A G F V S S R S P K D P K A  
3001 TAATATAAATTCAGATAATGAGAAGATAATTAAGCTGTCATCTGTGCTGGTTTATATCCCAAAGTTGCTAAAATCCGACTAAAATTTGGGTAAAAAGAGG  
813▶ N I N S D N E K I I K A V I C A G L Y P K V A K I R L N L G K K R  
3101 AAAATGGTAAAAGTTCACACGAAGTCTGACGGCCTGGTTCTATTCTCATCTAAGTCTGTCAACGTGGAGCAGACAGACTTCCACTACAACCTGGCTTATCT  
847▶ K M V K V H T K S D G L V S I H P K S V N V E Q T D F H Y N W L I

**BsrGI (3241)**

3201 ATCACCTGAAGATGAGAACAAGCAGTATATACTTGTACGACTGTACAGAAGTGTACCATACTGCCTCCTGTTCTTTGGAGGAGATATTTCCATCCAGAA  
880▶ Y H L K M R T S S I Y L Y D C T E V S P Y C L L F F G G D I S I Q K

**BsaBI (3301)**

3301 AGATAAGGATCAGGAAATTATTGCTGTAGATGAGTGGATTGTGTTTCAGTCTCCAGAAAAGTGGCCATCTTGTTAAGGGACTAAGAAAAGGAACTGGAT  
913▶ D K D Q E I I A V D E W I V F Q S P E R I A H L V K G L R K E L D

**XbaI (3486)**

3401 AGCCTTCTACAAGAGAAGATTGAAAGCCCTCATCTGTAGACTGGGACGACACTAAGTCAAGAGACTGTGCAGTACTGTACGATTTCTAGACTTGTATCA  
947▶ S L L Q E K I E S P H P V D W D D T K S R D C A V L S A I L D L I

**NheI (3589)**

3501 AAACCTCAAGAAAAGGCTACTCCAAGGAACCTGCCACCACGGTCCACAGGACGGATATTATAGCTGACAGCTTTCTGTGATGGCCTAAAACGCTAGCTGGCC  
980▶ K T Q E K A T P R N L P P R S Q D G Y Y S •  
3601 AGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTTATTGTGAAATTTGTGATGCTATTGCTTTATTT

**HpaI (3727) MfeI (3738)**

3701 GTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTCATTATTTATGTTTCAGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGT

**EcoRI (3823)**

3801 AAAACCTCTACAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAACCTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAAT

3901 AAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGTATTTTCCCAAGGTTTG

**SwaI (4076)**

4001 AACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAATAATTTAAATACATCATTGCAATGAA

**EcoO109I (4137)**

4101 AATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAG

4201 AAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTATCT

141▶ • N R T Y K L P I L E E I T T K V L K G N M E

**SacI (4337)**

**BstXI (4366)**

4301 CAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTA

118▶ 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 D S Y

4401 GGGGTGCCTGACAGCCCAATGGTGTCAAAGTCTTCTGCCCCTTGTCTACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATG

85▶ 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 R G I

**StuI (4501)**

4501 TAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTCTCATAGAGCATGGTATCTTCT

51▶ 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 T I K E

**XmnI (4643)**

4601 CAGTGGGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACT

18▶ T A V E V L E L D Q Q S I N F T K M

**AseI (4709)**

**SacI (4766)**

4701 ATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACAGC

**SpeI (4864)**

4801 CCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATGACGTCAATGGGG

**SnaBI (4992)**

4901 TGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGA

**NdeI (5097)**

5001 TGTACTGCCAAGTAGGAAAGTCCATAAGTTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCAT

5101 ATGATACACTTGTACTGCAAGTGGGCGTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATAC

**PaeI (5283)**

**PstI (5276)**

**SdaI (5275)**

**BspLU11I (5293)**

5201 GTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGACGGTTAAATTAAGAACTGTG

5301 AGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGC

5401 TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTA  
5501 CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCT  
ApaI (5607)  
5601 GGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG  
5701 GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAG  
5801 TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT  
5901 TGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGT  
EagI (6043)  
PacI (6023) SwaI (6032) NotI (6042)  
6001 TAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGA  
6101 ATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTCAAGTGCAGGTGCCAGAACATTTCTC  
6201 TATCGAA