



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

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCAGCGCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT **NgoMIV (441)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCAACATGTCGAGGGCGCAAGCAGCGGAAACCCAGCACATCAA **KasI (535)** **AgeI (552)**  
1 M S R R K Q A K P Q H I N

601 CTCGGAGGAGGACCAGGGCGAGCAGCAGCCGAGCAGACCCCGAGTTTGCAGATGCGGCCACAGCGCGCCCGGGGGAGCTGGGTGCTCCA **KasI (669)**  
13 S E E D Q G E Q Q P Q Q Q T P E F A D A A P A A P A A G E L G A P

701 GTGAACCCAGGGAATGACGAGGTGGCGAGTGAGGATGAAGCCACAGTAAAGCGGCTTCGTCGGGAGGAGACGACGTCTGTGAGAAATGCTGTGCGG **PshAI (771)**  
47 V N H P G N D E V A S E D E A T V K R L R R E E T H V C E K C C A

801 AGTTCCTCAGCATCTCTGAGTTCCTGGAACATAAGAAAAATTGCACTAAAAATCCACCTGTCCATCATGAATGACAGCGAGGGCCTGTGCCTTCAGA **BspHI (866)**  
80 E F F S I S E F L E H K K N C T K N P P V L I M N D S E G P V P S E

901 AGACTTCTCCGGAGCTGTACTGAGCCACCAGCCACCAGTCCCGGAGTAAGGACTGTCACAGGGAGAATGGCGGCAGCTCAGAGGACATGAAGGAGAAG **BspEI (907)**  
113 D F S G A V L S H Q P T S P G S K D C H R E N G G S S E D M K E K

1001 CCGATGCGGAGTCTGTGGTGTACCTAAAGACAGAGACAGCCCTGCCACCCACCCCCAGGACATAAGCTATTTAGCCAAAGGCAAAGTGCCAAACACTA **XcmI (1076)** **MscI (1088)**  
147 P D A E S V V Y L K T E T A L P P T P Q D I S Y L A K G K V A N T

1101 ACGTGACCTTGACGACTACGGGGACCAAGGTGGCGTGAATCAGCGGAGCGGATGCACCTCCCTGCCCGGTGCTGGTGCCAACAGCATCCCGTG **BstAPI (1159)**  
180 N V T L Q A L R G T K V A V N Q R S A D A L P A P V P G A N S I P W

1201 GGTCTCGAGCAGATCTTGTCTGCAGCAGCAGCTACAGCAGATCCAGCTCACCGAGCAGATCCGATCCAGGTGAACATGTGGGCCCTCCACGCC **XhoI (1204)** **BsaBI (1263)**  
213 V L E Q I L C L Q Q Q Q L Q Q I Q L T E Q I R I Q V N M W A S H A

1301 CTCCACTAAGCGGGGACGGGGCGACACTCTGAAGACCTTGGGAGCCACATGTCTCAGCAGTTTCTGAGCTGTGGCTTTGCTCAGCCAGAAAGCTG **BspI (1263)**  
247 L H S S G A G A D T L K T L G S H M S Q Q V S A A V A L L S Q K A

1401 GAAGCCAAGGTCTGTCTCTGGATGCCTTGAACAAAGCAAGCTACCTCAGCCAAACATCCCTTCTGCCACCAGTCCCTGTCCCAAGGCTGGCACCCCT **XmaI (1575)**  
280 G S Q G L S L D A L K Q A K L P H A N I P S A T S S L S P G L A P F

1501 CACTCTGAAGCCGGATGGGACCCGGGTGCTCCCGAACGTATGTCGCCCTCCCGAGCGCTTTGCTTCTCAGCCCCGGGCTCGGTGCTCTCCAGAGC **XmaI (1520)** **Eco47III (1555)** **Bsu36I (1567)**  
313 T L K P D G T R V L P N V M S R L P S A L L P Q A P G S V L F Q S

1601 CCTTCTCCACTGTGGCACTAGACACATCCAAGAAAGGGAAGGGGAAGCCACCGAACATCTCCGCGGTGGATGTCAAACCCAAAGACGAGGGCGCCCTCT **ScaI (1715)** **BsrBI (1764)**  
347 P F S T V A L D T S K K G K G K P P N I S A V D V K P K D E A A L

1701 ACAAGCACAAGTGAAGTACTGTAGCAAGGTTTTGGGACTGATAGCTCCTTGCAGATCCACCTCCGCTCCACACTGGAGAGAGACCCTTCTGTGCTC **ScaI (1715)** **BsrBI (1764)**  
380 Y K H K C K Y C S K V F G T D S S L Q I H L R S H T G E R P F V C S

1801 TGTCTGTGGTTCATCGCTTACCACCAAGGGCAACCTCAAGGTGCACTTCCAGCAGATCCCGAGTGAAGGCAACCCCGAGCTTTGCGGAGTTCCAG **ApaLI (1840)**  
413 V C G H R F T T K G N L K V H F H R H P Q V K A N P Q L F A E F Q

1901 GACAAAGTGGCGCGGCAATGGATCCCTATGCACTCTGTACTGACCCATAGATGAACCGAGTCTTTCTTTAGACAGCAAACCTGTCCCTGTAA **NgoMIV (1912)** **EagI (1910)**  
447 D K V A A G N G I P Y A L S V P D P I D E P S L S L D S K P V L V

2001 CCACCTCTGAGGCTACCTCAGAATCTTTCTCGGGACTAATCCAAAGGACCTCACGGTGGCTCCTTGGCCGGTACCTGCAGCCTGGGCCCTTCTCC **BstEII (2074)**  
480 T T S V G L P Q N L S S G T N P K D L T G G S L P G D L Q P G P S P

2101 AGAAAGTGAAGGTGGACCCACACTCCCTGGGGTGGGACCAAATAATCCCAAGGCTGGTGGCTTCCAAGGAGTGGGACCCCTGAGCCAGGGTCA **BstEII (2074)**  
513 E S E G G P T L P G V G P N Y N S P R A G G F Q G S G T P E P G S

2201 GAGACCCTGAAATTGAGCAGTTGGTGGAGAACATTGACAAGGCCACCACTGATCCCAACGAATGTCTCATTTGCCACCGAGTCTTAAGTGTGAGAGCT **BstEII (2380)**  
547 E T L K L Q L V E N I D K A T T D P N E C L I C H R V L S C Q S

2301 CCCTCAAGATGCATTATCGACCCACACCGGGGAGAGACCGTTCCAGTGAAGATCTGTGGCCGAGCCTTTTCTACCAAAGGTAACCTGAAGACACACCT **NsiI (2308)** **BstEII (2380)**  
580 S L K M H Y R T H T G E R P F Q C K I C G R A F S T K G N L K T H L

2401 TGGGGTTCACCGAACCAACATCCATTAAGACGAGCATTGTCGCCCATCTGCCAGAAGAAGTTCACTAATGCCGTGATGCTGCAGCAACATATCCG **BstEII (2380)**  
613 G V H R T N T S I K T Q H S C P I C Q K K F T N A V M L Q Q H I R

2501 ATGCACATGGGCGGTGAGATTCCCAACACGCCCCGTCAGAGAATCCCTGTGACTTTACGGGTTCTGAGCCAATGACCGTGGGTGAGAACGGCAGCACCG **SgrAI (2595)**  
647 M H M G G Q I P N T P L P E N P C D F T G S E P M T V G E N G S T

**ClaI (2632)**

2601 GCGCTATCTGCCATGATGATGTCATCGAAAGCATCGATGTAGAGGAAGTCAAGTCCAGGAGGCTCCAGCAGCTCCTCCAAGGTCACCGCCTCTTCC  
680▶ G A I C H D D V I E S I D V E E V S S Q E A P S S S S K V P T P L P  
2701 CAGCATCCACTCGGCATACCCACGCTAGGGTTTCCATGATGGCTTCCCTTAGATGCCCCAGGAAAGTGGTCTGCCCTTTTAACTGCAGCGCCAG  
713▶ S I H S A S P T L G F A M M A S L D A P G K V G P A P F N L Q R Q **EcoRV (2893)**

2801 GGCAGCAGAGAAAACGGTTCCTGGAGAGCGATGGCTTGACCAACGACTCATCCTCGCTGATGGGAGACCAGGAGTATCAGAGCCGAAGCCAGATATCC  
747▶ G S R E N G S V E S D G L T N D S S S L M G D Q E Y Q S R S P D I  
2901 TGGAAACCATCTTCCAGGCACTCTCCCGGCAATAGTCAAGCCGAAAGCATCAAGTCAAGTCTCCGATGCTGGGAGCAAAGCAGAGAGCTCCGA  
780▶ L E T T S F Q A L S P A N S Q A E S I K S K S P D A G S K A E S S E **PshAI (3063)**

3001 GAACAGCCGCACTGAGATGGAAGGTCGGAGCAGTCTCCCTTCCACGTTTATCCGAGCCCCGCCGACCTATGTCAAGGTTGAAGTTCCTGGCACATTTGTG  
813▶ N S R T E M E G R S S L P S T F I R A P P T Y V K V E V P G T F V **DraIII (3179)**

**Tth111I (3108)** **DraIII (3177)**

3101 GGACCCTCGACATTGTCCCAGGGATGACCCCTTGTAGCAGCCAGCCAGCCGACAGGCCAAGCAACATGGCTGCACACGGTGTGGGAAGAAGTCT  
847▶ G P S T L S P G M T P L L A A Q P R R Q A K Q H G C T R C G K N F **Eco47III (3208)**

**NheI (3205)** **BsrBI (3226)**

3201 CCTCTGCTAGCGCTCTTCCAGATCCACGAGCGGACTCACACTGGAGAGAAGCCTTTGTGTGCAACATTTGTGGGCGAGCTTTTACCACCAAAGGCAACT  
880▶ S S A S A L Q I H E R T H T G E K P F V C N I C G R A F T T K G N L **MscI (3359)** **NcoI (3374)**

3301 AAAGTTCACTACATGACACACGGGGCAACAATAACTCAGCCCCCGTGAAGGAAGTTGGCCATCGAGAACCACCATGGCTCTGTTAGGTACGGACGGA  
913▶ K V H Y M T H G A N N N S A R R G R K L A I E N T M A L L G T D G **SphI (3482)**

**XcmI (3422)**

3401 AAAAGAGTCTCAGAAATCTTCCCAAGGAAATCCTGGCCCTTCAAGTGAATGTGGACCCTGTTGTGTGGAACAGTACACGACATGCTCAATGGCGGT  
947▶ K R V S E I F P K E I L A P S V N V D P V V W N Q Y T S M L N G G  
3501 TGGCCGTGAAGACCAATGAGATCTCTGTGATCCAGAGTGGGGGGTCTACCTCCCGTTTCTTGGGGGCCACCTCCGTTGTAATAACGCCACTGT  
980▶ L A V K T N E I S V I Q S G G V P T L P V S L G A T S V V N N A T V  
3601 CTCAAGATGGATGGCTCCAGTCGGGTATCAGTGCAGATGTGAAAAACCAAGTGTACTGACGGGTTCCCAAACACAGTTTCTCACTTCTGGAA  
1013▶ S K M D G S Q S G I S A D V E K P S A T D G V P K H Q F P H F L E **NheI (3756)**

**EcoRI (3750)**

3701 GAAAACAAGATTGCGGTCACTAAGGGAGAAGTTCGTTGGAAGGAGCAATGAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAAACC  
1047▶ E N K I A V S •

3801 ACAACTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT

---

**HpaI (3919)** **MfeI (3930)**

3901 TATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTC

---

**EcoRI (4015)**

4001 TACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATA  
4101 GGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTGCT

---

**SspI (4254)** **Swal (4268)**

4201 CTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGATAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATG  
4301 TTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGA  
4401 CAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGGGGGATGAGTTCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGC  
1141▶ N R T Y K L P I L E E I T T K V L K G N M E I L  
4501 ACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCC  
1154▶ 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 P H R **StuI (4693)**

4601 TGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGTCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTC  
824▶ 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  
4701 AATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCG  
494▶ 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 T A  
4801 ACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTTTCATGGTGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGAT  
154▶ V E V L E L D Q Q S I N F T K M **AseI (4901)**

4901 GATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGC

---

**SpeI (5056)**

5001 CCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCATGGGGTGGAGAC

---

**SnaBI (5184)**

5100 TTGGAATCCCCGTGAGTCAAACCCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTACTG

---

**NdeI (5289)**

5200 CCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGCATAATGCCAGGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATAC

5300 ACTTGATGTAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTA

5400 TTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTG C A G G T T A A T T A A G A A C A T G T G A G C A A  
5498 AAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACA AAAATCGACGCTCAAG  
5598 TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGCTTACCGGA  
5698 TACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCT  
ApaLI (5799)  
5798 GTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC  
5898 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTT  
5998 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTT  
6098 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGG  
EagI (6235)  
6198 GATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGT  
PacI (6215) SwaI (6224) NotI (6234)  
6298 AACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCG  
6398 AA