



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

101 GAGAAAGTGGCGGGGTAACCTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCCGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTCCGCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTGTTTCGTTT **NgoMIV (441)**

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTAGGAGGGCCAGCATGCTACGAGCTCTTTGGCTCTTCTGGATC **KasI (535)** **AgeI (552)** **SphI (568)** **SapI (587)**  
1▶ M L R A L W L F W I

601 TTGGTGGCATAACAGTCTCTTTCAGCAACGCTGTTCTGCTCAGGAGTCTGTGTATGATGCTTCTGGGTGTGTATGGCCGCTCAGGCTTTTCA **BsrBI (683)**  
11▶ L V A I T V L F S K R C S A Q E S L S C D A S G V C D G R S R S F

701 CCTCTATTCCTCCGGACTCACAGCAGCCATGAAAAGCCTTGACCTGTCTTTCAACAAGATCACCTACATTGGCCATGGTGACCTCCGAGCGTGTGCGAA **BstEII (777)**  
44▶ T S I P S G L T A A M K S L D L S F N K I T Y I G H G D L R A C A N **NcoI (773)**

801 CCTCCAGGTTCTGATTTTGAAGTCCAGCAGAATCAATAACAATAGAGGGAGACGCTTTTATTCTCTGGCAGTCTTGAACATTTGGATTTGTCTGATAAT **Bsp120I (930)**  
77▶ L Q V L I L K S S R I N T I E G D A F Y S L G S L E H L D L S D N **EcoO109I (930)**

901 CACCTATCTAGTTTATCTTCTCCTGTTCCGGCCCTTTCTCTTTGAAATACTTAACTTAATGGGAAATCCTTACCAGACACTGGGGTAACATCGC **Bsu36I (1028)**  
111▶ H L S S L S S S W F G P L S S L K Y L N L M G N P Y Q T L G V T S

1001 TTTTTCCCAATCTCACAAATTTACAAACCTCAGGATAGGAAATGTAGAGACTTTCAGTGAAGATAAGGAGAATAGATTTTGTGGGCTGACTTCTCTCAA **HindIII (1361)**  
144▶ L F P N L T N L Q T L R I G N V E T F S E I R R I D F A G L T S L N  
1101 TGAACCTGAAATTAAGGCATTAAGTCTCCGGAATTACAGTCCCAAAGTCTAAAGTCGATCCGCGACATCCATCAGCTTCTCACTTAAAGCGAGTCT  
177▶ E L E I K A L S L R N Y Q S Q S L R I R D I H L T L H L S E S  
1201 GCTTTCCTGCTGGAGATTTTGCAGATATTCTGAGTTCGTGAGATATTTAGAACTAAGAGATACTAAGTGGCCAGGTTCCAGTTTTCCACTGCCCG  
211▶ A F L L E I F A D I L S S V R Y L E L R D T N L A R F Q F S P L P

1301 TAGATGAAGTCAGCTCACCGATGAAGAAGCTGGCATTCCGAGGCTCGGTTCTCACTGATGAAAGCTTAAACGAGCTCCTGAAGCTGTTGCGTTACATCTT **HindIII (1361)**  
244▶ V D E V S S P M K K L A F R G S V L T D E S F N E L L K L L R Y I L  
1401 GGAAGTGTGCGAGGTAGAGTTCGACGACTGTACCTCAATGGGCTCGGCGATTCAACCCCTCGGAGTCAGACGTAGTGAAGGAGCTGGGTAAAGTAGAA  
277▶ E L S E V E F D D C T L N G L G D F N P S E S D V V S E L G K V E

1501 ACAGTCACTATCCGGAGGTTGCATATCCCCAGTTCATTTGTTTATGACCTGAGTACTGTCTATTCCCTCCTGGAGAAGTGAAGCGAATCACAGTAG **ScaI (1554)**  
311▶ T V T I R R L H I P Q F Y L F Y D L S T V Y S L L E K V K R I T V

1601 AGAACAGCAAGGCTTCCCTGGTTCCCTGCTCGTTCTCCAGCATTAAATCATTAGAAATCTTAGACCTCAGCGAAAATCTGATGGTTGAAGAATATTT **BbsI (1611)** **EcoRI (1656)** **SspI (1693)**  
344▶ E N S K V F L V P C S F S Q H L K S L E F L D L S E N L M V E E Y L  
1701 GAAGAACTCAGCTGTAAGGGAGCCTGGCCTTCTACAAACCTTAGTTTGGCCAGAACTATTGAGATCAATGCAAAAAACAGGAGAGATTTTGTCTG  
377▶ K N S A C K G A W P S L Q T L V L S Q N H L R S M Q K T G E I L L

1801 ACTCTGAAAAACCTGACCTCCCTTGACATCAGCAGGAACACTTTTCATCCGATGCCCGACAGCTGTGAGTGGCCAGAAAAGATGCGCTTCTGAAATTTGT **PshAI (1857)**  
411▶ T L K N L T S L D I S R N T F H P M P D S C Q W P E K M R F L N L

1901 CCAGTACAGGGATCCGGGTGGTAAAAACGTGCATTCTCAGACGCTGGAGGTGTTGGATGTTAGTAACAACAATCTTGACTCATTCTTTCTTTGTTCTTGGC **BamHI (1909)**  
444▶ S S T G I R V V K T C I P Q T L E V L D V S N N N L D S F S L F L P

2001 TCGGCTGCAAGAGCTCTATATTTCCAGAAATAAGCTGAAAACACTCCCAGATGCTTCTGTTGTTCCCTGTGTTGCTGATGAAAATCAGAGAGAATGCA **BspHI (2077)**  
477▶ R L Q E L Y I S R N K L K T L P D A S L F P V L L V M K I R E N A

2101 GTAAGTACTTTCTCTAAAGACCAACTTGGTTCTTTTCCAAACTGGAGACTCTGGAAGCAGGCGACAACCACTTTGTTTGTCTCCTGCGAACTCCTATCCT **ScaI (2103)**  
511▶ V S T F S K D Q L G S F P K L E T L E A G D N H F V C S C E L L S  
2201 TTACTATGGAGACGCCAGCTCTGGCTCAAATCCTGGTTGACTGCGCAGACAGTACCTGTGTGACTCTCCGCCTCGCCTGCACGCGCCACAGGCTTCAGGA  
544▶ F T M E T P A L A Q I L V D W P D S Y L C D S P P R L H G H R L Q D  
2301 TGCCCGCCCTCCGTCTTGAATGTCCAGGCTGCAGTGGTGTCTGGAGTCTGCTGTGCCCTTCTCCTGTTGATCTTGCTCGTAGGTCCTGTGCCAC  
577▶ A R P S V L E C H Q A A L V S G V C C A L L L L I L L V G A L C H

2401 CATTTCCACGGAAGTGGTACCTGAGAATGATGTGGCGTGGCTCCAGGCCAAGAGGAAGCCCAAGAAAGCTCCCTGCAGGGACGTTTGTATGATGCCT **Acc65I (2416)** **PstI (2474)**  
611▶ H F H G L W Y L R M M W A W L Q A K R K P K K A P C R D V C Y D A **SdaI (2473)**

2501 TTGTTTCTACAGTGAGCAGGATCCCATTTGGGTGGAGAACCCTCATGTTCCAGCAGCTGGAGAAGTCTGACCCGCCCTTTAAGCTGTGTCTCCACAAGCG  
644▶ F V S Y S E Q D S H W V E N L M V Q Q L E N S D P P F K L C L K R

2601 GGACTTCTTCCGGGCAAAATGGATCATTGACAACATCATCGATTCCATCGAAAAGAGCCACAAAAGTGTGTTCTGTTGCTTCTGAGAAGTTCGTACGGAGC **ClaI (2637)** **BsiWI (2690)**  
677▶ D F V P G K W I I D N I I D S I E K S H K T V F V L S E N F V R S

2701 GAGTGGTGAAGTACGAAGTGGACTTCTCCACTTCAGGCTCTTTGACGAGAACAACGACGGCCATCCTTGTGTTTGTGGAGCCCATTGAGAGGAAAG  
711▶ E W C K Y E L D F S H F R L F D E N N D A A I L V L L E P I E R K

2801 CCATTCCCCAGCGCTTCTGCAAAGTGCAGCAAGATAATGAACACCAAGACCTACCTGGAGTGGCCCTTGGATGAAGGCCAGCAGGAAGTGTGTTGGGTAAA  
744▶ A I P Q R F C K L R K I M N T K T Y L E W P L D E G Q Q E V F W V N

2901 TCTGAGAACTGCAATAAAGTCTAGGTTCTCCACCCAGTTCCTGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC  
777▶ L R T A I K S •

3001 GAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACACAACAATG

3101 CATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTCTAAATAACAGCATAG  
EcoRI (3179)

3201 CAAAACCTTAACTCCAATCAAGCCTCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTCATTAGCTGTT

3301 TGCAGCCTCACCTCTTTTCATGGAGTTTAAAGATATAGTATTTTTCCCAAGGTTTGAAGTACTAGCTCTTTCATTCTTTATGTTTTAAATGCACTGACCTCCC  
SapI (3361)

3401 ACATTCCCTTTTATAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCT  
SspI (3418) SwaI (3432) EcoO109I (3493)

3501 TCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGT  
141▶ • N R T

3601 GTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCT  
137▶ Y K L P I L E E I T T K V L K G N M E I L V F C D P A Y D S I L E

3701 CTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGT  
103▶ R C M G C P S V V R I S R D V E D S Y P H R V A V I T D F D K Q G N

3801 TGCTCACAGCAGACCAATGGCAATGGCTTCCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGT  
70▶ S V A S G I A I A E A C V T V R G I Y A E I H V A S I I E G T K T

3901 CCTGATGGCCGCCGACATGGTCTTGTGTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTG  
37▶ R I A A G V H H K N D E Y L M T I K E T A V E V L E L D Q Q S I N

4001 AAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAG  
34▶ F T K M

4101 CTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATT

4201 TTGGAAAGTCCCCTGATTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCTGAGTCAAACCGCTATCCACGCC  
SpeI (4220)

4301 ATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAGGTCATGTAAGTGGCA  
SnaBI (4348)

4401 TAATGCCAGGCGGCCATTTACCCTGATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCCTGAAAT  
NdeI (4453)

4501 ACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGACCC  
PstI (4632) SdaI (4631) PacI (4639) BspLU11I (4649)

4601 AGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGT  
4701 TGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA  
4801 GCGGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTT  
ApaLI (4963)

4901 TCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCAAGCTGGGTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCT  
5001 TATCCGGTAACTATCGTCTTGTAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG  
5101 CGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAA  
5201 AGAGTTGGTAGCTCTTGTACCGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAG  
EagI (5399)

5301 AAGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCAATGGCTAGTTAATTAACATTTAAATC AGC  
PacI (5379) SwaI (5388) NotI (5398)

5401 GGCCGCAATAAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAA  
5501 ACAAACTAGCAAAATAGGCTGTCCCAAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA