



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

101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**Psp1406I (203)** 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC **HindIII (245)** **EcoNI (287)**

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT **NgoMIV (441)**

501 TCTGTTTGTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTA **KasI (535)** **AgeI (552)** CTGAGATCACCGGTCAGCATGCCGGAAAACCTCCGTAGTGACGCTGGTTTGAATC **SphI (560)**

601 AGACACCGCAATGAAAAAGGGGAGACACTGCGAAAGCAAACCGAGGAGAAAAGAGAAAAAGAGAAGCCAAAATCTGATAAGACTGAAGAGATAGCAGAA **1 M P G K L R S D A G L E S**  
13 D T A A M K K G E T L R K Q T E E K E K K E K P K S D K T E E I A E  
701 GAGGAAGAACTGTTTTCCCAAAGCTAAACAAGTAAAAAGAAAAGCAGAGCCTTCTGAAGTTGACATGAATTTCTCTAAATCCAAAAAGGCAAAAAAGA **1**  
47 E E E T V F P K A K Q V K K A E P S E V D M N S P K S K K A K K  
801 AAGAGGAGCCATCTCAAATGACATTTCTCTAAAACAAAAGTTTGAGAAAAGAAAAGGAGCCATTGAAAAGAAAAGTGGTTTCTTAAAACAAAAA **1**  
80 K E E P S Q N D I S P K T K S L R K K K E P I E K K V V S S K T K K  
901 AGTGACAAAAATGAGGAGCCTTCTGAGGAAGAAATAGATGCTCCTAAGCCCAAGAAGATGAAGAAAAGAAAAGAAAATGAATGGAGAACTAGAGAGAAA **1**  
113 V T K N E E P S E E E I D A P K K M K K E K E M N G E T R E K  
1001 AGCCCAAACCTGAAGAATGGATTTCTCATCTGAACCCGACTGAACCCAGTGAAGCTGCCAGTGAAGAAAAGTAACAGTGAGATAGAGCAGGAAATAC **1**  
147 S P K L K N G F P H P E P D C N P S E A A S E E S N S E I E Q E I  
1101 CTGTGAACAAAAAGAAGGCGCTTCTCTAATTTCCCATATCTGAAGAACTATTAACCTTCTCAAAGCCGAGGAGTGACCTTCTATTCTCTATACA **1**  
180 P V E Q K E G A F S N F P I S E E T I K L L K G R G V T F L F P I Q

1201 AGCAAAGACATTCATCATGTTTACAGCGGGAAGGACTTAATTGCACAGGCACGGACAGGAACTGGGAAGACATTCTCCTTTGCCATCCCTTTGATTGAG **XmnI (1266)**  
213 A K T F H H V Y S G K D L I A Q A R T G T G K T F S F A I P L I E  
1301 AAACCTCATGGGAACTGCAAGACAGGAAGAGAGGCCGTGCCCTCAGTACTGTTCTTGACCTACAAGAGAGTTGGCAAATCAAGTAAGCAAAGACT **1**  
247 K L H G E L Q D R K R G R A P Q V L V L A P T R E L A N Q V S K D

1401 TCAGTGACATCACAAAAAGCTGTGAGTGGCTTGTTTTTATGGTGAACCTCCCTATGGAGGTCATTTGAACGCATGAGGAATGGGATTGATATCCTGGT **EcoRV (1489)**  
280 F S D I T K K L S V A C F Y G G T P Y G G Q F E R M R N G I D I L V **BsaBI (1485)**

1501 TGGAACACCAGGTCGTATCAAAGACCACATACAGAATGGCAAACCTAGATCTCACCAAACCTAAGCATGTTGTCCTGGATGAAGTGACCAGATGTTGGAT **XcmI (1524)** **BglII (1545)**  
313 G T P G R I K D H I Q N G K L D L T K L K H V V L D E V D Q M L D  
1601 ATGGGATTGCTGATCAAGTGAAGAGATTTTAAAGTGTGCATACAAGAAAGATTCTGAAGACAATCCCAAACATTGCTTTTTCTGCAACTTGCCTC **1**  
347 M G F A D Q V E E I L S V A Y K K D S E D N P Q T L L F S A T C P

1701 ATTGGGTATTTAATGTTGCCAAGAAATACATGAAATCTACATATGAACAGGTGGACCTGATTGGTAAAAAGACTCAGAAAACGGCAATAACTGTGGAGCA **NdeI (1739)**  
380 H W V F N V A K K Y M K S T Y E Q V D L I G K K T Q K T A I T V E H  
1801 TCTGGCTATAAGTGCCACTGGACTCAGAGGGCAGCAGTTATTGGGGATGTCATCCGAGTATATAGTGGTCATCAAGGACGCATATCATCTTTTGTGAA **1**  
413 L A I K C H W T Q R A A V I G D V I R V Y S G H Q G R T I I F C E  
1901 ACCAAGAAAGAAGCCAGGAGCTGTCCAGAATTCAGCTATAAAGCAGGATGCTCAGCTTGCATGGAGACATTCCACAGAAGCAAAGGAAATCACCC **1**  
447 T K K E A Q E L S Q N S A I K Q D A Q S L H G D I P Q K Q R E I T

2001 TGAAGGTTTTAGAAATGGTAGTTTTGGAGTTTTGGTGGCAACCAATGTTGCTGCACGTGGGTAGACATCCCTGAGGTTGATTTGGTTATACAAAGCTC **BbrPI (2054)**  
480 L K G F R N G S F G V L V A T N V A A R G L D I P E V D L V I Q S S

2101 TCCACCAAAGGATGTAGAGTCTACATTCATCGATCCGGGCGGACAGGCAGAGCTGGAAGGACGGGGTGTGCATCTGCTTTTATCAGCACAAGGAAGAA **ClaI (2129)**  
513 P P K D V E S Y I H R S G R T G R A G R T G V C I C F Y Q H K E E

2201 TATCAGTTAGTACAAGTGGAGCAAAAAGCGGGAATTAAGTTCAAACGAATAGGTGTTCTTCTGCAACAGAAAATAATAAAAGCTTCCAGCAAAGATGCCA **HindIII (2279)** **XcmI (2297)**  
547 Y Q L V Q V E Q K A G I K F K R I G V P S A T E I I K A S S K D A  
2301 TCAGGCTTTTGGATTCCGTGCTCCCACTGCCATTAGTCACTTCAAACAATCAGCTGAGAAGCTGATAGAGGAGAAGGGAGCTGTGGAAGCTCTGGCAGC **1**  
580 I R L L D S V P P T A I S H F K Q S A E K L I E E K G A V E A L A A

2401 AGCACTGGCCATATTTTCAGGTGCCACGTCCTGATGACCACTCAATGTGGTTTTGTGACCATGATCTTGCAGTGTCTCAATTGAA **Eco47III (2438)**  
613 A L A H I S G A T S V D Q R S L I N S N V G F V T M I L Q C S I E **AfeI (2438)**

2501 ATGCCAAATATTAGTTATGCTTGAAAAGAACTTAAAGAGCAGCTGGGCGAGGAGATTGATTTCAAAGTGAAGGAATGGTTTTCTCAAAGGAAAGCTGG **SspI (2506)**  
647 M P N I S Y A W K E L K E Q L G E E I D S K V K G M V F L K G K L

2601 GTGTTTCTTTGATGTACCTACCGCATCAGTAACAGAAATACAGGAGAAATGGCATGATTACGACGCTGGCAGCTCTGTGGCCACAGAGCAACCAGA **MscI (2681)**  
680 G V C F D V P T A S V T E I Q E K W H D S R R W Q L S V A T E Q P E  
2701 ACTGGAAGGACCAGGGAAGGATATGGAGGCTTCAAGGGACAGCGGGAAGGCAGTTCAGGCTTCAAGGGACAGCGGACGGAACAGAAGATTCAAGGGA **1**  
713 L E G P R E G Y G G F R G Q R E G S R G F R G Q R D G N R R F R G

2801 CAGCGGGAAGGCAGTAGAGGCCCGAGAGGACAGCGATCAGGAGGTGGCAACAAAAGTAAACAGATCCAAAACAAAGCCAGAAGCGGAGTTTCAGTAAAG  
747▶ Q R E G S R G P R G Q R S G G G N K S N R S Q N K G Q K R S F S K  
MscI (2920)  
2901 CATTGGTCAATAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTGTGA  
780▶ A F G Q •  
HpaI (3052)  
3001 AATTTGTGATGCTATTGCTTTATTGTAAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAATTGCATTATTTATGTTTCAGGTTTCAGGGGGAG  
3101 GTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTAC  
3201 TTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAG  
SspI (3387)  
3301 ATATAGTGTATTTTCCCAAGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGTAATAATTCAGAAAT  
SwaI (3401) EcoO109I (3462)  
3401 AATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACT  
3501 TAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTGATGTTCTGGTGTACTTGGAGGGGATGAGTTCCTCAATGGTG  
141▶ • N R T Y K L P I L E E I T  
SacI (3662) BstXI (3691)  
3601 GTTTTGACCAGCTTGGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCCCTGA  
126▶ 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 I  
3701 TGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTC  
93▶ 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  
StuI (3826)  
3801 AGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTG  
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  
XmnI (3968)  
3901 TCCTCATAGAGCATGGTGTCTTCTCAGTGGGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAG  
26▶ 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  
AseI (4034) SacI (4091)  
4001 TCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACGCTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGC  
SpeI (4189)  
4101 TTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAA  
4201 CAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATG  
SnaBI (4317)  
4301 GTAATAGCGATGACTAATACGTAGATGTAAGTCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTACCCTGATTGA  
NdeI (4422)  
4401 CGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCC  
4501 TATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGCTGTTGGGCGGTACAGCAGGCGGGCCATTACCCTGAAAGTTATGTAACG  
PacI (4608)  
PstI (4601)  
SdaI (4600) BspLU11I (4618)  
4601 CCTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCC  
4701 TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCC  
4801 TCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTT  
ApaLI (4932)  
4901 CGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC  
5001 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCT  
5101 AACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAA  
5201 CCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA  
EagI (5368)  
PacI (5348) SwaI (5357) NotI (5367)  
5301 CGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAATATCTTTATTTTCATTAC  
5401 ATCTGTGTGGTTTTTTTGTGTGAATCGTAACATACGCTCTCCATCAAAAACAAAACGAAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGC  
5501 AAGTGCAGGTGCCAGAACATTTCTCTATCGAA