



**PvuI (7)**  
**SgfI (6)**  
1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA

101 GAGAAGGTGGCGCGGGTAAACTGGAAAAGTATGTCGTGTACTGGCTCCGCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**Psp1406I (203)** **HindIII (245)**  
201 GTGAACGTTCTTTTTCGAACGGGTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGCTCAACTCTACGCTTTTGTTCGTTT

**AgeI (552)** **BspHI (560)** **BspEI (574)** **SandI (583)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTATCATGAAGTCTGGCTCCGGAGGAGGGTCCCCGACCTCGCT

601 GTGGGGGCTCCTGTTTCTCTCCGCCGCTCTCTCGCTCTGGCCGACAGTGGAGAAATCTGCGGCCAGGCATCGACATCCGCAACGACTATCAGCAGCTG  
13> W G L L F L S A A L S L W P T S G E I C G P G I D I R N D Y Q Q L  
701 AAGCGCTGGAGAACTGCACGGTATCGAGGGTACCTCCACATCCTGCTCATCTCCAAGGCCGAGGACTACCGCAGTACCGCTTCCCAAGCTCACGG  
47> K R L E N C T V I E G Y L H I L L I S K A E D Y R S Y R F P K L T

**ScaI (809)** **XhoI (838)** **SacII (879)**  
801 TCATTACCGAGTACTTGTGCTGTTCCGAGTGGCTGGCCTCGAGAGCCTCGGAGACCTTCCCCAACCTCACGGTATCCGCGGCTGGAACCTTTCTA  
80> V I T E Y L L L F R V A G L E S L G D L F P N L T V I R G W K L F Y  
901 CAACTACGCCCTGGTTCATCTTCGAGATGACCAATCTCAAGGATATTGGGCTTACAACCTGAGGAACATTACTCGGGGGCCATCAGGATTGAGAAAAAT  
113> N Y A L V I F E M T N L K D I G L Y N L R N I T R G A I R I E K N

**PshAI (1099)**  
1001 GCTGACCTCTGTTACCTCTCCACTGTGGACTGGTCCCTGATCTGGATGCGGTGTCCAATAACTACATTGTGGGAATAAGCCCCAAAGGAATGTGGGG  
147> A D L C Y L S T V D W S L I L D A V S N N Y I V G N K P P K E C G

**NcoI (1115)**  
1101 ACCTGTGTCCAGGACCATGGAGGAGAAGCCGATGTGTGAGAAGACCACCATCAACAATGAGTACAACCTACCCTGCTGGACCACAAACCGCTGCCAGAA  
180> D L C P G T M E E K P M C E K T T I N N E Y N Y R C W T T N R C Q K

**BbrPI (1212)** **ApaLI (1230)** **BssHII (1280)**  
1201 AATGTGCCAAGCACGTGTGGGAAGCGGCGTGACCCGAGAACAATGAGTGTGCCACCCGAGTGCCTGGGAGTGCAGCGCGCTGACAACGACACG  
213> M C P S T C G K R A C T E N N E C C H P E C L G S C S A P D N D T  
1301 GCCTGTGTAGCTTCCCGCACTACTACTATGCCGGTGTCTGTGTGCCTGCCTGCCGCCAACACCTACAGGTTTGGGGCTGGCGCTGTGTGGACCGTG  
247> A C V A C R H Y Y A G V C V P A C P P N T Y R F E G W R C V D R

**SphI (1469)**  
1401 ACTTCTGCGCAACATCCTCAGCGCCGAGAGCAGCAGCTCCGAGGGGTTGTGATCCACGACGGGAGTGCATGCAGGAGTGCCTCCGCGCTTATCCG  
280> D F C A N I L S A E S S D S E G F V I H D G E C M Q E C P S G F I R

**BstAPI (1514)**  
1501 CAACGGCAGCCAGAGCATGACTGCATCCCTTGTGAAGGTCCTTGCCCGAAGGTCCTGTGAGGAAGAAAAGAAAACAAAGACCATTGATTCTGTTACTTCT  
313> N G S Q S M Y C I P C E G P C P K V C E E E K K T K T I D S V T S  
1601 GCTCAGATGCTCAAGGATGCACCATCTCAAGGGCAATTTGCTCATTAAACATCCGACGGGGAATAACATTGCTTCAGAGCTGGAGAACTTCATGGGGC  
347> A Q M L Q G C T I F K G N L L I N I R R G N N I A S E L E N F M G

**AvrII (1785)**  
1701 TCATCGAGGTGGTACGGGCTACGTGAAGATCCGCCATTCTCATGCCTTGGTCTCCTTGTCTTCTAAAAACCTTCGCTCATCTAGGAGAGGAGCA  
380> L I E V V T G Y V K I R H S H A L V S L S F L K N L R L I L G E E Q  
1801 GCTAGAAGGGAATTACTCCTTACGTCTCGACAACCAGAATTGACGAACTGTGGGACTGGGACCACCGCAACCTGACCATCAAAGCAGGGAAAATG  
413> L E G N Y S F Y V L D N Q N L Q Q L W D W D H R N L T I K A G K M  
1901 TACTTTGCTTTCAATCCCAAATTATGTTTCCGAAATTTACCGCATGGAGGAAGTACGGGGACTAAAGGGCCAAAGCAAAGGGGACATAAACACCA  
447> Y F A F N P K L C V S E I Y R M E E V T G T K G R Q S K G D I N T

**Acc65I (2093)**  
**AgeI (2090)**  
2001 GGAACAACGGGGAGAGAGCCTCCTGTGAAAGTACGTCCTGCATTTACCTCCACCACCACGTCGAAGAATCGCATCATCATAACCTGGCACCCTGACCG  
480> R N N G E R A S C E S D V L H F T S T T T S K N R I I I T W H R Y R  
2101 GCCCCTGACTACAGGATCTCATCAGCTTACCGTTTACTACAAGGAAGCACCTTTAAGAATGTACAGAGTATGATGGGAGGATGCCTGCGGCTCC  
513> P P D Y R D L I S F T V Y Y K E A P F K N V T E Y D G Q D A C G S

**XcmI (2234)**  
2201 AACAGCTGGAACATGGTGGACGTGGACCTCCCGCCCAACAAGGACGTGGAGCCGGCATCTTACTACATGGGCTGAAGCCCTGGACTCAGTACGCCGTTT  
547> N S W N M V D V D L P P N K D V E P G I L L H G L K P W T Q Y A V

**BsrGI (2366)**  
**NcoI (2321)** **BglII (2360)**  
2301 ACGTCAAGGCTGTGACCTCACCATGGTGGAGAAGCACCATATCCGTGGGGCCAAAGAGTGGAGTGTGATCATTTCGACCAATGCTTCAGTTCCTTCCAT  
580> Y V K A V T L T M V E N D H I R G A K S E I L Y I R T N A S V P S I  
2401 TCCCTGGACGTTCTTTCAGCATCGAACTCCTTCTCAGTTAATCGTGAAGTGGAAACCTCCCTCTGCCCCAACGGCAACCTGAGTTACTACATTGTG  
613> P L D V L S A S N S S S Q L I V K W N P P S L P N G N L S Y Y I V  
2501 CGCTGGCAGCGGACGCTCAGGACGGCTACCTTACCAGGACAATTAAGTCTCCTCAAGACAAAATCCCATCAGGAAGTATGCCGACGGCACCATCGACA  
647> R W Q R Q P Q D G Y L Y R H N Y C S K D K I P I R K Y A D G T I D

2601 TTGAGGAGGTACAGAGAACCCCAAGACTGAGGTGTGTGGTGGGGAGAAAGGGCCTTGTGCGCTGCCCCAAAAGTGAAGCCGAGAAGCAGGCCGAGAA  
680▶ I E E V T E N P K T E V C G G E K G P C C A C P K T E A E K Q A E K  
2701 GGAGGAGCTGAATACCGCAAAGTCTTTGAGAATTTCTGCACAACTCCATCTTCGTGCCAGACCTGAAAGGAAGCGGAGAGATGTCATGCAAGTGGCC  
713▶ E E A E Y R K V F E N F L H N S I F V P R P E R K R R D V M Q V A  
EagI (2834) SapI (2866)  
2801 AACACCACCATGTCCAGCCGAAGCAGGAACACCACGGCCGAGACACCTACAACATCACIGACCCGGAAGAGCTGGAGACAGAGTACCCTTTCTTTGAGA  
747▶ N T T M S S R S R N T T A A D T Y N I T D P E E L E T E Y P F F E  
EcoRV (2965)  
ClaI (2962)  
2901 GCAGAGTGGATAACAAGGAGAGAAGTGTCAATTTCTAACCTTCGGCCTTTACATTGTACCCGCATCGATATCCACAGCTGCAACCACGAGGCTGAGAAGCT  
780▶ S R V D N K E R T V I S N L R P F T L Y R I D I H S C N H E A E K L  
3001 GGGCTGCAGCGCTCCAACCTTGTCTTTGCAAGGACTATGCCCGCAGAAGGAGCAGATGACATTCCTGGGCCAGTGACCTGGGAGCCAAGGCCTGAAAAC  
813▶ G C S A S N F V F A R T M P A E G A D D I P G P V T W E P R P E N  
3101 TCCATCTTTTTAAAGTGGCCGGAACCTGAGAATCCCAATGGATTGATTCTAATGTATGAAATAAAATACGGATCACAAGTTGAGGATCAGCGAGAATGTG  
847▶ S I F L K W P E P E N P N G L I L M Y E I K Y G S Q V E D Q R E C  
XmaI (3252)  
3201 TGTCCAGACAGGAATACAGGAAGTATGGAGGGGCAAGCTAAACCCGGCTAAACCCGGGGAACCTACACAGCCCGGATTGAGCCACATCTCTCTCTGGAA  
880▶ V S R Q E Y R K Y G G A K L N R L N P G N Y T A R I Q A T S L S G N  
3301 TGGTCTGTGGACAGATCCTGTCTTCTATGTCCAGGCCAAAACAGGATATGAAAATTCATCCATCTGATCATCGCTCTGCCCGTCTGCTGTCTGTG  
913▶ G S W T D P V F F Y V Q A K T G Y E N F I H L I I A L P V A V L L  
ScaI (3497)  
3401 ATCGTGGGAGGTTGGTATTGCTGTACGTCTTCCATAGAAAAGAGAATAACAGCAGGCTGGGGAATGGAGTGTGTATGCCTCTGTGAACCCGGAGT  
947▶ I V G G L V I M L Y V F H R K R N N S R L G N G V L Y A S V N P E  
Eco47III (3505)  
3501 ACTTCAGCGCTGTGTGTACGTTCTGATGAGTGGGAGGTGGCTCGGAGAAGATCACCATGAGCCGGGAACCTGGGAGGGTCTTTGGGATGGT  
980▶ Y F S A A D V Y V P D E W E V A R E K I T M S R E L G Q G S F G M V  
SphI (3680)  
3601 CTATGAAGGAGTTGCCAAGGTTGGTGAAGATGAACCTGAAACCAGAGTGGCCATTAACAGTGAACGAGGCCGCAAGCATGCGTGAAGAGGATTGAG  
1013▶ Y E G V A K G V V K D E P E T R V A I K T V N E A A S M R E R I E  
HindIII (3710) DraIII (3739)  
3701 TTTCTCAACGAAGCTTGTGATGAAGGAGTCAATTGTCCATGTGGTGGGATTGCTGGTGTCCCAAGGCCAGCCAACTGGTTCATCATGG  
1047▶ F L N E A A S V M K E F N C H H V V R L L G V V S Q G Q P T L V I M  
3801 AACTGATGACACGGGCGATCTCAAAGTTATCTCCGCTCTGAGGCCAGAAATGGAGAATAATCCAGTCCTAGCACCTCAAGCCTGAGCAAGATGAT  
1080▶ E L M T R G D L K S Y L R S L R P E M E N N P V L A P P S L S K M I  
3901 TCAGATGGCCGAGAGATTGACAGCGCATGGCATACTCAACGCCAATAAGTTCGTCACAGAGACCTTGTGCCCGAATTGCATGGTAGCCGAAGAT  
1113▶ Q M A G E I A D G M A Y L N A N K F V H R D L A A R N C M V A E D  
EcoRV (4036)  
4001 TTCACAGTCAAAATCGGAGATTTTGGTATGACGCGAGATATCTATGAGACAGACTATTACCGAAAAGGAGGAAAGGGCTGCTGCCGTGCGCTGGATGT  
1147▶ F T V K I G D F G M T R D I Y E T D Y Y R K G G K G L L P V R W M  
4101 CTCCTGAGTCCCTCAAGGATGGAGTCTTACCACCTTACTCGGACGTCTGGTCTTCGGGGTCTGCTCTGGGAGATCGCCACACTGGCCGAGCAGCCCTA  
1180▶ S P E S L K D G V F T T Y S D V W S F G V V L W E I A T L A E Q P Y  
PshAI (4264) FspI (4295)  
4201 CCAGGGCTTGTCCAACGAGCAAGTCTTCGCTTCGTCATGGAGGGCGCCTTCTGGACAAGCCAGACAACCTGTCTGACATGCTGTTGAACTGATGCGC  
1213▶ Q G L S N E Q V L R F V M E G G L L D K P D N C P D M L F E L M R  
4301 ATGTGCTGGCAGTATAACCCCAAGATGAGGCCTTCTTCTGGAGATCATCAGCAGCATCAAAGAGGAGATGGAGCCTGGCTTCCGGAGGTTCTCTTCT  
1247▶ M C W Q Y N P K M R P S F L E I I S S I K E E M E P G F R E V S F  
4401 ACTACAGCGAGGAGAACAAGCTGCCCGAGCCGAGGAGCTGGACCTGGAGCCAGAGAATGGAGAGCGTCCCCTGGACCCCTCGGCTCCTCGTCCTC  
1280▶ Y Y S E E N K L P E P E E L D L E P E N M E S V P L D P S A S S S S  
4501 CCTGCCACTGCCGACAGACACTCAGGACACAAGGCCGAGAAGGCCCGGCCCTGGGTGCTGGTCTCCCGCCAGCTTCGACGAGAGACAGCCTTAC  
1313▶ L P L P D R H S G H K A E N G P G P G V L V L R A S F D E R Q P Y  
BsrBI (4627) NheI (4671)  
4601 GCCACATGAACGGGGCCGCAAGAACGAGCGGCCTTGGCGCTGCCAGTCTTCGACCTGTGATCCTTGTAGCTGGCCAGACATGATAAGATACAT  
1347▶ A H M N G G R K N E R A L P L P Q S S T C •  
4701 TGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGC  
HpaI (4809)  
4801 AATAAACAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTG  
EcoRI (4905)  
4901 GTATGGAATCTAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTGAACTCTTTCTGAGGATGAATAAGGCATAGGCATCAGGG  
SapI (5087)  
5001 GCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGTGTGAACTAGCTCTTCATTTCT  
SspI (5144) SmaI (5158)  
5101 TTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTA  
5201 GGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATGGACAGCAAGAAA

5301 GCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGT  
 141 • N R T 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  
 SacI (5419)

5401 CAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCAC  
 112 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 V A V

5501 AATGGTGTCAAAGTCCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTACAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACA  
 79 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 I H V

5601 GCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCA  
 45 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 V E V L  
 AseI (5791)

5701 GCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCCTATAGTGAGTGTATTATACTATGCCGATATACTATGCCGATGATTAATTGT  
 12 E L D Q Q S I N F T K M  
 SacI (5848)

5801 CAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTTGCGT  
 SpeI (5946)

5901 CAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCC  
 SnaBI (6074)

6001 CGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAA  
 NdeI (6179)

6101 AGTCCATAAGGTCATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAA  
 6201 TGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGACGTCAAT  
 PacI (6365)

6301 GGGCGGGGTCGTTGGGCGTCCAGCCAGCGGGCCATTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAA  
 SdaI (6357) BspLU11I (6375)

6401 GGCCAGGAACCGTAAAAAGCCGCTTGGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGA  
 6501 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCCT  
 ApaLI (6689)

6601 TTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACC  
 6701 CCCCCTTCCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAAC  
 6801 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTC  
 6901 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGAT  
 7001 TACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATG  
 EagI (7125)

7101 GCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGC  
 PacI (7105) SwaI (7114) NotI (7124)

7201 TCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA