



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

**Psp1406I (203)** 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC  
**Bsu36I (291)**  
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT

**BspHI (560)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCCTACCTGAGATCACCGGTATCATGAGCTGGTACCGTCCCTCCCAACCCAGACATGTGG  
**AgeI (552)** 1► M S W S P S L P T Q T C G  
**BstEII (569)**  
**BspLU11I (592)**  
**BsaBI (681)**

601 AGCCTGGAAATGAAAGAACGCCTGGGACCGGGGATTTGAAACGTCATCCGGTGGCACAATCAGGCGACAGGTGAACAGATCGCCATCAAGCAATGC  
13► A W E M K E R L G T G G F G N V I R W H N Q A T G E Q I A I K Q C

**BspHI [m] (750)** 701 CGACAGGAGCTCAGCCAAAGAACAGAGACCGCTGGTGCCTCGAAATCCAGATCATGAGAAGGCTGAACCATCCCAATGTGGTGGCTGCCCGGATGTCC  
47► R Q E L S P K N R D R W C L E I Q I M R R L N H P N V V A A R D V  
**XmaI (788)**

**NcoI (845)** 801 CAGAGGGGATGCAGAACCTGGCACCCAATGATTTGCCACTGCTGGCCATGGAGTACTGCCAAGGAGGATCTCCGAAGATACTGAACCAAGTTCGAGAA  
80► P E G M Q N L A P N D L P L L A M E Y C Q G G D L R R Y L N Q F E N  
901 CTGCTGTGGCCTGCGGAAGGAGCTGTCTTACCTGTGAGTGACATCGCATCGGCTCTTAGATACCTTACGAAAACAGAATCATCCATCGAGACCTG  
113► C C G L R E G A V L T L S D I A S A L R Y L H E N R I I H R D L

**AseI (1038)** 1001 AAGCCAGAAAACATCGTTCTGCAGCAAGGAGAGAAAAGATTAACACAAAATTATTGATCTAGGATATGCCAAGGAGCTGGATCAGGGCAGTCTGTGCA  
147► K P E N I V L Q Q G E K R L I H K I I D L G Y A K E L D Q G S L C  
1101 CGTCATTTGTGGGACTCTGCAATACCTGGCGCAGAGCTTCTGGAGCAGCAGAAGTACACCGTGAACCGTGGACTACTGGAGCTTCCGCCACCTGGCCTT  
180► T S F V G T L Q Y L A P E L L E Q Q K Y T V T V D Y W S F G T L A F  
**SapI (1272)**

1201 CGAGTGCATCACTGGCTCCGGCCCTTCTCCCTAACTGGCAGCTGTGCACTGCACTCCAAAGTCCGGCAGAGAGGAAGTGGACATCGTTGTTAGT  
213► E C I T G F R P F L P N W Q P V Q W H S K V R Q K S E V D I V V S  
1301 GAAGACTTGAATGGAGCAGTGAAGTTTTCAAGTTCGCTACCCTTCCCAATAATCTTAACAGTGTCTTGGCTGAACCGCTGGAGAAGTGGCTGCAGCTGA  
247► E D L N G A V K F S S S L P F P N N L N S V L A E R L E K W L Q L

**BamHI (1431)** 1401 TGCTTATGTGGCACCTCGGCAAAGGGCACGGATCCCCAGTATGGCCCAACGGCTGCTTCCAGACCCCTGGATGACATCTTGAACCTGAAGCTGGTTCA  
280► M L M W H P R Q R G T D P Q Y G P N G C F R A L D D I L N L K L V H  
**BamHI (1595)**

1501 TGTCTGAACATGGTCACAGGCACCGTTACACATACCCCGTACCGGAGGATGAGAGTCTGCAGAGCTAAAAACCAGAATCCAGGAAGACACGGGATC  
313► V L N M V T G T V H T Y P V T E D E S L Q S L K T R I Q E D T G I  
1601 CTGAGACAGACCAGGAGCTGCTCAAGAGGCAGGGCTGGTGTGCTCCCTGACAAGCCTGCTACTCAGTGCATCTCAGACAGCAAGACAAACGAGGGCC  
347► L E T D Q E L L Q E A G L V L L P D K P A T Q C I S D S K T N E G  
1701 TCACGTTGGACATGGATGTTTCTTTTGTGACAAACAGTAAACTCAACTAGAGATCAGATCACCCCGACCCCAACCGGAAAGTGTCAAGTGTGAT  
380► L T L D M D L V F L F D N S K I N Y E T Q I T P R P Q P E S V S C I  
1801 CCTCAGGAGCCCAAGCGGAACCTCTCCTTCTCCAGTGAAGAAAGTGTGGGCCAAGTCTGGCACAGCATCCAGACGCTGAAGGAAGACTGTAACCGG  
413► L Q E P K R N L S F F Q L R K V W G Q V W H S I Q T L K E D C N R

**SfiI (1977)**  
**NcoI (1973)**  
1901 CTGCAGCAGGGACAGCGAGCAGCCATGATGAGTCTCCTCCGAATAACAGTGCCTCTCTAAGATGAAGAAGCCATGGCCTCCACGGCCAGCAGCTCA  
447► L Q Q G Q R A A M M S L L R N N S C L S K M K N A M A S T A Q Q L  
2001 AGGCCAAGCTGGACTTCTCAAACCAGCATCCAGATCGACCTGGAGAAGTATAAAGAGCAGACCGAGTTTGGGATCACCTCAGATAAATGTGCTGGC  
480► K A K L D F F K T S I Q I D L E K Y K E Q T E F G I T S D K L L L A

**BsrBI (2164)**  
2101 TTGGCGGAGATGGAGCAGGCTGTGGAGCAGTGTGGCGGGAGAATGACGTGAAGCATCTAGTAGAGCGGATGATGGCACTGCAGACTGACATTGTGGAC  
513► W R E M E Q A V E Q C G R E N D V K H L V E R M M A L Q T D I V D  
2201 CTGAGAGGAGCCGATGGGTCGGAAGCAGGGGGCACCTGGATGACCTAGAGGAACAAGGAGGAGCTCTACCGAAGACTCAGGGAGAAGCCAAAG  
547► L Q R S P M G R K Q G G T L D D L E E Q A R E L Y R R L R E K P R  
2301 ACCAAAGGACAGAAGGTGACAGCCAGGAGATGGTACGGCTGCTGCTTCAAGCAATCCAAAGCTTTGAGAAGAAAGTTCGGGTGATTTATACAGCTCAG  
580► D Q R T E G D S Q E M V R L L L Q A I Q S F E K K V R V I Y T Q L S

**PshAI (2484)**  
2401 TAAGACCGTGGTTTGAAGCAGAAGGCACTGGAGTTGCTGCCAAGTGAAGAGGTAGTGGACCTTATGAACGAGGACGAGAGGACCGTGGTCCGGCTT  
613► K T V V C K Q K A L E L L P K V E E V V S L M N E D E R T V R L  
2501 CAGGAGAAGCGGCAAGGAACCTGGAACCTCCTGAAGATCGCTGTAGCAAAGTCCGAGGTCCTGAGTGGAAAGCCAGACAGCATGAATGTGTCT  
647► Q E K R Q K E L W N L L K I A C S K V R G P V S G S P D S M N V S

**Tth111I (2600)**  
2601 GACTCAGTACCCTGGTCAGCTAATGTCCAGCCTTCCAGTGCCTGTGACAGCTTACCTGAATCAGACAAGAAAGTGAAGAAGTGGTGGCCGAAGCCCA  
680► R L S H P G Q L M S Q P S S A C D S L P E S D K K S E E L V A E A H

2701 CGCCCTCTGCTCCCGCTAGAAAAGTGCCTGCAGGACACTGTGAAGGAGCAAGACAGAAGCTTACGACTCTAGACTGGAGCTGGTTACAGATGGAGGAT  
713▶ A L C S R L E S A L Q D T V K E Q D R S F T T L D W S W L Q M E D  
XbaI (2769)

2801 GAAGAAAGGTGTAGCCTGGAGCAGGCCTGTGACTGAGGTGCCATGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC  
747▶ E E R C S L E Q A C D •  
StuI (2822) XcmI (2841) MseI (2853)

2901 TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTTAAACAACAAT  
HpaI (2985) MfeI (2996)

3001 TGCATTCATTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCAT  
EcoRI (3081)

3101 AGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTG

3201 TTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAC TAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTC  
SapI (3263)

3301 CCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC  
SspI (3320) SwaI (3334)

3401 CTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCTG  
141 • N R

3501 GTGTA CTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCT  
137▶ 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 P A Y D S I L E  
3601 CTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCC  
104▶ 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

3701 GTTGCTCACAGCAGCCCAATGGCAATGGCTTACGACAGACAGTGCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTG  
71▶ 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 A S I I E G T K  
3801 GTCCTGATGGCCGCCGACATGGTGTGTTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGT  
37▶ 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 E L D Q Q S I N

3901 TGAAGTCTTCATGGTGGCCCTCTATAGTGTGATTATACTATGCGGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCC  
4▶ F T K M  
XmnI (3901) AseI (3967)

4001 AGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACA

4101 TTTTGAAAGTCCC GTTGATTTACTAGTCAAAACAAACTCCCATTTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGC  
SpeI (4122)

4201 CCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGTAGCTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGG  
SnaBI (4250)

4301 CATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGACTGCCAAGTGGGCGAGTTTACCGTAA  
NdeI (4355)

4401 ATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAG

4501 CCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGAAAAGGCCAGGAACCGTAAAAAGGCCG  
PacI (4541) SdaI (4533) BspLU11I (4551)

4601 GTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC  
4701 CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC

4801 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGC  
ApaLI (4865)

4901 CTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA  
5001 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA  
5101 AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCA  
PacI (5281) SwaI (5290)

5201 AGAAGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC A  
EagI (5301)  
NotI (5300)

5301 GCGGCCGCAATAAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACA  
5401 AAACAACTAGCAAAATAGGCTGTCCCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA