



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

**Psp1406I (203)** **HindIII (245)** **Bsu36I (291)**  
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

**KasI (535)** **AgeI (552)** **BspHI (560)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTTCATGAGCTCCCAAAGCCATCCAGATGGACTTTCTGGCCG  
1 M S S Q S H P D G L S G R

**Eco47III (669)**  
601 AGACCAGCCTGTGGAGCTGCTGAATCCTGCCCGTGTGAACCACATGCCAGCACGGTGGACGTGGCTACAGCGCTGCCTCTGCAAGTGGCCCTACAGCA  
13 D Q P V E L L N P A R V N H M P S T V D V A T A L P L Q V A P T A  
701 GTACCCATGGACCTGCGCTTGACCACCAAGTTCCTACTGCCCTTGAACCTGCATTGCGGGAGCAGCAACTGCAGCAGGAACCTCTAGCACTGAAACAGA  
47 V P M D L R L D H Q F S L P L E P A L R E Q Q L Q Q E L L A L K Q  
801 AGCAGCAGATCCAGCGGCAGATACTCATTGCAGAGTTCAGCGCTCAACATGAGCAGTTGTCCCGACAGCATGAGGCACAGTTGCATGAACATATCAAGCA  
80 K Q Q I Q R Q I L I A E F Q R Q H E Q L S R Q H E A Q L H E H I K Q

**BsrBI (961)**  
901 GCAGCAGGAGATGCTGGCCATGAAGCACCAGCAGGAGCTGCTGGAGCACCAGCGGAAACTGGAGCGGACCCGGAAGAGCAGGAGCTGGAGAAGCAGCAC  
113 Q Q E M L A M K H Q Q E L L E H Q R K L E R H R Q E Q E L E K Q H  
1001 CGTGAGCAGAAGCTGCAGCAGCTCAAGAACAAGGAGAAGGGCAAGAGAGTGTGTGGCGAGCACAGAGGTGAAGATGAAGCTCAGGAGTTTGTCTCA  
147 R E Q K L Q Q L K N K E K G K E S A V A S T E V K M K L Q E F V L  
1101 ACAAGAAGAAGGCTTAGCCACCGAACCTGAACCCTGCATTTCCAGCGATCCCGCTACTGGTATGGGAAGACACAGCAGCTCCCTTGACCAGAG  
180 N K K K A L A H R N L N H C I S S D P R Y W Y G K T Q H S S L D Q S  
1201 CTCTCCACCCAGAGTGGGGTGTACGCTCCTACAACCACCCCGTCTTGGAAATGTACGACGCCAAAGATGACTTCCCTCTTAGGAAAACAGCTTCTGAA  
213 S P P Q S G V S A S Y N H P V L G M Y D A K D D F P L R K T A S E

**FspI (1364)**  
1301 CCTAACCTGAAATTACGCTCAAGGCTTAAGCAGAAAAGTAGCTGAGAGACGGAGTAGCCCCCTGTTGCGCAGGAAAGATGGCCCTGTGGCCACTGCTCTAA  
247 P N L K L R S R L K Q K V A E R R S S P L L R R K D G P V A T A L

**SphI (1429)**  
1401 AAAAGCGACCCCTGGATGTTACAGACTCCGCATGCAGCAGCGCCCTGGCTCCGGTCCAGCTCTCCAAATAGCAGCTCTGGCAACGTGAGCACTGAGAA  
280 K K R P L D V T D S A C S S A P G S G P S S P N S S S G N V S T E N

**XhoI (1562)**  
1501 TGGCATCGACCCACTGTGCCAGTGTCCAGCTGAGACGAGCTTGGCACACAGACTTGTGACTCGAGAAGGCTCAGTCGCCCACTTCTCTCTACAG  
313 G I A P T V P S A P A E T S L A H R L V T R E G S V A P L P L Y T

**XcmI (1613)**  
1601 TCACCATCCTTACCACATCACCTTGGGACTTCTGCCACTGCCCTGCCGCTGGTGCAGGAGTGCAGCAGGATGCTGAGAGGCTTGTCTCCAGCTC  
347 S P S L P N I T L G L P A T G P A A G A A G Q Q D A E R L A L P A

**SaI (1723)**  
1701 TCCAGCAGCGGATCTTGTCCCTGGGACCCACCTCACCCGTACCTGAGCACCTCGCCCTGGAGAGGGACGGTGCAGCAGCTCACACCCCTCTGCA  
380 L Q Q R I L F P G T H L T P Y L S T S P L E R D G A A A H N P L L Q  
1801 GCACATGGTCTGTGGAGCAGCCACCCACAGACCCCTTGTACAGGCTGGGGGCGCTGCCCTCCACTCACAGTCCCTGGTGGTGGGACAGG  
413 H M V L L E Q P P T Q T P L V T G L G A L P L H S Q S L V G A D R

**DraIII (1990)**  
1901 GTGTCCCATCCATTACAAGCTGCGGCAGCACCCGCTCTGGGGCGCACGAGTGCAGCACCCCTGCCGAGAACGCACAGGCCCTGCAGCAGCTGGTGA  
447 V S P S I H K L R Q H R P L G R T Q S A P L P Q N A Q A L Q H L V  
2001 TCCAGCAGCAGCACCAGCAGTTCCTGGAGAAGCACAAGCAACAGTTCAGCAGCAGCAGTGCACCTCAGCAAGATAATCTCCAAACCTAGTGAGCCACC  
480 I Q Q Q H Q Q F L E K H K Q Q F Q Q Q L H L S K I I S K P S E P P  
2101 TCGCAGCCTGAGAGCCACCCAGAGGAGACAGAGGAGGAGTCCGTGAGCACCAGGCTTGTGGATGAGCCCTACCTAGATCGGCTACCTGGGAGAAAG  
513 R Q P E S H P E E T E E E L R E H Q A L L D E P Y L D R L P G Q K

**XhoI (2270)**  
2201 GAGCCCTCCCTGGCTGGTGTGCAGGTGAAGCAGGAGCCATTGAGAGTGGAGGAGAAGAGCGGAGGCCACTCGAGAGACAGAGCCCGCCAGCGCCAG  
547 E P S L A G V Q V K Q E P I E S E E E A E A T R E T E P G Q R P

**BamHI (2354)**  
2301 CCACTGAGCAGGAGTGTCTTTCAGACAGCAAGCCCTCCTACTGGAGCAGCAGGATCCACAGTAAAGAACTACCAGGCATCTATGGAGGCTGCTGG  
580 A T E Q E L L F R Q Q A L L L E Q Q R I H Q L R N Y Q A S M E A A G

**PshAI (2426)**  
2401 CATCCCTGTGTCATTTGGCAGCCACAGACTCTGTCTCGGGCACAGTCTCCCGAGCATCTGCCACCTTCCCATGTCAGTCCAGGAGCCCCACCAAG  
613 I P V S F G S H R P L S R A Q S S P A S A T F P M S V Q E P P T K

**SphI (2583)** **BamHI (2594)**  
2501 CCAAGTTTACCACAGGTCTTGTGTATGACACACTGATGTTGAAGCATCAGTGCACCTGTGGGAACACCAACAGCCACCCGGAGCATGCTGGGAGGATCC  
647 P R F T T G L V Y D T L M L K H Q C T C G N T N S H P E H A G R I

**SdaI (2616)**  
2601 AGAGCATCTGGTCCCGCTGCAGGAGACTGGACTCCGTGGCAAGTGTGAGTGCATCCGTGGACGCAAGGCCACATTGGAGGAGCTGCAGACAGTGCACCT  
680 Q S I W S R L Q E T G L R G K C E C I R G R K A T L E E L Q T V H S  
2701 GGAGGCCACACACTCCTCTACGGCACAAATCCTCTCAACAGACAGAACTGGACAGCTCGTGCACCTCAGTGTTCGTCAGGCTTCTTGTGGTGGTGT  
713 E A H T L L Y G T N P L N R Q K L D S S L T S V F V R L P C G G V

NdeI (2817)  
2801 GGGGTGGATAGCGACACCATATGGAATGAGGTGCACTCGTCTGGGGCAGCCCGCTGGCTGTAGGCTGTGTAGTGGAGCTGGTCTTCAAGGTGGCCACGG  
747▶ G V D S D T I W N E V H S S G A A R L A V G C V V E L V F K V A T  
2901 GAGAGCTAAAGAATGGCTTTGCTGTGGTTCGTCCTCCAGGACACCATGCCGAGGAGACACACCATGGGTTTCTGCTACTTTAACTCCGTGGCAGTTGC  
780▶ G E L K N G F A V V R P P G H H A E E S T P M G F C Y F N S V A V A

BsrGI (3057)  
3001 AGCCAAACTTCTCCAGCAGAGGCTGAATGTGAGCAAGATCCTCATTGTAGACTGGGATGACATCGGGAATGGGACCCAGCAGGCCTTCTACAATGAC  
813▶ A K L L Q Q R L N V S K I L I V D W D V H H G N G T Q Q A F Y N D

BspLUII (3113)  
3101 CCCAATGTTCTACATGTCCTGCACCGCTATGACGATGGGAACCTTCTCCAGGAAGTGGAGCACCAGATGAGGTGGGCACAGGGCCAGGCCTGGGTT  
847▶ P N V L Y M S L H R Y D D G N F F P G S G A P D E V G T G P G V G

ScaI (3254)  
3201 TCAATGTCAACATGGCTTTCACGGGTGGCTCGAACCCCATGGGAGACGCTGAGTACTTGGCAGCCTTCAAACCGGTGGTTATGCCTATCGAAATGA  
880▶ F N V N M A F T G G L E P P M G D A E Y L A A F K T V V M P I A N E  
3301 GTTTGCCCCAGATGTGGTACTGGTGTATCGGGCTTCCGATGCTGTGGAGGGCCACCCACACCTTGGAGGGTACAATCTCTCTGCCAAATGTTTTGGG  
913▶ F A P D V V L V S S G F D A V E G H P T P L G G Y N L S A K C F G

NgoMIV (3434)  
3401 TACTTGACAAAACAGCTGATGGGCTTAGCTGGTGGCCGCTTGTGCTGGCCCTTGGAGGAGGCCATGACCTGACAGCCATCTGTGATGCTTCTGAAGCCT  
947▶ Y L T K Q L M G L A G G R L V L A L E G G H D L T A I C D A S E A

BsaBI (3597)  
3501 GCGTGTCTGCTGCTGGAAACGAGCTTGAGCCTCTGCCAGAAAAGTTCTACATCAGAGACCAATGCCAATGCTGTCCACTCCATGGAGAAAAGTGAT  
980▶ C V S A L L G N E L E P L P E K V L H Q R P N A N A V H S M E K V M

ScaI (3614)  
3601 GGACATCCACAGCAAGTACTGGCGCTGCCTGCAGCGTCTGTCCTCCACGGTGGGCACTCTCTGATTGAGGCGAAAAGTGTGAGAAGGAAGAAGCCGAG  
1013▶ D I H S K Y W R C L Q R L S S T V G H S L I E A Q K C E K E E A E

BglII (3751)  
3701 ACAGTACC GCCATGGCCTCGTGTCTGTAGGCGTCAAACCTGCTGAGAAGAGATCTGAGGAGGAGCCATGGAGGAGGAACCACCACTGTAGCCGAAG  
1047▶ T V T A M A S L S V G V K P A E K R S E E E P M E E E P P L •

NheI (3820)  
3801 CTGCTGTTCTCTCTTTGTTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTAT

HpaI (3958) MfeI (3969)  
3901 TTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAGTTAAACAACAACAATTGCATTCAATTTATGTTTCAGGTTCCAG

EcoRI (4054)  
4001 GGGGAGGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGC  
4101 CTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAG

SspI (4293)  
4201 TTTAAGATATAGTATTTTTCCCAAGGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTC

SwaI (4307)  
4301 AGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGT  
4401 TGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCA  
4501 ATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAAGGATGATGAGTCTGAGATGAGCTCTGACATGCCACAGGGGCTGACCA  
128▶ 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 R C M G C P S V V  
4601 CCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAAT  
95▶ 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 S V A S G I A I  
4701 GGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGC  
62▶ 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 R I A A G V H H  
4801 TTGTTGCTCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCACTCCAGATCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCCTAT  
28▶ 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 F T K M

AseI (4940)  
4901 AGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTIATCTGACGGTTCCTAAACGAG

SpeI (5095)  
5001 CTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTGTTGATTTACTAG  
5101 TCAAAACAAACTCCCATTGACGTC AATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATC

SnaBI (5223)  
5201 ATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCCATAGGTGATGACTGGGCATAATGCCAGGGCCATTTACCGT

NdeI (5328)  
5301 CATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAA  
5401 AGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTGCTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTAT

SdaI (5506) PacI (5514) BspLUII (5524)  
5501 GTAACGCCTGCAGGTTAAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG

5601 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTC  
5701 GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC  
5801 TCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCAGCCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC  
5901 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG  
6001 TGGCCTAACTACGGCTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTTTGATCCGGCA  
6101 AACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGG

6201 GTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC **EagI (6274)**  
**NotI (6273)**  
PacI (6254) SwaI (6263)  
6301 CATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCC  
6401 CAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA