



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGGATCGCTCCGGTGCCCGTCAGTGGGACAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
**MfeI (82)** **EcoNI (96)**  
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

**KasI (535)** **AgeI (552)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAACATGTCCACCACGGTCAATGTAGATTCCTTGCGAATA  
1 M S T T V N V D S L A E Y  
601 TGAGAAGAGCCAGATCAAGAGAGCCCTGGAGCTGGGACGGTGTGACTGTGTTGAGCTTCCGCAAGTCCACCCCGAGCGGAGAACCCTCCAGGTGATC  
13 E K S Q I K R A L E L G T V M T V F S F R K S T P E R R T V Q V I

**EcoRV (757)**  
701 ATGGAGACGGCGCAGGTGGCCTGGAGCAAGACCCGCGACAAGATCGAGGGCTTCTGGATATCATGGAATAAAAAGAAATCCGCCAGGGAAGAACTCCA  
47 M E T R Q V A W S K T A D K I E G F L D I M E I K E I R P G K N S  
801 AAGATTTGAGCGAGCAAAGCAGTTCCGCCAGAAAGACTGCTGCTTACCATCCTATATGGCACTCAGTTCGCTCCTCAGCACGCTCAGCTTGGCAGC  
80 K D F E R A K A V R Q K E D C C F T I L Y G T Q F V L S T L S L A A

**HpaI (919)**  
901 TGA CTCTAAAGAGGATGTCAGTTAACTGGCTCTCTGGCTTAAAACTTACACCAGGAAGCGATGAATGCTCCACGCCACCATTATCGAGAGTTGGCTG  
113 D S K E D A V N W L S G L K I L H Q E A M N A S T P T I I E S W L  
1001 AGAAAGCAGATATATTCTGTGGATCAAACCAGAAGAACAGCATCAGTCTCCGAGAGTTGAAGACCATCTTGGCCCTGATCAACTTTAAAGTGAAGCAGTG  
147 R K Q I Y S V D Q T R R N S I S L R E L K T I L P L I N F K V S S  
1101 CCAAGTTCCTAAAGATAAGTTTGTGGAATAGGAGCACAAAGATGAGCTCAGCTTTGAACAGTTCATCTCTTATAAAAACTTATGTTTGAACA  
180 A K F L K D K F V E I G A H K D E L S F E Q F H L F Y K K L M F E Q

**ClaI (1206)** **EcoRI (1219)**  
1201 GCAAAAATCGATTCTCGATGAATTCAAAAAGGATTCGTCGGTTCATCTGGGGAACACTGACAGGCCGGATGCCTCTGCTGTTTACCTGCGTGACTTC  
213 Q K S I L D E F K K D S S V F I L G N T D R P D A S A V Y L R D F  
1301 CAGAGGTTTCTACATACGAACAGCAGGAGCATTGGGCTCAGGATCTGAACAAAGTCCGTGAGCGGATGACAAAGTTCATTGATGACACCATCGGTGAAA  
247 Q R F L I H E Q Q E H W A Q D L N K V R E R M T K F I D D T M R E  
1401 CTGCTGAGCCTTTCTGTTTGTGGATGAGTTTCTCACGTACCTGTTTTACGAGAAAAACAGCATCTGGGATGAGAAGTATGACGCGGTGGACATGCAGGA  
280 T A E P F L F V D E F L T Y L F S R E N S I W D E K Y D A V D M Q D

**BstEII (1558)** **HindIII (1589)**  
1501 CATGAACAACCCCTGTCTCATTACTGGATCTCCTCGTACATAACACGTACCTTACAGGTGACCAGCTGCGGAGCGAGTCGTCGCCAGAAGCTTACATC  
313 M N N P L S H Y W I S S S H N T Y L T G D Q L R S E S S P E A Y I

**FspI (1607)** **Bsp120I (1648)** **NcoI (1674)**  
1601 CGCTGCCTGCGCATGGGCTGTGCTGCTGATTGAACTGGACTGCTGGGACGGGCCGATGGGAAGCCGGTCACTACCATGGCTGGACGCGGACTACCAAGA  
347 R C L R M G C R C I E L D C W D G P D G K P V I Y H G W T R T T K

**Tth111I (1708)**  
1701 TCAAGTTTGACGACGTCGTGACGGCCATCAAAGACCACGCCTTTGTTACGTCGAGCTTCCAGTGATCCTGTCCATCGAGGAGCACTGCAGCGTGGAGCA  
380 I K F D D V V Q A I K D H A F V T S S F P V I L S I E E H C S V E Q

**StuI (1817)**  
1801 ACAGCGTCACATGGCCAAGGCTTCAAGGAAGTATTTGGCGACCTGCTGTTGACGAAGCCACGGAGGCCAGTGTGACCAGCTGCCCTCGCCAGCCAG  
413 Q R H M A K A F K E V F G D L L L T K P T E A S A D Q L P S P S Q

**BsaBI (1911)** **Bsp120I (1935)**  
1901 CTGCGGGAGAAGATCATCATCAAGCATAAGAAGCTGGGCCCGGAGCGATGTGGATGTCAACATGGAGGACAAGAAGGACGAACACAAGCAACAGGGGG  
447 L R E K I I I K H K K L G P R G D V D V N M E D K K D E H K Q Q G

**BsrGI (2003)**  
2001 AGCTGTACATGTGGGATTCCATTGACCAGAAATGGACTCGGCACTACTGCGCCATTGCCGATGCCAAGCTGTCCTTCACTGATGACATTGAACAGACTAT  
480 E L Y M W D S I D Q K W T R H Y C A I A D A K L S F S D D I E Q T M

**XmnI (2151)**  
2101 GGAGGAGGAAGTGCCCGAGATATACCCCTACAGAACTACATTTTGGGGAGAAATGGTTCCACAAGAAGGTGGAGAAGAGGACGAGTCCCGAGAAGTTG  
513 E E E V P Q D I P P T E L H F G E K W F H K K V E K R T S A E K L  
2201 CTGAGGAATACTGCATGGAGACGGGGGCAAGGATGGCACCTTCTGTTTGGGAGAGCGAGACCTTCCCAATGACTACACCTGTCTTCTGGCGGT  
547 L Q E Y C M E T G G K D G T F L V R E S E T F P N D Y T L S F W R

**NcoI (2333)** **SandI (2344)**  
2301 CAGGCCGGTCCAGCACTGCCGGATCCGCTCCACCATGGAGGGCGGGACCCTGAAATACTACTTACTGACTGACAACCTCACCTTACGAGCATCTATGCCCT  
580 S G R V Q H C R I R S T M E G G T L K Y Y L T D N L T F S S I Y A L  
2401 CATCCAGCACTACCGGAGACGCACCTGCGCTGCGCCGAGTTCGAGCTGCGGCTCACGGACCCTGTGCCAACCCCAACCCCAAGTCCAAGCCGTGG  
613 I Q H Y R E T H L R C A E F E L R L T D P V P N P N P H E S K P W

**SacII (2517)** **XmaI (2552)** **BspEI (2573)**  
2501 TACTATGACAGCCTGAGCCGCGGAGAGGAGGACATGCTGATGAGGATTCCTCCGGGACGGGCTTCTGATCCGGAAGCGAGAGGGGAGCGACTCCT  
647 Y Y D S L S R G E A E D M L M R I P R D G A F L I R K R E G S D S

**DraIII (2662)**  
**NgoMIV (2657)**  
**EagI (2655)**  
2601 ATGCCATCACCTTACGGGCTAGGGGCAAGGTAAGCATTGTGCGATCAACCGGGACGGCGGCACCTTTGTGCTGGGACCTCCGCTATTTTGAAGTCT  
680 Y A I T F R A R G K V K H C R I N R D G R H F V L G T S A Y F E S L

Eco47III (2783)

2701 GGTGGAGCTCGTCACTACTACGAGAAGCATTCACTCTACCGAAAGATGAGACTGCGCTACCCCGTGACCCCGAGCTCCTGGAGCGCTACAATATGGAA  
713▶ V E L V S Y Y E K H S L Y R K M R L R Y P V T P E L L E R Y N M E  
2801 AGAGATATAAACTCCTCTACGACGTCAGCAGAATGTATGTGGATCCCAAGTAAATCAATCCGTCATGCCTCAGAGAACCGTAAAAGCTCTGTACT  
747▶ R D I N S L Y D V S R M Y V D P S E I N P S M P Q R T V K A L Y D

XmaI (2968)

2901 ACAAAGCCAAGCGAAGCGATGAGCTGAGCTTCTGCCGTGGTGCCTCATCCACAATGTCTCCAAGGAGCCCGGGGCTGGTGGAAAGGAGACTATGGAAC  
780▶ Y K A K R S D E L S F C R G A L I H N V S K E P G G W W K G D Y G T

ScaI (3011)

3001 CAGGATCCAGCAGTACTTCCCATCCAACCTACGTCGAGGACATCTCAACTGCAGACTTCGAGGAGCTAGAAAAGCAGATTATTGAAGACAATCCCTTAGGG  
813▶ R I Q Q Y F P S N Y V E D I S T A D F E E L E K Q I I E D N P L G

SspI (3114)

3101 TCTCTTTGAGAGGAATATTGGACCTCAATACCTATAACGTCGTGAAAGCCCTCAGGGAAAAAACAGAAAGTCTTTGTCTTTCATCCTGGAGCCCAAGC  
847▶ S L C R G I L D L N T Y N V V K A P Q G K N Q K S F V F I L E P K

XcmI (3227)

3201 AGCAGGGCTATCCTCCGGTGGAGTTTGCCACAGACAGGGTGGAGGAGCTCTTTGAGTGGTTTCAGAGCATCCGAGAGATCACCTGGAAGATTGACACCAA  
880▶ Q Q G Y P P V E F A T D R V E E L F E W F Q S I R E I T W K I D T K

ScaI (3314)

3301 GGAGAACAACATGAAGTACTGGGAGAAGAACCAGTCCATCGCCATCGAGCTCTCTGACCTGGTTGTCTACTGCAAACCAACCAGAAAACCAAGGACAAC  
913▶ E N N M K Y W E K N Q S I A I E L S D L V V Y C K P T S K T K D N

SalI (3475)

3401 TTAGAAAATCCTGACTTCCGAGAATCCGCTCCTTTGTGGAGACGAAGGCTGACAGCATCATCAGACAGAAGCCCGTCCGACTCTGAAGTACAATCAAA  
947▶ L E N P D F R E I R S F V E T K A D S I I R Q K P V D L L K Y N Q

PshAI (3507)

3501 AGGGCCTGACCCGCTACCCAAAGGGACAAAGAGTTGACTCTTCAAACACGACCCCTCCGCTCTGGCTGTGCGGTTCTCAGATGGTGGCACTCAA  
980▶ K G L T R V Y P K G Q R V D S S N Y D P F R L W L C G S Q M V A L N

3601 TTTCCAGACGGCAGATAAGTACATGCAGATGAATCAGCATTGTTTTCTCAACGGGCGCACGGGCTACGTTCTGCAGCCTGAGAGCATGAGGACAGAG  
1013▶ F Q T A D K Y M Q M N H A L F S L N G R T G Y V L Q P E S M R T E

3701 AAATATGACCCGATGCCACCCGAGTCCCAGAGGAAGATCTGATGACGCTGACAGTCAAGGTTCTCGGTGCTGCCATCTCCCAAACCTGGACGAAGTA  
1047▶ K Y D P M P P E S Q R K I L M T L T V K V L G A R H L P K L G R S

BglII (3824)

3801 TTGCCTGTCCCTTTGTAGAAGTGGAGATCTGTGGAGCCGAGTATGACAACAACAAGTTCAAGACGACGGTTGTGAATGATAATGGCCTCAGCCCTATCTG  
1080▶ I A C P F V E V E I C G A E Y D N N K F K T T V V N D N G L S P I W

3901 GGCTCAAACACAGGAGAAGGTGACATTTGAAATTTATGACCCAAACCTGGCATTCTCGCCTTTGTGGTTTATGAAGAAGATATGTTACAGCATCCCAAC  
1113▶ A P T Q E K V T F E I Y D P N L A F L R F V V Y E E D M F S D P N

4001 TTTCTTGCTCAGTCCACTTACCCATTAAGCAGTCAAATCAGGATTCAGTCCGTTCTGAAGAATGGGTACAGGAGGACATGAGCTGGCTTCCC  
1147▶ F L A H A T Y P I K A V K S G F R S V P L K N G Y S E D I E L A S

4101 TCCTGGTTTTCTGTGAGATGCGGCCAGTCTGGAGAGCGAAGAGGAACTTACTCCTCTGTGCGCCAGCTGAGGAGCGGCAAGAAGAACTGAACAACCA  
1180▶ L L V F C E M R P V L E S E E E L Y S S C R Q L R R R Q E E L N N Q

FspI (4229)

4201 GCTCTTTCTGTATGACACACACCAGAATTTGCGCAATGCCAACCCGGATGCCCTGGTTAAAGAGTTTCAGTGTAAATGAGAACCAGCTCCAGCTGTACCAG  
1213▶ L F L Y D T H Q N L R N A N R D A L V K E F S V N E N Q L Q L Y Q

MscI (4391)

4301 GAGAAATGCAACAAGAGGTTAAGAGAGAAGAGAGTCAAGCAACAGCAAGTTTTACTCATAGAAGCTGGGGTATGTGTGAAGGGTAGCTAGCTGGCCAGAC  
1247▶ E K C N K R L R E K R V S N S K F Y S •

NheI (4385)

4401 ATGATAAGATACATTGATGAGTTTGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAA

HpaI (4523)

MfeI (4534)

4501 CCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTGCATTTATGTTTCAGGTTACAGGGGAGGTGTGGAGTTTTTTAAAGCAAGTAAAA

EcoRI (4619)

4601 CCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGG

4701 CATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTATTTTCCCAAGTTTGAAC

SspI (4858)

SwaI (4872)

4801 AGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATA

4901 AATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAT

5001 TGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCTCAAT

141

• N R T Y K L P I L E E I T T K V L K G N M E I

BstXI (5162)

5101 GAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGG  
117▶ L V F C D 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

StuI (5297)

5201 TGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGG  
83▶ H R V A V 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

5301 CCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTGGTCTGATGGCCGCCCGACATGGTCTTGTGCTCCTCATAGAGCATGGTATCTTCTCAGT  
50▶ E I H V 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

BspHI (5447)

XmnI (5439)

5401 GGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGC  
17▶ A V E V L E L D Q Q S I N F T K M

5501 **AseI (5505)**  
CGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTA

5601 **SpeI (5660)**  
CCGCCATTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGG

5700 **SnaBI (5788)**  
AGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGT

5800 **NdeI (5893)**  
ACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATG

5900  
ATACACTTGTACTGCCAAGTGGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTC

6000 **SdaI (6071)****PacI (6079)**  
ATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCCCTGCAGGTTAA TTAAGAACATGTGA

6098  
GCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGGTGGCGTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCT

6198  
CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTAC

6298  
CGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTG

6398 **ApaLI (6403)**  
GGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG

6498  
CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGT

6598  
ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTT

6698  
GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTT

6798 **EagI (6839)**  
**PacI (6819)** **SwaI (6828)** **NotI (6838)**  
AAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAA

6898  
TCGTAACATAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCT

6998  
ATCGAA