



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

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGGGGTAACGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245) **EcoNI (287)**

201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCGCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT

NcoI (560)
BstEII (555)
KasI (535) **AgeI (552)**

501 TCTGTTCTGCGCGCTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCACCATGAAAGAAATGTTCTTACAACATTTTACAGGAAAT
601 GTCCCAGTTAATTTGAATGAAATGCCAAAAGCTGAATATCCAGTTTATCAATGATTTTGTGAATCTGAATTTTTTTGATTGATGGGGATTCATTA
13▶ S Q L I L N E M P K A E Y S S L F N D F V E S E F F L I D G D S L
701 CTTATCAGATGTATCTGTGAGATATCATTAAAGCCTGGCAGAACCTCCATTTCTTCTATCTGTTGAACGCTATCTTGGATCTTATTAGCAAAGGAG
47▶ L I T C I C E I S F K P G Q N L H F F Y L V E R Y L V D L I S K G
801 GACAATTCACCATAGTTTTCTTCAAGGATGCCAGTATGCGTATTTCAACTCCCTGAAGTCTTTCTTTGAGAACTGCTTAAATCTTCATCTTCAGAA
80▶ G Q F T I V F F K D A E Y A Y F N F P E L L S L R T A L I L H L Q K

BsaBI (935)

901 GAATACCACCATTGATGTTGCAACAACATTTTTCGAGATGCTTATCAAAAGAGTGGGGAAGTTTCTTGAAGAGAGTTACCCATATTTCTGATAGTTGCA
113▶ N T T I D V R T T F S R C L S K E W G S F L E E S Y P Y F L I V A
1001 GACGAAGGCTGAACGATCTACAAACACAGCTTTTCAACTTTTAAATCATTCTTCTGGGCAAGGAAGTCAACGTTGTACTTTCCTCAGGGCAAGAAT
147▶ D E G L N D L Q T Q L F N F L I H S W A R K V N V V L S S G Q E

NsiI (1118) **BsrGI (1139)**

1101 CTGATGTTCTTTGCTTTTATGCATACCTTCTTCCAAGCATGTACAGACACCAGATTTTTTCTGGAAGAATAAGCAGAACATTAAGATGCTTATACAAC
180▶ S D V L C L Y A Y L L P S M Y R H Q I F S W K N K Q N I K D A Y T T

HindIII (1225)

1201 CCTGTTAACAGTGGAAAGATTAAAGCTTTCAGCATTAGCACCTCTTTTTGGAAGTTTAAATGGAATAATATTACGGAAGAGGCACACAAGACTGTA
213▶ L L N Q L E R F F K L S A L A P L F G S L K W N N I T E E A H K T V

MseI (1318)

1301 TCTCTGTTACACAAGTCTGGCCAGAAGGATCTGACATTCGGCGTGTCTTTTGTGTTACTTTCATGCTCATTATCTTTGAGAATGTACCATCGCTTTTTAG
247▶ S L L T Q V W P E G S D I R R V F C V T S C S L S L R M Y H R F L

PstI (1465)
SdaI (1464)

1401 GAAACAGAGAGCCCTCCTCTGGTCAGGAAACTGAGATCCAACAGGTGAACAGTAATTGCTTAAACCCTGAGGAGATGGAAGATTTGTGTTAAACTGCATTG
280▶ G N R E P S S G Q E T E I Q Q V N S N C L T L Q E M E D L C K L H C
1501 TCTCACTGTGTTTTTCTACTCCATCGCTCTTTCTCAAAGACTTGTGCTAGAGTCATCACTTCACTTGGGCTGAGGACATGAAGCCTTTATTACAA
313▶ L T V V F L L H L P L S Q R A C A R V I T S H W A E D M K P L L Q
1601 ATGAAAAAGTGGTGTGAATATTTTCATCTTAAAGAAATATACATACTTTTGAATTTTGAATCTGAATTTAATTCACCTTCTGACTTAAATGATGAGCTTT
347▶ M K K W C E Y F I L R N I H T F E F W N L N L I H L S D L N D E L
1701 TGTTGAAGAATATTGCTTTTACTATGAAAATGAAAATGTAAGGACCTACATTTTGAATTTGGGAGATACCATTATGAAAGATTATGAATATCTCTGGAA
380▶ L L K N I A F Y Y E N E N V K G L H L N L G D T I M K D Y E Y L W N
1801 TACCGTACAAAGTTGGFCAGAGACTTTGAGTTGGAGCCATTTCTCTGAGAACAACAAAAGTTTGTCTTGGAAAGAACCATCAACATCAAA
413▶ T V S K L V R D F E V G Q P F P L R T T K V C F L G K K P S P I K
1901 GACAGCTCCAATGAAATGGTGCCTAATTTGGGTTTTATCCAACGTCATTTTTGTGTTGATAAATTTGCTGGAGATATTTTGAAGATTTGCTTTTTC
447▶ D S S N E M V P N L G F I P T S S F V V D K F A G D I L K D L P F

ApaLI (2059)

2001 TAAAGAGTGATGATCCTATTGTTACTTCACTGGTTAAACAAAAGGAATTTGATGAACCTTGTGCACTGGCATTCTCATAAACCCCTGAGTGATGATTATGA
480▶ L K S D D P I V T S L V K Q K E F D E L V H W H S H K P L S D D Y D

BglII (2149) **XcmI (2175)** **EcoRI (2190)**

2101 CAGGTCCAGGTGTGAGTTTGTGATGAAAAATCTAGAGACCCTCGTGTCTTAGATCTGTGCAAAAAGTATCATGTTTTCCAACGGTTTTATGGGAATTCATTA
513▶ R S R C Q F D E K S R D P R V L R S V Q K Y H V F Q R F Y G N S L

Bsp120I (2260)

2201 GAAACAGTCTCTTCGAAAATCATCGTGAAGTCAAACTATTAAGTCAAAGAAGGATTTTGTGGCCCAAGAGCAAAAAGGCACACGAGACCAAGGCTGAAA
547▶ E T V S S K I I V T Q T I K S K K D F S G P K S K K A H E T K A E
2301 TAATTGCTAGAGAGAATAAGAAAAGTTATTTGCCAGGGAAGAACAAAAGGAAGCAAAAAGTGAATGCTTTGTGCTTTTCTATTGAAGAGCAATTTGAA
580▶ I I A R E N K K R L F A R E E Q K E E Q K W N A L S F S I E E Q L K

HpaI (2490)

2401 AGAAAAATTTACACTCTGGAATAAAGAGCCCTGGAAGATTTTTTGAATCCTGTAAGTGTGTAAGTGTGAAACTCAGGTTGAAATGGTGGGGTTAACTGCT
613▶ E N L H S G I K S L E D F L K S C K S S C V K L Q V E M V G L T A

BamHI (2585)

2501 TGCTTGAAGCCTGGAAAGAACATTGCCGAAGTGAAGAAGGTAAGTAAACCACGAAAGATTTAAGTATAGCTGTTCAAGGTGATGAAAAGGATCCACTCCTTGA
647▶ C L K A W K E H C R S E E G K T T K D L S I A V Q V M K R I H S L
2601 TGGAAAAATACTCAGAAGCTTTTACAAGAAGATGATCGGCAACTCATAGCCAGATGCCTTAAAGTATTTAGGATTTGATGAGTTGGCAAGTTCTTTACATCC
680▶ M E K Y S E L L Q E D D R Q L I A R C L K Y L G F D E L A S S L H P
2701 AGCCAGGATGCAGAAAATGATGTAAGTGAAGAAGGAAATAAATTTCAATTGGCATTGGGCCAGCTCGGTTCCAAGTGAATACATGSGCCATTAT
713▶ A Q D A E N D V K V K K R N K Y S I G I G P A R F Q L Q Y M G H Y

2801 TTGATACGAGATGAGAGAAAAGACCCAGATCCCAGGGTCCAGGATTTTATCCCGACACATGGCAGCGAGAGCTCCTTGATGTTGGATAAGAATGAGT
747▶ L I R D E R K D P D P R V Q D F I P D T W Q R E L L D V V D K N E
2901 CAGCAGTGATTGTTGCCCACTGCTCAGGCAAAACAATATGCCTCCTACTACTGTATGGAGAAAGTGCCTGAAGGAGAGCGACGACGGGGTGGTCGTGTA
780▶ S A V I V A P T S S G K T Y A S Y Y C M E K V L K E S D D G V V V Y
EcoO109I (3014)
3001 CGTTGCACCCACAAGGCCCTTGTAAATCAAGTGGCAGCAACTGTTTCAAGTGGTGAAGTCTCTGTGGTGTTC
813▶ V A P T K A L V N Q V A A T V Q N R F T K N L P S G E V L C G V F
3101 ACCAGGAGTACGTATGATGCCTTAACTGTCAAGTACTTATTACAGTGCCTGCCTGCTTTGAAATTTCTGTGCTTGTCTCCTCAGCCAAAAGTGGG
847▶ T R E Y R H D A L N C Q V L I T V P A C F E I L L L A P H R Q N W
3201 TGA AAAAGATCAGATATGTTATATTTGATGAGGTTTCTGTCTTGGTGGAGAAATGGAGCAGAAATCTGGAAACATCTCCTTGTGATGATCCGATGTC
880▶ V K K I R Y V I F D E V H C L G G E I G A E I W E H L L V M I R C P
DraIII (3342)
3301 CTTTTGGCTCTTTCAGCTACCATAAGTAATCCTGAACATCTCACCGATGGCTACAATCGGTAAAATGGTACTGGAACAAGAAGACAAAATAATTGAA
913▶ F L A L S A T I S N P E H L T E W L Q S V K W Y W K Q E D K I I E
NgoMIV (3436)
3401 AATAATACCGCTTCTAAAAGACATGTGGGTCGTGAGGCCGGCTTTCCCAAAGACTACTTGCAAGTAAAACAATCGTATAAAGTTAGACTTGTGCTCTATG
947▶ N N T A S K R H V G R Q A G F P K D Y L Q V K Q S Y K V R L V L Y
3501 GAGAGAGGTATAATGATCAGAGAAGCATGTATGTTCAATAAAACATGGTGACATTCATTTTGATCATTTCACCCATGTGCTGCACTAACACAGATCA
980▶ G E R Y N D L E K H V C S I K H G D I H F D H F H P C A A L T T D H
XhoI (3641) Bsp120I (3698)
3601 TATTGAAAGGTATGGATCCCTCCTGATCTTACCCTTTACCTCGAGAAAGCATCCAGCTGTATGATGCCATGTTTCAAATTTGGAAAAGTTGGCCTCGG
1013▶ I E R Y G F P P D L L T C A S P R E S I Q L Y D A M F Q I W K S W P R
3701 GCCCAGAACTGTGCCAGAAAACCTCATTCTTTAAACAAATTAAGTCTATAAAAAGATGGATGAGGAAATGAGAGAGTCTAAAAGCAGAAAT
1047▶ A Q E L C P E N F I H F N N K L V I K K M D A R K Y E E S L K A E
3801 TAACAAGTTGGATTA AAAATGGCAACGTAGAGCAGGCCAGAATGGTACTTCAAGTCTTGAAGCAGATTTGAGTCCAGAAAACATGATCACCAT
1080▶ L T S W I K N G N V E Q A R M V L Q N L S P E A D L S P E N M I T M
3901 GTTCCACTTCTAGTTGAAAAACTAAGGAAAATGGAGAAGTTACCTGCACTATTTTTTTTATTCAAGTTAGGAGCTGTAGAAAACGCAGCTGAAAGTGTG
1113▶ F P L L V E K L R K M E K L P A L F F L F K L G A V E N A A E S V
4001 AGCACTTTCTAAAGAAAAAGCAGGAGACAAAAAGCCCTCCCAAAGCTGATAAAGAAGCCCATGTGATGGCTAACAAAACCTCGAAAAGTTAAAAAATCCA
1147▶ S T F L K K K Q E T K R P P K A D K E A H V M A N K L R K V K K S
4101 TAGAGAAAACAAAGATCATAGATGAAAAGAGCCAGAAAAAACAGAAATGTGGATCAAAGCCTAATACATGAAGCTGAACATGATAATCTAGTGAAGTG
1180▶ I E K Q K I I D E K S Q K K T R N V D Q S L I H E A E H D N L V K C
4201 TCTAGAGAAGAACCTGAAAATCCACAGGACTGCACATATGCTGATCAAAAAGCAGTGGACACTGAGACTTTG CAGAAGGTATTTGGTCGAGTAAATTT
1213▶ L E K N L E I P Q D C T Y A D Q K A V D T E T L Q K V F G R V K F
4301 GAAAGAAAAGGTGAAGAATTGAAAGCCTTGGCAGAAAAGGGTATTGGATATCATCACAGTGCATGAGTTTCAAAGAAAACAATTAGTTGAAATCCTCT
1247▶ E R K G E E L K A L A E R G I G Y H H S A M S F K E K Q L V E I L
4401 TTAGAAAAGGATATCTTAGGVTGGTGACAGCTACTGGAACACTTGCTTAAAGTGTCAACATGCCTTGTAAATCTGTGGTTTTTGTCAAACACTCAGTCTA
1280▶ F R K G Y L R V V T A T G T L A L G V N M P C K S V V F A Q N S V Y
PshAI (4520)
4501 TCTGGATGCGTTGAATTATAGACAGATGTCTGGCCGTGCTGGAAGAAGAGGTCAAGACCTGATGGGAGATGTATATTTCTTTGATATTCCATCCCAA
1313▶ L D A L N Y R Q M S G R A G R R G Q D L M G D V Y F F D I P F P K
4601 ATAGAAAACCTATAAAATCCAATGTTCTGAGCTGAGAGGACACTTCCCTCTCAGCATAACCTGGTCTGCGACTCATGCTGCTGGCTTCCAAGGGAG
1347▶ I G K L I K S N V P E L R G H F P L S I T L V L R L M L L A S K G
BstAPI (4743)
4701 ATGACCCAGAGGATCAAGGCAAAAGGTGCTATCAGTGCTAAAGCATTATTGCTGTCCTTCAAGCAACCCAGAGTCATGGACATGTTAAAACCTTACT
1380▶ D D P T E D T T K A K V L S V L L K H S L L T S F G K Q P R V M D M L K L L Y F
4801 CCTGTTTTCTTGCAGTCTCTGTTGAAAGAGGGCTATTTAGATCAAGAAGGTAATCCTATGFGGTTTGTGCTGGACTGTATCACATTTGCATATCATGAA
1413▶ L F S L Q F L V K E G Y L D Q E G N P M G F A G L V S H L H Y H E
4901 CCTCTAATCTTGTGTTTGTGAGTTTCTGTAATGGACTCTCCATGATCTCTGTGAGCAACCCAGGAAAAGGCTCAAACATTTTTCTCAAGACGTTA
1447▶ P S N L V F V S F L V N G L F H D L C Q P T R K G S K H F S Q D V
5001 TGGAAAAGCTAGTATTAGTATTGGCACATCTCTTTGGAAGAAGATATTTCCACCAAAGTCCAAGATGCACACTTCGAGTTTTATCAATCAAAGGTGTT
1480▶ M E K L V L V L A H L F G R R Y F P P K F Q D A H F E F Y Q S K V F
XmnI (5183)
5101 CTTGATGATCTCCTGAGGATTTTAGTGATGCTTTAGATGAATATAACATGAAAATTATGGAGGACTTTACCCTTTCTACGAATTGTTTCCAACCTG
1513▶ L D D L P E D F S D A L D E Y N M K I M E D F T T F L R I V S K L
BsaBI (5203)
5201 GCTGATATGAATCAGGAATATCAACTCCATTGTCAAAAATCAAATTCACAGGTAAAGAATGTGAAGACTCTCAACTCGTATCTCATTGATGAGCTGCA
1547▶ A D M N Q E Y Q L P L S K I K F T G K E C E D S Q L V S H L M S C
5301 AGGAAGGAAGAGTAGCAATTTACCATTGTTTGTCTGTCTGGAACTTTGATGATGATTTGCTTCGACTAGAACTCCAACCATGTTACTCTAGGCAC
1580▶ K E G R V A I S P F V C L S G N F D D D L L R L E T P N H V T L G T
5401 AATCGGTGCAATCGCTCAGGCTCAGTGTCTGTGACAGAAATTTGATAACCGAGGAAGAAAATGTGCTTAAATGCCTATGCACTGGATTTCTAC
1613▶ I G V N R S Q A P V L L S Q K F D N R G R K M S L N A Y A L D F Y
5501 AAACATGGTTCCTTGATAGGATAGTCCAGGATAACAGGATGAATGAAGGAGATGCTTATTATTTGTTGAAGGATTTGCACTCACCATTAAATCTATCA
1647▶ K H G S L I G L V Q D N R M N E G D A Y Y L L K D F A L T I K S I
5601 GTGTTTCTTGCCTGAGCTATGTGAAAATGAAGACGACAACGTTGCTTAAAGCCTTTGAACAACTGAGTACAACCTTTTGGGAAAAGTTAAACAAAAGTCTA
1680▶ S V S L R E L C E N E D D N V V L A F E Q L S T T F W E K L N K V •
MscI (5730)
NheI (5724)
5701 AAAACAAAGTCTATGCAAAACCTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTAAAAAATGCT
1713▶
HpaI (5862)
5801 TTATTTGAAAATTTGTGATGCTATTGCTTTATTTGTAACATTATAAGCTGCAATAAACAAGTTAACAACAACAAATGCATTCAATTTATGTTTCAGGT

5901 TCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTC EcoRI (5958)
6001 AAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCAT
6101 GGAGTTTAAAGATATAGTGTATTTTCCCAAGTTTGAAC TAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAT

6201 ATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAG SwaI (6211) EcoO109I (6272)
6301 TAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTC
6401 CTCAATGGTGGTTTTGACCAGTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGTCTCTGCACATGCCACAGGGGCTG 141 • N R T Y K L P I L E
130 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 R C M G C P S
6501 ACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGG
96 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 N S V A S G I A
6601 CAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATG
63 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 T R I A A G V H
XmnI (6778)
6701 GTGCTTGTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTC
30 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 F T K M
AseI (6844)
6801 CTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAA
SpeI (6999)
6901 CGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTGCGTCAATGGGGCGGAGTTGTACGACATTTTGAAAAGTCCCGTTGATTTA
7001 CTAGTCAAAACAACTCCATTGACGTCAATGGGTGGAGACTTGAAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCG

7101 CATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTA CTGGGCATAATGCCAGGCGGGCCATTTA SnaBI (7127)
7201 CCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTA CTGCCAAGTGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAAT
7301 GGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGGCGGGTCTGTTGGCGGT CAGCCAGGCGGGCCATTTACCCTAAG

7401 TTATGTAACGCCTGCAGGTTAATTAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGC PstI (7411) SdaI (7410) PacI (7418)
7501 TCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAAGCTC
7601 CCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGG

7701 TATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTG ApaLI (7742)
7801 AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAA
7901 GTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCC
8001 GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA

8101 CGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTA PacI (8158) SwaI (8167) EagI (8178) NotI (8177)
8201 TTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACATAACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTG
8301 TCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA