



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) **PvuII (239)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACACGCCGCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGTTGCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) **AgeI (552)** **BspHI (560)**
501 TCTGTTCTGGCGGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCATCATGAACCACCAGCAGCAGCAGCAGCAGCAGAAAGCGGG
601 CGAGCAGCAGTTGAGCGAGCCCGAGGACATGGAGATGGAAGCGGGAGATACAGATGACCCACCAAGAATTACTCAGAACCCTGTGATCAATGGGAATGTG
13> E Q Q L S E P E D M E M E A G D T D D P P R I T Q N P V I N G N V
1> M N H Q Q Q Q Q Q Q K A G

SacII (722) **Eco47III (788)**
701 GCCCTGAGTGATGGACACAACACCGCGGAGGAGGACATGGAGGATGACACCAGTTGGCGCTCCGAGGCAACCTTTTCAGTTCAGTGTGGAGCGCTTCAGCA
47> A L S D G H N T A E E D M E D D T S W R S E A T F Q F T V E R F S
801 GACTGAGTGAGTCGGTCTTAGCCCTCCGTGTTTTGTGCGAAATCTGCCATGGAAGATTATGGTGATGCCACGCTTTTATCCAGACAGACCACCAAAAA
80> R L S E S V L S P P C F V R N L P W K I M V M P R F Y P D R P H Q K
901 AAGCGTAGGATCTTTCTCCAGTGAATGCTGAATCTGATTCCACGTCATGGTCTTGCCATGCACAAGCAGTGTGAAGATAATAAATACAGAGATGAT
113> S V G F F L Q C N A E S D S T S W S C H A Q A V L K I I N Y R D D
1001 GAAAAGTCGTTACGTCGTCGATTTAGTCAATTTGTTCTCCATAAAGAAAATGATTGGGGATTTTCCAATTTTATGGCCTGGAGTGAAGTGACCGATCCTG
147> E K S F S R R I S H L F F H K E N D W G F S N F M A W S E V T D P

BsrGI (1143)
1101 AGAAAGGATTTATAGATGATGACAAAGTTACCTTTGAAGTCTTTGTACAGGCGGATGCTCCCATGGAGTTGCGTGGGATTCAAAGAAGCACACAGGCTA
180> E K G F I D D D K V T F E V F V Q A D A P H G V A W D S K K H T G Y

BsrGI (1286)
1201 CGTCGGCTTAAAGAATCAGGGAGCGACTTGTACATGAACAGCCTGCTACAGACGTTATTTTTACGAATCAGCTACGAAAGGCTGTGATCATGATGCCA
213> V G L K N Q G A T C Y M N S L L Q T L F F T N Q L R K A V Y M M P

HpaI (1395)
1301 ACCGAGGGGGATGATTCGTCTAAAAGCGTCCCTTTAGCATTACAAAGAGTGTTCTATGAATTACAGCATAGTGATAAACCTGTAGGAACAAAAAGTTAA
247> T E G D D S S K S V P L A L Q R V F Y E L Q H S D K P V G T K K L
1401 CAAAGTCATTTGGTGGAACCTTTAGATAGCTTCATGAACATGATGTTCCAGGAGCTTTGTCCGAGTGTTCGATGATAATGTGAAAAAATAGATGAAAAG
280> T K S F G W E T L D S F M Q H D V Q E L C R V L L D N V E N K M K G

SacII (1533)
1501 CACCTGTGTAGAGGGCACCATAACCAAATTTATCCGCGGCAAAATGGTGTCTATATCCAGTGTAAGAAGTAGACTATCGGTCTGATAGAAGAGAAGAT
313> T C V E G T I P K L F R G K M V S Y I Q C K E V D Y R S D R R E D

EcoRV (1606)
1601 TATTATGATATCCAGCTAAGTATCAAAGGAAAGAAAAATATTTGAATCATTGTGGATTATGTGGCAGTAGAACAGCTCGATGGGGACAATAAATACG
347> Y Y D I Q L S I K G K K N I F E S F V D Y V A V E Q L D G D N K Y
1701 ACGCTGGGGAACATGGCTTACAGGAAGCAGAGAAAAGGTGTGAAATTCCTAACATTGCCACCAGTGTACATCTACAACCTGATGAGATTTATGTATGACCC
380> D A G E H G L Q E A E K G V K F L T L P P V L H L Q L M R F M Y D P

Bsu36I (1885)
1801 TCAGACGGACCAAAATATCAAGATCAATGATAGTTTGAATCCAGAGCAGTTACCCTTGATGAATTTTTGCAAAAAACAGATCCTAAGGACCCTGCA
413> Q T D Q N I K I N D R F E F P E Q L P L D E F L Q K T D P K D P A
1901 AATTATATCTTCATGCAGTCTGGTTCATAGTGGAGATAATCATGGTGGCATTATGTGGTTTATCTAAACCCAAAGGGGATGGCAAATGGTGTAAAT
447> N Y I L H A V L V H S G D N H G G H Y V Y L N P K G D G K W C K
2001 TTGATGACGACGTGGTGTCAAGGTGACTAAAGAGGAAGCAATTGAGCACAATTATGGGGTCCAGATGACGACCTGTCTGTTGACACTGCACTAATGC
480> F D D D V V S R C T K E E A I E H N Y G G H D D D L S V R H C T N A

BstEII (2151)
2101 TTACATGTTAGTCTACATCAGGGAATCAAACCTGAGTGAAGTTTTACAGGCGGTCACCGACCATGATATTCTCAGCAGTTGGTGGAGCGATTACAAGAA
513> Y M L V Y I R E S K L S E V L Q A V T D H D I P Q Q L V E R L Q E

BsrBI (2230) **MscI (2289)**
2201 GAGAAAAGGATCGAGGCTCAGAAGCGGAAGGAGCGGCAGGAAGCCCATCTCTATATGCAAGTGAGATAGTCGAGAGGACCAGTTTTGTGGCCACCAAG
547> E K R I E A Q K R K E R Q E A H L Y M Q V Q I V A E D Q F C G H Q
2301 GGAATGACATGACGATGAAGAAAAGTGAATACACTGTGTTCAAAGTATTGAAGAATCCTCGCTTGTGAGTTGTTGAGGCTCTCAGACCAT
580> G N D M Y D E E K V K Y T V F K V L K N S S L A E F V Q S L S Q T M
2401 GGGATTTCCACAAGATCAAATTCGATTGTGGCCATGCAAGCAAGGAGTAATGGAACAAAACGACCAGCAATGTTAGATAATGAAGCCGACGGCAATAAA
613> G F P Q D Q I R L W P M Q A R S N G T K R P A M L D N E A D G N K

SspI (2538) **BsaBI (2599)**
2501 ACAATGATTGAGCTCAGTGATAATGAAAACCTTTGGACAATATTCCTGGAAACAGTTGATCCCGAGCTGGCTGCTAGTGGAGCGACCTTACCCAAGTTG
647> T M I E L S D N E N P W T I F L E T V D P E L A A S G A T L P K F

BspHI [m] (2605)
2601 ATAAAGATCATGATGTAATTTTTGAAGATGATGATGCCAAAACCGGAGCTTGAATTACTGTGGGCATATCTACACACCAATATCCTGTAAAT
680> D K D H D V M L F L K M Y D P K T R S L N Y C G H I Y T P I S C K I
2701 ACGTGACTTGTCCAGTTATGTGTGACAGAGCAGGATTTATTCAAGATACTAGCCTTATCCTCTATGAGGAAGTTAAACCGAATTTAACAGAGAGAATT
713> R D L L P V M C D R A G F I Q D T S L I L Y E E V K P N L T E R I

2801 CAGGACTATGACGTGTCTCTTGATAAAGCCCTTGATGAACTAATGGATGGTGACATCATAGTATTTTCAGAAGGATGACCTGAAAATGATAACAGTGAAT
747▶ Q D Y D V S L D K A L D E L M D G D I I V F Q K D D P E N D N S E

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

3001 ATCAAATAGAATGAATTATTTTCAGGTTGCAAAGACAGTTGCACAGAGGCTCAACACAGATCCAATGTTGCTGCGATTTTCAAGTCTCAAGGTTATAGG
813▶ S N R M N Y F Q V A K T V A Q R L N T D P M L L Q F F K S Q G Y R

3101 GATGGCCAGGTAATCCTCTTAGACATAATTATGAAGTACTTTAAGAGATCTTCTACAGTTCTTCAAGCCTAGACAACCTAAGAAAATTTACTATCAGC
847▶ D G P G N P L R H N Y E G T L R D L L Q F F K P R Q P K K L Y Y Q

3201 AGCTTAAGATGAAAATCACAGACTTTGAGAACAGCGAAGTTTTAAATGTATATGGTTAAACAGCCAATTTAGGGAAGAGGAAAATAACACTATATCCAGA
880▶ Q L K M K I T D F E N R R S F K C I W L N S Q F R E E E I T L Y P D

3301 CAAGCATGGTGTGTCGGGACCTGTTAGAAGAATGTA AAAAGGCCGTGGAGCTTGGGAGAAAGCATCAGGGAACCTTAGGCTGCTAGAAAATGTAAGC
913▶ K H G C V R D L L E E C K K A V E L G E K A S G K L R L L E I V S

3401 TACAAAATCATTGGTGTTCATCAAGAAGATGAAC TATTAGAATGTTTATCTCCTGCAACGAGCCGACGTTTCGAATAGAGGAAAATCCCTTTGGACCAGG
947▶ Y K I I G V H Q E D E L L E C L S P A T S R T F R I E E I P L D Q

3501 TGGACATAGACAAAGAGAATGAGATGCTTGT CACAGTGGCGCATTTCATAAAGAGGCTTCGGAACTCCGTTTGTGAGGATACACCA
980▶ V D I D K E N E M L V T V A H F H K E V F G T F G I P F L L R I H Q

3601 GGGCGAGCATTTTCGAGAAGTGAAGCGAATCCAGAGCCTGCTGGACATCCAGGAGAAGGAGTTT GAGAAGTTAAATTTGCAATTGTAATGATGGC
1013▶ G E H F R E V M K R I Q S L L D I Q E K E F E K F K F A I V M M G

3701 CGACACCAGTACATAAATGAAGACGAGTATGAAGTAAATTTGAAAGACTTTGAGCCACAGCCCGGTAATATGTCTCATCTCGCCCTTGCTAGGGCTCG
1047▶ R H Q Y I N E D E Y E V N L K D F E P Q P G N M S H P R P W L G L

3801 ACCACTTCAACAAAGCCCAAAGAGGAGTGCCTACACTTACCTTGAAAAGGCCATTA AAAATCCATAACTAGGGCCTCATGGCTAGCTGGCCAGACATGAT
1080▶ D H F N K A P K R S R Y T Y L E K A I K I H N •

3901 AAGATACATTGATGAGTTTGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATT

4001 ATAAGCTGCAATAAAACAAGTTAAACAACA CAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCT

4101 ACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAG

4201 GCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTG CAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTGCT

4301 TTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGT

4401 TTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTT CATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGAC

4501 AGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGT GACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCTCAATGAGCA
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

4601 CAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCT
115▶ 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 H R

4701 GACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCA
82▶ 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 E

4801 ATGTGGACAGCAGAGATGATCTCCCGTCTGGTCTGATGGCCGCCCGACATGGTCTTGTGTCCTCATAGAGCATGGTATCTTCTCAGTGGCGA
48▶ 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 A V

4901 CCTCCACAGCTCCAGATCTGCTGAGAGATGTTGAAGGCTTCTCATGGTGGCCCTCTATAGT GAGTCGATTATACTATGCCGATATACTATGCCGATG
15▶ E V L E L D Q Q S I N F T K M

5001 ATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTIATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCC

5101 CATTTGCGTCAATGGGGCGAGTTGTTACGACATTTTGAAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTT

5201 GGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATG TACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCC

5301 AAGTAGGAAAGTCCATAAGGTCATG TACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACAC

5401 TTGATG TACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATT

5501 GACGTCAATGGGGGGGGTCTGGTGGCGGTG CAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTGCAGTTAAATTAAGAACATGTGAGCAAAGG

5601 CCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAG
5701 AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACC
5801 TGCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGT
5901 GCACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
6001 ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTA
6101 TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAA
6201 GCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATT

ApaLI (5898)

EagI (6334)

PacI (6314) SwaI (6323) **NotI (6333)**

6301 TTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACT
6401 AACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA