



PvuI (7)
SgfI (6)
MfeI (82) **EcoNI (96)**

1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
 101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245) **Bsu36I (291)**

Psp1406I (203) **PvuII (239)** **EcoNI (287)**

201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCCGCCGCCCTACCTGAGGCC
 301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)

401 GGGCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGTCTTTGTTTCGTTT

BspHI (560)

KasI (535) **AgeI (552)**

501 TCTGTTCTGCGCGCTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCATCATGAGTCATGTGGCAGTGAAAAATGCGCTCGGGCTGGA
1▶ M S H V A V E N A L G L D

XcmI (600)

601 CCAGCAGTTTGTGGCCTAGACCTGAACCTTTAGATAATCAGAGTGGAGGAAGTACAGCCAGCAAAGGGCGCTATATTCCTCTCATTTAAGGAACCGA
 13▶ Q Q F A G L D L N S S D N Q S G G S T A S K G R Y I P P H L R N R
 701 GAAGCTACTAAAGTTTCTACGATAAAGACAGTTCAGGTTGGAGTTCTAGCAAAGATAAGGATGCGTATAGCAGTTTTGGATCTCGTAGTGATTAAGAG
 47▶ E A T K G F Y D K D S S G W S S S K D K D A Y S S F G S R S D S R
 801 GGAAGTCTAGCTTCTCAGTGATCGTGAAGTGGATCAAGGGGAAGTTTGTATGATCGTGGACGGAGTGATTACGATGGCATTGGCAGCCGTGGTGACAG
 80▶ G K S S F F S D R G S G S R G R F D D R G R S D Y D G I G S R G D R

DraIII (983)

901 AAGTGGCTTTGGCAAATTTGAACGTGGTGGAAACAGTCGCTGGTGTGACAAATCAGATGAAGATGATTGGTCAAACCACTCCCAACCAAGTGAACGCTTG
 113▶ S G F G K F E R G G N S R W C D K S D E D D W S K P L P P S E R L

AseI (1033)

1001 GAACAGAACTCTTTTCTGGAGGCAACACTGGGATTAATTTGAGAAATACGATGACATTCCAGTTGAGGCAACAGGCAACAACCTGCTCTCCACATATTG
 147▶ E Q E L F S G G N T G I N F E K Y D D I P V E A T G N N C P P H I

SphI (1185)

1101 AAAGTTTCAGTGATGTTGAGATGGGAGAAATTATCATGGAAACATTGAGCTTACTCGTTATACTCGCCCAACTCCAGTGCAAAAAGCATGCTATTCTAT
 180▶ E S F S D V E M G E I I M G N I E L T R Y T R P T P V Q K H A I P I

XcmI (1233) **PstI (1253)**

1201 TATCAAAGAGAAAAGAGACTTGATGGCTTGTGCCAAAACAGGGTCTGGAAAAACTGCAGCATTCTGTTGCCATCTTGAGTCAGATTTATTCAGATGGT
 213▶ I K E K R D L M A C A Q T G S G K T A A F L L P I L S Q I Y S D G

KasI (1343)

1301 CCAGGCGAGGCTTTGAGGGCCATGAAGGAAAATGGAAGGTATGGGCGCCGCAAAAATACCCAATCTCCTTGGTATTAGCACCAACGAGAGAGTTGGCAG
 247▶ P G E A L R A M K E N G R Y G R R K Q Y P I S L V L A P T R E L A

BglII (1403) **XmnI (1421)** **XbaI (1438)**

1401 TACAGATCTACGAGGAAGCCAGAAAATTTTCATACCGACTAGAGTTCGCTCTGCTGGTTTATGGTGGTGCCGATATTGGTCAGCAGATTCGAGACTT
 280▶ V Q I Y E E A R K F S Y R S R V R P C V V Y G G A D I G Q Q I R D L
 1501 GGAACGTGGATGCCATTTGTTAGTAGCCACTCCAGGACGTCTAGTGGATATGATGAAAGAGGAAAGATTGGATTAGACTTTTCAAATACTTGGTGTTA
 313▶ E R G C H L L V A T P G R L V D M M E R G K I G L D F C K Y L V L
 1601 GATGAAGCTGATCGGATGTTGGATATGGGGTTTGGCCTCAGATTCGTAAGAATAGTGAACAAGATACTATGCCTCAAAGGGTGTCCGCCCACTATGA
 347▶ D E A D R M L D M G F E P Q I R R I V E Q D T M P P K G V R H T M

Bsu36I (1717)

1701 TGTTTGTGCTACTTTTCTAAGGAAATACAGATGCTGGCTCGTGATTTCTTAGATGAATATATCTTCTTGGCTGTAGGAAGAGTTGGCTCTACCTCTGA
 380▶ M F S A T F P K E I Q M L A R D F L D E Y I F L A V G R V G S T S E
 1801 AAACATCACACAGAAAGTATGTTGGTGGAAAGATCAGACAAACGGTCATTTCTGCTTACCTCCTAAATGCAACAGGCAAGGATTCACTGACCTTAGTG
 413▶ N I T Q K V V W V E E S D K R S F L L D L L N A T G K D S L T L V

NcoI (1974)

1901 TTTGTGGAGACAAAAAGGGTGCAGATTCTCTGGAGGATTTCTTATACCATGAAGGATACGCATGTACCAGCATCCATGGAGACCGTTCTCAGAGGGATA
 447▶ F V E T K K G A D S L E D F L Y H E G Y A C T S I H G D R S Q R D

EcoO109I (2006) **BsrBI (2022)** **BspLU11I (2094)**

2001 GAGAAGAGGCCCTTACCAGTTCCGCTCAGGAAAAAGCCCAATTTTATGTTGCTACAGCAGTAGCAGCAAGAGGACTGGACATTTCAAATGTGAAACATGT
 480▶ R E E A L H Q F R S G K S P I L V A T A V A A R G L D I S N V K H V

BsrGI (2136) **BsiWI (2152)**

2101 TATCAATTTTACTGCTGCAAGTATGATTTGAAGAATATGTACATCGTATTGGTCTGACGGGACGTGTAGGAAACCTTGGCCTGGCAACCTCATTCTTTAAC
 513▶ I N F D L P S D I E E Y V H R I G R T G R V G N L G L A T S F F N

SspI (2212)

2201 GAGAGGAACATAAATATTACTAAGGATTTGTTGGATCTTCTTGTGGAAGTAAACAAGAAGTGGCGTCTTGGTTAGAAAAACATGGCTTATGAACACCACT
 547▶ E R N I N I T K D L L D L L V E A K Q E V P S W L E N M A Y E H H
 2301 ACAAGGGTAGCAGTCGTGGACGTTCTAAGAGCAGATTTAGTGGAGGGTTTGGTGCAGAGACTACCGACAAAGTAGCGGTGCCAGCAGTTCAGCTTCAG
 580▶ Y K G S S R G R S K S R F S G G F G A R D Y R Q S S G A S S S S F S

MscI (2436)

2401 CAGCAGCCGCGCAAGCAGCAGCCGAGTGGCGGAGTGGCCACGGTAGCAGCAGAGGATTTGGTGGAGGTGGCTATGGAGGCTTTTACAACAGTGATGGA
 613▶ S S R A S S S R S G G G G H G S S R G F G G G G Y G G F Y N S D G

MscI (2587)

BstEII (2563) **NheI (2581)**

2501 TATGGAGGAAATATAACTCCCAGGGGTTGACTGGTGGGTAAGTGAAGCTGCTTTGCAGTAGGTCACCCTGCCAAACAAGCTAGCTGGCCAGACATGA
 647▶ Y G G N Y N S Q G V D W W G N •

2601 TAAGATACATTGATGAGTTTGGACAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT

2701 TATAAGCTGCAATAACAAGTTAAACAACAACAAATTGCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTC

2801 TACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATA

2901 GGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGTTTGAAGTAGCT

3001 CTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATG

3101 TTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGA

3201 CAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGC
 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

3301 ACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCC
 115 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 H R

3401 TGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTC
 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

3501 AATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGTAGGCCGCCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCG
 49 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

3601 ACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGAT
 15 V E V L E L D Q Q S I N F T K M

3701 GATTAATTGTCAAACAGCGTGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGC

3801 CCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCATGGGGTGGAGAC

3900 TTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTG

4000 CCAAGTAGGAAAGTCCATAAGGTCACTGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCATAGGGGGCGTACTTGGCATATGATAC

4100 ACTTGATGTACTGCCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTA

4200 TTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTG C A G G T T A A T T A A G A A C A T G T G A C A A
 SdaI (4267) PaeI (4275) BspLU11I (4285)

4298 AAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG

4398 TCAGAGGTGGCGAAACCCGACAGACTATAAAGATACCAGCGTTTTCCCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCGACCCTGCCGCTTACCGGA

4498 TACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCAAGCTGGGCT

4598 GTGTGCACGAACCCCGTTGACCCGACCGTGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC

4698 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTT

4798 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCCGCTGGTAGCGGTGTTTTTTTGT

4898 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGG

4998 GATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTTGTGTAATCGT

5098 AACTAACATACGCTCTCCATCAAACAAACGAAACAAACAAACTAGCAAAATAGGCTGCCCCAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCG

5198 AA