



PvuI (7)
SgfI (6)
1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
EcoNI (96)

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
PvuII (239)
Bsu36I (291)
201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC
EcoNI (287)

301 GCCATCCACGCGGTTGAGTCGCGTTCGCCGCTCCCGCCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCTTGGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGTTGCTCAACTCTACGTCTTTGTTTCGTTT

BspLU11I (560)
AgeI (552)
501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCAACATGTCTCAGTGGAAATCAAGTCCAACAGTTAGAAATCAA
1 M S Q W N Q V Q Q L E I K

601 GTTTTTGAGCAGGTGGATCAATTCTATGATGACAACCTTTCCCATGGAAATTCGGCATCTGTTGGCCCAATGGATTGAAAATCAAGACTGGGAGGCGACT
13 F L E Q V D Q F Y D D N F P M E I R H L L A Q W I E N Q D W E A A

701 TCTAACAATGAAACCATGGCAACGATTCTTCTTCAAACTTGTAAATACAACCTGGATGAACAGTTAGGTCGTGTTTCCAAAGAGAAAACTACTCTTGA
47 S N N E T M A T I L L Q N L L I Q L D E Q L G R V S K E K N L L L

NsiI (856)
801 TACACAATCTAAAAAGAATTAGGAAGGTCCTTCAGGGAAAATTTCAATGGAAATCCAATGCATGTAGCTGTGGTTATTTCAAACCTGTTTAAAGGGAAGAGAG
80 I H N L K R I R K V L Q G K F H G N P M H V A V V I S N C L R E E R

XbaI (939)
901 GAGAATATTGGCTGCAGCCAACATGCCTGTCCAGGGCCCTCTAGAGAAATCCTTACAAAGTTCTTCAGTTTCAGAAAGACAGAGGAATGTGGAGCACAAA
113 R I L A A A N M P V Q G P L E K S L Q S S S V S E R Q R N V E H K

DraIII (994)
1001 GTGGCTGCCATTAATAAACAGTGTGCAGATGACAGAAACAAGATACCAAATACTTAGAAGATCTGCAAGACGAATTTGACTACAGGTATAAAAACAAATTCAGA
147 V A A I K N S V Q M T E Q D T K Y L E D L Q D E F D Y R Y K T I Q

BstAPI (1155)
1101 CAATGGATCAGAGTGACAAGAATAGTGCCATGGTGAATCAGGAAGTTTTGACACTGCAGGAAATGCTTAAACAGCCTCGATTTCAAGAGAAAGGAGGCTCT
180 T M D Q S D K N S A M V N Q E V L T L Q E M L N S L D F K R K E A L

SapI (1258)
1201 CAGTAAAAATGACCCAAATCATCCATGAGACAGACTGTTAATGAACACCATGCTCATAGAAGAGCTGCAAGACTGGAAGCGGGCGAGCAAAATCGCCTGC
213 S K M T Q I I H E T D L L M N T M L I E E L Q D W K R R Q Q I A C

1301 ATCGGGGGTCCACTCCACAATGGGCTCGACCAGCTTCAGAACTGCTTTACACTATTGGCAGAAAGTCTTTTCCAACAGGAAAGCAATGGGAGAACTAG
247 I G G P L H N G L D Q L Q N C F T L L A E S L F Q L R R Q L G E K L

1401 AGGAGCAATCTACAAAATGACATATGAAGGTGATCCCATTCCAATGCAAAGAACTCACATGCTAGAAAGAGTACCTTCTTGATCTACAACCTTTTCAA
280 E E Q S T K M T Y E G D P I P M Q R T H M L E R V T F L I Y N L F K

1501 GAACCTATTTGTGGTTGAGCGACAGCCATGTATGCCAACCCACCTCAGAGGCGGTTGGTACTTAAAACCCCTAATTCAGTTCACTGTAACCTAAGGCTA
313 N S F V V E R Q P C M P T H P Q R P L V L K T L I Q F T V K L R L

1601 CTAATAAAAATGGCAGAACTAACTATCAGGTAAGGTTAAGGCATCAATTGACAAAGAATGTTTCAACTTAAGCAACCGAAGATTTGTACTTTGTGGAA
347 L I K L P E L N Y Q V K V K A S I D K N V S T L S N R R F V L C G

1701 CTAATGTCAAAGCCATGTCTATTGAAGAATCTTCCAATGGGAGTCTCTCAGTAGAATTTGCAGATTTGCAACCAAGGAAATGAAGTCCAGTGTGGAGG
380 T N V K A M S I E E S S N G S L S V E F R H L Q P K E M K S S A G G

DraIII (1819)
Psp1406I (1848)
BglIII (1862)
1801 TAAAGAAATGAGGGCTGCACATGGTACTGAAGAACTTCATCCATAACGTTTGAACACAGATCTGCCTCTATGGCTGACCATAGATTTGGAGACC
413 K G N E G C H M V T E E L H S I T F E T Q I C L Y G L T I D L E T

1901 AGCTCATTGCCTGTGGTGTGATGATTTCCAATGTCAGTCAGTTACCTAATGCTTGGGCATCCATCATTTGGTACAACGTGTCAACCAACGATTCCAGAACT
447 S S L P V V M I S N V S Q L P N A W A S I I W Y N V S T N D S Q N

BsiWI (2072)
2001 TGGTTTTCTTAAATAATCCTCCACCTGCCACATTGAGTCAACTACTGGAGGTGATGAGCTGGCAGTTTTTCATCGTACGTTGGTGGTCTTAACTCAGA
480 L V F F N N P P P A T L S Q L L E V M S W Q F S S Y V G R G L N S D

HindIII (2122)
BstEII (2152)
2101 TCAACTCCATATGCTGGCAGAGAAGCTTACAGTCCAATCTAGCTACAGTGTGGTCACTCACCTGGGCCAAGTTCTGCAAGGAACATTTACCTGGTAAA
513 Q L H M L A E K L T V Q S S Y S D G H L T W A K F C K E H L P G K

2201 TCATTTACCTTTTGGACATGGCTTGAAGCAATATTGGATCTAATTAAGAAACACATTCTTCCCTTTGGATTGATGGGTATGTCATGGGCTTTGTTAGCA
547 S F T F W T W L E A I L D L I K K H I L P L W I D G Y V M G F V S

2301 AAGAGAAGGAACGGCTGTTGCTAAAGGATAAAAATGCCTGGCACCTTTTTATTAAGATTGAGTAAAGCCATCTCGGAGGAATAACTTTACCTGGGTGGA
580 K E K E R L L L K D K M P G T F L L R F S E S H L G G I T F T W V D

XcmI (2400)
2401 CCATTCTGAAAGTGGGAAGTGAGATTCCACTCTGTAGAACCCTACAATAAAGGCCGGTGTCTGCTGCTGCCATTGCTGACATCCTGCGAGACTACAAA
613 H S E S G E V R F H S V E P Y N K G R L S A L P F A D I L R D Y K

2501 GTTATTATGGCTGAAAACATTCCTGAAAACCTCTGAAGTACCTATATCCTGACATCCCAAAGACAAAGCCTTCGGTAAACACTACAGCTCTCAGCCTT
647 V I M A E N I P E N P L K Y L Y P D I P K D K A F G A K H Y S S Q P

2601 GCGAAGTTTCAAGACCAACAGAAAAGGGTGACAAAGGTTATGTTCTTCTGTTTTATCCCATCTCAACAATCCGAAGTGATTCAACAGAGCCACATTC
680 C E V S R P T E R G D K G Y V P S V F I P I S T I R S D S T E P H S

2701 TCCATCAGACCTTCTCCATGTCTCCAAGTGTGATGCGGTGTTGAGAGAAAACCTGAGTCCCAACAACATGAAACTGCAATGAAGTCTCCTTATTCT
713 P S D L L P M S P S V Y A V L R E N L S P T T I E T A M K S P Y S

MscI (2837)

2801 GCTGAATGACAGGATAAACTCTGACGCACCAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAA
747▶ A E •

HpaI (2969)

2901 AAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTTCATTTTATGT

EcoRI (3065)

3001 TTCAGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAAACTTTAACCT
3101 CCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCT

SapI (3247)

3201 CTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTAA

SwaI (3318)

3301 GTAAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCC
3401 AGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGA
141◀ • N R T Y K L P I

SacI (3579)

3501 TGAGTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCACTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACA
132◀ L E 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
3601 GGGGTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGCTCTTCTGCCGTTGCTCACAGCAGAC
99◀ P S 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

StuI (3743)

3701 CCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCC
65◀ G I A 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

BbsI (3889)

3801 CGACATGGTGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGT
32◀ V H 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 (3951)

3901 GGCCCTCCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTT

SacI (4008)

4001 CACTAAACGAGCTCTGTTATATAGACCTCCACCGTACACGCCTACCGCCATTGCGTCAATGGGCGGAGTTGTTACGACATTTGGAAAGTCCCGT

SpeI (4106)

4101 TGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCC

SnaBI (4234)

4201 AAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTATGACTGGGCATAATGCCAGGCGGG
4301 CCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTGCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGA
4401 CGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTA

PacI (4525) **SdaI (4517)** **BspLU11I (4535)**

4501 CCGTAAGTTATGTAACGCCTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCCTTTTC
4601 CATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
4701 GAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCAG

ApaLI (4849)

4801 CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTAT
4901 CGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGT
5001 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTC
5101 TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATC

EagI (5285) **PacI (5265)** **SwaI (5274)** **NotI (5284)**

5201 TTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAAT

5301 ATCTTTATTTTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAACGAAACAAACAAACTAGCAAAA
5401 TAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA