



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGCTCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)
AgeI (552) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACGGTACCCATGGAGCTGGCGGCTTGTGCCGTGGGGCTCCTCCT
1▶ M E L A A L C R W G L L L
601 CGCCTCTTGCCCCCGAGCCGCGAGCACCAAGTGTGACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACCTGGACATGCTC
13▶ A L L P P G A A S T Q V C T G T D M K L R L P A S P E T H L D M L

SdaI (780) EcoRV (787)
701 CGCCACCTCTACCAGGGCTGCCAGTGGTGCAGGAAACCTGGAACCTCACCTACCTGCCACCAATGCCAGCCTGTCTTCTGCAGGATATCCAGGAGG
47▶ R H L Y Q G C Q V V Q G N L E L T Y L P T N A S L S F L Q D I Q E
801 TGCAGGGCTACGTGCTCATCGCTCACAACCAAGTGAAGCAGGCTCCACTGCAGAGGCTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACATGC
80▶ V Q G Y V L I A H N Q V R Q V P L Q R L R I V R G T Q L F E D N Y A

BstBI (986) **Bsp119I (986)**
901 CCTGGCCGTGCTAGACAATGGAGACCCGCTGAACAATACCACCCCTGTACAGGGGCTCCCGAGGCTCGGGGAGCTGCAGCTTCAAGCCTCACA
113▶ L A V L D N G D P L N N T T P V T G A S P G G L R E L Q L R S L T

BglII (1001)
1001 GAGATCTTGAAGGAGGGTCTTGATCCAGCGGAACCCAGCTCTGCTACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACAGCTGG
147▶ E I L K G G V L I Q R N P Q L C Y Q D T I L W K D I F H K N N Q L
1101 CTCTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAGGGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTGTCA
180▶ A L T L I D T N R S R A C H P C S P M C K G S R C W G E S S E D C Q
1201 GAGCCTGACGCGCACTGTCTGTGCCGGTGGCTGTGCCGCTGCAAGGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGTGCTGCCGGCTGCACGGG
213▶ S L T R T V C A G G C A R C K G P L P T D C C H E Q C A A G C T G
1301 CCAAGCACTCTGACTGCCTGCCTCAACTTCAACCACAGTGGCATCTGTGAGTGCACCTGCCAGCCCTGGTCACTACAACAGACAGACTTTG
247▶ P K H S D C L A C L H F N H S G I C E L H C P A L V T Y N T D T F

Bst1107I (1425)
1401 AGTCCATGCCAATCCCGAGGGCTATACATTCCGGCCAGCTGTGTGACTGCCTGTCCCTACAACCTACCTTTCTACGGACGTGGGATCCTGCACCCT
280▶ E S M P N P E G R Y T F G A S C V T A C P Y N Y L S T D V G S C T L
1501 CGTCTGCCCTGCACAACCAAGAGGTGACAGCAGGAGTGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGAGTGTGCTATGGTCTGGGC
313▶ V C P L H N Q E V T A E D G T Q R C E K C S K P C A R V C Y G L G

BglII (1667)
1601 ATGGAGCACTTGCAGAGGTGAGGGCAGTTACCAGTGCCAATATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTCTGCCGGAGA
347▶ M E H L R E V R A V T S A N I Q E F A G C K K I F G S L A F L P E

SauI (1709)
1701 GCTTTGATGGGGACCCAGCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGTGAGACTCTGGAAGAGATCACAGTTACCTATACATCTC
380▶ S F D G D P A S N T A P L Q P E Q L Q V F E T L E E I T G Y L Y I S

BbsI (1831)
1801 AGCATGGCCGGACAGCCTGCCTGACCTCAGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACTCGCTGACCTGCAA
413▶ A W P D S L P D L S V F Q N L Q V I R G R I L H N G A Y S L T L Q

DraIII (1993)
1901 GGGCTGGGCATCAGCTGGCTGGGGTGCCTCACTGAGGGAAGTGGGAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTGCACAGCG
447▶ G L G I S W L G L R S L R E L G S G L A L I H H N T H L C F V H T
2001 TGCCCTGGGACCACTCTTTTCGAAACCCGACCAAGCTCTGCTCCACTGCAACCCGAGAGGACGAGTGTGGGGCAGGGCCTGGCCTGCCACCA
480▶ V P W D Q L F R N P H Q A L L H T A N R P E D E C V G E G L A C H Q

DraIII (2136) **ScaI (2196)**
2101 GCTGTGCGCCCGAGGGCACTGCTGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCCGGGGCCAGGAGTGCCTGGAGGAATGCCGAGTA
513▶ L C A R G H C W G P G P T Q C V N C S Q F L R G Q E C V E E C R V
2201 CTGAGGGGCTCCCGAGGAGTATGTAATGCCAGGCACTGTTTGGCGTGCACCCCTGAGTGTGAGCCCAAGTGGCTCAGTGCCTGTTTGGACCGG
547▶ L Q G L P R E Y V N A R H C L P C H P E C Q P Q N G S V T C F G P
2301 AGGCTGACCAAGTGTGGCTGTGCCACTATAAGGACCTCCCTTCTGCGTGGCCCGCTGCCAGCGGTGTGAAACCTGACCTCTCTACATGCCCCAT
580▶ E A D Q C V A C A H Y K D P P F C V A R C P S G V K P D L S Y M P I

SphI (2425)
2401 CTGGAAGTTTCCAGATGAGGAGGGCGCATGCCAGCCTTCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGTGGCCCGCCGAGCAG
613▶ W K F P D E E G A C Q P C P I N C T H S C V D L D D K G C P A E Q

BsaBI (2577)
2501 AGAGCCAGCCTCTGACGTCATCTCTCTGCGGTGGTGGCATTCTGCTGGTGTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGC
647▶ R A S P L T S I V S A V V G I L L V V V L G V V F G I L I K R R Q

BspEI (2606) **BsaBI (2694)**
2601 AGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGAGCGATGCCAACCAGGGCGCAGATGCG
680▶ Q K I R K Y T M R R L L Q E T E L V E P L T P S G A M P N Q A Q M R
2701 GATCCTGAAAGAGACGGCTGAGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAGAATGTG
713▶ I L K E T E L R K V K V L G S G A F G T V Y K G I W I P D G E N V

2801 AAAATTCCAGTGGCCATCAAAGTGTGAGGGAAAAACACATCCCCAAAGCCAACAAAGAAATCTTAGACGAAGCATACTGATGGCTGGTGTGGGCTCCC
747▶ K I P V A I K V L R E N T S P K A N K E I L D E A Y V M A G V G S
NdeI (2900) Tth111I (2983)
2901 CATATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTACACAGCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAA
780▶ P Y V S R L L G I C L T S T V Q L V T Q L M P Y G C L L D H V R E N
SacII (3000)
3001 CCGCGGACGCTGGGCTCCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGGCTCGTACACAGGGACTTG
813▶ R G R L G S Q D L L N W C M Q I A K G M S Y L E D V R L V H R D L
3101 GCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAATTTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGCAG
847▶ A A R N V L V K S P N H V K I T D F G L A R L L D I D E T E Y H A
SgrAI (3247)
3201 ATGGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGGTTCCACCACAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTG
880▶ D G G K V P I K W M A L E S I L R R R F T H Q S D V W S Y G V T V W
XmaI (3344)
3301 GGAGCTGATGACTTTTGGGGCCAAACCTTACGATGGGATCCCAGCCCGGAGATCCCTGACCTGCTGGAAAAGGGGGAGCGGCTGCCCCAGCCCCCATC
913▶ E L M T F G A K P Y D G I P A R E I P D L L E K G E R L P Q P P I
3401 TGCACCATTGATGCTACATGATCATGTCAAATGTTGGATGATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTTCCCGCATGG
947▶ C T I D V Y M I M V K C W M I D S E C R P R F R E L V S E F S R M
SmaI (3503) Eco47III (3512)
3501 CCAGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCCTCACTGCTGGAGGACGATGA
980▶ A R D P Q R F V V I Q N E D L G P A S P L D S T F Y R S L L E D D D
Acc65I (3633) XmaI (3669)
3601 CATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACAGCAGGGCTTCTTCTGTCCAGACCTGCCCCGGGCGCTGGGGGCATGGTCCACCACAGG
1013▶ M G D L V D A E E Y L V P Q Q G F F C P D P A P G A G G M V H H R
3701 CACCGCAGCTCATCTACCAGGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGCTCCACTGGCACCTCCGAAG
1047▶ H R S S S T R S G G G D L T L G L E P S E E E A P R S P L A P S E
3801 GGGCTGGTCCGATGATTTGATGGTACCTGGGAATGGGGGACCCAAAGGGCTGCAAAGCTCCACACATGACCCAGCCCTTACAGCGGTACAG
1080▶ G A G S D V F D G D L G M G A A K G L Q S L P T H D P S P L Q R Y S
3901 TGAGGACCCACAGTACCCTGCCCTCTGAGACTGATGGTACGTTGCCCCCTGACCTGCAGCCCCAGCCTGAATATGTGAACCAGCCAGATGTTCCG
1113▶ E D P T V P L P S E T D G Y V A P L T C S P Q P E Y V N Q P D V R
4001 CCCCAGCCCCCTCGCCCCGAGAGGGCCCTCTGCCTGTGCCGACCTGCTGGTCCACTCTGGAAAGGCCAAAGACTCTCTCCCAGGGAAGAAATGGGG
1147▶ P Q P P S P R E G P L P A A R P A G A T L E R P K T L S P G K N G
ScaI (4145)
4101 TCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCAGGGAGGAGCTGCCCTCAGCCCCACCCTCCTCTGCCTT
1180▶ V V K D V F A F G G A V E N P E Y L T P Q G G A A P Q P H P P P A F
4201 CAGCCCAGCCTTCGACAACCTCTATTACTGGGACCAGGACCCACCAGAGCGGGGGCTCCACCAGCACCTTCAAAGGGACACCTACGGCAGAGAACCCA
1213▶ S P A F D N L Y Y W D Q D P P E R G A P P S T F K G T P T A E N P
NheI (4347)
4301 GAGTACCTGGGTCTGGACGTGCCAGTGTGAACCAAGGCCAAGTCCGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC
1247▶ E Y L G L D V P V •
HpaI (4485) MfeI (4496)
4401 TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAAACAACAATA
4501 TGCATTCATTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCAT
4601 AGCAAACCTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTG
SapI (4763)
4701 TTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAACCTAGCTCTTCATTCTTTTATGTTTTAAATGCACTGACCTC
SspI (4820) SmaI (4834)
4801 CCACATTCCTTTTTAGTAAATATTAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC
4901 CTTCATAATATCCCCAGTTTAGTGTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCTG
1414 • N R
SacI (5095)
5001 GTGTAATTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCT
1374▶ T Y K L P I 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
5101 CTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCC
1044▶ R C M G C 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
StuI (5259)
5201 GTTGCTCACAGCAGACCAATGGCAATGGCTTACGACAGACAGTACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTG
714▶ N S V A S 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
5301 GTCCTGATGGCCGCCCCGACATGGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGT
374▶ T R I A A G 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
BbsI (5405) XmnI (5401) AseI (5467)
5401 TGAAGTCTTCATGGTGGCCCTCTATAGTGTGATCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCC
44▶ F T K M
SacI (5524)
5501 AGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACA
SpeI (5622)
5601 TTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTTGACGTCATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCCTATCCACGC

5701 CCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAAGGCATGACTGGG
SnaBI (5750)

5801 CATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAA
NdeI (5855)

5901 AACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAG

6001 CCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCG
SdaI (6033) PacI (6041) BspLU11I (6051)

6101 GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC
6201 CAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCG
6301 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCCTGCGC
6401 CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA
6501 GCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA
6601 AAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCA
PacI (6781) SwaI (6790)

6701 AGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC A
EagI (6801)
NotI (6800)

6801 GCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACA
6901 AAACAAACTAGCAAATAGGCTGTCCCCAGTGAAGTGCAAGTGCCAGGTCAGAACATTTCTCTATCGAA