



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

KasI (535) **AgeI (552)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTACCTGAGATCACCGGTCCATCATGACTAGCATCTTCCATTTTGCATTATCTTCATGT
1 M T S I F H F A I I F M
601 TAATACTTCAGATCAGAATAACAATTATCTGAAGAAAGTGAATTTTGTAGTTGATAGGTCAAAAAACGGTCTCATCCAGTTCCTAAAGCCTATCCCAGAA
13 L I L Q I R I Q L S E E S E F L V D R S K N G L I H V P K D L S Q K
701 AACAACAATCTTAAATATATCGCAAAATATATATCTGAGCTTTGGACTTCTGACATCTTATCACTGTCAAAACTGAGGATTTTGATAATTTCTCATAAT
46 T T I L N I S Q N Y I S E L W T S D I L S L S K L R I L I I S H N
R80 (802) **EcoRV (815)**
801 AGAATCCAGTATCTTGATATCAGTGTTTTCAAATTCACCAGGAATTGGAATACTGGATTTGTCCACAACAAGTTGGTGAAGATTTCTTGCCACCCTA
80 R I Q Y L D I S V F K F N Q E L E Y L D L S H N K L V K I S C H P
901 CTGTGAACCTCAAGCACTTGGACCTGTCAATTTAATGCATTTGATGCCCTGCCTATATGCAAAGAGTTTGGCAATATGTCTCAACTAAAATTTCTGGGGTT
113 T V N L K H L D L S F N A F D A L P I C K E F G N M S Q L K F L G L
1001 GAGCACCACACACTTAGAAAAATCTAGTGTGCTGCAATTTGCTCATTGAAATATCAGCAAGGCTTGTGCTTGTAGGAGAGACTTATGGGAAAAAGAA
146 S T T H L E K S S V L P I A H L N I S K V L L V L G E T Y G E K E

EcoO109I (1107) **Bsu36I (1103)** **DraIII (1139)**
1101 GACCCTGAGGGCCTTCAAGACTTTAACTGAGAGTCTGCACATTGTGTTCCCAACAACAAGAATTCATTTTATTTGGATGTGTCAGTCAAGACTG
180 D P E G L Q D F N T E S L H I V F P T N K E F H F I L D V S V K T
1201 TAGCAATCTGGAATCTAATATCAAATGTGTGCTAGAAGATAACAATGTTCTTACTTCCAAAGTATTCTGGCGAAACTTCAAACAAATCCAAAGTT
213 V A N L E L S N I K C V L E D N K C S Y F L S I L A K L Q T N P K L
BamHI (1350)
1301 ATCAAATCTTACCTTAAACAACATTGAAACAACCTGGAATTTCTTTCATTAGGATCCTCCAGTGGTTTGGCATACAACTGTATGGTATTTCTCAATTTCA
246 S N L T L N N I E T T W N S F I R I L Q L V W H T T V W Y F S I S
StuI (1461)
1401 AACGTGAAGTACAGGGTCTGACTTCCAGAGATTTTGATTATTCTGGCATTCTTGAAGCCTTGTCTATACACCAAGTTGTCAGCGATGTGTTTCG
280 N V K L Q G Q L D F R D F D Y S G T S L K A L S I H Q V V S D V F
BstBI (1531)
1501 GTTTTCCGCAAGTTATATCTATGAAATCTTTTGAATATGAACATCAAAAATTTTACAGTGTCTGGTACACGCATGGTCCACATGCTTTGCCATCCAA
313 G F P Q S Y I Y E I F S N M N I K N F T V S G T R M V H M L C P S K
1601 AATTAGCCCGTCTCGCATTGGATTTTCCAATAATCTTAAACAGACACGGTTTTTGAATAATGGGGCACCTTACTGAGTTGGAGACACTTATTTTA
346 I S P F L H L D F S N N L L T D T V F E N C G H L T E L E T L I L
1701 CAAATGAATCAATTAAGAAGCTTTCAAAAATAGCTGAAATGACTACACAGATGAAGTCTCTGCAACAATTGGATATTAGCCAGAATTCTGTAAGCTATG
380 Q M N Q L K E L S K I A E M T T Q M K S L Q Q L D I S Q N S V S Y
1801 ATGAAAAGAAAGGAGACTGTTCTTGACTAAAAGTTTATTAAGTTTAAATATGCTTCAAATATACTTACTGACACTATTTTTCAGATGTTTACCTCCAG
413 D E K K G D C S W T K S L L S L N M S S N I L T D T I F R C L P P R
HindIII (1965)
1901 GATCAAGTACTTGTATCTTACAGCAATAAAATAAAGAGCATTCTAAACAAGTCGTAATAACTGGAAGCTTTGCAAGAAGTCAATGTTGCTTTCAATTCT
446 I K V L D L H S N K I K S I P K Q V V K L E A L Q E L N V A F N S
2001 TTAAGTACCTTCTGGATGTGGCAGCTTTAGCAGCCTTTCTGTATTGATCATTGATCACAATTCAGTTTCCCACCCATCGGCTGATTTCTCCAGAGCT
480 L T D L P G C G S F S L S V L I I D H N S V S H P S A D F F Q S
2101 GCCAGAAGATGAGTCAATAAAGCAGGGGACAATCCATTCCAATGTACTCTGTGAGCTAGGAGAATTTGTCAAAAATATAGACCAAGTATCAAGTGAAGT
513 C Q K M R S I K A G D N P F Q C T C E L G E F V K N I D Q V S S E V
BspLU11I (2274)
2201 GTTAGAGGGCTGGCTGATTCTTATAAGTGTGACTACCCGAAAAGTTATAGAGGAACCTACTAAAGGACTTTTCCAGTCTGAATATCCTGCAACATA
546 L E G W P D S Y K C D Y P E S Y R G T L L K D F H M S E L S C N I
XcmI (2328) **BstXI (2325)** **PstI (2362)**
2301 ACTCTGCTGATCGTACCATCGTTGCCACCATGCTGGTGTGGCTGTGACTGTGACCTCCCTCTGCAGCTACTTGGATCTGCCCTGGTATCTCAGGATGG
580 T L L I V T I V A T M L V L A V T V T S L C S Y L D L P W Y L R M
NsiI (2475)
2401 TGTGCCAGTGGACCCAGACCCGCGCAGGGCCAGGAACATACCCTTAGAAGAAGTCCAAAGAAATCTCCAGTTTCATGCATTTATTTTCATATAGTGGCA
613 V C Q W T Q T R R A R N I P L E E L Q R N L Q F H A F I S Y S G H
2501 CGATCTTTCTGGGTGAAGAATTATTGCCAAACTAGAGAAAGAAGGATGAGATTTGCTTTCATGAGAGAAACTTTGTTCTGGCAAGAGCATT
646 D S F W V K N E L L P N L E K E G M Q I C L H E R N F V P G K S I
2601 GTGAAAAATATCATCACCTGCATTGAGAAGAGTTACAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAGTGAATGGTCCATTATGAAGTCTACT
680 V E N I I T C I E K S Y K S I F V L S P N F V Q S E W C H Y E L Y
SeaI (2769)
2701 TTGCCATCAAACTCTTTTCATGAAGGATCTAATAGCTTAATCCTGATCTTGTGGAACCCATTCCGAGTACTCCATTTCTAGCAGTTATCAAGCT
713 F A H H N L F H E G S N S L I L I L L E P I P Q Y S I P S S Y H K L

2801 CAAAAGTCTCATGGCCAGGAGGACTTATTTGGAATGGCCCAAGGAAAAGAGCAACAGTGGCCTTTTTGGGCTAACTTAAGGGCAGCCATTAATATTAAG
746▶ K S L M A R R T Y L E W P K E K S K R G L F W A N L R A A I N I K

2901 CTGACAGAGCAAGCAAAGAAAAGATCCTATCCCTATGATGTGCCAGACTATGCTGGCTATCCATATGATGTTCCCTGATTATGCTGGATACCCCTATGATG
780▶ L T E Q A K K R S Y P Y D V P D Y A G Y P Y D V P D Y A G Y P Y D

3001 TGCCAGACTATGCCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTT
813▶ V P D Y A •

3101 GTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGG
3201 GGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCT
3301 CTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTT

3401 TAAGATATAGTGATTTTTCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGAAAAATATTCAG

3501 AAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTG
3601 GACCTAGGGAACAAAGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGATGAGTTCCTCAAT
141◀ • N R T Y K L P I L E E I

3701 GGTGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCAC
128◀ 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 V V

3801 CTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGG
94◀ 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 I A

3901 CTTGACGACAGACAGTGACCCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCCCGACATGGTGCTT
61◀ 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 H K

4001 GTTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTTCATGATGGCCCTCTATAG
28◀ N D E Y L M T I K E T A V V L E L D Q Q S I N F T K M

4101 TGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGC

4200 TCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAG
4299 TCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATC

4399 ATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGT

4499 CATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTCAATGGGGCGGAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAA
4599 AGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGCGGGGTCTTGGGGGTCAGCCAGGCGGGCCATTTACCGTAAAGTTAT

4699 GTAACGCCTGCAGGTTAA TTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTC
4797 CGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC
4897 TCGTGCCTCTCCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGCCGCTTTTCATAGCTCACGCTGTAGGTA

4997 TCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGG
5097 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGT
5197 GGTGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGG
5297 CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG

5397 GGGTCTGACGCTCAGTGGAAACGAAAACCTACGTTAAGGGATTTTGGTTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTAT
5496 TTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAACGAAACAAACAACTAGCAAATAGGCTGT
5596 CCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA