



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

Bsu36I (291)
Psp1406I (203) **PvuII (239)** **EcoNI (287)**
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCGCCCGCCGCTACCTGAGGGCC
301 GCCATCCACGCGGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGTCTTTGTTTCGTTT

KasI (535) **AgeI (552)** **BstEII (575)**
501 TCTGTTTGTCCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTAGGAGGGCCAACATGTGGGTGACCAAACCTGCCAGCCCTG
601 CTGCTGCAGCATGTCCTCTGCATCTCCTCTGCTCCCATCGCCATCCCTATGCAGAGGGACAAAGGAAAAGAAGAAATACAATTCATGAATTCAAAA
111 L L Q H V L L H L L L L P I A I P Y A E G Q R K R R N T I H E F K
701 AATCAGCAAAGACTACCTAATCAAAATAGATCCAGCACTGAAGATAAAAAACCAAAAAGTGAATACTGCAGACCAATGTGCTAATAGATGTACTAGGAA
441 K S A K T T L I K I D P A L K I K T K K V N T A D Q C A N R C T R N
801 TAAAGGACTTCCATTCACTTGAAGGCTTTGTTTTGATAAAGCAAGAAAACATGCCTCTGGTTCCCTTCAATAGCATGTCAAGTGGAGTAAAAAA
771 K G L P F T C K A F V F D K A R K Q C L W F P F N S M S S G V K K
901 GAATTTGGCCATGAATTTGACCTCTATGAAAACAAAGACTACATTAGAAAAGTGCATCATTGGTAAAGGACGCAGCTACAAGGGAACAGTATCTATCACTA
111 E F G H E F D L Y E N K D Y I R N C I I G K G R S Y K G T V S I T
1001 AGAGTGGCATCAAAATGTAGCCCTGGAGTTCATGATACCACGAAACACAGCTTTTTGCTTCGAGCTATCGGGGTAAAGACCTACAGGAAAACACTAGT
144 K S G I K C Q P W S S M I P H E H S F L P S S Y R G K D L Q E N Y C

XhoI (1108) **SmaI (1124)**
1101 TCGAAATCCTCGAGGGGAAGAAGGGGGACCTGGTTCACAAAGCAATCCAGAGGTACGCTACGAAAGTCTGTGACATTCCTCAGTGTTCAGAAGTTGAA
177 R N P R G E E G G P W C F T S N P E V R Y E V C D I P Q C S E V E
1201 TGCATGACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCATACAGAATCAGGCAAGATTTGTGACGCTGGGATCATCAGACACCACCCGGCACA
211 C M T C N G E S Y R G L M D H T E S G K I C Q R W D H Q T P H R H

EcoRV (1315) **NcoI (1374)**
1301 AATCTTGGCTGAAAGATATCCCGACAAGGCTTTGATGATAATTATTGCCGCAATCCCGATGGCCAGCCGAGGCCATGGTGTATACTCTTGACCCTCA
244 K F L P E R Y P D K G F D D N Y C R N P D G Q P R P W C Y T L D P H

ScaI (1411) **Ppu10I (1481)** **NsiI (1481)**
1401 CACCCGCTGGGAGTACTGTGCAATTAACATGCGCTGACAATACTATGAATGACACTGATGTTCTTTGAAACAACATGAATGCATCCAAGGTCAAGGA
277 T R W E Y C A I K T C A D N T M N D T D V P L E T T E C I Q G Q G
1501 GAAGGCTACAGGGGCACTGTCAATACCATTGGAAATGGAATCCATGTGAGCGTTGGGATTCTCAGTATCCTCAGGAGCATGACATGACTCCTGAAAAAT
311 E G Y R G T V N T I W N G I P C Q R W D S Q Y P H E H D M T P E N
1601 TCAAGTGAAGGACTACGAGAAAATTACTGCCGAAATCCAGATGGGTCTGAATCACCCCTGGTGTTTTACCAGTATCCAACATCCGAGTTGGCTACTG
344 F K C K D L R E N Y C R N P D G S E S P W C F T T D P N I R V G Y C

BsaBI (1737) **BglII (1788)**
1701 CTCCCAAATTCAAAATGTGATATGTCACATGGACAAGATTGTTATCGTGGGAATGGCAAAAATTATATGGGCAACTTATCCCAAACAAGATCTGGACTA
377 S Q I P N C D M S H G Q D C Y R G N G K N Y M G N L S Q T R S G L

BbsI (1827)
1801 ACATGTTCAATGTGGGACAAGAATGGAAGACTTACATCGTCATATCTTCTGGGAACAGATGCAAGTAAGCTGAATGAGAATTACTGCCGAAATCCAG
411 T C S M W D K N M E D L H R H I F W E P D A S K L N E N Y C R N P
1901 ATGATGATGCTCATGGACCTGGTGTACACGGGAAATCCACTCATTCTGGGATTATTGCCCTATTTCTGTTGTGAAGGTGATACACACACTACAAT
444 D D D Y T G C Y T G G N P L I P W D Y C P I S R C E G D T T P T I
2001 AGTCAATTTAGACCATCCCGTAATATCTTGTGCCAAAACGAAACAATTGCGAGTTGTAATGGGATTCCAACGAAACAACATAGGATGGATGGTTAGT
477 V N L D H P V I S C A K T K Q L R V V N G I P T R T N I G W M V S

XhoI (2182)
2101 TTGAGATACAGAAATAAATATCTGCGGAGGATCATTGATAAAGGAGAGTTGGGTTCTTACTGCACGACAGTGTTCCTTCTCAGACTTGAAGATT
511 L R Y R N K H I C G G S L I K E S W V L T A R Q C F P S R D L K D

PvuII (2274)
2201 ATGAAGCTTGGCTTGAATTCATGATGTCCACGGAAGAGGAGATGAGAAATGCAACAGGTTCTCAATGTTTCCAGCTGGTATATGGCCCTGAAGGATC
544 Y E A W L G I H D V H G R G D E K C K Q V L N V S Q L V Y G P E G S

BglII (2305) **SmaI (2325)**
2301 AGATCTGGTTTTAATGAAGCTTGCAGGCTGTGCTGCTGGATGATTTTGTAGTACGATTGATTTACCTAATATGGATGCACAATTCCTGAAAAGACC
577 D L V L M K L A R P A V L D D F V S T I D L P N Y G C T I P E K T
2401 AGTTGAGTGTATGGCTGGGGTACACTGGATTGATCAACTATGATGGCCTATTACGAGTGGCACATCTCTATATAATGGGAAATGAGAAATGCAGCC
611 S C S V Y G W G Y T G L I N Y D G L R V A H L Y I M G N E K C S
2501 AGCATCATCGAGGAGGTGACTCTGAATGACTGAAATATGTGCTGGGGCTGAAAAGATTGGATCAGGACCATGTGAGGGGATTATGGTGGCCACT
644 Q H H R G K V T L N E S E I C A G A E K I G S G P C E G D Y G G P L
2601 TGTTTGTGAGCAACATAAATGAGAATGGTTCTTGGTGTATTGTTCTGGTGTGATGTCATTCCAATCGTCTGGTATTTTTGTCCGAGTAGCA
677 V C E Q H K M R M V L G V I V P G R G C A I P N R P G I F V R V A

NheI (2779) **Eco4VII (2776)**
2701 TATTATGCAAAATGGATACAAAATATTTTAAACATATAAGGTACCACAGTCAAGTGAAGTAAAGTGTCTGAAGCGCTAGCTGGCCAGACATGATA
711 Y Y A K W I H K I I L T Y K V P Q S •

2801 AGATACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA

2901 TAAGCTGCAATAAACAAGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTA

3001 CAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGG

3101 CATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTGCTCT

3201 TCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTT

3301 TTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAATAGAAATTGGACA

3401 GCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGTTGCCATTATCTCAATGAGCAC

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

3501 AAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCGACATGCCACAGGGGCTGACCCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTG

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

3601 ACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAA

81 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 I

3701 TGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTTCCTCATAGAGCATGGTATCTTCTCAGTGGCGAC

48 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

3801 CTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGA

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

3901 TTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCC

4000 CATTTCGTCATGGGGCGGAGTTGTTACGACATTTTGAAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACT

4099 TGGAAATCCCGTGAGTCAAACCGCTATCCAGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGGATGACTAATACGTAGATGTACTGC

4199 CAAGTAGGAAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACA

4299 CTTGATGTACTGCCAAGTGGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTAT

4399 TGACGTCAATGGGCGGGGTCGTTGGGCGTCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTGCAGGTTAA TTAAGAACATGTGAGCAAA

4497 AGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGT

4597 CAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGCTTACCGGAT

4697 ACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCTTCCGCTCAAGCTGGGCTG

4797 TGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA

4897 GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTG

4997 GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTG

5097 CAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGG

5197 ATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTTGTGTAATCGTA

5297 ACTAACATACGCTCTCCATCAAACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGA

5397 A