



PvuI (7) MfeI (82)
 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
 101 GAGAAAGTGGCGGGGTAACGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
 Psp1406I (203) HindIII (245) Bsu36I (291)
 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
 301 GCCATCCAGCGCGGTTGAGTCGCGTTTCTGCCGCCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC
 NgoMIV (441)
 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCGCTTGTCTCAACTCTACGCTTTTGTTCGTTT
 NcoI (560) BstEII (555) AgeI (552)
 501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTACCATTGGAAGATCCGCGTAGAAGGACGACGGCGCCACGGCG
 1 M E D P R R R T T A P R A
 Bsp120I (647) PvuI (685)
 601 TAAGAAGCGTCCGCGAAGCGGCCCGACGCAGCCAGCAGGACCAGGCCCGGAAAGCTGCGGCCCGCAAAGGGGGCTCGATCGCGCGGGCGG
 13 K K P S A K R A P T Q P S R T R A H A E S C G P Q R G A R S R R A
 BbrPI (728) Bsp120I (739) NcoI (797)
 701 GAGCGTGACGGGGACACCACGGAGAAGCCACGTGCCCGAGGGCCCGAGTGCATCCAGCAAGGGCCACTGAGCTCACCAAGATGCACAGCCCTCGGCCA
 47 E R D G D T T E K P R A P G P R V H P A R A T E L T K D A Q P S A
 BspEI (864) BamHI (861) BsiWI (879) SacII (899)
 801 TGGACGCGCAGGAGCCACCGCGCGGCTGCCGTCCGGGTGCCCGACGAGCAGGCCATCTGGATCCGGAGCTGCCCGCGTACGGGAGCCCGAGCCGCC
 80 M D A A G A T A R P A V R V P Q Q Q A I L D P E L P A V R E P Q P P
 BamHI (903)
 901 CGCGGATCCCGAGGCGGGAAAGTCTGAAGGGGACCTAGCCACAGAAGGGGCGCGCTCCACCGGGCAGCCAGAGCGCCGCGAGGGTCCAGGAAGGAA
 113 A D P E A R K V V R G P S H R R G A R S T G Q P R A P R G S R K E
 EcoRV (1048) EagI (1061) FspI (1097)
 1001 CCGGACAAGCTAAAGAAGGTGCTGGACAAATTGAGATTGAAACGCAAAGATATCTCGGAGCGCCGAGACGGTGAATAAAGTTGTGGAACCGCTGCTGC
 147 P D K L K K V L D K L R L K R K D I S E A A E T V N K V V E R L L
 1101 GCAGAATGCAGAAACGGGAGTCCGAGTTCAAAGGTGTGGAGCAGCTGAACACTGGCAGCTACTATGAACATGTGAAGATTTCTGCTCCTAATGAATTTGA
 180 R R M Q K R E S E F K G V E Q L N T G S Y Y E H V K I S A P N E F D
 PmeI (1206) SspI (1241) EcoRI (1283)
 1201 TGTTATGTTAAACTGGAAGTCCCAGGATTGAGCTACAAGAATATTATGAAACAGGTGCTTTCTATCTTGTGAAATTCAAAAGAAATCCACAGGAAAT
 213 V M F K L E V P R I E L Q E Y Y E T G A F Y L V K F K R I P R G N
 1301 CCGCTGAGTCATTTCTTAGAAGGGGAAGTATTATCAGTACCAAGATGCTGTCAAAGTTTAGGAAAATCATTAAAGAAGAAGTAAAGAAATCAAAGATA
 247 P L S H F L E G E V L S A T T K M L S K F R K I I K E E V K E I K D
 1401 TAGATGTCAGTGTGGAGAAGGAAAAACAGGAAGCCCTGCTGTAACTTCTTATCAGGAACCCCTGAAGAAATCTCTGTGGATATAATCTGCCTTTGGA
 280 I D V S V E K E K P G S P A V T L I R N P E E I S V D I I L A L E
 1501 GTCAAAGGCAGCTGGCCTATTAGTACCAAAGAAGGACTACCTATTCAAGGCTGGCTGGGCACAAAAGTGAGGACCAATCTAAGACGAGAGCCGTTTTAT
 313 S K G S W P I S T K E G L P I Q G W L G T K V R T N L R R E P F Y
 1601 CTGTAACCAAGAATGCAAAGGATGAAATAGTTTTCAAGGAGAGACCTGGCGCTCTTTCTCTCACACTGAAAAGTACATTTTGAATAATCACGGGA
 347 L V P K N A K D G N S F Q G E T W R L S F S H T E K Y I L N N H G
 BspEI (1723) PstI (1737) AseI (1758)
 1701 TAGAGAAAACATGCTGTGAATCTTCCGGAGCAAATGCTGCAGAAAAGATGTTTAAATTAATGAAATACCTTTTGGAACAGTTGAAAAAGAGTTTCA
 380 I E K T C C E S S G A K C C R K E C L K L M K Y L L E Q L K K E F Q
 NsiI (1808) SmaI (1880)
 1801 AGAGCTGGATGCATTCTGTTCTACCATGTGAAAAGTCCATCTTTCACATGTGGACCCAGGACCCGAGGACAGTCAGTGGGACCCAGGAACCTCAGC
 413 E L D A F C S Y H V K T A I F H M W T Q D P Q D S Q W D P R N L S
 Bsu36I (1938)
 1901 TCCTGCTTCGATAAGTTGTTAGCATTCTTTCTGAGTGCCTCAGGACAGAGAAAAGTGGATCATTATTTTATCCAAAAGTTCAATCTATTCTCTCAAGAAC
 447 S C F D K L L A F F L E C L R T E K L D H Y F I P K F N L F S Q E
 HindIII (2077)
 2001 TAATTGACCGAAAAAGTAAAGAATTTCTATCGAAGAAAATGAATATGAAAGAAAATAGGGTTTCCAATTTTGGCAAGCTTTGAAACTGTATTTGTGT
 480 L I D R K S K E F L S K K I E Y E R N N G F P I F D K L •
 MscI (2107) NheI (2101)
 2101 TGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCT
 HpaI (2239) MfeI (2250)
 2201 ATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGGGAGTTT
 EcoRI (2335)
 2301 TTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTGAACTCTTTTC
 2401 TGAGGGATGAATAAGGCATAGGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGATTT
 SapI (2517) SspI (2574) SmaI (2588)
 2501 TCCCAAGTTTGAAGTACCTCTTCAATTTTATGTTTAAATGCACTGACCTCCACATTCCTTTTGTAGAAAATATTCAGAAATAATTTAAATACAT

2601 CATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGG

2701 AACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCT
 141 • N R T Y K L P I L E E I T T K V L K

2801 TGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGTGACCACCCTGATGGATCTGTCCAC
 122 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 R I S R D V

2901 CTCATCAGAGTAGGGGTGCTGCAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTG
 89 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 E A C V T

3001 ACCCTGCCAATGTAGGCCCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTGTCTCATAGAGCA
 55 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 N D E Y L M

3101 TGGTGATCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCCTATAGTGAGTCGTATTATACTA
 22 T I K E T A V E V L E L D Q Q S I N F T K M

3201 TGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACAGACTCTGCTTATATAGACCTC

3301 CCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTGATTACTAGTCAAAACAACTCCCATTG
 3376

3401 ACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGA

3501 CTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTACTGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGG

3601 CGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGTACT

3701 ATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGACCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTGCAGGTTAAAT
 3788
 3787

3801 TAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAC

3901 AAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGA

4001 CCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGT

4101 TCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTACAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGAC
 4119

4201 TTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACA

4301 CTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACACCGCTGGTAG

4401 CGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAAC

4501 GAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGG
 4555
 4554
 4535
 4544

4601 TTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCC

4701 AGAACATTTCTCTATCGAA