



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

101 GAGAAAGTGGCGGGGTAACGGAAAGTGTGTCTGTACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTAAAGTCAAGTGCAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) **AgeI (552)** **BstEII (555)**
501 TCTGTTTGTGCGCGTTACAGATCCAAGCTGTGACCGGGCCCTAGCTGAGATCACCGGTACCATGGGTGACGTTAAAAATTTCTGTATGCCTGGTGTGG

601 CAAAAGGAAGATGACCCCATCTATGAAATTAGAGCAGTGGGGAACAAAAACAGGCAGAAATTCATGTGTGAGTTTCAGTGGAAAGTTATAATTACCT

13> K R K M T P S Y E I R A V G N K N R Q K F M C E V Q V E G Y N Y T

HpaI (757)
701 GGCATGGGAAATCCACCAATAAAAAAGATGCACAAAGCAATGCTGCCAGAGACTTTGTTAACTATTTGGTTCAATAAATGAAATAAAGAGTGAAGAAG

47> G M G N S T N K K D A Q S N A A R D F V N Y L V R I N E I K S E E

801 TTCAGCTTTTGGGTAGCATCTCCGCCCCACTTACTGATACCTGACACTACAGCAAATGCTGAAGGAGATTTACCAACAACCATGGGAGACCTCT

80> V P A F G V A S P P P L T D T P D T T A N A E G D L P T T M G G P L

Bsp120I (967)
901 TCCTCCACATCTGGCTCTCAAAGCAGAAAAATAATTCTGAGGTAGGGCCCTCTGGCTATGGTGTTCCTGGGCCACCTGGGACCGAGGAGCCAACCTGAAG

113> P P H L A L K A E N N S E V G A S G Y G V P G P T W D R G A N L K

1001 GATTACTACTCAAGAAAGGAAGAACAAGAGTGCAGGCACTTAGAATCAGAAGAAGTGATTTAAATGCTGGGCTTCATGGAACTGGACCTTGAAAA

147> D Y Y S R K E E Q E V Q A T L E S E E V D L N A G L H G N W T L E

1101 ATGCTAAAGCTCGTCTAAACCAATATTTTCAGAAAGAAAAAGATCCAAGGAGAATAAAGTACACCCAAGTGGTCTGATCACAACAGGAGCTTTATTGC

180> N A K A R L N Q Y F Q K E K I Q G E Y K Y T Q V G P D H N R S F I A

1201 AGAAATGACCATTTATATCAAGCAGCTGGGCGAAGGATTTTGCACGAGAACATGGATCAAATAAGAAATGGCAGCACAGTCTGTGCCCTGTCACCT

213> E M T I Y I K Q L G R R I F A R E H G S N K K L A A Q S C A L S L

BspEI (1339)
HindIII (1331)
1301 GTCAGACAACGTACCATCTGGAGTGGTTGAAGCTTACTCCGGACTTACAAAGAAGAAGGAAGGAGAGACAGTGGAGCCTTACAAAGTAAACCTCTCTC

247> V R Q L Y H L G V V E A Y S G L T K K K E G E T V E P Y K V N L S

1401 AAGATTTAGAGCATCAGCTGCAAAACATCATTCAAGAGCTAAATCTTGAATTTTGCCTCCGCTGAAGATCCTTCTGTGCCAGTTGCACTCAACATTGG

280> Q D L E H Q L Q N I I Q E L N L E I L P P E D P S V P V A L N I G

BstEII (1556) **SpeI (1588)**
1501 CAAATGGCTCAGTTCGAACCATCTCAGCGACAAAACCAAGTGGGTGTGGTTCCTTGGTACCTCCACAATCCAACCTGGAATCCTTGGACTAGTAGCAAC

313> K L A Q F E P S Q R Q N Q V G V V P W S P P Q S N W N P W T S S N

1601 ATTGATGAGGGGCTCTGGCTTTTGTACTCCAGAGCAATAAGCATGGACCTCAAGAATGAATTGATGTACCAGTTGGAACAGGATCATGATTTGCAAG

347> I D E G P L A F A T P E Q I S M D L K N E L M Y Q L E Q D H D L Q

1701 CAATCTTGAGGAGAGAGTACTGCCTGTGAAGAAATTTGAAAGTGAAGTCTGGAAGCAATCAGCCAAAATTCAGTTGTCTATTAGAGGGGCTAC

380> A I L Q E R E L L P V K K F E S E I L E A I S Q N S V V I I R G A T

1801 TGGATGTGGGAAAACACACAGTTCCCCAGTTCATTCTAGATGACTTTATCCAGAATGACCGAGCAGAGAGTGAACATCGTAGTAACTCAGCCGAGA

413> G C G K G T T T Q V P Q F I L D D F I Q N D R A A E C N I V V T Q P R

1901 AGAATCAGTGCCTTCTGTGGCAGAGGATTTGCATTTGAAAGAGGAGAAGAGCTGGAAAAAGCTGTGGTCAAGCTGATGATTTGATGCTATACCTC

447> R I S A V S V A E R V A F E R G E E P G K S C G Y S V R F E S I L

2001 CTCGCTCATGCCAGTATAATGTTTTGACTGTAGGTGTGCTCCTGAGAAAATAGAAGCAGGCATTCGAGGAATCAGTCATGTAATTGTAGATGAAAT

480> P R P H A S I M F C T V G V L L R K L E A G I R G I S H V I V D E I

Scal (2136)
2101 ACATGAAAGAGATATAAATACTGACTTCCTTTTGGTAGTACTGCGTGATGTTGTTTCAGGCTTATCCTGAAGTTCGCATTGTTCTTATGTCTGCTACTATT

513> H E R D I N T D F L L V V L R D V V Q A Y P E V R I V L M S A T I

BbsI (2287)
2201 GATACCAGCATGTTTTGTGAATATTTCTCAATTGCCCATCATTGAAGTTTATGGGAGGACTTACCCAGTTCAAGAATATTTCTGGAAGACTGCATTC

547> D T S M F C E Y F F N C P I I E V Y G R T Y P V Q E Y F L E D C I

2301 AGATGACCCACTTTGTTCTCCACAAAAGACAAAAGAAAGGATAAGGATGATGATGGTGGTGGAGGATGATGATGCAAAATGCAACTTGATCTGTGG

580> Q M T H F V P P P K D K K K K D K D D D G G E D D D A N C N L I C G

2401 TGATGAATATGGTCCAGAAACAAGGTTGAGCATGTCTCAATTGAACGAAAAGGAACTCCTTTTGAACATCAGAGGCTCTACTTAAGTACATTGAAACC

613> D E Y G P E T R L S M S Q L N E K E T P F E L I E A L L K Y I E T

2501 CTTAATGTTCTGGAGCTGTGTTGGTTTTTTCCTGGCTGGAATCTGATTTACTATGCAGAAGCATTGGAAATGAATCCACATTTTGAAGCCATC

647> L N V P G A V L V F L P G W N L I Y T M Q K H L E M N P H F G S H

XhoI (2633)
2601 GGTATCAGATTCTACCCCTGCATTCTCAGATTCCTCGAGAGGAACAGCGCAAAGTGTGTTGATCCAGTACCAGTTGGAGTAACCAAGGTTATTTGTCCAC

680> R Y Q I L P L H S Q I P R E E Q R K V F D P V P V G V T K V I L S T

2701 AAATATTGCTGAACAAGCATAAACCATGTTGTTTATGTACTTGCCTGCAAGCAGAAAAGTGAACCTGTTACTGCTCACAACAATATGACC

713> N I A E T S I T I N D V V Y V I D S C K Q K V K L F T A H N N M T

2801 AACTATGCTACCGTATGGGCATCAAAAACAACCTTGAGCAACGGAAAGGGCAGCTGGCCGAGTACGGCCTGGATTCTGCTTTCACCTGTGCAGCCGAG

747> N Y A T V W A S K T N L E Q R K G R A G R V R P G F C F H L C S R

2901 CTCGTTTTGAGAGACTTGAACCCACATGACACCAGAGATGTTCCGAACACCATTGCATGAAATGCTCTTAGCATAAAAATCTGCGTCTAGGAGGAAT

780> A R F E R L E T H M T P E M F R T P L H E I A L S I K L L R L G G I

NsiI (3080)

3001 TGGCCAATTTCTGGCCAAAGCAATTGAACCTCCCCCTTTGGATGCTGTGATTGAAGCAGAACACACTCTTAGAGAGCTTGATGCATTAGATGCCAATGAT
813> G Q F L A K A I E P P P L D A V I E A E H T L R E L D A L D A N D
3101 GAGTTGACTCCTTTGGGACGAATCCTGGCTAAACTCCCATTTAGAGCCTCGTTTTGGCAAAATGATGATAATGGGGTGTATTTTCTACGTGGGAGATGCTA
847> E L T P L G R I L A K L P I E P R F G K M M I M G C I F Y V G D A
3201 TCTGTACCATTGCTGCTGCTACCTGCTTTCCAGAGCCTTTTCATCAATGAAGGAAAGCGGCTGGGCTATATCCATCGAAATTTTGTGGAAACAGATTTTC
880> I C T I A A A T C F P E P F I N E G K R L G Y I H R N F A G N R F S
3301 TGATCAGTAGCCCTTTTATCAGTATTTCAAGCCTGGGATGATGCTAGAATGGTGGAGAAGAAGCAGAGATACGTTTTTGTGAGCACAAAAGACTTAAT
913> D H V A L L S V F Q A W D D A R M G G E E A E I R F C E H K R L N
3401 ATGGCTACACTAAGAATGACCTGGGAAGCCAAAGTTTCAGCTCAAAGAGATTTTGATTAATTTCTGGGTTTCCAGAAGATTGTTTGTGACACAAGTGTTTA
947> M A T L R M T W E A K V Q L K E I L I N S G F P E D C L L T Q V F
3501 CTAACACTGGACCAGATAATAATTTGGATGTTGTTATCTCCCTCCTGGCCTTTGGTGTGTACCCCAATGTATGCTATCATAAGGAAAAGAGGAAGATTCT
980> T N T G P D N N L D V V I S L L A F G V Y P N V C Y H K E K R K I L
3601 CACGACTGAAGGGCGTAATGCACCTTCCACAAATCATCTGTTAATTTGCTTTTAGTAGCCAAGACATGAAGTACCCATCTCCCTTCTTTGTATTTGGT
1013> T T E G R N A L I H K S S V N C P F S S Q D M K Y P S P F F V F G

XhoI (3713)

PshAI (3738)

PstI (3754)

BstXI (3786)

3701 GAAAAGATTCGAACCTCGAGCCATCTCTGCTAAAGGCATGACTTTAGTACACCCCTGCAGTTGCTTCTTTGCCTCCAAGAAAGTCCAAATCTGATGGGC
1047> E K I R T R A I S A K G M T L V T P L Q L L F A S K K K V Q S D G
3801 AGATTGTGCTTGTAGACTGGATTAACCTGCAAAATCTCATGAAGCTGCTGCTGTATCACTGGTCTCCGGGCAGCCATGGAGGCTTTGGTTGTTGA
1080> Q I V L V D D W I K L Q I S H E A A A C I T G L R A A M E A L V V E

BglII (3971)

3901 AGTAACCAACAACCTGCTATCATCAGCCAGTTGGACCCCGTAAATGAACGTATGCTGAACATGATCCGTCAGATCTCTAGACCCCTCAGCTGCTGGTATC
1113> V T K Q P A I I S Q L D P V N E R M L N M I R Q I S R P S A A G I
4001 AACCTTATGATTGGCAGTACACGGTATGGAGATGGTCCACGTCCTCCCAAGATGGCCCGATACGACAATGGAAGCGGATATAGAAGGGGAGGTTCTAGTT
1147> N L M I G S T R Y G D G P R P P K M A R Y D N G S G Y R R G G S S
4101 ACAGTGGTGGAGGCTATGGCGGTGGCTATAGCAGTGGAGGCTATGGTAGCGGAGGCTATGGTGGCAGCGCCAACCTCTTTCCGGCAGGATATGGTGCAGG
1180> Y S G G G Y G G G Y S S G G Y G S G G Y G G S A N S F R A G Y G A G
4201 TGTTGGTGGAGGCTATAGAGGAGTTTCCCGAGGTGGCTTTAGAGGCAACTCTGGAGGAGACTACAGAGGCGCTAGTGGAGGCTACAGAGGCTTGGGGGA
1213> V G G G Y R G V S R G G F R G N S G G D Y R G P S G G Y R G S G G
4301 TTCCAGCGAGGAGGTGGTAGGGGGCTATGGAAGTGGCTACTTTGGACAGGGAAGAGGAGGTGGCGGCTATTAACCTGGTTATGTCAGTTCTCTGTG
1247> F Q R G G G R G A Y G T G Y F G Q G R G G G G Y •

NheI (4400)

4401 GCTAGCTGGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTA

HpaI (4538)

4501 TTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTT

EcoRI (4634)

4601 TTAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAGCCTCTACTTGAATCCTTTTCT

4701 GAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTTCATGGAGTTTAAAGATATAGTGATTTTT

4801 CCCAAGGTTTGAAGTACTCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAAATATTTAAATACATC

4901 ATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTGTAGTGTGGACTTAGGGAACAAGGA

5001 ACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTT

1414 • N R T Y K L P I L E E I T T K V L K

BstXI (5177)

5101 GCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACC

1224 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

5201 TCATCAGATAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCACCCCAATGGVCAATGGCTTACAGCAGACAGTGA

884 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 V

StuI (5312)

5301 CCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGCTCATAGAGCAT

554 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

BbsI (5458)

5401 GGTGATCTTCTCAGTGGCGACTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTAT

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

5501 GCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTATCTGACGGTTCCTAAACGAGCTCTGCTTATATAGACCTCC

SpeI (5675)

5601 CACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGA

5701 CGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGAC

SnaBI (5803)

5801 TAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAATGCCAGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGC

NdeI (5908)

5901 GTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTA

PacI (6094)

PstI (6087)

SdaI (6086)

6001 TGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGACAGGTTAAAT

BspLU11I (6104)

6101 AAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACA

6201 AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGAC

6301 CTGCGCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT

ApaI (6418)

6401 CGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT

6501 TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACAC

6601 TAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTTTGATCCGGCAAACAAACCACCGCTGGTAGC

6701 GGTGGTTTTTTTGGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACG

EagI (6854)

PacI (6834)

NotI (6853)

6801 AAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGT

6901 TTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCA

7001 GAACATTTCTCTATCGAA