



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203)
 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCGCCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

KasI (535) **AgeI (552)** **BspHI (560)** **BstAPI (575)** **Acc65I (596)**
 501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTATCATGAAGATGGCAAGCACCCGCTGCAAGCTGGCCAGGTA

601 CCTGGAGGACCTGGAGGATGTGGACTTGAAGAAATTTAAGATGCACCTTAGAGGACTATCCTCCCAGAAGGGTGCATCCCCCTCCGAGGGGTGAGACA

13> L E D L E D V D L K K F K M H L E D Y P P Q K G C I P L P R G Q T

701 GAGAAGGCAGACCATGTGGATCTAGCCACGCTAATGATCGACTTCAATGGGGAGGAGAAGCGCTGGGCCATGGCCGTGTGGATCTTCGCTGCGATCAACA

47> E K A D H V D L A T L M I D F N G E E K A W A M A V W I F A A I N

Neol (767)
BstBI (864)
Bsp119I (864)
BbrPI (857)
 801 GGAGAGACCTTTATGAGAAAGCAAAAAGAGATGAGCCGAAGTGGGGTTCAGATAATGCACGTGTTTTCGAATCCCCTGTGATATGCCAGGAAGACAGCAT

80> R R D L Y E K A K R D E P K W G S D N A R V S N P T V I C Q E D S I

901 TGAAGAGGAGTGGATGGGTTTACTGGAGTACCTTTTCGAGAATCTATTTGTAATAAGAAAGATTACCGTAAAGATACGAAAGTACGTGAGAAGC

113> E E E W M G L L E Y L S R I S I C K M K K D Y R K K Y R K Y V R S

Tth111I (1070)
 1001 AGATTCCAGTGCATTGAAGACAGGAATGCCCGTCTGGGTGAGAGTGTGAGCCTCAACAAACGCTACACACGACTGCGTCTCATCAAGGAGCACCGGAGCC

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

1101 AGCAGGAGAGGGAGCAGGAGCTTCTGGCCATCGGCAAGACCAAGAGCTGTGAGAGCCCGTGTGAGTCCCATTAGATGGAGTGTCTGTTTGACCCCGATGA

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

DraIII (1219)
 1201 TGAGCATTCTGAGCCTGTGCACACCGTGTGTTCCAGGGGGCGGAGGGATTGGGAAAACAATCTGGCCAGGAAGATGATGTTGGACTGGGCATCGGGG

213> E H S E P V H T V V F Q G A A G I G K T I L A R K M M L D W A S G

BspHI [m] (1386)
 1301 AACTCTACCAAGACAGGTTTACTATCTGTTCTATATCCACTGTCGGGAGGTGAGCCTGTGACACAGAGGAGCCTGGGGGACTGATCATGAGCTGCT

247> T L Y Q D R F D Y L F Y I H C R E V S L V T Q R S L G D L I M S C

1401 GCCCGACCAAAACCCACCCATCCACAAGATCGTGAGAAAACCTCCAGAATCCTTCTCATGGACGGCTTCGATGAGCTGCAAGGTGCCTTTGACGA

280> C P D P N P P I H K I V R K P S R I L F L M D G F D E L Q G A F D E

BsrBI (1510) **BsrBI (1537)** **StuI (1589)**
 1501 GCACATAGGACCGCTCTGCACTGACTGGCAGAAGGCGGAGCGGGAGACATTCTCCTGAGCAGCCTCATCAGAAAAGAGCTGCTTCCCGAGGCCCTCTCTG

313> H I G P L C T D W Q K A E R G D I L L S S L I R K K L L P E A S L

XcmI (1607) **BstAPI (1635)**
 1601 CTCATCACCAGAGACTGTGGCCCTGGAGAACTGCAGCACTTGTGGACCATCCTCGGCATGTGGAGATCCTGGGTTTCTCCGAGGCCAAAAGGAAAG

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

ScaI (1700) **ScaI (1712)** **DraIII (1779)**
 1701 AGTACTTCTTCAAGTACTTCTCTGATGAGGCCAAGCCAGGGCAGCCTTCACTGATTCAGGAGAAGAGGTCCTTCCACCATGTGCTTCATCCCCCT

380> E Y F F K Y F S D E A Q A R A A F S L I Q E N E V L F T M C F I P L

SacII (1880)
 1801 GGTCTGCTGGATCGTGTGCACTGGACTGAAACAGCAGATGGAGAGTGGCAAGAGCCTTGGCCAGACATCTAAGACCACCACCGGGTGTACGTTCTTCTC

413> V C W I V C T G L K Q Q M E S G K S L A Q T S K T T T A V Y V F F

XmaI (1919)
 1901 CTTCCAGTTTGTGAGCCCGGGGAGGGAGCCAGGACACGGCCTCTGCGCCACCTCTGGGGCTCTGCTCTTTGGCTGCAGATGGAATCTGGAACC

447> L S S L L Q P R G G S Q E H G L C A H L W G L C S L A A D G I W N

2001 AGAAAATCCTGTTTGGAGAGTCCGACCTCAGGAATCATGGACTGCAGAAGGGGATGTGTCTGCTTCTGAGGATGAACCTGTTCCAAAAGGAAGTGA

480> Q K I L F E E S D L R N H G L Q K A D V S A F L R M N L F Q K E V D

Psp1406I (2194)
 2101 CTGCGAGAAGTTCTACAGCTTCATCCACATGACTTTCCAGGAGTCTTTGCCCCATGTACTACCTGCTGGAAGAGGAAAAGGAAGGACGAACTT

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

BstBI (2263)
Bsp119I (2263)
 2201 CCAGGGAGTCGTTTGAAGCTTCCAGCCGAGACGTGACAGTCTTCTGGAAAATCGAAAATTCGAAAAGGGGTATTTGATTTTTGTTGTACGTTTCC

547> P G S R L K L P S R D V T V L L E N Y G K F E K G Y L I F V V R F

BglIII (2357)
 2301 TCTTTGGCCTGGTAAACCAGGAGAGGACCTCCTACTTGGAGAAGAAATTAAGTTGCAAGATCTCTCAGCAATCAGGCTGGAGCTGCTGAAATGGATTGA

580> L F G L V N Q E R T S Y L E K K L S C K I S Q Q I R L E L L K W I E

Neol (2498)
 2401 AGTGAAAGCCAAAGCTAAAAAGCTGCAGATCCAGCCAGCCAGCTGGAATTTGTTACTGTTTGTACGAGATGCAGGAGGAGGACTTCGTGCAAAGGGCC

613> V K A K A K K L Q I Q P S Q L E L F Y C L Y E M Q E E D F V Q R A

BsaBI (2517)

2501 ATGGACTATTTCCCAAGATTGAGATCAATCTCTCCACCAGAATGGACCACATGGTTTCTTCTTTTGCATTGAGAACTGTCATCGGGTGGAGTCACTGT
647▶ M D Y F P K I E I N L S T R M D H M V S S F C I E N C H R V E S L
2601 CCTGGGGTTTCTCATAACATGCCAAGGAGGAGGAGGAAAAGGAAAGCCGACACCTTGATATGGTGCAGTGTGTCCTCCCAAGCTCCTCTCA
680▶ S L G F L H N M P K E E E E E K E G R H L D M V Q C V L P S S S H
2701 TGTCGCCTGTTCTCATGGATTGGTGAACAGCCACCTCACTTCCAGTTTTTGGCCGGGCTCTTTTTCAGTCTGAGCACCAGCCAGAGTCTAACTGAATTG
713▶ A A C S H G L V N S H L T S S F C R G L F S V L S T S Q S L T E L

SmaI (2822)

2801 GACCTCAGTGACAATTCTCTGGGGACCCAGGGATGAGAGTGTGTGTGAAACGCTCCAGCATCTGGCTGTAACATTCGGAGATTGTGGTTGGGGCGCT
747▶ D L S D N S L G D P G M R V L C E T L Q H P G C N I R R L W L G R
2901 GTGGCCTCTCGCATGAGTGTCTCGACATCTCTTGGTCTCAGCAGCAACAGAAGCTGGTGGAGCTGGACCTGAGTGACAACGCCCTCGGTGACTT
780▶ C G L S H E C C F D I S L V L S S N Q K L V E L D L S D N A L G D F
3001 CGGAATCAGACTTCTGTGTGGGACTGAAGCACCCTGTTGTGCAATCTGAAGAAGCTCTGGTTGGTCACTGCTGCCTCACATCAGCATGTTGTCAGGAT
813▶ G I R L L C V G L K H L L C N L K K L W L V S C C L T S A C C Q D

XcmI (3134)

3101 CTTGCATCAGTATTGAGCACCAGCCATTCCTGACCAGACTCTATGTGGGGGAGAATGCCTTGGGAGACTCAGGAGTCGCAATTTTATGTGAAAAAGCCA
847▶ L A S V L S T S H S L T R L Y V G E N A L G D S G V A I L C E K A

EcoRI (3237)

3201 AGAATCCACAGTGTAACTGCAGAAACTGGGGTGGTGAATTCGGCCTTACGTCAGTCTGTTGTTTCAGCTTTGCTCCTCGGTAAGTCACTAATCAGAA
880▶ K N P Q C N L Q K L G L V N S G L T S V C C S A L S S V L S T N Q N
3301 TCTCACGCACCTTACCTGCGAGGCAACTCTCGGAGACAAGGGGATCAAACTACTCTGTGAGGGACTCTTGACCCCGACTGCAAGCTTCAGGTGTTG
913▶ L T H L Y L R G N T L G D K G I K L L C E G L L H P D C K L Q V L
3401 GAATTAGACAACCTCAACCTCACACTGCTGGGATCTTTCCACACTTCTGACCTCCAGCCAGCAGCTCGAAAGCTGAGCCTGGGCAACAATG
947▶ E L D N C N L T S H C C W D L S T L L T S S Q S L R K L S L G N N

BspHI (3518)

3501 ACCTGGGCGACCTGGGGTGCATGATGTTCTGTGAAGTGTGAAACAGCAGAGCTGCCTCCTGCAGAACCTGGGGTGTCTGAAATGTATTTCAATTATGA
980▶ D L G D L G V M M F C E V L K Q Q S C L L Q N L G L S E M Y F N Y E

PshAI (3645)

NheI (3686)

3601 GACAAAAAGTGCCTTAGAAAACACTTCAAGAAGAAAAGCCTGAGCTGACCGTCTTTGAGCCTTCTTGGTAGGAGTGAAACGGGGCTAGCTGGCCAGA
1013▶ T K S A L E T L Q E E K P E L T V V F E P S W •
3701 CATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTA

HpaI (3824) MfeI (3835)

3801 ACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAA

EcoRI (3920)

3901 ACCTCTACAAATGTGGTATGGAATTCTAAATAACAGCATAGCAAAAACCTTAACTCCAATCAAGCCTCTACTGAACTCTTTCTGAGGGATGAATAAG
4001 GCATAGGCATCAGGGCTGTGCCAATGTGCATTAGCTGTTTGACGCTCACCTCTTTCATGGAGTTAAGATATAGTGATTTTCCAAGGTTTGAAC

SapI (4102)

SspI (4159)

SwaI (4173)

4101 TAGCTCTTCATTTCTTTATGTTTAAATGCAGTACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAAT

4201 AAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAA

4301 TTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGATGAGTTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAA
1114 ▶ • N R T Y K L P I L E E I T T K V L K G N M E I

SacI (4434)

4401 TGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGG
1174 ▶ 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 E D S Y P

StuI (4598)

4501 GTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAG
841 ▶ 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 R G I Y

4601 GCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGCTCATAGAGCATGGTATCTTCTCAG
504 ▶ A E I H V A S I E I G K T R I A A G V H H K N D E Y L M T I K E T

4701 TGGCGACCTCCACAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCTGATATTATACTATGCCGATATACTATG
174 ▶ A V E V L E L D Q Q S I N F T K M

AseI (4806)

SacI (4863)

4801 CCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCT

SpeI (4961)

4901 ACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGG

SnaBI (5089)

5001 AGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATG

NdeI (5194)

5101 ACTGCCAAGTAGGAAAGTCCATAAGGTGATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATTGACGTCAATAGGGGGCGTACTTGGCATATG

5201 ATACACTTGATGACTGCCAAGTGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTC

PacI (5380)

SdaI (5372)

BspLU11I (5390)

5301 ATTATTGACGTCAATGGGGCGGGTCTGTTGGCGGTACAGCCAGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATAAGAACATGTGAGC

5401 AAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCA
5501 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGCTTACCG
5601 GATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGG
5701 CTGTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA
5801 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTAT
5901 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGT
6001 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACTCACGTTAA

EagI (6140)

PacI (6120) SmaI (6129) **NotI (6139)**

6101 GGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATC
6201 GTAACAAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTAT
6301 CGAA