



PvuI (7)
SgfI (6)
1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGAAAAGTATGTCGTGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **Bsu36I (291)**
201 GTGAACGTTCTTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGGCCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGGCTTTCGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552) **SphI (560)** **NcoI (596)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAGCATGCAGTATCTAAATATAAAAGAGGACTGCAATGCCAT
601 GGCCTTCTGTGCTAAAATGAGGAGCTTCAAGAAGACTGAGTGAAGCAGGTGGTCCCTGAGCCTGGAGTGGAGGTGACTTTCTATCTGTTGGACAGGGAG
13▶ A F C A K M R S F K K T E V K Q V V P E P G V E V T F Y L L D R E

Bst1107I (723)
701 CCCCTCCGCTGGGACGCGAGAGTATACAGCCGAGGAGCTGTGCATCAGGGCCGCCAGGAGTGCAGTATCTCTCCTCTGTGCACAACTCTTCGCC
47▶ P L R L G S G E Y T A E E L C I R A A Q E C S I S P L C H N L F A
801 TGTACGATGAGAGCACCAGCTCTGGTACGCTCCGAACCGAATCATCACTGTGGATGACAAAACGCTCTCCGGCTCCACTACCGCATGAGGTTCTACTT
80▶ L Y D E S T K L W Y A P N R I I T V D D K T S L R L H Y R M R F Y A
901 TACCAACTGGCAGCAACCAATGACAACGACAGCTGTATGGCGACATTCTCAAAGAGCAGAAAACGGCTATGAGAAGAAAAGGTTCCAGAAGCA
113▶ T N W H G T N D N E Q S V W R H S P K K Q K N G Y E K K R V P E A
1001 ACCCACTCCTTGATGCCAGTTCAGTGGAGTATCTGTTTGACAGGGACAGTATGATTTGATCAAATGCCTGGCTCCCATTCGGGACCCCAAGACGGAGC
147▶ T P L L D A S S L E Y L F A Q G Q Y D L I K C L A P I R D P K T E
1101 AAGACGGACATGATATTGAAAATGAGTCCCTGGGCATGGCGGTCTGGCCATCTCCCACTATGCCATGATGAAGAAGTGCAGTTGCCGGAACCTCCCAA
180▶ Q D G H D I E N E C L G M A V L A I S H Y A M M K K M Q L P E L P K
1201 AGACATCAGTACAACGATATATCCAGAAACATTGAATAAATCCATCAGACAGAGGAACCTTCTTACCAGGATCGGAATAAATAATGTTTTCAAGGAT
213▶ D I S Y K R Y I P E T L N K S I R Q R N L L T R M R I N N V F K D
1301 TTCTTGAAGGAATTTAAACAAGACCATCTGTGACAGCAGTGTGTCTACACATGACCTGAAGGTGAAATACCTGGTACCTTGGAAACTTTGACAAAAC
247▶ F L K E F N N K T I C D S S V S T H D L K V K Y L A T L E T L T K

SspI (1413) **BstBI (1470)**
1401 ATTATGGAGCTGAAATATTTGAGACTTCTATGCTACTGATTTTCATCAGAAAATGAATTGAGTCGATGCCATTGCAATGACAGTGGCAATGTTCTCTATGA
280▶ H Y G A E I F E T S M L L I S S E N E L S R C H S N D S G N V L Y E

BamHI (1523)
1501 GGCATGGTGACTGGAATCTCGGGATCCAGTGGCGGCAGAAACCAATGTTGTTCTGTTGAAAAGGAAAAATAAAGTGAAGCGGAAAAAAGTGGAA
313▶ V M V T G N L G I Q W R Q K P N V V P V E K E K N K L K R K K L E
1601 TATAATAAACACAGAAGGATGATGAGAGAAACAACTCCGGGAAGAGTGAACAATTTTTCTATTTCCCTGAAATCACCCACATTTGAATAAAGGAGT
347▶ Y N K H K K D D E R N K L R E E W N N F S Y F P E I T H I V I K E

XhoI (1754)
1701 CTGTGGTCAAGCTTAAACAAACAGGACAACAAAAACATGGAAGTCAAGCTCTTCTCGAGAGGAAGCCTTGTCTTTGTGTCCTGGTGGATGGTACTT
380▶ S V V S I N K Q D N K N M E L K L S S R E A L S F V S L V D G Y F
1801 CCGGCTCACTGCAGATGCCACCACTTCTGTACTGATGTGGCTCCCCACTGATTGTCCACAATACAGAAGCAGGCTGCCACGGTCCAATCTGCACA
413▶ R L T A D A H H Y L C T D V A P P L I V H N I Q N G C H G P I C T
1901 GAATATGCCATCAATAAGCTGCGGCAGGAAGGGAGTGAAGAGGGGATGACGTGCTGAGGTGGAGCTGCACCGACTTTGACAACATTTATGACTGTCA
447▶ E Y A I N K L R Q E G S E E G M Y V L R W S C T D F D N I L M T V

XmnI (2034)
2001 CCTGCTTTGAAAAGTCTGAGGTATTGGGTGGCCAGAAGCAGTTCAAGAATTTAGATTGAGGTACAGAAGGGCCGCTACAGCCTGCATGGCTCTATGGA
480▶ T C F E K S E V L G G Q K Q F K N F Q I E V Q K G R Y S L H G S M D

FspI (2150)
2101 CCACTTTCCAGCCTGCGAGACCTCATGAACCACCTCAAGAAGCAGATCCTGCGCACGGACAACATAAGCTTTGTGCTGAAACGATGCTGTCAGCCTAAG
513▶ H F P S L R D L M N H L K K Q I L R T D N I S F V L K R C C Q P K

XhoI (2201) **BamHI (2288)**
2201 CCTCGAGAAATCTCCAATCTGCTCGTAGCCACTAAGAAAGCCAGGAGTGGCAGCCTGTCTACTCCATGAGCCAGCTGAGCTTTGATCGGATCCTTAAGA
547▶ P R E I S N L L V A T K K A Q E W Q P V Y S M S Q L S F D R I L K
2301 AAGATATTATAAAGGTGAGCACCTTGGCAGAGGCACAAGAACACATATCTATTCTGGACCTGTGGACTACAAGGATGAGGAAGGAATGCTGAAGA
580▶ K D I I Q G E H L G R G T R T H I Y S G T L L D Y K D E E G I A E E

NheI (2473)
2401 GAAGAAGATAAAAAGTATCCTCAAAGTCTAGACCCACGCCACCGGACATCTCTGGCCTTCTTTGAGGCTGCTAGCATGATGAGACAGGTTTCCCAC
613▶ K K I K V I L K V L D P S H R D I S L A F F E A A S M M R Q V S H
2501 AAACATATAGTGTACCTCTACGGCGTGTGTGTCGAGATGTGGAAAATATCATGGTGAAGAGTGTGGAGGGGGGGCGGTTGGATCTTTCATGCACC
647▶ K H I V Y L Y G V C V R D V E N I M V E E F V E G G P L D L F M H
2601 GGAAAAGTATGCGCTTACTACCCCTGGAAGTTCAAGGTTGCCAAACAGCTGGCCAGTGCCTGAGTTACTTGAAGATAAAGACCTGGTTTCATGAAA
680▶ R K S D A L T T P W K F K V A K Q L A S A L S Y L E D K D L V H G N

ApaLI (2703)
2701 TGTGTGCACTAAAAACCTCTTCTGGCCCGTGAGGGCATTGACAGTGACATTGGCCCGTTCATCAAGCTTAGTGACCTGGCATCCCAGTCTCTGTGCTG
713▶ V C T K N L L L A R E G I D S D I G P F I K L S D P G I P V S V L
2801 ACCAGGCAAGAGTGCATAGAGCGAATCCCCTGGATCGCTCCTGAGTGTGTTGAAGACTCCAAGAACCTGAGTGTGGCTGCTGACAAGTGGAGCTTTGGAA
747▶ T R Q E C I E R I P W I A P E C V E D S K N L S V A A D K W S F G

StuI (2989)
2901 CCACGCTCTGGGAAATCTGCTACAACGGAGAGATTCTCTCAAAGACAAGACCCTCATTGAGAAAGAGAGGTTTTATGAAAGCCGCTGCAGGCCTGTGAC
780▶ T T L W E I C Y N G E I P L K D K T L I E K E R F Y E S R C R P V T

NheI (3015)
3001 TCCATCTTGAAGGAGCTAGCTGACCTCATGACTCGTGCATGAACTATGACCCCAACCAGAGACCCTTCTTCCGAGCCATCATGAGGGACATTAACAAG
813▶ P S C K E L A D L M T R C M N Y D P N Q R P F F R A I M R D I N K
3101 CTGGAGGAGCAGAATCCAGACATTGTTTCAGAAAAGCAGCCAAACAACAGAGGTGGACCCCACTCATTGAAAAGCGGTTCTGAAGAGGATTCGTGACT
847▶ L E E Q N P D I V S E K Q P T T E V D P T H F E K R F L K R I R D

Bsu36I (3244)
3201 TGGGAGAGGGTCACTTTGGGAAGTTGAGCTCTGCAGATATGATCCTGAGGGAGACAACAGGGGAGCAGGTAGCTGTCAAGTCCCTGAAGCCTGAGAG
880▶ L G E G H F G K V E L C R Y D P E G D N T G E Q V A V K S L K P E S

BstEII (3304) BglII (3338) BsaBI (3333)
3301 TGGAGGTAACCACATAGCTGATCTGAAGAAGGAGATAGAGATCTTACGGAACCTCTACCATGAGAACATTGTGAAGTACAAGGAATCTGCATGGAAGAC
913▶ G G N H I A D L K K E I E I L R N L Y H E N I V K Y K G I C M E D
3401 GGAGGCAATGGTATCAAGCTCATCATGGAGTTTCTGCCTTCGGGAAGCCTAAAGGAGTATCTGCCAAAGAATAAGAACAAAATCAACCTCAAACAGCAGC
947▶ G G N G I K L I M E F L P S G S L K E Y L P K N K N K I N L K Q Q
3501 TAAAATATGCCATCCAGATTTGTAAGGGGATGGACTACTTGGGTTCTCGGCAATACGTTCCACGGGACTTAGCAGCAAGAAATGTCCTTGTGAGAGTGA
980▶ L K Y A I Q I C K G M D Y L G S R Q Y V H R D L A A R N V L V E S E

ScaI (3656)
3601 GCATCAAGTGAAGATCGGAGACTTTGGTTAAACCAAAGCAATTGAAACCGATAAGGAGTACTACACAGTCAAGGACGACCGGGACAGCCAGTGTCTGG
1013▶ H Q V K I G D F G L T K A I E T D K E Y Y T V K D D R D S P V F W

BspEI (3705)
3701 TACGCTCCGGAATGTTTAAATCCAGTGTAATTTTATATCGCCTCTGATGTCTGGTCTTTTGGAGTGACACTGCACGAGCTGCTCACTTACTGTGACTCAG
1047▶ Y A P E C L I Q C K F Y I A S D V W S F G V T L H E L L T Y C D S

NcoI (3809)
3801 ATTTTAGTCCCATGGCCTTGTCTGAAAATGATAGGCCCAACTCATGGCCAGATGACAGTGACACGGCTTGTGAATACTCTGAAAGAAGGAAAGCGTCT
1080▶ D F S P M A L F L K M I G P T H G Q M T V T R L V N T L K E G K R L

EcoRI (3955)
3901 GCCATGTCCACCAACTGTCCTGATGAGGTTTATCAGCTTATGAGAAAATGCTGGGAATTCACCACTCTAACCGGACAACCTTTTCAAGAACCTTATTGAA
1113▶ P C P P N C P D E V Y Q L M R K C W E F Q P S N R T T F Q N L I E

XbaI (4049)
4001 GGATTTGAAGCACTTTTAAAATAAGAAGCATGAACAACATTTAAATTCCTCTAGACTAGCTGGCCAGACATGATAAGATAACATTGATGAGTTTGGACAAA
1147▶ G F E A L L K •

HpaI (4192)
4101 CCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAA

EcoRI (4288)
4201 CAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAATGTGGTATGGAATTCATAAAT
4301 ACAGCATAGCAAAACTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCA

SapI (4470)
4401 TTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGATTTTTCCCAAGGTTTGAAGTAGCTCTTCAATTTCTTTATGTTTTAAATGCAC

SspI (4527)
4501 TGACCTCCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCT
4601 CAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTA
141▶ •

4701 GTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAG
140▶ 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 F C D P A Y D S
4801 ATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCT
106▶ 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 V A V I T D F D K

StuI (4966)
4901 TCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCC
73▶ 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 H V A S I I E G
5001 AGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGA
40▶ 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 E V L E L D Q Q

XmnI (5108) AseI (5174)
5101 GAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTGTATTACTATGCGGATATACTATGCGGATGATTAATTGTCAAACACAGCGTGGATGG
6▶ S I N F T K M
5201 CGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATGGGGCGGAGTTGT

5301 TACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTA
SpeI (5329) ←

5401 TCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATG
SnaBI (5457)

5501 TACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTT
NdeI (5562)

5601 ACCGTAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCTGTTGGG

5701 CGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAAITTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
SdaI (5740) PacI (5748) BspLU11I (5758) ←

5801 AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA

5901 AAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGC

6001 GTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGACC
ApaLI (6072)

6101 GCTGCGCCTTATCCGGTAACTATCGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC

6201 GTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC

6301 TTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAG

6401 GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATT
PacI (6488)

EagI (6508)
NotI (6507)

6501 TAAATCAGCGGCCGAATAAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAA
6601 CGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA