



PvuI (7)
SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGTGCCTA
MfeI (82)
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGTCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACACGCCGCCGCCCTACCTGAGGCC
PvuII (239)
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
Bsu36I (291)

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGTTGCTCAACTCTACGTCTTTGTTTCGTTT

BspHI (560)
AgeI (552) 501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTCATCATGATGGATATTTACGTATGTTTGAACGACCATCCTG
1 M M D I Y V C L K R P S W
601 GATGGTGGACAATAAAGAATGAGGACTGCTTCAAATTTCCAGTGGCTGTTATCAACATTTATCTTCTATATCTAATGAATCAAGTAAATAGCCAGAAA
13 M V D N K R M R T A S N F Q W L L S T F I L L Y L M N Q V N S Q K
BspHI (711)
701 AAGGGGCTCCTCATGATTTGAAGTGTGTAACAACTTCAAGTGTGAACTGTTCTTGAAAGCACCTCTGGAACAGGCCGTGGTACTGATTATG
47 K G A P H D L K C V T N N L Q V W N C S W K A P S G T G R G T D Y
DraIII (872)
801 AAGTTTGATTGAAAACAGTCCCGTTCTGTTATCAGTTGGAGAAAACAGTATTAATCCAGCTCTTTCCATGGTATTGAAATAACAATAAA
80 E V C I E N R S R S C Y Q L E K T S I K I P A L S H G D Y E I T I N
Psp1406I (952) **BglII (980)**
901 TTCTCTACATGATTTTGAAGTTCTACAAGTAAATTCACACTAAATGAACAAAACGTTTCTTAATCCAGATACTCCAGAGATCTTGAATTTGTCTGCT
113 S L H D F G S S T S K F T L N E Q N V S L I P D T P E I L N L S A
BsrBI (1059)
1001 GATTTCTCAACCTCTACATTATACCTAAAGTGAACGACAGGGGTTGAGTTTTCCACACCGCTCAAATGTTATCTGGGAAATTAAGTTCTACGTAAG
147 D F S T S T L Y L K W N D R G S V F P H R S N V I W E I K V L R K
1101 AGAGTATGGAGCTCGTAAATAGTGACCCACAACACAACCTCTGAATGGCAAAGATACACTTCATCACTGGAGTTGGGCTCAGATATGCCCTTGAATG
180 E S M E L V K L V T H N T T L N G K D T L H H W S W A S D M P L E C
XcmI (1202)
1201 TGCCATTCATTTTGTGGAATAGATGCTACATTGACAATCTTCATTTTTCTGGTCTCGAAGAGTGGAGTGACTGGAGCCTGTGAAGAACATTTCTGG
213 A I H F V E I R C Y I D N L H F S G L E E W S D W S P V K N I S W
1301 ATACCTGATTTCTCAGACTAAGTTTTCTCAAGATAAGTGAATCTGTAGGCTCAGACATAACATTTGTTGTGTGAGTCAAGAAAAAGTGTATCAG
247 I P D S Q T K V F P Q D K V I L V G S D I T F C C V S Q E K V L S
MscI (1407) **BstXI (1431)**
1401 CACTGATGGCCATACAACTGCCCTTGTATCCATCTTGTGAGGAAAATGTTGCAATCAAGATTCGTAATATTTCTGTTTCTGCAAGTAGTGAACAAA
280 A L I G H T N C P L I H L D G E N V A I K I R N I S V S A S S G T N
EcoRV (1550)
1501 TGTAGTTTTTACAACCGAAGATAACATATTTGGAACCGTTATTTTGTGTTGATATCCACAGATACTCCTCAACAACGAATTTGTGAGACACATGATTTA
313 V V F T T E D N I F G T V I F A G Y P P D T P Q Q L N C E T H D L
BbrPI (1652)
Bsp120I (1647)
1601 AAAGAAATATATGTAGTTGAAATCCAGGAAGGGTGACAGCGTGGTGGGCCACGCTACAAGCTACACTTTAGTTGAAAGTTTTTTCAGGAAAATATG
347 K E I I C S W N P G R V T A L V G P R A T S Y T L V E S F S G K Y
1701 TTAGACTTAAAAGAGCTGAAGCACCTACAAACGAAAGCTATCAATTTATTTCAAATGTTCCAAATCAAGAAATATATAATTTTACTTTGAATGCTCA
380 V R L K R A E A P T N E S Y Q L L F Q M L P N Q E I Y N F T L N A H
BsaBI (1814)
1801 CAATCCGCTGGGTCGATCACAATCAACAATTTTGTGTAATATAACTGAAAAGTTTATCCCATACTCTACTTCATTCAAAGTGAAGGATATTAATTC
413 N P L G R S Q S T I L V N I T E K V Y P H T P T S F K V K D I N S
PvuII (1901)
1901 ACAGCTGTTAAACTTTCTGGCATTACCAGGCAACTTTGCAAAGATTAATTTTTTATGTGAAATGAAATTAAGAAATCTAATTCAGTACAAGAGCAGC
447 T A V K L S W H L P G N F A K I N F L C E I E I K K S N S V Q E Q
2001 GGAATGTCACAATCAAAGGAGTAGAAAATCAAGTTATCTTGTGCTCTGGACAAGTTAAATCCATACACTCTATATACTTTTTCGATTGCTGTTCTAC
480 R N V T I K G V E N S S Y L V A L D K L N P Y T L Y T F R I R C S T
2101 TGAACCTTTCTGGAATGGAGCAATGGAGCAATAAAAACAACATTTAAACAACAAGCCAGTCTTCAAAGGGGCCTGATACTTGGAGAGAGTGGAGT
513 E T F W K W S K W S N K K Q H L T T E A S P S K G P D T W R E W S
2201 TCTGATGAAAAATTAATAATCTATTGGAAGCCTTTACCCTAATGAAGCTAAATGGAAAAATACCTTCTCAATGTATCTGTTTCAATCAGATGAGG
547 S D G K N L I I Y W K P L P I N E A N G K I L S Y N V S C S S D E
2301 AAACACAGTCCCTTTCTGAAATCCCTGATCCTCAGCACAAAGCAGAGATACGACTTGATAAGAATGACTACATCATCAGCGTAGTGGCTAAAAATCTGT
580 E T Q S L S E I P D P Q H K A E I R L D K N D Y I I S V V A K N S V
2401 GGGCTCATACCACCTTCCAAAATAGCGAGTATGAAATTTCAAATGATGATCTCAAATAGAACAAGTTGTTGGGATGGGAAAGGGGATTCTCCTCACC
613 G S S P P S K I A S M E I P N D D L K I E Q V V G M G K G I L L T
Tth111I (2527)
2501 TGGCATTACGCCCAACATGACTTGGACTACGTCAATTAAGTGGTGAACCTGCTCGGTCGGAACCATGCCTTATGGACTGGAGAAAAGTTCCCTCAA
647 W H Y D P N M T C D Y V I K W C N S S R S E P C L M D W R K V P S
EcoRV (2684)
2601 ACAGCACTGAAACTGTAATAGAATCTGATGAGTTTCGACCAGGTATAAGATATAATTTTTTCTGATGGATGCAGAAATCAAGGATATCAATTTATTACG
680 N S T E T V I E S D E F R P G I R Y N F F L Y G C R N Q G Y Q L L R

2701 CTCCATGATTGGATATATAGAAGAATTGGCTCCATTGTTGCACCAAATTTACTGTTGAGGATACTTCTGCAGATTGATATTAGTAAAATGGGAAGAC
713▶ S M I G Y I E E L A P I V A P N F T V E D T S A D S I L V K W E D
2801 ATTCCTGTGGAAGAACTTAGAGGCTTTTTAAGAGGATATTTGTTTTACTTTGGAAAAGGAGAAAAGAGACACATCTAAGATGAGGGTTTTAGAATCAGGTC
747▶ I P V E E L R G F L R G Y L F Y F G K G E R D T S K M R V L E S G
2901 GTTCTGACATAAAAGTTAAGAATATTACTGACATATCCAGAAGACACTGAGAATTGCTGATCTTCAAGGTAAAACAAGTTACCACCTGGTCTTGCAGC
780▶ R S D I K V K N I T D I S Q K T L R I A D L Q G K T S Y H L V L R A

Bsp120I (3018)

3001 CTATACAGATGGTGGAGTGGCCCGGAGAAAGATATGTATGTGGTGACAAAGGAAAATCTGTGGGATTAATTATTGCCATTCTCATCCCAGTGGCAGTG
813▶ Y T D G G V G P E K S M Y V V T K E N S V G L I I A I L I P V A V
3101 GCTGTCATTGTTGGAGTGGTGACAAGTATCCTTTGCTATCGGAAACGAGAATGGATTAAGAAACCTTCTACCCTGATATTCAAATCCAGAAAACCTGTA
847▶ A V I V G V V T S I L C Y R K R E W I K E T F Y P D I P N P E N C
3201 AAGCATTACAGTTTCAAAAAGTGTCTGTGAGGGAAGCAGTGTCTTAAAACATTGGAATGAATCCTTGTACCCCAATAATGTTGAGGTTCTGGAAAC
880▶ K A L Q F Q K S V C E G S A L K T L E M N P C T P N N V E V L E T
3301 TCGATCAGCATTCTCTAAAATAGAAGATACAGAATAATTTCCCCAGTAGCTGAGCGTCTGAAGATCGCTCTGATGCAGAGCCTGAAAACCATGTGGTT
913▶ R S A F P K I E D T E I I S P V A E R P E D R S D A E P E N H V V
3401 GTGTCCTATTGTCCACCCATCATTGAGGAAGAAATACCAAACCCAGCCGAGATGAAGCTGGAGGGACTGCACAGGTTATTTACATTGATGTTGAGTCGA
947▶ V S Y C P P I I E E E I P N P A A D E A G G T A Q V I Y I D V Q S
3501 TGTATCAGCCTCAAGCAAAACGAGAAGAAACAAGAAAATGACCCTGTAGGAGGGGAGCTATAAGCCACAGATGCACCTCCCATTAATTACTACTGT
980▶ M Y Q P Q A K P E E E Q E N D P V G G A G Y K P Q M H L P I N S T V

Bsu36I (3649)

3601 GGAAGATATAGTGCAGAAGAGGACTTAGATAAAAACGCGGTTACAGACCTCAGGCAATGTAATACATGGAATTTAGTGTCTCCAGACTCTCCTAGA
1013▶ E D I A A E E D L D K T A G Y R P Q A N V N T W N L V S P D S P R
3701 TCCATAGACAGCAACAGTGAAGTGTCTCATTGGAAGTCCATGCTCCATTAATTTCCGACAATTTTTGATTCTCTAAAGATGAAGACTCTCCTAAAT
1047▶ S I D S N S E I V S F G S P C S I N S R Q F L I P P K D E D S P K

NheI (3888)

XcmI (3882) MscI (3894)

3801 CTAATGGAGGAGGGTGGTCTTTACAACTTTTTTCAAAACAAACCAACGATTAACAGTGTACCCGTGTCACTTCAGTCAGCCATCTGCTAGCTGGCCA
1080▶ S N G G G W S F T N F F Q N K P N D •
3901 GACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG

HpaI (4026) MfeI (4037)

4001 TAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGGGGAGGTTTTTTAAAGCAAGTA

EcoRI (4122)

4101 AAACCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATA
4201 AGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGA

SapI (4304)

SwaI (4375)

4301 ACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTGTAAAATATTCAGAAAATATTTAAATACATCATTGCAATGAAA

4401 ATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGA

4501 AATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTC
141▶ • N R T Y K L P I L E E I T T K V L K G N M E

BstXI (4665)

4601 AATGAGCACAAGCAGTCCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAG
118▶ 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 E D S Y
4701 GGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTTCAGCAGACAGTGCACCTGCCAATGT
84▶ 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 R G I Y

StuI (4800)

4801 AGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTCTGTTGTCTCATAGAGCATGGTATCTTCTC
51▶ 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 T I K E

XmnI (4942)

4901 AGTGGCGACTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTA
18▶ T A V E V L E L D Q Q S I N F T K M
5001 TGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCTACTAAACGAGCTCTGCTTATATAGACTCCACCGTACACGC

SpeI (5163)

5101 CTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAAGTCCCCTTGGATTTACTAGTCAAAAACAACTCCCATGACGTCAATGGGGT
5201 GGAGACTTGAAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGAT

NdeI (5396)

5301 GTAAGTCAAGTAGGAAAGTCCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATA

5401 TGATACACTTGATGTAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTTACTATGGGAACATACG

PaeI (5582)

SdaI (5574)

BspLU11I (5592)

5501 TCATTATTGACGTCAATGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACAATGTGA
5601 GCAAAAAGCCAGCAAAAGCCAGGAACCGTAAAAGGCCGCTGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCT

5701 CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTAC

5801 CGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTG

ApaLI (5906)

5901 GGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG

6001 CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGT

6101 ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTT

6201 GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT

EagI (6342)

PacI (6322) SmaI (6331) **NotI (6341)**

6301 AAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAA

6401 TCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCT

6501 ATCGAA