



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82) 101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
EcoNI (96)

HindIII (245) 201 **Psp1406I (203)** GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC
PvuII (239) 301 **Bsu36I (291)** GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
EcoNI (287)

NgoMIV (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BstEII (555) 501 **AgeI (552)** **NcoI (560)** TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTACCATTGGTAAAGTCCCTCTGGGATAGCCTCTGCAACATGAG
1 M V K S L W D S L C N M S
601 CCAAGACAGAAAACCCATCGTGGGGAGTTCCACTTTGTTTGGCCCTGGCCTAATAGTCGGAAGCATGACCCCGTTCTCTAATGAACCTGAGTCTATG
13 Q D R K P I V G S F H F V C A L A L I V G S M T P F S N E L E S M
701 GTAGACTATTCAACAGGAACCTTACTCATGTCCCAAAGACCTGCCACCAAGAACAAAAGCCCTGAGTCTGTCTCAAAACTATATCTGAGCTTCGGA
47 V D Y S N R N L T H V P K D L P P R T K A L S L S Q N S I S E L R
801 **EcoRV (805)** TGCCTGATATCAGCTTCTGTGAGAGCTGAGAGTTCTGAGACTCTCCACAACAGGATACGGAGCCTTGATTTCCATGTATTCTTGTTCATCAGGACTT
80 M P D I S F L S E L R V L R L S H N R I R S L D F H V F L F N Q D L
Bsu36I (960) **XbaI (969)**
901 AGAATACCTGGATGTCTCACACAATCGGTTGCAAAACATCTCTTGTGTCCTATGGCGAGCCTGAGGCATCTAGACCTCTCATTCAATGACTTTGATGTA
113 E Y L D V S H N R L Q N I S C C P M A S L R H L D L S F N D F D V
1001 CTGCTGTGTGAAGGAATTTGGCAACCTGACGAAGCTGACTTTCCTGGGATTAAGTGTGCAAAAGTCCGACAACCTGGATCTGCTCCAGTTGCTCACT
147 L P V C K E F G N L T K L T F L G L S A A K F R Q L D L L P V A H
1101 TGCATCTAAGCTGCATTCTTCTGGACTAGTGAGTTATCATATAAAAGCGGGGAAACAGAAAGTCTTCAAGATCCCAATACCACCGTTCTCCATTTGTT
180 L H L S C I L L D L V S Y H I K G G E T E S L Q I P N T T V L H L V
1201 CTTTCATCCAAATAGCTTGTCTCTGTTCAAGTGAACATGTCTGTAAACGCTTTAGGACATTTACAACCTGAGTAAATTTAAATGAATGATGAAAAGTGT
213 F H P N S L F S V Q V N M S V N A L G H L Q L S N I K L N D E N C
1301 CAAAGGTTAATGACATTTTTATCAGAACTCACCAGAGGTTCAACCTTATTGAATGTGACCTCCAGCAGACATAGAAAACCACTGGAAAGTCTCGGTTAAAC
247 Q R L M T F L S E L T R G P T L L N V T L Q H I E T T W K C S V K
AgeI (1422)
1401 TTTTCCAATTTCTTTGGCCCCGACCGGTGGAGTACCTCAATATTTACAACCTAACGATAACTGAGAGAATCGACAGGGAAGAATTTACTTACTCGGAGAC
280 L F Q F F W P R P V E Y L N I Y N L T I T E R I D R E E F T Y S E T
BsaBI (1584)
1501 AGCACTGAAGTCACTGATGATAGAGCAGCTCAAAAACCAAGTGTCTCTTTTCAAAGGAGCGCTATACTCGGTGTTTGTGAGATGAACATCAAGATG
313 A L K S L M I E H V K N Q V F L F S K E A L Y S V F A E M N I K M
DraIII (1627)
1601 CTCTCTATCTCAGACACCCCTTTTCATCCACATGGTGTGCCCGCCATCCCAAGCTCATTACATTTCTGAACTTTACCCAGAATGTTTTTACTGACAGTG
347 L S I S D T P F I H M V C P P S P S S F T F L N F T Q N V F T D S
BspHI (1784)
1701 TTTTTCAAGGCTGTTCCACCTTAAAGAGATTGCAGACCTTATCTTACAAAGGAATGGTTTGAAGAATTTTTTAAAGTAGCTCTCATGACTAAGAATAT
380 V F Q G C S T L K R L Q T L I L Q R N G L K N F F K V A L M T K N M
NdeI (1850)
EcoRI (1830) **NsiI (1847)**
1801 GTCCTCTCTGAAACTTTGGATGTTAGTTTGAATTCTTTGAACTCTCATGCATATGACAGGACATGCGCCTGGGCTGAGAGCATATTGGTGTGAAATTTG
413 S S L E T L D V S L N S L N S H A Y D R T C A W A E S I L V L N L
BstBI (1902)
1901 TCTTCGAATATGCTTACAGGCTCTGTCTTCCAGATGCTTACCTCCCAAGGTCAAGGTCCTTACCTTACCAACAACAGGATAATGAGCATCCCTAAAGATG
447 S S N M L T G S V F R C L P P K V K V L D L H N N R I M S I P K D
BstAPI (2010)
PstI (2008)
SdaI (2007)
2001 TCACCCACCTGCAGGCTTTGAGGAACTCAATGTAGCATCCAACCTCCTTAACTGACCTTCTGGGTGTGGGGCCTTACGACGCTTTTCTGTGCTGGTCAT
480 V T H L Q A L Q E L N V A S N S L T D L P G C G A F S S L S V L V I
2101 CGACCATAACTCAGTTCCCATCCCTCTGAGGATTTCTTCCAGAGCTGTGAGAATATTAGATCCCTAACAGCGGGAAACCAACCCATTCCAATGCATGT
513 D H N S V S H P S E D F F Q S C Q N I R S L T A G N N P F Q C T C
Tth111I (2209)
2201 GAGCTGAGGGACTTTGTCAAGAACATAGGCTGGGTAGCAAGAGAAGTGGTGGAGGGCTGGCCTGACTCTTACAGGTGTGACTACCCAGAAAGCTTAAGG
547 E L R D F V K N I G W V A R E V V E G W P D S Y R C D Y P E S S K
BstXI (2372)
2301 GAACTGCACTGAGGGACTTCCACATGTCTCCACTGTCTGTGATACTGTTCTGCTGACTGTCCACATCGGGGCCACTATGCTGGTGTGGCTGTCACTGG
580 G T A L R D F H M S P L S C D T V L L T V T I G A T M L V L A V T G
2401 GGCTTTCCCTGTCTCTACTTTGACCTGCCCTGGTATGTGAGGATGCTGTGTCAGTGGACACAGACCAGGCACAGGGCCAGGCACATCCCTTAGAGGAA
613 A F L C L Y F D L P W Y V R M L C Q W T Q T R H R A R H I P L E E
2501 CTCCAGAGAACCTCAGTTCCATGCTTTGTCTCATACAGTGAAGCATGATTCTGCGGGTGAAGAACGAATTAACCCCAACCTAGAGAAAGATGACA
647 L Q R N L Q F H A F V S Y S E H D S A W V K N E L L P N L E K D D
2601 TCCGGGTTTGCCTCCATGAGAGGAACTTTGTCCTGGCAAGAGCATTGTGGAGAACATCATCAATTTTATTGAGAAGAGTTACAAGGCCATCTTTGTGCT
680 I R V C L H E R N F V P G K S I V E N I I N F I E K S Y K A I F V L

2701 GTCTCCCCACTTCATCCAGAGTGAGTGGTGCCATTATGAACTCTATTTTGCCATCATAATCTCTCCATGAAGGCTCTGATAACTTAATCCTCATCTTG
713▶ S P H F I Q S E W C H Y E L Y F A H H N L F H E G S D N L I L I L
2801 CTGGAACCCATTCTACAGAACCAACATCCAGTAGATACCACAAGCTGCGGGCTCTCATGGCACAGCGGACTTACTTGGAAATGGCCTACTGAGAAGGGCA
747▶ L E P I L Q N N I P S R Y H K L R A L M A Q R T Y L E W P T E K G
2901 AACGTGGGCTGTTTTGGGCCAACCTTAGAGCTTCATTTATTATGAAGTAGCCTTAGTCAATGAGGATGATGTAAAACCTGAAAACCTGGGTTCTAACT
780▶ K R G L F W A N L R A S F I M K L A L V N E D D V K T •

MscI (3030)
NheI (3024)

3001 TAATAAACTGTCAACCTGGGCTCTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCT

HpaI (3162) MfeI (3173)

3101 TTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAACAACATTCATTCTTTATGTTTCAGGT

EcoRI (3258)

3201 TCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTTAACCTCAAATC
3301 AAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCAT

SapI (3440)

3401 GGAGTTTAAGATATAGTGTATTTTCCCAAGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCTTTTTAGTAAAT

SwaI (3511)

3501 ATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAG
3601 TAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGGATGAGTTC
141◀ • N R T Y K L P I L E

SacI (3772)

3701 CTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTG
130◀ E I T T K V L K 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

BstXI (3801)

3801 ACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGG
96◀ V V R I S R D V 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

StuI (3936)

3901 CAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATG
63◀ I A E A C V T 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
4001 GTGCTTGTTCCTCATAGACATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTC
30◀ H K N D E Y L M T I K E T A V E V L E L D Q Q S I N F T K M ◀

AseI (4144)

4101 CTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAA

SacI (4201) SpeI (4299)

4201 CGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTA
4301 CTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCG

SnaBI (4427)

4401 CATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTTA

NdeI (4532)

4501 CCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTGGCAGTTTACCGTAAATACTCCACCATTGACGTCAAT
4601 GGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGGGGGGTCTTGGCGGTCAGCCAGGCGGGCCATTTACCGTAAG

PstI (4711) SdaI (4710) PacI (4718)

4701 TTATGTAACGCTGAGGTTAAITAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCTTGTGGCGTTTTTCCATAGGC
4801 TCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTC
4901 CCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGG

ApaLI (5042)

5001 TATCTCAGTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
5101 AGTCCAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAA
5201 GTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCC
5301 GGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA

EagI (5478) NotI (5477)
PacI (5458) SwaI (5467)

5401 CGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTA
5501 TTTTCATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTG
5601 TCCCAGTCAAGTGCAGGTGCCAGAACATTTCTATCGAA