



PvuI (7)
SgfI (6) **EcoNI (96)**
1 GGATCTGCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGGGGTAACGGGAAAGTGTGCTGTACTGGCTCCGCCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Bsu36I (291)
Psp1406I (203) **PvuII (239)** **EcoNI (287)**
201 GTGAACGTTCTTTTTGCGAACGGGTTTCCGCCAGAACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCCACGGCCCGCCCTACCTGAGGGC
301 GCCATCCACGCCGGTTGAGTCGCGTTCGCCGCCCTCCCGCTGTGGTGCCTCTGAAGTCCGCTCCGCCCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMI (441)
NaeI (441)
401 GGGCCTTTGTCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCTCCACGCTTTCGCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

AgeI (552)
501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGTAGGAGGCCATCATGACCAAAGACAAAGAACCTATTGTAA
601 AGCTTCCATTTTGTTCCTTATGATCATAATAGTTGGAACCAAGTCCAGTTCCTCCGACGAAATGAATTTGCAGTAGACAAGTCAAAAAGAGGTCCTA
110 S F H F V C L M I I I V G T R I Q F S D G N E F A V D K S K R G L
701 TTCATGTTCCAAAAGACCTACCGCTGAAAACCAAGTCTTAGATATGTCTCAGAACTACATCGCTGAGCTCAGGCTCTGACATGAGCTTTCTATCAGA
44 I H V P K D L P L K T K V L D M S Q N Y I A E L Q V S D M S F L S E
801 GTTGACAGTTTTGAGACTTTCCATAACAGAATCCAGCTACTTGAATTAAGTGTTCGAAATTCACCAAGGATTAGAAATTTGGATTATCTCATAAT
77 L T V L R L S H N R I Q L L D L S V F K F N Q D L E Y L D L S H N

EcoRV (911) **BglII (947)** **EcoO109I (970)**
901 CAGTTGCAAAAAGATCTGCGCATCTATTGTGAGTTCAGGCATTTAGATCTCTCATTCAATGATTTCAAGGCCCTGCCATCTGAAGGAATTTGGCA
111 Q L Q K I S C H P I V S F R H L D L S F N D F K A L P I C K E F G
1001 ACTTATCACAAGTAAATTTCTGGGATTGAGTGTCTGAAGTGCAAAATTTAGATTTGCTGCAATTTGCTCACTTGCATCTAAGTTATATCCTTCTGGGA
144 N L S Q L N F L G L S A M K L Q K L D L L P I A H L H L S Y I L L D
1101 TTTAAGAAATTTATATAAAAAGAAAATGAGACAGAAAGTCTACAAATTTCTGAATGCAAAAACCTTACCTTGTTCACCCAAGTATTTATTCGCT
177 L R N Y Y I K E N E T E S L Q I L N A K T L H L V F H P T S L F A

PshAI (1266)
1201 ATCCAAGTGAACATATCAGTTAATACTTTAGGGTCTTACAAGTACTGACTAATATAAATGAATGATGACAAGTCTCAAGTTTTCATTAATTTTATCAG
211 I Q V N I S V N T L G C L Q L T N I K L N D D N C Q V F I K F L S
1301 AACTCACCAGAGTTCAACCTTACTGAATTTTACCCTCAACCACATAGAAAAGCAGCTTGGAAATGCCTGGTCAGAGTCTTTCAATTTCTTTGGCCAAACC
244 E L T R G S T L L N F T L N H I E T T W K C L V R V F Q F L W P K P
1401 TGTGGAATATCTCAATTTTACAATTTAACAATAATTGAAAGCATTCTGTAAGAAAGATTTTACTTATCTAAAACGACATTGAAAGCATTGACAAATAGAA
277 V E Y L N I Y N L T I E S I R E E D F T Y S K T T L K A L T I E

BsrGI (1540) **HpaI (1574)**
1501 CATATCACGAACCAAGTTTTTCTGTTTTACAGACAGCTTTGTACACCGTGTTCGAGATGAACATTATGATGTTAACCATTTTCAGATACACCTTTTA
311 H I T N Q V F L F S Q T A L Y T V F S E M N I M M L T I S D T P F

Psp1406I (1656)
1601 TACACATGCTGTCTCATGCACCAAGCAGTCAAGTTTTGAACTTTACCAGAACGTTTTACAGATAGTATTTTGAATAATGTTCCAGGTTAGT
344 I H M L C P H A P S T F K F L N F T Q N V F T D S I F E K C S T L V
1701 TAAATGGAGACACTTATCTTACAAAAGATGGATTAAGAACCTTTCAAAGTAGGCTCATGACGAAGGATATGCCTTCTTTGAAATGACTGGATGTT
377 K L E T L I L Q K N G L K D L F K V G L M T K D M P S L E I L D V

EcoRI (1805)
1801 AGCTGGAATCTTTGGAATCTGGTAGACATAAAGAAAAGTGCAGTTCGGGTTGAGAGTATAGTGGTGTAAATTTGTCTTCAAATGCTTACTGACTCTG
411 S W N S L E S G R H K E N C T W V E S I V V L N L S S N M L T D S
1901 TTTTCAGATGTTTACCTCCAGGATCAAGTACTTGTATCTTACAGCAATAAAAATAAGAGCGTTCCTAAACAAGTGTAAAAGTGGAAAGCTTTGCAAGA
444 V F R C L P P R I K V L D L H S N K I K S V P K Q V V K L E A L Q E
2001 ACTCAATGTTGCTTTCAATCTTTAAGTACTGACTTCTGGATGTCGAGCTTTAGCAGCCTTTCTGTATTGATCATTGATCACAATTCAGTTCCTCCACCA
477 L N V A T L N F G S L T D L P G C G S F S S L S V L I D H N S V S H P
2101 TCGGCTGATTTCTCCAGAGCTGCCAAGATGAGGTCAATAAAGAGCAGGGGCAATCCATTCCAATGTACCTGTGAGCTAAGAGAATTTGTCAAAAATA
511 S A D F F Q S C Q K M R S I K A G D N P F Q C T C E L R E F V K N

PciI (2296)
BspLU11I (2296)
2201 TAGACCAAGTATCAAGTGAAGTGTAGAGGGCTGGCCTGATCTTATAAGTGTGACTACCCAGAAAAGTTATAGAGGAAGCCCACTAAAGGACTTTACAT
544 I D Q V S S E V L E G W P D S Y K C D Y P E S Y R G S P L K D F H M

XcmI (2350)
2301 GTCTGAATTATCCTGCAACATAACTCTGCTGATCGTACCATCGGTGCCACCATGCTGGTGTGGCTGTGACTGTGACCTCCCTCTGCATCTACTGGAT
577 S E L S C N I T L L I V T I G A T M L V L A V T V T S L C I Y L D
2401 CTGGCCTGGTATCTCAGGATGGTGTGCGAGTGGACCCAGACTCGGCGCAGGGCCAGGAACATACCTTAGAAGAACTCCAAGAAAACCTCCAGTTTCATG
611 L P W Y L R M V C Q W T Q T R R R A R N I P L E E L Q R N L Q F H

Asp718 (2549)
Acc65I (2549)
2501 CTTTTATTTTCATATAGTGAACATGATTTCTGCCTGGGTGAAAAGTGAATTTGGTACCTTACCTAGAAAAGAAGATATACAGATTTGTCTTCATGAGAGAAA
644 A F I S Y S E H D S A W V K S E L V P Y L E K E D I Q I C L H E R N
2601 CTTTGTCCCTGGCAAGAGCATTGTGGAATAATCATCAACTGCATTGAGAAGAGTTACAAGTCCATCTTTGTTTTGTCTCCCACTTTGTCCAGAGTGAG
677 F V P G K S I V E N I N C I E K S Y K S I F V L S P N F V Q S E
2701 TGGTGCCATTACGAECTATTTTCCCATCACAATCTCTTTCATGAAGGCTAATAACTTAATCCTCATCTTACTGGAACCCATTCCACAGAACAGCA
711 W C H Y E L Y F A H H N L F H E G S N N L I L I L L E P I P Q N S
2801 TTCCAACAAGTACCACAAGTGAAGGCTCATGACGCGAGGACTTATTTGCAAGTGGCCCAAGGAGAAAAGCAACGTGGGCTCTTTTGGGCTAACAT
744 I P N K Y H K L K A L M T Q R T Y L Q W P K E K S K R G L F W A N I

BamHI (2958) **NdeI (2998)**
2901 TAGAGCCGCTTTTAAATATGAAATTAACACTAGTCACTGAAAACAATGATGTGAAATCTGGATCCCTATCCCTATGATGTGCCAGACTATGCTGGCATCCCA
777 R A A F N M K L T L V T E N N D V K S G S Y P Y D V P D Y A G Y P

MscI (3061)

3001 TATGATGTTCTGATTATGCTGGATACCCCTTATGATGTGCCAGACATATGCCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAA
811 ▶ Y D V P D Y A G Y P Y D V P D Y A •

HpaI (3193)

3101 ACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTTAAACA

EcoRI (3289)

3201 ACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAAATCTAAAA

3301 TACAGCATAGCAAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGC

3401 ATTAGCTGTTTGCAGCCTCACCTTCTTTATGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAGCTAGCTCTTCTTTATGTTTAAATGCA

SwaI (3542)

3501 CTGACCTCCACATTCCCTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGC

EcoO109I (3603)

3601 TCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTT
141 ◀ •

3701 AGTTCCTGGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCACTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGA
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

SacI (3803)
EclI36II (3803)

3801 GATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCC
107 ◀ 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

StuI (3967)

3901 TTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAAGCAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCC
73 ◀ K 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

4001 CAGTCTTGGTCTGATGGCCGCCGACATGGTCTGTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCCTGCTG
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

AseI (4175)

4101 AGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATG
7 ◀ S I N F T K M

SacI (4232)
EclI36II (4232)

4201 GCGTCTCCAGCTTATCTGACGGTCACTAACGAGCTCTGCTTATATAGACCTCCACCCTACAGCCTACCGCCATTGCGTCAATGGGCGGAGTTG

4301 TTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCT

SnaBI (4458)

4401 ATCCAGCCCATGATGTAAGTCCAAAACCGCATCATGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAAAGTCCAT

NdeI (4563)

4501 GTACTGGGCATAATGCCAGCGGGCCATTACCGTCAATGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTT
4601 TACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGG

PacI (4749)

PstI (4742)
SdaI (4741)
SbfI (4741)

PciI (4759)
BspLU11I (4759)

4701 GCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCGGAACCGTAA
4801 AAGGCCGCTTGTGGGTTTTCCATAGGCTCCGCCCTGACGAGCATCAGCAAAATCGCGTCAAGTCAAGTGGCGAAACCCGACAGGACTAT
4901 AAAGATACCAGGCTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCGCTTTCTCCCTTGGGAA

ApaLI (5073)

5001 CGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTCGCTCCAAGCTGGGCTGTGTGACGAAACCCCGTTCAGCCCGAC
5101 CGCTGCGCTTATCCGGTAACTATCGTCTTGTGAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA
5201 GGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC
5301 CTTGCGAAAAAGATTGGTAGCTCTTGTATCCGGCAAAACACCCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAA

PacI (5489) SwaI (5498)

5401 GGATCTCAAGAAGTCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACAT

EagI (5509)
NotI (5508)

5501 TTAATCAGCGGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAA
5601 ACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA