



PvuI (7)
SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAAGTATGTCGTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGGCCCGCCCGCCCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCCGGTTGAGTCGGCTTTCGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT
NgoMIV (441)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGGACGCGTCACCATGGGTTTCTGCCGAGCGCCCTGCACCCGCT
MluI (558) **NcoI (566)**
1 M G F C R S A L H P L

601 GTCTCTCCTGGTGCAGGCCATCATGCTGGCCATGACCTGGCCCTGGGTACCTTGCTGCCTTCTACCTGTGAGCTCCAGCCCCACGGCCTGGTGAAC
Acc65I (646)
11 S L L V Q A I M L A M T L A L G T L P A F L P C E L Q P H G L V N
NcoI (737) **DraIII (746)**
701 TGCAACTGGCTGTTCTGAAGTCTGTGCCCCACTTCTCCATGGCAGCACCCCGTGGCAATGTCAACGCTTTCCTTGTCTCAACCGCATCCACCACC
45 C N W L F L K S V P H F S M A A P R G N V T S L S L S S N R I H H
801 TCCATGATTCTGACTTTGCCACCTGCCAGCCTGGCGCATCTCAACTCAAGTGGAACTGCCCGCGTTGGCTCAGCCCCATGCATTTCCCTGCCA
78 L H D S D F A H L P S L R H L N L K W N C P P V G L S P M H F P C H

901 CATGACCATCGAGCCAGCACCTTCTGGCTGTGCCACCCTGGAAGAGCTAAACCTGAGCTACAACAACATCATGACTGTGCTGCGCTGCCAAATCC
SapI (943) **BspHI (971)**
111 M T I E P S T F L A V P T L E E L N L S Y N N I M T V P A L P K S

1001 CTCATATCCCTGTCCCTCAGCCATACCAACATCCTGATGCTAGACTCTGCCAGCCTCGCCGGCCTGCATGCCCTGCGCTTCTATTATGAGCGGCAACT
NgoMIV (1057)
145 L I S L S L S H T N I L M L D S A S L A G L H A L R F L F M D G N

1101 GTTATTACAAGAACCCTGCAGGCAGGCACTGGAGGTGGCCCCGGTGCCTCCTGGCTGGGCAACCTCACCCACCTGCTCAAGTACAACAACCT
SdaI (1115) **XmaI (1140)**
178 C Y Y K N P C R Q A L E V A P G A L L G L G N L T H L S L K Y N N L

1201 CACTGTGGTGCCCGCAACCTGCCTTCCAGCCTGGAGTATCTGCTGTTGCTTACAACCGCATCGTCAAAGTGGCGCTGAGGACCTGGCCAATCTGACC
Bsu36I (1276)
211 T V V P R N L P S S L E Y L L L S Y N R I V K A L A P E D L A N L T
1301 GCCCTGCGTGTGCTCGATGTGGCGGAAATTGCCGCCCTGCGACCAGCTCCAACCCCTGCATGGAGTGCCTCGTCACTTCCCCAGCTACATCCCG
245 A L R V L D V G G N C R R C D H A P N P C M E C P R H F P Q L H P
1401 ATACCTTCAGCCACCTGAGCCGCTTGAAGGCTGGTGTGAAGGACAGTTCTCTCTGCTGATGCCAGTTGGTTCCGTGGGCTGGGAAACCTCCG
278 D T F S H L S R L E G L V L K D S S L S W L N A S W F R G L G N L R

1501 AGTGTGGACCTGAGTGAGAATTCCTCTACAAATGCATATAAACCAAGGCTTCCAGGGCCTAACACAGCTGCGCAAGCTTAACCTGTCCTTCAAT
NsiI (1533) **HindIII (1579)**
311 V L D L S E N F L Y K C I T K T K A F Q G L T Q L R K L N L S F N
1601 TACCAAAAGAGGGTGTCTTTGCCACCTGTCTTGGCCCTTCTTCCGGAGCCTGGTGCCTGAAAGAGCTGGACATGCACGGCATCTTCTCCGCT
345 Y Q K R V S F A H L S L A P S F G S L V A L K E L D M H G I F F R
1701 CACTCGATGAGACCAGCTCCGGCCACTGGCCGCTGCCATGCTCCAGACTCTGCGTCTGCAGATGAACTCATCAACCAGGCCAGCTCGGCATCTT
378 S L D E T T L R P L A R L P M L Q T L R L Q M N F I N Q A Q L G I F

1801 CAGGGCCTTCCCTGGCCTGCGCTACGTGGACCTGTCCGACAACCGCATCAGCGGAGCTTCGGAGCTGACAGCCACCATGGGGGAGGCAGATGGAGGGAG
NcoI (1874)
411 R A F P G L R Y V D L S D N R I S G A S E L T A T M G E A D G G E

1901 AAGGTCTGGTGCAGCCTGGGGACCTTGTCCGGCCCCAGTGGACACTCCCAGCTCTGAAGACTTCAGGCCAACTGCAGCACCTCAACTTCACCTTGG
BbsI (1957)
445 K V W L Q P G D L A P A P V D T P S S E D F R P N C S T L N F T L

2001 ATCTGTCACGGAACAACCTGGTGAACCTGACCGGAGATGTTTGGCCAGCTCTCGCACCTGCAGTGCCTGCGCCTGAGCCACAACCTGCATCTCGCAGGC
BstEII (2019)
478 D L S R N N L V T V Q P E M F A Q L S H L Q C L R L S H N C I S Q A

2101 AGTCAATGGCTCCAGTCTGCGCCTGACCGGTCTGCAGGTGCTAGACTGTCCCACAATAAGTGGACCTTACCACGAGCACTCATTACGGAGCTA
AgeI (2128)
511 V N G S Q F L P L T G L Q V L D L S H N K L D L Y H E H S F T E L
2201 CCAGACTGGAGCCCTGGACCTCAGTACAACAGCCAGCCCTTTGGCATGCAGGGCTGGGCCACAACCTCAGCTTCTGGCTCACCTGCGCACCCCTGC
545 P R L E A L D L S Y N S Q P F G M Q G V G H N F S F V A H L R T L
2301 GCCACCTCAGCCTGGCCACAACAACATCCACAGCCAAGTGTCCAGCAGCTGTCAGTACGTCGCTGCGGGCCCTGGACTTCAGCGGCAATGCAGTGGG
578 R H L S L A H N N I H S Q V S Q Q L C S T S L R A L D F S G N A L G

2401 CCATATGTGGCCGAGGGAGACCTCTATCTGCACCTTCTCCAAGGCTGAGCGGTTTGTCTGGTGGACTTGTCCAGAACCCTGCACACCCCTCTG
NdeI (2401)
611 H M W A E G D L Y L H F F Q G L S G L I W L D L S Q N R L H T L L
2501 CCCCAAACCTGCGCAACCTCCCAAGAGCCTACAGGTGCTGCGTCTCCGTGACAATTACCTGGCCTTCTTAAGTGGTGGAGCCTCACTTCTGCCCA
645 P Q T L R N L P K S L Q V L R L R D N Y L A F F K W W S L H F L P

BspEI (2675)

2601 AACTGGAAGTCCCTCGACCTGGCAGGAAACCAGCTGAAGGCCCTGACCAATGGCAGCCTGCCTGCTGGCACCCGGCTCCGGAGGCTGGATGTCAGCTGCAA
678▶ K L E V L D L A G N Q L K A L T N G S L P A G T R L R R L D V S C N
2701 CAGCATCAGCTTCGTGGCCCCGGCTTCTTTTCAAGGCCAAGGAGCTGCAGAGCTCAACCTTAGCGCCAACGCCCTCAAGACAGTGGACCACCTCTGG
711▶ S I S F V A P G F F S K A K E L R E L N L S A N A L K T V D H S W
2801 TTTGGGCCCTGGCAGTGCCTGCAAATACTAGATGAAGGCCAACCCCTGCACTGCGCCTGTGGGGCGGCTTTATGGACTTCTGCTGGAGGTGC
745▶ F G P L A S A L Q I L D V S A N P L H C A C G A A F M D F L L E V
2901 AGGCTGCCGTGCCGGTCTGCCAGCCGGTGAAGTGTGGCAGTCCGGGCCAGCTCCAGGGCCTCAGCATCTTTGCACAGGACCTGCGCCTCTGCCTGGA
778▶ Q A A V P G L P S R V K C G S P G Q L Q G L S I F A Q D L R L C L D
3001 TGAGGCCCTCTCCTGGGACTGTTTCGCCCTCTCGCTGCTGGCTGTGGCTCTGGCCTGGGTGTGCCATGCTGCATCACCTCTGTGGCTGGGACCTCTGG
811▶ E A L S W D C F A L S L L A V A L G L G V P M L H H L C G W D L W

NheI (3177)

3101 TACTGTTCCACCTGTGCTGGCTGGCTTCCCTGGCGGGGGCGCAAAGTGGCGAGATGAGGATGCCCTGCCCTAGCTAGCTGGCCAGACATGATAAG
845▶ Y C F H L C L A W L P W R G R Q S G R D E D A L P •
3201 ATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATA

HpaI (3315) MfeI (3326)

3301 AGCTGCAATAAACAGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACA

EcoRI (3411)

3401 AATGTGGTATGGAATTCTAAATACAGCATAGCAAACCTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCA

SapI (3593)

3501 TCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAACACTAGCTCTTC

SspI (3650) SwaI (3664)

3601 ATTTCTTTATGTTTTAAATGCACTGACCTCCCACATCCCTTTTTAGTAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTT

3701 TTATTAGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGC

3801 AAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCACTCAATGAGCACAA

141▶ • 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

3901 AGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGAC

114▶ 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 H R V

4001 AGCCACAATGGTGTCAAAGTCTTCTGCCCGTGTCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATG

81▶ 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 A E I

4101 TGGACAGCAGAGATGATCTCCCACTTGGTCTGATGGCCGCCGACATGGTGTCTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCT

47▶ 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 T A V E

BspHI (4239)

BbsI (4235)

XmnI (4231)

AseI (4297)

4201 CCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATT

14▶ V L E L D Q Q S I N F T K M

4301 AATTGTCAAACACGCGTGATGGCGTCTCCAGCTTATCTGACGGTTACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCCAT

SpeI (4452)

4401 TTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGGA

SnaBI (4580)

4501 AATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCGCTGCTTATATAGACCTCCACCGTACACGCTACCGCCCAT

NdeI (4685)

4601 TAGGAAAGTCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTG

4701 ATGTAAGTCCCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGAC

SdaI (4863) PaeI (4871) BspLU11I (4881)

4801 GTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCATTACCCTAAGTTATGTAACGCCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCA

4901 GCAAAAAGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGG

5001 TGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGT

ApaLI (5195)

5101 CCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCTTCGCTCCAAGCTGGGCTGTGTGCA

5201 CGAACCCCCGTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT

5301 GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGACAGTATTTGGTATCT

5401 GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGATTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCA

5501 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACTCACGTTAAGGGATTTTG

EagI (5631)

Pacl (5611) SwaI (5620) **NotI (5630)**

5601 GTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAAC

5701 ATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA