



EagI (1)
NotI (-1)
1 GCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACAAACATACGCTCTCCATCAAACAAAACGAAACA
PvuI (172)
SgfI (171)
101 AAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAAGGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCA
EcoNI (261)
201 GAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGGAGGGTCCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAAGTATG
Psp1406I (368)
301 TCGTGTACTGGCTCCGCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTCAGATAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTCCCGCCAG
HindIII (410) Bsu36I (456)
PvuII (404) EcoNI (452)
401 AACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCCCATCCACGCCGTTGAGTCCGCTTCTGCCGCT
501 CCCGCTGTGGTGCCTCTGAAGTCCGCTCCGCGCTAGGTAAAGTTAAAGCTCAGTTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGCCTACTTA
NgoMI (606)
NaeI (606)
601 GACTCAGCCGGCTCTCCACGCTTTGCCTGACCCTGCTTGTCAACTCTACGCTTTTGTGTTTTCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACC
NcoI (725)
BstEII (720)
AgeI (717) StuI (762)
701 GCGCCCTACCTGAGATCACCGGTCAACCATGGAGAGTAAATACAAGGAGATACTTCTGCTAACAGGCTGGATAACATCACTGATGAGGAACTGGATAGGT
1 M E S K Y K E I L L L T G L D N I T D E E L D R
SspI (826)
801 TTAAGTCTTTCTTTTCAGACGAGTTAATTGCCACAGGCAAACTACATACTGCAAAACAGAATACAAGTAGCTACCTTGATGATTCAAAATGCTGGGGC
25 F K F F L S D E F N I A T G K L H T A N R I Q V A T L M I Q N A G A
901 GGTGTCTGAGTGAAGACCATTCGTATTTTTCAGAAGTTGAATTATATGCTTTTGGCAAAACGCTTCAGGAGGAGAAGGAGAAGTTGATAAGCAA
58 V S A V M K T I R I F Q K L N Y M L L A K R L Q E E K E K V D K Q
1001 TACAAATCGGTAACAAAACCAAGCCACTAAGTCAAGCTGAAATGAGTCTGTCAGCATCTGCAGCCATCAGAAATGATGTGCAAAAGCAACGCTGCAC
92 Y K S V T K P K P L S Q A E M S P A A S A A I R N D V A K Q R A A
Eco47III (1178)
1101 CAAAAGTCTCTCCTCATGTTAAGCCTGAACAGAAAACAGATGGTGGCCAGCAGGAATCTATCAGAGAAGGGTTTCAGAAAGCGCTGTTTGCAGTTATGGT
125 P K V S P H V K P E Q K Q M V A Q Q E S I R E G F Q K R C L P V M V
EcoRI (1282)
1201 ACTGAAAGCAAAGAAGCCCTTTCACGTTTGAGACCCAAAGGCAAGCAGGAGATGTTTCATGCTACAGTGGCTACAGAAAAGGAATTCCTTTGTAAAA
158 L K A K K P F T F E T Q E G K Q E M F H A T V A T E K E F F F V K
1301 GTTTTAATACACTGCTGAAAGATAAATTCATTCCAAAGAGAATAAATATAATAGCAAGATATTATCGGCACAGTGGTTTCTAGAGGTAATAGCGCCT
192 V F N T L L K D K F I P K R I I I I A R Y Y R H S G F L E V N S A
BbrPI (1400)
1401 CACGTGTGTTAGATGCTGAATCTGACCAAAAGGTTAATGTCGCCGTGAACATTATCAGAAAAGCTGGTGAACCCCGAAGATCAACACGCTTCAAACCTCA
225 S R V L D A E S D Q K V N V P L N I I R K A G E T P K I N T L Q T Q
1501 GCCCTTGGAAACAATTGTGAATGGTTTGTGTAGTCCAGAAGGTAACAGAAAAGAAACATATTATTTGACCTAAGTGACAACTGGGAAAATG
258 P L G T I V N G L F V V Q K V T E K K K N I L F D L S D N T G K M
ScaI (1602) PvuII (1699)
1601 GAAGTACTGGGGTTAGAAAACGAGGACACAATGAAATGAAGGAAGGAGATAAGGTTTCGACTTACATTCTTACACTGTCAAAAATGGAGAAAACTAC
292 E V L G V R N E D T M K C K E G D K V R L T F F T L S K N G E K L
BamHI (1756) NdeI (1796)
1701 AGCTGACATCTGAGTTCATAGCACCATAAAGGTTATTAAGGCCAAAAAAGCAAGGATCCTATCCCTATGATGTGCCAGACTATGCTGGCTATCCATA
325 Q L T S G V H S T I K V I K A K K K T G S Y P Y D V P D Y A G Y P Y
MscI (1859)
NheI (1853)
1801 TGATGTTCTGATTGCTGGATACCCCTTATGATGTGCCAGACTATGCCTAAGCTAGCTGGCAGACATGATAAGATACATTGATGAGTTGGACAAAC
358 D V P D Y A G Y P Y D V P D Y A •
HpaI (1991)
1901 CACAAC TAGAATGCAGTGA AAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTTAAACA
EcoRI (2087)
2001 AACAAATGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCCTAAATA
2101 CAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCAT
2201 TAGCTGTTTGCAGCCTCACCTTCTTCATGAGTAAAGATAGTGTATTTTCCCAAGGTTTGAAGTACTCTTCAATTTCTTATGTTTAAATGCAT
SspI (2326) SmaI (2340)
2301 GACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTC
EcoO109I (2401)
2401 AAGGCCCTTCATAATATCCCCAGTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAG
141 •
2501 TTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAGCAGTCAAGGAGCAGTCAAGGAGCAGT
139 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 I
SacI (2601)
EclI36II (2601) BstXI (2630)
2601 TGAGCTCTCTGCACATGCCACAGGGCTGACCACTGATGATCTGTCCACTCATCAGAGTAGGGTGCCTGACGCCACAATGGTGTCAAAGTCTT
106 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 (2765)

2701 CTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCA
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 E G
2801 GTCTTGTCCTGATGGCCGCCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAG
39 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 S

BspHI (2915) AseI (2973)

2901 AGATGTTGAAGGCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGC
6 I N F T K M

SacI (3030)
Ecl136II (3030)

3001 GTCTCCAGCTTATCTGACGGTTCACAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTT

SpeI (3128)

3101 ACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGAGACTTGAAATCCCGTGAGTCAAACCGCTAT

SnaBI (3256)

3201 CCACGCCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCATAAGGTCATGT

NdeI (3361)

3301 ACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTCCCAAGTGGGCGAGTTTA
3401 CCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGGC

PacI (3547) PciI (3557)
SdaI (3539) BspLU11I (3557)
SbfI (3539)

3501 GGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAA
3601 GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
3701 AGATAACAGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG

ApaLI (3871)

3801 TGGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCG
3901 CTGCGCCTTATCCGTAATCTGCTTGTAGTCAACCCGGTAAGACAGCACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
4001 TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT
4101 TCGGAAAAAGAGTTGGTAGCTCTTATCGGCAAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTTTGTTTGAAGCAGCAGATTACGGCAGAAAAAAGG

PacI (4287) SwaI (4296)

4201 ATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTT
4301 AAATCA