



NotI (-1)
1 GCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACA
PvuI (172)
SgfI (171)
101 AAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAAGGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCA
AgeI (254)EcoNI (261)
201 GAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCCGCAATTGAACCGGTGCC TAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGATG
Psp1406I (368)
301 TCGTGTACTGGCTCCGCTTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAG
Bsu36I (456)
PvuII (404) EcoNI (452)
401 AACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCGCCATCCACGCCGGTTGAGTCGCGTTCTCGCCCT
501 CCCGCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTA
NaeI (606)
601 GACTCAGCCGGCTCCACGCTTTGCTGACCTGCTTGCTCAACTCTACGCTTTTGTGTTTTCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACC
AgeI (717)
701 GCGCCCTACCTGAGATCACCGTAGGAGGCCATCATGACCAAAGACAAAACCTATTGTTAAAAGCTTCCATTTTGTTCGCTTATGATCATAATAGT
1 M T K D K E P I V K S F H F V C L M I I I V
801 TGAACCAAGTCCAGTTCTCCGACGAAATGAATTTGCAGTAGACAAGTCAAAAAGAGGCTTATTTCATGTTCCAAAAGACCTACCGCTGAAAACAAA
22 G T R I Q F S D G N E F A V D K S K R G L I H V P K D L P L K T K
901 GTCTTAGATATGCTCAGAATCATCGTGTGAGCTTCAAGTCTGACATGAGCTTTCTATCAGAGTTGACAGTTTTGAGACTTTCCCAATAACAGAAATCC
56 V L D M S Q N Y I A E L Q V S D M S F L S E L T V L R L S H N R I
EcoRV (1076)
1001 AGCTACTTGATTAAGTGTTCCTCAAGTTCACACAGGATTTAGAATATTTGGATTTATCTCATAATCAGTTGCAAAAGATATCCTGCCATCCTATTGTGAG
89 Q L L D L S V F K F N Q D L E Y L D L S H N Q L Q K I S C H P I V S
BglII (1112) EcoO109I (1135)
1101 TTTGAGCATTAGATCTCTCATTCAATGATTTCAAGGCCCTGCCATCTGTAAGGAATTTGGCAACTTATCACAACCTGAATTTCTGGGATGAGTGT
122 F R H L D L S F N D F K A L P I C K E F G N L S Q L N F L G L S A
1201 ATGAAGCTGCAAAAATAGATTTGCTGCAATGCTCACTTGCATCTAAGTTATATCCTTCTGGATTTAAGAAATATTATATAAAAAGAAAATGAGACAG
156 M K L Q K L D L L P I A H L H L S Y I L L D L R N Y Y I K E N E T
1301 AAAGTCTACAAATCTGAATGCAAAAACCTTACCTGTTTTCCACCAACTAGTTTATTCGCTATCCAAGTGAACATATCAGTTAATACTTTAGGGTG
189 E S L Q I L N A K T L H L V F H P T S L L F A I Q V N I S V N T L G C
PshAI (1431)
1401 CTTACAACGACTAATATTAATGAATGATGACAACCTGTCAAGTTTTTCAATAATTTTATCAGAACCTACCAGAGGTTCAACCTTACTGAATTTTACC
222 L Q L T N I K L N D N C Q V F I K F L S E L T R G S T L L N F T
BstXI (1558)
1501 CTCAACCACATAGAAACGACTTGGAAATGCCTGGTCCAGAGTCTTCAATTTCTTTGGCCAAACCTGTGGAATATCTCAATATTTACAATTTAACAATA
256 L N H I E T T W K C L V R V F Q F L W P K P V E Y L N I Y N L T I
1601 TTGAAAGCATTGCGAAGAAGATTTTACTTATCTAAAACGACATTGAAAGCATTGACAATAGAACATATCAGAACCAAGTTTTCTGTTTTCACAGAC
289 I E S I R E E D F T Y S K T T L K A L T I E H I T N Q V F L F S Q T
BsrGI (1705) HpaI (1739)
1701 AGCTTTGTACACCGTGTTCCTGAGATGAACATTATGATGTTAACATTTGATACACCTTTTATACACATGCTGTGCTCATGCACCAAGCACATTC
322 A L Y T V F S E M N I M M L T I S D T P F I H M L C P H A P S T F
Psp1406I (1821)
1801 AAGTTTTTGAACCTTACCAGAACGTTTTTACAGATAGTATTTTTGAAAATGTTCCACGTTAGTTAAATTTGGAGACACTTATCTTACAAAAGAAATGGAT
356 K F L N F T Q N V F T D S I F E K C S T L V K L E T L I L Q K N G
1901 TAAAAGCCTTTCAAAGTAGGCTCATGACGAAAGGATATGCTTCTTTGAAATACTGGATGTTAGCTGGAATCTTTGGAATCTGGTAGACATAAAGA
389 L K D L F K V G L M T K D M P S L E I L D V S W N S L E S G R H K E
2001 AAAGTGCATTTGGGTTGAGAGTATAGTGGTGTAAATTTGCTTCAAATGCTTACTGACTCTGTTTTGAGATGTTTACCTCCAGGATCAAGTACTT
422 N C T W V E S I V V L N L S S N M L T D S V F R C L P P R I K V L
2101 GATCTTCACAGCAATAAAAAGAGCGTTCTTAAACAAGTCGTAACAACTGGAAGCTTTGCAAGAATCAATGTTGCTTTCAATTTTAACTGACCTTC
456 D L H S N K I K S V P K Q V V K L E A L Q E L N V A F N S L T D L
2201 CTGGATGTGGCAGCTTTAGCAGCCTTTCTGTATTGATCATTGATCACAATTCAGTTTCCACCCATCGGCTGATTTCTCCAGAGTCCGAGAAGATGAG
489 P G C G S F S L S V L I I D H N S V S H P S A D F F Q S C Q K M R
2301 GTCAATAAAAGCAGGGGACAATCCATTTCAATGTACCTGTGAGCTAAGAGAATTTGCAAAAATATAGACCAAGTATCAAGTGAAGTGTAGAGGGCTGG
522 S I K A G D N P F Q C T C E L R E F V K N I D Q V S S E V L E G W
BspLU11I (2461)
2401 CCTGATTCTTAAAGTGTGACTACCCAGAAAAGTTATAGAGGAAGCCCACTAAAGGACTTTACATGTCTGAATTTATCTGCAACATAACTCTGCTGATCG
556 P D S Y K C D Y P E S Y R G S P L K D F H M S E L S C N I T L L I
XcmI (2515)
BstXI (2512)
2501 TCACCATCGGTGCCACCATGCTGGTGTGGCTGTGACTGTGACCTCCCTCTGCATCTACTTGGATCTGCCCTGGTATCTCAGGATGGTGTGCCAGTGGAC
589 V T I G A T M L V L A V T V T S L C I Y L D L P W Y L R M V C Q W T
2601 CCAGACTCGGCGCAGGGCCAGGAACATCCCTTAGAAGAAGCTCAAAAAGAACTCCAGTTTTCATGCTTTTATTTTATATAGTGAACATGATTTCTGCTGG
622 Q T R R R A R N I P L E E L Q R N L Q F H A F I S Y S E H D S A W
Acc65I (2714)
2701 GTGAAAAGTGAATTTGGTACCTTACCTAGAAAAAGATATACAGATTTGTCTTCATGAGAGAAACTTTGTCCTGGCAAGAGCATTGTGAAAATATCA
656 V K S E L V P Y L E K E D I Q I C L H E R N F V P G K S I V E N I
2801 TCAACTGCATTGAGAAGATTAAGTCCATCTTTGTTGTCTCCCACTTTGTCCAGAGTGAAGTGGTCCATTACGAACCTTATTTTGGCCATCACAA
689 I N C I E K S Y K S I F V L T S P N F V Q S E W C H Y E L Y F A H N
2901 TCTCTTTCATGAAGGATCTAATAACTTAATCCTCATCTTACTGGAACCCATTCCACAGAACAGCATTCCCAACAAGTACCACAAGCTGAAGGCTCTCATG
722 L F H E G S N N L I L I L L E P I P Q N S I P N K Y H K L K A L M
3001 ACGACGCGACTTATTTGAGTGGCCCAAGGAGAAAACGTTGGGCTTTTGGGCTAACATTAGAGCCGCTTTTAAATATGAAATTAACACTAGTCA
756 T Q R R T Y L Q W P K E K S K R G L F W A N I R A A F N M K L T L V
XcmI (3154)
3101 CTGAAAACATGATGTGAAATCTTAAAAAATTTAGGAAATCAACTTAAGAAACCATTTTACTTGGATGATGGTGAATAGTACAGTCGTAAGTNACT
789 T E N N D V K S •
BallI (3222)
NheI (3216)
3201 GTCTGGAGGTGCCTCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAACCAACAATAAGATGCAGTGAATAAATGCTTTATTTGT
HpaI (3354)
3301 GAAATTTGTGATGCTATTGCTTTATTTGTAACCATATAAGCTGCAATAAACAAAGTTAACAAACAATTCATTCTTTTATGTTTCAGGTTTCAGGGG
3401 AGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTTCAAATACAGCATAGCAAACTTTAACTCCAATCAAGCCTCT

AlwNI (3540)
3501 ACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTTA
3601 AGATATAGTGTATTTTCCCAAGGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATAGTAAAATATTCAGAA

Swal (3703) EcoO109I (3764)
3701 ATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCCAGTTTAGTAGTTGGA
3801 CTTAGGGAACAAGGAACCTTTAATAGAATGGACAGCAAGAAAGCGAGCTTCTAGCGAATTCGACTCATTCTTTGCCCTCGGACGAGTGTGGGG
342 • E K A R P R T S P

ScaI (3921)
3901 CGTCGGTTTCCACTATCGGCGAGTACTTCTACACAGCCATCGGTCCAGACGGCCGCGCTTCTGCGGGCGATTTGTGTACGCCGACAGTCCCGGCTCCGG
331 R R N G S D A L V E V C G D T W V A A S R R A I Q T R G V T G A G S
4001 ATCGGACGATTGCGTCGCATCGACCTCGCCCAAGCTGCATCATGAAATGCCGTCAACCAAGCTCTGATAGAGTTGGTCAAGACCAATGCGGAGCAT
298 R V I A D C R G Q A W A A D D F N G D V L S Q Y L Q D L G I R L M

SaeII (4112) BsrBI (4142)
4101 ATACGCCCGGAGCCGCGGATCCTGCAAGCTCCGGATGCCCTCCGCTCGAAGTAGCGCGTCTGCTGCTCCATACAAGCCAACCACGGCTCCAGAAGAAG
265 Y A R L R P S G A L E P H R R E F Y R T Q Q E M C A L W P R W F F
4201 ATGTTGGCAGCTGTATTGGGAATCCCGAACATCGCTCGCTCCAGTCAATGACCGCTGTTATGCGCCATTGTCGCTCAGGACATTGTTGGAGCCGA
231 I N A V E Y Q S D G F M A E S W D I V A T I R G N D T L V N N S G F

DraIII (4310) Tth11I (4385)
4301 AATCCGCTGCACGAGTCCGGGACTCGGGGAGTCTCGGCCAAAGCATCGCTCATCGAGAGCTGCGGACGGACGCACTGACGGTGTCTGCTCAT
198 D A H V L H R V E P C D E A W L M L E D L A Q A V S A S V T D D M

NdeI (4440) RsrII (4481)
4401 CACAGTTTCCAGTGATACACATGGGGATCGCAATCGCCGATATGAAATCACGCCATGTAGTGTATTGACCGATTCTTTCGGTCCGAATGGGCCGAAC
165 V T Q W H Y V H P D A I A C I F D R W T T Y Q G I G Q P G F P G F

BsrBI (4500) SgfI (4527) SaeI (4541)
4501 CCGCTCGTCTGGCTAAGATCGGCCGAGCGATCGCATGAGCTCCGCGACGGGTTGAGAACAGCGGGCAGTTCGGTTTCAGGAGGTCCTTGAACG
131 G S T Q S L D A A A I A D M L E A V P Q L V A P L E T E P L D Q L T

DraIII (4603)
4601 TGACACCCTGTGCACGCGGGAGATGCAATAGGTCAGGCTCTCGCTGAATTCCTCAATGTCAAGCACTCCGGAATCGGGAGCGCGCCGATGCAAAGTG
98 V G Q A R R S I C Y T L S E S F E G I D L V E P I P L A A S A F H
4701 CCGATAAACATAACGATCTTTGTAGAAACCATCGGCGAGCTATTTACCCGAGGACATATCCACGCCCTCTACATCGAAGCTGAAAGCACGAGATTTCT
65 R Y V Y R D K Y F G D A C S N V R L V Y G R G G V D F S F A R S E

PshAI (4862) NruI (4858)
4801 TCGCCCTCCGAGAGTGCATCAGGTCGGGACGCTGTGAACTTTTCGATCAGAACTTCGCGACAGAGCTCGCGGTGAGTTCAGGCTTTTTTCATGATGG
31 E G E S L Q M L D S V S D F K E I L F K A V S T A T L E P K K M

AseI (4949)
4901 CCCTCCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCA

SacI (5006)
5001 CTAACGAGCTCTGTTATATAGACCTCCCACGTACACGCCATCCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTGGAAAGTCCCGTTG

SnaBI (5232)
5201 AACCGCATCATGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAGGTCATGTACTGGGCATAATGCCAGCGGGCC

NdeI (5337)
5301 ATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTGTACTGCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACC
5401 TCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTCAGCCAGCGGGCCATTACC

PacI (5523) PstI (5516) SdaI (5515) BspLU11I (5533)
5501 GTAAGTTATGTAACGCTGCAGGTTAATTAGAACAATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTGTGCGGTTTTTCCA
5601 TAGGCTCCGCCCTGACGAGCATCAGAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGGTTTCCCTGGA
5701 AGCTCCCTCGTGCCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCT
5801 GTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCAGAACCCCGCTTACGCCGACCGCTGCGCCTTATCCGGTAACATCG

AlwNI (5944)
5901 TCTTGAGTCCAACCCGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC
6001 TTGAAGTGGTGGCTAACTACGGCTACACTAGAAAGCAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTT
6101 GATCCGGCAAACAACCCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACGCCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
6201 TTCTACGGGTCTGACGCTCAGTGAACGAAAACCTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCA

PacI (6263) Swal (6272)