



NotI (-1)
1 GCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACA
PvuI (172)
SgfI (171)
101 AAACAACTAGCAAAATAGGCTGTCCCCAGTGCAGAGTGCAGGTGCCAGAACATTTCTCTATCGAAGGATCTGCGATCGCTCCGGTCCCGTCACTGGGCA
EcoNI (261)
201 GAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGTCCGCAATTGAACGGTGCCTAGAGAAGTGGCGGGGTAAACTGGGAAAGTGATG
Psp1406I (368)
301 TCGTGTAAGTCCGCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCCGCTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAG
Bsu36I (456)
PvuII (404)
EcoNI (452)
401 AACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCCCATCCACGCCGTTGAGTCCGCTTCTGCCCT
501 CCGCCTGTGGTCCCTCTGAAGTCCGCTCCGCTCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGCCCTTTGTCGGCGCTCCCTTGGAGCTACCTA
NaeI (606)
601 GACTCAGCCGGCTCTCCAGCTTTGCTGACCCTGCTTCAACTCTACGCTTTTGTTCGTTTTCTGTTCTGCCGCTTACAGATCCAAGCTGTGACC
KasI (700) AgeI (717) NcoI (733)
701 GCGCCCTACCTGAGATCACCGGTAGGAGGGCCACCTGGTGTTCATGTGGACTGAAGAGACTAATTCTTATCCTTTTTAACATAATCCTAATTT
1 M V F P M W T L K R L I L I L F N I I L I S
801 CAAACTCCTGGGGCTAGATGTTTCCATAAACTCTGCCCTGTGATGCTACTCTGGATGTTCCAAAGAACATGTGATCGTGACTGCACAGACAAGCAT
22 K L L G A R W F P K T L P C D V T L D V P K N H V I V D C T D K H
901 TTGACAGAAATTCCTGGAGGTATCCACGAACACACGAACTCACCTACCATTAACACATACCAGACATCTCCCGAGCTCCTTTACAGAGCTGG
56 L T E I P G G I P T N T T N L T L T I N H I P D I S P A S F H R L
ClaI (1014)
1001 ACCATCTGGTAGAGATCGATTTACAGATGCAACTGTGTACCTATCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAACCCAGAAG
89 D H L V E I D F R C N C V P I P L G S K N N M C I K R L Q I K P R S
1101 CTTTAGTGGACTCACTTATTTAAATCCCTTTACCTGGATGGAACACAGCTACTAGAGATACCGCAGGGCTCCCGCTAGCTTACAGCTTCTCAGCCTT
122 F S G L T Y L K S L Y L D G N Q L K L E I P Q G L P P S L Q L S L
1201 GAGGCCAACACATCTTTCCATCAGAAAAGAGAACTAACAGAAGTGGCCAAACATAGAAATACTCTACCTGGGCAAAAAGTATTATTAGAAATCCTT
156 E A N N I F S I R K E N L T E L A N I E I L Y L G Q N C Y Y R N P
1301 GTTAGTTCATATTCATAGAGAAAGATGCCTTCTAACTGACAAAGTAAAGTGTCTCCCTGAAAGATAACAATGTCACAGCCGCTCCCTACTGT
189 C Y V S Y S I E K D A F L N L T K L K V L S L K D N N V T A V P T V
1401 TTTGCCATCTACTTAAACAGAACTATATCTACAACAACATGATTGAAAAATCAAAGAAGATGATTTAATAACCTCAACCAATTACAAATCTTGAC
222 L P S T L T E L Y L Y N N M I A K I Q E D D F N N L N Q L Q I L D
1501 CTAAGTGGAAATTCGCTCGTTTATAATGCCCATTTCTGTGCGCGGTAAAAATAATTCTCCCTACAGATCCCTGTAAGTCTTTTGTGCGG
256 L S G N C P R C Y N A P F P C A P C K N N S P L Q I P V N A F D A
1601 TGACAGAAATTAAGTTTTACGTCTACACAGTAACTCTTTCAGCATGTGCCCAAGATGGTTAAGAAACATCAACAACTCAGGAACTGGATCTGTC
289 L T E L K V L R L H S N S L Q H V P P R W F K N I N K L Q E L D L S
1701 CAAAACTTCTGGCCAAAGAAATGGGGATGCTAAATTTCTGCATTTTCTCCCGAGCTCATCAATGGATCTGTCTTTCAATTTTGAATTCAGGTC
322 Q N F L A K E I G D A K F L H F L P S L I Q L D L S F N F E L Q V
1801 TATCGTGCATCTAATCTATCAAGCATTTTCTTCACTGAAAAGCTGAAAATTCGCGGATCAGAGGATATGCTTTAAAGAGTGAAGCTTTA
356 Y R A S M N L S Q A F S L K L I L R I R G Y V F K E L G S F
PmeI (1982)
1901 ACCTCTCGCCATTACATAATCTTCAAAATCTTGAAGTCTTGATCTTGGCCTAACTTTAATAAAATGCTAACCTCAGCATGTTTAAACAATTTAAAG
389 N L S P L H N L Q A N L E V L D L G T N F I K I A N L S M F K Q F K R
BglII (2012)
2001 ACTGAAAGTCATAGATCTTTCAGTGAATAAAATACACCTCAGGAGATCAAGTGAAGTTGGCTTCTGCTCAAATGCCAGAATCTTCTGAGAAAGTTAT
422 L K V I D L S V N K I S P S G D S S E V G F C S N A R T S V E S Y
2101 GAACCCAGGTCCTGGAACAATTACATTTTTCAGATATGATAAGTATGCAAGGAGTTCAGATTCAAAAACAAGAGGCTTCTTTTCTGCTGTTAATG
456 E P Q V L E Q L H Y F R Y D K Y A R S C R F K N K E A S F M S V N
2201 AAAGCTGTACAAGTATGGGCAGACCTTGGATCTAAGTAAAAATAGTATATTTTGTCAAGTCTCTGATTTTCAGCATCTTTCTTCTCAAAATGCCT
489 E S C Y K Y G Q T L D L S K N S L I F F V K S S D F Q H L S F L K C L
XcmI (2324)
2301 GAATCTGTGAGAAATCTCATTAGCCAACTTAAATGGCAGTGAATCCAACTTTAGCAGAGCTGAGATATTTGGACTTCTCAACAACCGGCTTGAT
522 N L S G N L I S Q T L N G S E F Q P L A E L R Y L D F S N N R L D
XmnI (2421)
2401 TTACTCCATTCAACAGCATTTGAAGAGCTTCAAAAAGTCTGATATAAGCAGTAATAGCATTATTTTCAATCAGAAGGAATTACTCATATGC
556 L L H S T A F F E L H K L E V L D I S S N S H Y F Q S E G I T H M
XcmI (2569) NcoI (2578)
2501 TAACTTTACCAAGAACCTAAAGTCTCGAGAACTGATGATGAACGACAATGACATCTCTCCACCAGCAGGACCATTGGAGAGTGAGTCTCTTAG
589 L N F T K N L K V L Q K L M M N D N D I S S T S R T M E S E S L R
2601 AACTCTGGAATTCAGAGGAATCACTTAGATGTTTATGGAGAGAAGTGAACAGATACTTACAATTAATCAAGAATCTGCTAAAAATAGAGGAATTA
622 T L E F R G N H L D V L W R E G D N R Y L Q L F K N L L K L E E L
2701 GACATCTCAAAATCCCTAAGTTCTTCCCTTCTGGAGTTTGTATGGTATGCTTCAAAATCTAAAGAATCTCTTTTGGCCAAAATCGGCTCAAAAT
656 D I S K N S L S F L P S G V F D G M P P N L K N L S L A K N G L K
PshAI (2867)
2801 CTTTCAGTTGGAAGAACTCCAGTGTCTAAAGAACCTGAAACTTTGGACCTCAGCCACAACCAACTGACCATGTCCTGAGAGATTACCAACTGTT
689 S F S W K K L Q C L K N L E T L D L S H N Q L T T V P E R L S N C S
EcoRV (2983)
2901 CAGAAGCTCAAGAATCTGATTCTTAAGAATAATCAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCTTCCAGTGGCATATCTGGATCTCAGC
722 R S L K N L I L K N N Q I R S L T K Y F L Q D A F Q L R Y L D L S
3001 TCAAATAAATCCAGATGATCCAAAGAACAGCTTCCAGAAAAATGCTCAACAACTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCT
756 S N K I Q M I Q K T S F P E N V L N N L K M L L L H H N R F L C T
HpaI (3126)
3101 GTGATGCTGTGGTGTGTTGCTGTGGTGGGTTAACCATACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGG
789 C D A V W F V W V N H T E V T I P Y L A T D V T C V G P G A H K G
BsrGI (3223) BglII (3239)
3201 CCAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTTCTGTTCTCACTTTCCATATCTGTATCTCTTTCTCATGGT
822 Q S V I S L D L Y T C E L D L T N L I L F S L S I S V S L F L M V
3301 ATGATGACAGCAAGTCACTTATTTCTGGATGTGTGGATATAATACCTTTGTAAGGCCAAGATAAAGGGTATCAGCGTCTAATATCACAGACT
856 M M T A S H L Y F W D V W Y I Y H F C K A K I K G Y Q R L I S P D
PvuII (3439)
3401 GTTGCTATGATGCTTTTATTGTATGACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAACTGGAAGACCCAAAGAGAGAA
889 C C Y D A F I V Y D T K D P A V T E W V L A E L V A K L E D P R E K
XhoI (3516) BstEII (3532)
3501 ACATTTTAAATTTATGCTCGAGGAAAGGGACTGGTTACCGAGGACGCAAGTCTGAAAAACCTTTCCAGAGCATACAGCTTAGCAAAAAGACAGTGT
922 H F N L C L E E R D W L P G Q P V L E N L S Q S I Q L S K K T V F

3601 GTGATGACAGACAAGTATGCAAAGACTGAAAATTTAAGATAGCATTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGA
956▶ V M T D K Y A K T E N F K I A F Y L S H Q R L M D E K V D V I I L
3701 TATTTCTTGAGAAGCCCTTCAGAAAGTCCAGTTCCTCCAGCTCCGGAAGGCTCTGTGGGAGTTCTGTCTTGAAGTGGCAACAAACCCGCAAGCTCA
989▶ I F L E K P F Q K S K F L Q L R K R L C G S S V L E W P T N P Q A H
BstXI (3801)
3801 CCCATACCTTCTGGCAGTGTCTAAAGAAGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAAACGGTCTAGCCCTTCTTTGCAAAA
1022▶ P Y F W Q C L K N A L A T D N H V A Y S Q V F K E T V •
StuI (3925) NheI (3939)
3901 CACAACCTGCCTAGTTTACCAAGGAGAGGCTGGCTGTTTGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGC
HpaI (4077)
4001 AGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACACAACAATTGCATTCA
4101 TTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTAAAATACAGCATAGCAAAAC
4201 TTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGC
4301 CTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTACTGCTTCTATTCTTTATGTTTTAAATGCACTGACCTCCACATTC
SspI (4412) SmaI (4426)
4401 CCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAAGTCCAGATGCTCAAGGCCCTTCATAA
4501 TATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAATTTGGACAGCAAGAAAGCGAGCTTCTAGCGAATTCCTCGACTATTCTCT
3424 • E K
ScaI (4644)
4601 TTGCCCTCGGACGAGTGTGGGGCGTCGGTTCCACTACTCGGCGAGTACTTCTACACAGCCATCGGTCCAGACGGCCGCGCTTCTGGGGCGATTGTGT
339▶ A R P R T S P 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
4701 ACGCCGACAGTCCCGCTCCGATCGGATCGCTGCATCGACCTCGCCAAAGCTGCATCAGAAATGCGTCAACCAAGCTCTGATAGT
306▶ R G V T G A G S 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
SacII (4835) BsrBI (4865)
4801 TGGTCAAGACCAATCGGAGCATATACGCCCGGAGCGCGGATCGTCAAGCTCGGATGCTCCGCTCGAAGTAGCGGCTGCTGCTCCATACAAG
272▶ Q D L G I R L M 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
4901 CCAACACGGCTCCAGAAGAAGATGTTGGCGACTCGTATTGGGAATCCCGAACATCGCTCGCTCCAGTCAATGACCGCTGTTATGCGCCATTGTCT
239▶ L W P R W F F 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
DraIII (5033)
5001 CGTCAGGACATTGTTGGAGCCGAAATCCGCGTGCACGAGGTGCCGACTTCCGGGCGAGTCTCGGCCAAAGCATCAGCTCATCGAGAGCCTGCGCGACG
206▶ T L V N N S G F 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
Tth111I (5108)
5101 GACGACTGACGGTGTCTCCATCACAGTTTCCAGTGATACATGGGATCAGCAATCGCCATATGAAATCAGCCATGATGTATTGACCGATTCT
172▶ S A S V T D D M 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
PvuI (5251)
RsrII (5204) BsrBI (5223) SgfI (5250) SacI (5264)
5201 CTTGCGTCCGAATGGCCGAACCCGCTCGTCTGGTAAGATCGGCCGACGATCGCATCCATGAGCTCCGCGACGGTTCGAGAACAGCGGGCAGTTC
139▶ Q P G F P G F 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
DraIII (5326)
5301 GGTTTCAGGACGTTCTTCAACGTGACACCTGTGACGCGGGGAGATGCAATAGGTGAGGCTCTCGTGAATCCCAATGTCAAGCACTTCGGAATC
106▶ T E P L D Q L T 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
5401 GGGAGCGCGCGGATGCAAGTCCGATAAACATAACGATCTTTGTAGAACCATCGGCGAGTATTTACCGCAGGACATATCCAGCCCTCTACAT
72▶ P L A A S A F H 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
Tth111I (5552) PshAI (5585) NruI (5581)
5501 CGAAGCTGAAAGCACGAGATCTTCGCCCTCCGAGAGCTGCATCAGGTCCGAGAGCGTGTGCAACTTTTCGATCAGAAATTCGCGACAGAGCTCGCGGT
39▶ F S F A R S E 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
BspHI (5614) AseI (5672)
5601 GAGTTCAAGCTTTTTTCTGATGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATGTCAAACAGCGTGGATGGCG
6▶ L E P K K M
SacI (5729)
5701 TCTCCAGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTA
SpeI (5827)
5801 CGACATTTTGAAAGTCCCGTTGATTACTAGTCAAAACAACTCCCATTGACGTCAATGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATC
SnaBI (5955)
5901 CAGCCCATGATGACTGCCAAAACGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTA
6001 CTGGCATAATGCCAGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTAC
6101 CGTAAATACTCCACCCATTGACGTCAATGGAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGCGG
SdaI (6238)
6201 GTCAGCCAGGCGGCCATTTACCGTAAGTTATGTAACCGCTGACAGTTAATTAAGAACATGTGAGCAAAAGGCCAGAAAAGGCCAGGAACCGTAAAAAG
6301 GCCCGTTGCTGGCTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCAAAACCGACAGGACTATAAA
6401 GATACAGGCGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGT
6501 GGGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGC
6601 TGCCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGTAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
6701 ATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACCTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
6801 CGGAAAAAGAGTTGGTAGCTTTGATCCGGCAAAACACCGCTGGTAGCGGTGTTTTTTTTGTTGCAAGCAGCAGATTACCGCGAGAAAAAAGGA
SwaI (6995)
6901 TCTCAAGAAGATCCTTTGATCTTTTACGGGCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTGGTCATGGCTAGTTAATTAACATTTA
7001 AATCA