



PstI (6) Bsp120I (6)
SdaI (6) SpeI (13)
1 CCTGCAGGGCCACTAGTCTGTAAGCTGAAGACCTGGCAGTGTGAGCTGGTCAGCCCCAGGACCTCTTTTGTGCCACGAGTGAATCACCTTGGCAT
101 AGACATAATGGTCAGGGGTGGGCACGCAGCCTGCTTCCCGCTGTGCTCCAGGCCTCCTTCGATGCTTCCGAGAAGTCTATTGAGCTGGGAGCTTGTACT
SmaI (206) NdeI (262)
201 GCACCCGGGGCTGACATCCTGGCATCCTGGGATAAAGCAGCCACGGGGTGCCTTGGCCATATGCCTCACTGGCGGACAGAAACAAGGCTCTATTACG
301 CGAGTACCCTGGAGTAGACACCAGAAGCCCAAGCATGGGCAGAGGAAGGCAGGGGTTGGGGGAGCAGAGCTGTCTGTGTCCAGAAGCCCAAGGACACA
EcoNI (418)
401 GATGGCTAAGGCGCTGGGAGAGGGACCTGAGTGAAGAGATAGATGGCCTGAAGTCTCAAGCAGCAACAGCCTCTCCCCGCCATTGGTGAGGGTGGG
Acc65I (583)
501 GTTTGGTTTCCCGACCTACATATCCCTCAGAGGCTGGTGTGTAGGAATTTAAAGGGGTAAATCTCTGAGAGAATGAGGGGTACCCAGGAAGACGGG
SphI (628)
601 GTGTTACAGAAAGAAAGACTCCAGCATGCACAGCCAACCTATTCAAACACTCTGTGTCAGGGGCTGCCAGGGGCCAGGCTCGGGGTGGGGGTGGGGGGC
701 AACGAGAAGCTGGATCAGGGAGAAATGGCCCACTAGGCTGGATAAGAGGCCACAGAGGGGCTCAGGAATGAAGCCTGCTGTCTTACCCTATTAGGATCTG
801 CGTGCATACCTTCTGCCGTGCACTCTAAACACACAGCCAGAGGCTCAAGTTGACCTGGAGTACAGAGAGGGCTCAACCTTAGCCCTCCACTCTCTGAA
901 CTCAGGAATGAGAAGATAGAGTTGGAGAGATTACAGGGGAGAGGACTCTGTTGAGAATGGGGTCCACAGAAACTGTAATATAGTTGATCCCGGAGGAA
XmnI (1002)
1001 GGGAAATAGGTTCTCAAGTTCCTAGCATCTCACAGCCCCAGAGAAGGACAGAGTTGGGGTGGTCTGGCTTACAGGCTCTAAGAACTGGAAGCTGATT
1101 ACCCCACCGAGCTGTGCACTCTCTGTCTGTCTGTGTGTGCGCTGTGCACACTTATCACAAATGTTTATGTGTGCACATACATGTGTTGAGA
1201 CCAGAGGTCAACCTCAGGCACTGTTGCCTTGGTTTTCTGAGAGAGCATTCTCTCTGGATCTGGAACCTGCCAATTAGTGAGAGCCAGGAAGTCTGCTGA
PstI (1369)
1301 TTTTCACTGCCAGCACTGGAGTTTACAAGTATGCACTGTCAACCCAGGCCTTTTGTATTCACTTCTGCAGCTAGAAGTGGGTGGGTCTTTCATGCTTGAC
1401 AGGCAAGCAATTTATGACTAAGCTGTTCCCTCGGCCCTCTCTTGACCATTACCAGAAAGGGGTTCTTGTATCAATGGCGAAGCCAGGCTGGTGTTC
Bsp120I (1570)
1501 CCAAGAAAGCCTTGACTCTGGGTACAGTGACCTCAGTGGGGTGGAGAGGAGTCTCCCTAGCTGGGCTGGGCCCCAGCTCCACCCCTCAGGCTATTCA
NcoI (1694)
1601 ATGGGGGTGCTTCCAGGAAGTCAGGGGAGATTTAGTCCAACCCGTTCTCCATAAAGGCCCTGACATCCCAGGAGCCAGCAGAGGCAGGGCCACATGGG
M G
NheI (1732) Acc65I (1788)
1701 GGGTTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAATGGGTCGGGATCTGTACGACGATGACGATAAGGTACCTAAGGAT
2 G S H H H H H H G M A S M T G G Q Q M G R D L Y D D D K V P K D
1801 CAGCTTGGAGTGTATCCCGTCTTTTACAACGCTGCTGACTGGGAAAACCTGGCGTTACCCAACTAATCGCCTTGACAGCATTCCCTTTCGCCAGCT
36 Q L G V D P V V L Q R R D W E N P G V T Q L N R L A A H P P F A S
FspI (1948)
1901 GCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCTGTTTCCGGCACCAGAAGCGGT
69 W R N S E E A R T D R P S Q Q L R S L N G E W R F A W F P A P E A V
2001 GCCGAAAGCTGGCTGGAGTGGCATCTTCTGAGGCCGATACTGTGCTGCTCCCTCAAAGTGGCAGATGCACGGTACGATGGCCCATCTACACCAAC
102 P E S W L E C D L P E A D T V V V P S N W Q M H G Y D A P I Y T N
2101 GTAACCTATCCATTACGGTCAATCCGCGTGTGTTCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGG
136 V T Y P I T V N P P F V P T E N P T G C Y S L T F N V D E S W L Q
2201 AAGGCCAGACGCGAATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTCAACGGGCGCTGGGTGCGTTACGGCAGGACAGTCTGTTGCCGTC
169 E G Q T R I I F D G V N S A F H L W C N G R W V G Y G Q D S R L P S
2301 TGAATTTGACCTGAGCGCATTTCACGCGCCGGAGAAAACCGCTCGCGGTGATGTTGCTGCGTTGGAGTGACGGCAGTCTGGAAGATCAGGATG
202 E F D L S A F L R A G E E N R L A A V M V L R W S D G S Y L T L G E D Q D M
AatII (2429)
2401 TGGCGGATGAGCGGATTTCCGCTGACGCTCTGTTGCTGCATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTAAATGATGATTTC
236 W R M S G I F R D V S L L H K P T T Q I S D F H V A T R F N D D F
2501 GCCGCGCTGTACTGGAGGCTGAAGTTCAGATGTGCGCGAGTGTGCTGACTACCTACGGGTAACAGTTCCTTATGGCAGGGTAAAACGAGGTCGCCAG
269 S R A V L E A E V Q M C G E L R D Y L R V T V S L W Q G E T Q V A S
ClaI (2630)
2601 CGGCACCGCCTTTCCGGCGTGAATATCGATGAGCGTGGTGGTTATGCCGATCGGTCACACTACGCTGTAACGCTGAAAACCCGAAACTGTGGAGC
302 G T A P F G G E I I D E R G G Y A D R V T L R L N V E N P K L W S
2701 GCCGAAATCCCGAATCTCTATCGTGGTGGTGAAGTGCACACCGCCGACGGCAGCGTATTGAGCAGAAAGCCTGCGATGTGCGTTTCCGCGAGGTGC
336 A E I P N L Y R A V V E L H T A D G T L I E A E A C D V G F R E V
2801 GGATGAAAATGGTCTGCTGCTGAACGGCAAGCCGTTGCTGATTCAGGCGTTAACCGTACAGCAGCATCATCCTGTCATGTCAGGTCATGATGATGA
369 R I E N G L L L L N G K P L L I R G V N R H E H H P L H G Q V M D E
EcoRV (2919) DraIII (2996)
2901 GCAGACGATGGTGCAGGATATCTGCTGATGAAGCAGAACAATTTAACGCCGTGCGCTGTTTCGATTATCCGAACCATCCGCTGGTGTACACGCTGTGC
402 Q T M V Q D I L L M K Q N N F N A V R C S H Y P N H P L W Y T L C
SspI (3036)
3001 GACCGCTACGGCCTGTATGTGGTGGATGAAGCCAAATATTGAAACCCACGGCATGGTCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCA
436 D R Y G L Y V V D E A N I E T H G M V P M N R L T D D P R W L P A
BsaBI (3132)
3101 TGAGCGAACCGTAACGCGAATGGTGCAGCGGATCGTAATCACCCGAGTGTGATCATCTGGTCTGGGGAATGAATCAGGCCACCGGCTAATCACGA
469 M S E R V T R M V Q R D R N H P S V I I W S L G N E S G H G A N H D
3201 CGCGCTGTATCGTGGATCAAATCTGCTGATCCTTCCCGCCGGTGCAGTATGAAGCGGCGGAGCCGACACCAGCCACCGATATTATTTGCCAGG
502 A L Y R W I K S V D P S R P V Q Y E G G G A D T T A T D I I C P M

BssHII (3304)
3301 TACGCGCGCGTGGATGAAGACCAGCCCTTCCGGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTCCTACCTGGAGAGACGCCGCCGCTGATCCCTT
536▶ Y A R V D E D Q P F P A V P K W S I K K W L S L P G E T R P L I L
3401 GCGAATACGCCACCGGATGGTAACAGTCTTGGCGGTTTCGTAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTTCGTCTGGGA
569▶ C E Y A H A M G N S L G G F A K Y W Q A F R Q Y P R L Q G G F V W D
3501 CTGGGTGGATCAGTCGCTGATTAATATGATGAAAACGGCAACCCGTGGTGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTTCTGT
602▶ W V D Q S L I K Y D E N G N C P W S A Y G G D F G D T P N D R Q F C

Eco47III (3641)
3601 ATGAACGGTCTGGTCTTTGCCGACCGCACGCCGATCCAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTTCCAGTTCGGTTTATCCGGGCAAAACCA
636▶ M N G L V F A D R T P H P A L T E A K H Q Q Q F F Q F R L S G Q T

SacI (3746)
3701 TCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGCTCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCC
669▶ I E V T S E Y L F R H S D N E L L H W M V A L D G K P L A S G E V P
3801 TCTGGATGCTGCCACAAGGTAACAGTTGATGAAGTGCCTGAAGTACCGCAGCCGAGAGCGCCGGGCAACTCTGGCTCACAGTACCGTAGTGCAA
702▶ L D V A P Q S G K Q L I E L P E L P Q P E S A G Q L W L T V R V V Q
3901 CCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGCGCTGGCAGCAGTGGCTGTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCTCCACG
736▶ P N A T A W S E A G H I S A W Q Q W R L A E N L S V T L P A A S H
4001 CCATCCCGCATCTGACCACCGAAATGGATTTTGCATCGAGCTGGTGAATAAGCGTTGGCAATTAACGCCAGTCAGGCTTTCTTTCACAGATGTG
769▶ A I P H L T T S E M D F C I E L G N K R W Q F N R Q L S G F L A D S Q M W
4101 GATTGGCGATAAAAAACAAGTCTGACGCCGCTGGCGATCAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGATTGAC
802▶ I G D K K Q L L T P L R D Q F T R A P L D N D I G V S E A T R I D
4201 CCTAACGCTGGTGAACGCTGGAAGCGGGCGCCATTACCAGGCCAAGCAGCGTGTTCAGTGCACGGCAGATACACTGTGTADATGATGCTGA
836▶ P N A W P E R W K A A G H Y Q A E A A L L Q C T A D T L A D A V L
4301 TTACGACCGCTCAGCGTGGCAGCATCAGGGGAAAACCTATTATCAGCCGAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGA
869▶ I T T A H A W Q H Q G K T L F I S R K T Y R I D G S G Q M A I T V D
4401 TGTGAAGTGGCGAGCAGATACCCGCATCCGGCGGATTGGCTGAAGTCCAGCTGGCGCAGGTAGCAGAGCGGGTAACTGGCTCGATTAGGGCCG
902▶ V E V A S D T P H P A R I G L N C Q L A Q V A E R V N W L G L G P

BsiWI (4576)
Bst1107I (4568)
4501 CAAGAAAATATCCCGACCGCCTTACTGCCGCTGTTTTGACCGCTGGATCTGCCATTGTCAGACATGTATACCCGCTACGCTTCCCGAGCGAAAACG
936▶ Q E N Y P D R L T A A C F D R W D L P L S D M Y T P Y V F P S E N
4601 GTCTCGCTGCGGGACGCGCAATTGAATTATGGCCACACAGTGGCGCGGCGACTTCCAGTTCACATCAGCCGCTACAGTCAACAGCAACTGATGGA
969▶ G L R C G T R E L N Y G P H Q W R G D F Q F N I S R Y S Q Q Q L M E

NdeI (4763)
4701 AACCGCCATCGCCATCTGCTGCACCGGAAGAAGGCACATGGCTGAATATCGACGTTTCCATATGGGATTGGTGGCGAGACTCTCGAGCCCGTCA
1002▶ T S H R H L L H A E E G T W L N I D G F H M G I G G D D S W S P S

NheI (4885)
EcoRI (4879)
4801 GTATCGCGGAATTACAGCTGAGCGCGGTGCTACCATTACAGTGGTCTGGTGTCAAAAATAATAATCTAGTCGAGAATTCGCTAGCTCGACATGAT
1036▶ V S A E L Q L S A G R Y H Y Q L V W C Q K •
4901 AAGATACATTGATGAGTTTGGACAAACCACTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTTGTGAAATTT

MfeI (5059)
5001 GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTACAGGGGAGGTGTG

SwaI (5150)
5101 GGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAGATCCATTTAAATGTTAATTAAGTACGATGACCAAAATCCCTAACGTTGAGTTTTCGT
5201 TCCACTGAGCGTCAGACCCGTCAGAAAAGTCAAGGATCTTCTTGGATCCTTTTTTCTGCGGTAATCTGCTGCTTGCAAAACAAAAAACCCGCT
5301 ACCAGCGGTGTTTGTTCGCCGATCAAGAGCTACCAACTCTTTTTCCGAAGTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGTG
5401 TAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGT
5501 CGTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGTTCTGTCACACAGCCAGCTTGGAGCGAAC
5601 GACCTACACCGAACTGAGATACCTACAGCGTGGCTATGAGAAAGCGCCACGCTTCCGAAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTC
5701 GGAACAGGAGAGCGCAGGGGAGCTTCCAGGGGAAACGCGCTGTTATCTTATAGTCTGTGCGGTTTCCGCCACTCTGACTTGAGCGTCTGATTTTTGT
5801 GATGCTCGTCAAGGGGCGGAGCTATGAAAAACGCCAGCAACCGGCTTTTTACGGTTCCTGCGCTTTTGTGCGCTTTTGTCAATGTTCTTAAT

AseI (5926) **SfiI (5977)** **MscI (5988)**
5901 TAAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACC
6000 AGTGTGTCCAGTGTCTCACAGCCAGGGATGTGGCTGGAGCTGTTGAGTCTGGACTGACAGGTTGGGGTCTCCAGAGATTTTGTGGAGGATGACTTTG
6▶ S A V P V L T A R D V A G A V E F W T D R L G F S R D F V E D D F
1▶ M A K L T

SexAI (6147)
6100 CAGGTGTGGTCAGAGATGATGTCACCTGTTCTCAGCAGTCCAGGACCGGTTGGCTGACAACACCTGGCTTGGGTGTGGTGGAGGACTGGA
39▶ A G V V R D D V T L F I S A V Q D Q V V P D N T L A W V W V R G L D
6200 TGAGCTGTATGCTGAGTGGAGTGGTGGTCTCCACCAACTTCAGGGATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAG
72▶ E L Y A E W S E V V S T N F R D A S G P A M T E I G E Q P W G R E

DraIII (6338) **SfiI (6386)**
6300 TTTGCCCTGAGAGACCCAGGCAACTGTGTGCACTTTGTGGCAGAGGAGCAGGACTGAGGATAAGAAATTGAGTTTCAGAAAAGGGGCGCTGAGTGGCC
106▶ F A L R D P A G N C V H F V A E E Q D •
6400 CCTTTTTTCAACTTAATTA