



PvuI (7)
SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAAGTATGTCGTGTACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGCGCCCGCCCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552) **NcoI (560)**
501 TCTGTTTCTGCCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATGGATCTGACAAAAATGGGCATGATCCAGTGCAGAA
1▶ M D L T K M G M I Q L Q N
601 CCCTAGCCACCCACGGGCTACTGTGCAAGGCCAACCAGATGCGGCTGGCCGGGACTTTGTGCGATGGTGCATCATGGTGGACAGCCAGGAGTTCCAC
13▶ P S H P T G L L C K A N Q M R L A G T L C D V V I M V D S Q E F H
701 GCCACCGGACGGTGTGCCTGCACCAGCAAGATGTTTGAGATCCTCTCCACCGCAATAGTCAACTATACTTTGGACTTCTCTCGCCAAAGACCT
47▶ A H R T V L A C T S K M F E I L F H R N S Q H Y T L D F L S P K T

801 TCCAGCAGATTCTGGAGTATGCATATACAGCCACGCTGCAAGCCAAGCGGAGACCTGGATGACCTGCTGTATGCGGCCGAGATCCCTGGAGATCGAGTA
NsiI (818) **EagI (875)**
80▶ F Q Q I L E Y A Y T A T L Q A K A E D L D D L L Y A A E I L E I E Y

901 CCTGGAGGAACAGTGCCTGAAGATGCTGGAGACCATCCAGGCCTCAGACGACAATGACACGGAGGCCACCATGGCCGATGGCGGGCCGAGGAAGAAGAG
StuI (938) **SfiI (963)**
113▶ L E E Q C L K M L E T I Q A S D D N D T E A T M A D G G A E E E

1001 GACCGCAAGGCTCGGTCAAGAACATCTTCATCTCGAAGCATTCCAGCGAGGAGAGTGGGTATGCCAGTGTGGCTGGACAGAGCCTCCCTGGGCCCA
Acc65I (1013) **NcoI (1097)**
147▶ D R K A R Y L K N I F I S K H S S E E S G Y A S V A G Q S L P G P
1101 TGGTGGACAGAGCCCTTCAGTCTCCACTTCATTTGGTCTTTCAGCCATGAGTCCCACCAAGGCTGCAGTGGACAGTTTGTGACCATAGGACAGTCTCT
180▶ M V D Q S P S V S T S F G L S A M S P T K A A V D S L M T I G Q S L
1201 CCTGCAGGGAACCTTTCAGCCACCTGCAGGGCCGAGGAGCAACTCTGGCTGGGGTGGGGCCACCTGGGGTGGCTGAGGTGAAGACGGAGATGATG
213▶ L Q G T L Q P P A G P E E P T L A G G G R H P G V A E V K T E M M
1301 CAGGTGGATGAGTGGCCAGCCAGGACAGCCCTGGGGCAGCCGAGTCCAGCATCTCAGGAGGGATGGGGACAAGGTTGAGGAAAGAGGCAAGAGGGGC
247▶ Q V D E V P S Q D S P G A A E S S I S G G M G D K V E E R G K E G

1401 CTGGGACCCGACTCGAAGCAGCGTCATCACCAGTGTAGGGAGTACACTATGGGCGAGAGGAGAGTGGCAGCAGGTGCCACCCAGCTGAGGCTGG
SandI (1402) **MscI (1497)**
280▶ P G T P T R S S V I T S A R E L H Y G R E E S A E Q V P P P A E A G
1501 CCAGGCCCCACTGGCCGACCTGAGCACCCAGACCCCGCCTGAGAAGCATCTGGGCATCTACTCCGTGTTGCCAACCAAGGCTGACGCTGTATTG
313▶ Q A P T G R P E H P A P P P E K H L G I Y S V L P N H K A D A V L

1601 AGCATGCCGCTTCCGTGACCTCTGGCCTCCAGTGCAGCCTGCCCTGGCTGTCTCCATGGACTTCAGCACCTATGGGGGGCTGCTGCCCCAGGGCTTCA
SphI (1601) **BbrPI (1630)** **NcoI (1655)**
347▶ S M P S S V T S G L H V Q P A L A V S M D F S T Y G G L L P Q G F
1701 TCCAGAGGGAGCTGTTAGCAAGCTGGGGAGCTGGCTGTGGGCATGAAGTCAAGAGCCGACCATCGGAGAGCAGTGCAGCGTGTGTGGGTGCGAGCT
380▶ I Q R E L F S K L G E L A V G M K S E S R T I G E Q C S V C G V E L

1801 TCCTGATAACGAGGCTGTGGAGCAGCACAGGAAGCTGCACAGTGGGATGAAGACGTACGGGTGCGAGCTCTGCGGAAGCGGTTCTGGATAGTTGCGG
BsiWI (1853)
413▶ P D N E A V E Q H R K L H S G M K T Y G C E L C G K R F L D S L R

1901 CTGAGAATGCACTTACTGGCTCATTGAGCGGGTCCAAAGCCTTTGTCTGTGATCAGTGGGTGCACAGTTTTTGAAGGAGGATGCCCTGGAGACACACA
BstBI (1971)
447▶ L R M H L L A H S A G A K A F V C D Q C G A Q F S K E D A L E T H

2001 GGCAGACCCATACTGGCACTGACATGGCCGCTTCTGTCTGTGTGGGAAGCGCTTCCAGGCGCAGAGCGCACTGCAGCAGCACATGGAGGTCCACGC
Eco47III (2051)
480▶ R Q T H T G T D M A V F C L L C G K R F Q A Q S A L Q Q H M E V H A

2101 GGGCGTGCAGCTACATCTGCAAGTGAAGTGAACCGCACCTTCCAGCCACACGGCTCTCAAACGCCACCTGCGCTCACATACAGGCGACCCCTAC
FspI (2105) **BstAPI (2120)**
513▶ G V R S Y I C S E C N R T F P S H T A L K R H L R S H T G D H P Y
2201 GAGTGTGAGTTCTGTGGCAGTCTTCCGGGATGAGAGCACACTCAAGAGCCAAACGCATCCACACGGGTGAGAAACCTACGAGTGAATGGCTGTG
547▶ E C E F C G S C F R D E S T L K S H K R I H T G E K P Y E C N G C

2301 GCAAGAAGTTCAGCCTCAAGCATCAGCTGGAGACGCACTATAGGGTGCACACAGGTGAGAAGCCCTTTGAGTGAAGCTCTGCCACCAAGCGCTCCCGGGA
XmaI (2393) **Eco47III (2387)**
580▶ G K K F S L K H Q L E T H Y R V H T G E K P F E C K L C H Q R S R D

2401 CTAAGTGCACAGTACTGCCCCAGCCTCTCCTCCATGCAG
BstAPI (2456) **ScaI (2471)**
613▶ Y S A M I K H L R T H N G A S P Y Q C T I C T E Y C P S L S S M Q

NheI (2590)

2501 AAGCACATGAAGGGCCACAAGCCCGAGGAGATCCCGCCGACTGGAGGATAGAGAAGACGTACCTCTACCTGTGCTATGTGTGAAGGGAGGCTAGCTGGC
 647▶ K H M K G H K P E E I P P D W R I E K T Y L Y L C Y V • —
 2601 CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATT

HpaI (2728) MfeI (2739)

2701 TGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAG

EcoRI (2824)

2801 TAAAACCTCTACAAATGTGGTATGGAATTCAAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAA

2901 TAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTT

SapI (3006)**SspI (3063)**

SwaI (3077)

3001 GAACTAGCTCTTCATTTCTTTATGTTTTAAATGCAGTACCTCCCACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGA

3101 AAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATA

3201 GAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTATC

141◀ • N R T Y K L P I L E E I T T K V L K G N M

BstXI (3367)

3301 TCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGT

118◀ E I L V F 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

3401 AGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTTCAGCACAGACAGTGCACCTGCCAAT

85◀ P H R V 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

StuI (3502)

3501 GTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGTATCTTC

52◀ Y A E I 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

3601 TCAGTGGCAGCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATAC

18◀ E T A V E V L E L D Q Q S I N F T K M

AseI (3710)

3701 TATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACAC

SpeI (3865)

3801 GCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGG

SnaBI (3993)

3901 GTGGAGACTTGAAATCCCCTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAG

NdeI (4098)

4001 ATGTAAGTCCAAAGTAGGAAAGTCCATAAAGTCAATGTAAGTGGGCGGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAG

4101 TATGATACTTGTGTAAGTGGGCGGAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATA

PacI (4284) **BspLU11I (4294)**

4201 CGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACCGCTGCAGGTTAATTAAGAACATGT

4301 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGAGC

4401 CTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGCTT

4501 ACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGC

4601 TGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGTGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACAGACTTATCGCCACT

4701 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACA

4801 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTAGTCCAACCCGGTAAGACAGACTTATCGCCACT

4901 TTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAACGAAAACTCACG

EagI (5044)

PacI (5024) SwaI (5033) **NotI (5043)**

5001 TTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTGTGTG

5101 AATCGTAACATAACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCT

5201 CTATCGAA