



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGCGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGTCAAGTGCAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT

NcoI (560)
501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCCTACCTGAGATCACCGTACCATGGCACATACTTTAGAGGATGCTCATTGGCATTAT
1 M A H T F R G C S L A F M

BbrPI (609) **BamHI (691)**
601 GTTTATAATCACGTGGCTGTTGATTAAGCAAAAATAGATGCGTGAAGAGAGGCGATGTGACTGTGAAGCCTTCCCATGTAATTTTACTTGGATCCACT
13 F I I T W L L I K A K I D A C K R G D V T V K P S H V I L L G S T

SspI (703) **BsrGI (770)**
701 GTCATATTACATGCTCTTTGAAGCCCAGACAAGGCTGCTTTCACTATCCAGACGTAACAAGTTAATCCTGTACAAGTTTGACAGAAGAATCAATTTTC
47 V N I T C S L K P R Q G C F H Y S R R N K L I L Y K F D R R I N F

NcoI (801)
801 ACCATGGCCACTCCCTCAATTCTCAAGTACAGGTCTTCCCCTTGGTACAACCTGTTTGTCTGCAAAGTGGCCTGTATCAATAGTGTGAAATTCAAAT
80 H H G H S L N S Q V T G L P L G T T L F V C K L A C I N S D E I Q I

BglII (911)
901 ATGTGGAGCAGAGATCTTCGTTGGTGTGCTCCAGAACAGCCTCAAAATTTATCCTGCATACAGAAGGGAGAACAGGGGACTGTGGCCTGCACCTGGAA
113 C G A E I F V G V A P E Q P Q N L S C I Q K G E Q G T V A C T W E

Bst1107I (1029)
1001 AGAGGACGAGACACCCACTTATACACTGAGTATACTCTACAGTAAGTGACCAAAAAATTTAACCTGGCAGAAGCAATGTAAGACATTTATTGTGACT
147 R G R D T H L Y T E Y T L Q L S G P K N L T W Q K Q C K D I Y C D
1101 ATTTGGACTTTGGAATCAACCTCACCCCTGAATCACCTGAATCCAATTTACAGCCAAGGTTACTGCTGTCAATAGTCTTGAAGCTCCTTCACTTCC
180 Y L D F G I N L T P E S P E S N F T A K V N S L G S S S L P
1201 ATCCACATTCACATTTCTTGACATAGTGAAGCCTCTTCTCCGTGGGACATTAGAATCAAATTTCAAAGGCTTCCGTGAGCAGATGTACCCCTTTATTGG
213 S T F T F L D I V R P L P P W D I R I K F Q K A S V S R C T L Y W

EcoRV (1334)
1301 AGAGATGAGGACTGGTACTGCTTAATCGACTCAGATACCGCCAGTAACAGCAGGCTCTGGAATATGGTTAATGTTACAAAGGCCAAAGGAAGACATG
247 R D E G L V L L N R L R Y R P C P S N S R L W N M V N V T K A K G R H
1401 ATTTGCTGGACTGAAACCATTACAGAATATGAATTTAGATTTCTCTAAGCTACATCTTTATAAGGGAAGTTGGAGTGATTGGAGTGAATCATTGAG
280 D L L D L K P F T E Y E F Q I S S K L H L Y K G S W S D W S E S L R

SapI (1516)
1501 AGCACAAACACCAGAAGAAGCCTACTGGGATGTTAGATGTCTGGTACATGAAACGGCACATTGACTACAGTAGACAACAGATTTCTCTTTTCTGGAAG
313 A Q T P E E E P T G M L D V W Y M K R H I D Y S R Q Q I S L F W K
1601 AATCTGAGTGTCTCAGAGGCAAGAGGAAAAATTTCTCCACTATCAGGTGACCTTGACAGGAGCTGACAGGAGGAAAGCCATGACACAGAACATCACAGGAC
347 N L S V S E A R G K I L H Y Q V T L Q E L T G G K A M T Q N I T G

PshAI (1710)
1701 ACACCTCTGGACCACAGTCAATCCTAGAACCGGAAATTTGGGCTGTGGCTGTGTCTGACGAAATTTCAAAGGAGTTCTCTGCCACTCGTATTAACAT
380 H T S W T T V I P R T G N W A V A V S A A N S K G S S L P T R I N I
1801 AATGAACCTGTGTGAGGCAGGGTGTGCTGCTCCTGCCACGTCTCTGCAAAGTCAAGGGCATGGACAACATTCTGGTACTTGGCAGCCTCCAGGAAA
413 M N L C E A G L L A P R H V S A N S E G M D N I L V T W Q P P R K

EcoO109I (1964) **SalI (1988)**
1901 GATCCCTCTGCTGTTTCAAGAGTACGTGGTGAATGGAGAGAGCTCCATCCAGGGGTTGACACACAGGTCCTTAAACTGGTACGGAGTCGACCCATCA
447 D P S A V Q E Y V V E W R E L H P G G D T Q V P L N W L R S R P Y
2001 ATGTGTCTGCTGATTTTCAAGAACATAAAATCCTACATCTGTTATGAATCCGTGTGTATGCACTCTCAGGGGATCAAGGAGGATGCAGCTCCATCCT
480 N V S A L I S E N I K S Y I C Y E I R V Y A L S G D Q G G C S S I L

DraIII (2123) **AseI (2140)**
2101 GGGTAACTCTAAGCACAAAGCACCACTGAGTGGCCCACTTAATGCCATCACAGAGGAAAAGGGGAGCATTTTAATTTTATGGAACAGCATTCCAGTC
513 G N S K H K A P L S G P H I N A I T E E K G S I L I S W N S I P V
2201 CAGGAGCAAAATGGGCTGCCTCCTCATTATAGGATATACTGGAAGGAAACGGGACTCCTCACTCCAGCCTCAGCTCTGTGAAATTTCCCTACAGAGTCTCCC
547 Q E Q M G C L L H Y R I Y W K E R D S N S Q P Q L C E I P Y R V S

NdeI (2339)
2301 AAAATTCACATCCAATAAACAGCCTGCAGCCCCGAGTGACATATGTCCTGTGGATGACAGCTCTGACAGCTGCTGGTGAAGTTCCACGGAAATGAGAG
580 Q N S H P I N S L Q P R V T Y V L W M T A L T A A G E S S H G N E R

PshAI (2337)
2401 GGAATTTTGTCTGCAAGGTAAGCCAATTTGGATGGCGTTTGTGGCACCAAGCATTGTCATTGCTATCATGTTGGGCAATTTCTCAACGCATTACTTC
613 E F C L Q G K A N W M A F V A P S I C I A I I M V G I F S T H Y F
2501 CAGCAAAAGGTGTTTGTCTCCTAGCAGCCTCAGACCTCAGTGGTGTAGCAGAGAAATTCAGATCCAGCAATAGCACTTGGCCTAAGAAATATCCCA
647 Q Q K V F V L L A A L R P Q W C S R E I P D P A N S T C A K K Y P

XmnI (2686)
2601 TTGCAGAGGAGAAGACAGCTGCCCTTGACAGGCTCCTGATAGACTGGCCACGCCTGAAGATCCTGAACCCTGGTTCATCAGTGAAGTCCCTTCATCA
680 I A E E K T Q L P L D R L L I D W P T P E D P E P L V I S E V L H Q

2701 AGTGACCCAGTTTTTCAGACATCCCCCTGCTCCAAGTGGCCACAAGGGAAAAAGGAATCCAAGGTCATCAGGCCTCTGAGAAAGACATGATGCACAGT
713▶ V T P V F R H P P C S N W P Q R E K G I Q G H Q A S E K D M M H S

2801 GCCTCAAGCCACACCTCCAAGAGCTCTCAAGCTGAGAGCAGACAAGTGGTGGATCTGTACAAGGTGCTGGAGAGCAGGGGCTCCGACCCAAAGCCAG
747▶ A S S P P P P R A L Q A E S R Q L V D L Y K V L E S R G S D P K P

2901 AAAACCCAGCCTGTCCCTGGACGGTCTCCAGCAGGTGACCTCCACCCATGATGGCTACTTACCCTCCAACATAGATGACCTCCCCTCACATGAGGC
780▶ E N P A C P W T V L P A G D L P T H D G Y L P S N I D D L P S H E A

3001 ACCTCTCGCTGACTCTGGAAGAACTGGAGCCTCAGCACATCTCCCTTTCTGTTTTCCCTCAAGTTCTTTCACCCACTCACCTTCTCCTGTGGTGT
813▶ P L A D S L E E L E P Q H I S L S V F P S S S L H P L T F S C G D

3101 AAGCTGACTCTGGATCAGTTAAAGATGAGGTGTGACTCCCTCATGCTCTGAGTGGTGGGCTTCTAGCTGGCCAGACATGATAAGATAACATTGATGAG
847▶ K L T L D Q L K M R C D S L M L •

3201 TTTGGACAAACCACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACATTATAAGCTGCAATAAAC

HpaI (3302) **EcoRI (3398)**
3301 AAGTTAACAAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGA

3401 ATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTG

3501 CCAATGTGCATTAGCTGTTTCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAACCTAGCTCTTCAATTTCTTTATGTT

3601 TTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAA

3701 TCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCT

3801 TCTAGCTTTAGTTCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGC
141▶ • 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

3901 ATAGTCAGAGATGAGCTCTGTCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTG
110▶ 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 P H R V A V I T

4001 TCAAAGTCCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGA
76▶ 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 Y A E I H V A S I

4101 TGATCTCCCACTTGGTCTGATGGCCGCCGACATGGTGTGTTGCTTCATAGAGCATGGTGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAG
43▶ I E G 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

4201 ATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTACTATGCCGATATACTATGCCGATGATTAATTGTCAAACA
10▶ D Q Q S I N F T K M

4301 GCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTTCGTCATGGG

4401 GCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGT

4501 CAAACCGCTATCCAGGCCATTGATGTAAGTCCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCA

4601 TAAGTCACTGACTGGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAG

4701 TGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGG

4801 GGTCGTTGGCGGTGAGCCAGGCGGGCATTACCCTAAGTTATGTAACGCTGAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAGGCCAGG

4901 AACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAATCAGCGCTCAAGTCAGAGGTGGCGAAACCCGA

5001 CAGGACTATAAGATAACCAGGCGTTTTCCCTGGAAAGTCCCTCGTGGCTCTCTGTTCCGACCCTGCCCTTACGGATACCTGTCCGCTTTCTCC

5101 TTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTGCACGAACCCCGCTT

5201 CAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGTAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTA

5301 GCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAA

5401 GCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGC

5501 AGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTT

EagI (5618)

Swal (5607) NotI (5617)

5601 AATTAACATTTAAATC AGCGGCCGC AATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGAATCGTAACTAACATACGCTCTCCAT
5701 CAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA