



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGACAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **MfeI (82)**

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCGCTACCTGAGGCC **HindIII (245)** **Bsu36I (291)**

301 GCCATCCACGCGGGTTGAGTCGCGTTTTCGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)
BstEII (555)
KasI (535) 501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCCATGGAAGCGGGACCCCTGGGAGCAGCTGCGGGCGCTTA **AgeI (552)** **SmaI (570)**

601 CCTGCCCCCTGCAGCAGGTGTTCCAGGCACCTCGCCGCGCTGGCATTGGCACTGTGGGAAACCAATCAAGCTCCTGGCCAATTACTTTGAGGTGGAC **1 M E A G P S G A A A G A Y**

701 ATCCCTAAGATCGACGTGTAACACTACGAGGTGGACATCAAGCCGGATAAGTGTCCCGTAGAGTCAACCGGGAAGTGGTGAATAACATGGTCCAGCATT **13 L P P L Q Q V F Q A P R R P G I G T V G K P I K L L A N Y F E V D**

801 TCAAGCCTCAGATCTTTGGTGATCGCAAGCCTGTGTATGATGGAAGAAGAACATTTACTGTACAGCACTGCCATTGGCAACGAACGGGTGACTT **47 I P K I D V Y H Y E V D I K P D K C P R R V N R E V V E Y M V Q H**

BglIII (809) 801 TCAAGCCTCAGATCTTTGGTGATCGCAAGCCTGTGTATGATGGAAGAAGAACATTTACTGTACAGCACTGCCATTGGCAACGAACGGGTGACTT **80 F K P Q I F G D R K P V Y D G K K N I Y T V T A L P I G N E R V D F** **SalI (892)**

901 TGAGGTGACAATCCCTGGGAAGGAAGGATCGAATCTTTAAGGTCTCCATCAAGTGGCTAGCCATTGTGAGCTGGCAATGCTGCATGAGGCCCTGGTC **NheI (957)**

1001 AGCGGCCAGATCCCTGTTCCCTTGGAGTCTGTGCAAGCCCTGGATGTGGCCATGAGGCACCTGGCATCCATGAGGTACACCCCTGTGGCCGCTCCTTCT **113 E V T I P G E G K D R I F K V S I K W L A I V S W R M L H E A L V**

XcmI (1049) 1001 AGCGGCCAGATCCCTGTTCCCTTGGAGTCTGTGCAAGCCCTGGATGTGGCCATGAGGCACCTGGCATCCATGAGGTACACCCCTGTGGCCGCTCCTTCT **BsrBI (1089)**

1101 TCTCACCGCTGAGGGCTACTACCACCCGCTGGGGGTGGGCGCAGGCTGTGTTTCGGCTTTACCAGTCTGTGCGCCCTGCCATGTGGAAGATGATGCT **147 S G Q I P V P L E S V Q A L D V A M R H L A S M R Y T P V G R S F**

Bsu36I (1108) 1101 TCTCACCGCTGAGGGCTACTACCACCCGCTGGGGGTGGGCGCAGGCTGTGTTTCGGCTTTACCAGTCTGTGCGCCCTGCCATGTGGAAGATGATGCT **180 F S P P E G Y Y H P L G G G R E V W F G F H Q S V R P A M W K M M L**

1201 CAACATTGATGTCAGCAGCTGCCTTTTATAAGGCACAGCCAGTGATTGAGTTCATGTTGAGGTGCTGGACATCAGAACATAGATGAGCAGCCCAAG **213 N I D V S A T A F Y K A Q P V I E F M C E V L D I R N I D E Q P K**

1301 CCCCTCACGGACTCTCAGCGGTTTCGTTACCAAGGAGATCAAGGGCTGAAGTGGAAAGTCAACCCACTGTGGACAGATGAAGAGGAAGTACCGGCTGT **247 P L T D S Q R V R F T K E I K G L K V E V T H C G Q M K R K Y R V**

NheI (1420) 1401 GTAATGTTACCCGCTGCCCTGCTAGCCATCAGACATTCCTTACAGCTGGAGAGTGGACAGACTGTGGAGTGGACAGTGGCAGATTTTCAAGCAGAA **280 C N V T R R P A S H Q T F P L Q L E S G Q T V E C T V A Q Y F K Q K** **ApaLI (1470)**

1501 ATATAACCTTCAGCTCAAGTATCCCATCTGCCCTGCCTACAAGTGGCCAGGAACAAAAGCATACTACCTTCCCTAGAGGTCTGTAACTTGTGGCT **313 Y N L Q L K Y P H L P C L Q V G Q E Q K H T Y L P L E V C N I V A**

Eco47III (1604) 1601 GGCAGCGCTGTATTAATAAAGCTGACCGACAACCAGACCTCGACCATGATAAAGGCCACAGCTAGATCCGCTCCAGACAGACAGGAGGAGATCAGTCGCC **AfeI (1604)** **BsrBI (1667)**

1701 TGATGAAGAATGCCAGCTACAACCTAGATCCCTACATCCAGGAATTTGGGATCAAAGTGAAGGATGACATGACGGAGGTGACAGGGCGAGTGTGCCGC **380 L M K N A S Y N L D P Y I E F G I K V K D D M T E V T G R V L P A** **KasI (1797)**

EagI (1818) 1801 GCCATCTTGACGTACGGCGCCGGAACCGGCCATTGCCACACCCAATCAGGGTGTCTGGGACATGCGGGGGAACAGTCTTACAATGGGATTGAGATC **413 P I L Q Y G G R N R A I A T P N Q G V W D M R G K Q F Y N G I E I** **XmnI (1872)** **BsaBI (1890)**

1901 AAAGTCTGGCCATCGCTGCTTCCGACCCCAAAAAGTGTGAGAAGAGGTGCTCAAGAACTTACAGACAGCTGCGGAAGATTTCAAGGATCGGG **447 K V W A I A C F A P Q K Q C R E E V L K N F T D Q L R K I S K D A**

BsaBI (2001) 2001 GGATGCCTATCCAGGGTCAACCTTGTCTGCAAATATGCACAGGGGGCAGACAGCGTGGAGCCTATGTTCCGGCATCTCAAGAACACCTACTCAGGGCT **480 G M P I Q G Q P C F C K Y A Q G A D S V E P M F R H L K N T Y S G L**

SgrAI (2132) 2101 GCAGCTCATTATTGTCATCCTGCCAGGGAAGACCGGGTGTATGCTGAGGTGAAACGTGTCGGAGATACTCTTGGGAATGGCTACGCAGTGTGTGCAG **513 Q L I I V I L P G K T P V Y A E V K R V G D T L L G M A T Q C V Q**

Tth111I (2229) 2201 GTGAAGAAGCTGGTCAAGACCTCACCTCAGACTGTCCAACTCTGCCTCAAGTCAATGTCAAACCTTGGTGGCATTAAACAATCCTAGTCCCACACC **547 V K N V V K T S P Q T L S N L C L K I N V K L G G I N N I L V P H**

Eco47III (2300) 2301 AGCGCTCTGCCGTTTTTCAACAGCCAGTGTATTCCTGGGAGCAGATGTTACACACCCCCAGCAGGGGATGGGAAAAACCTTCTATCACAGCAGTGGT **580 Q R S A V F Q Q P V I F L G A D V T H P P A G D G K K P S I T A V V** **AfeI (2300)**

2401 AGGCAGTATGGATGCCACCCAGCGGATACTGTGCTACTGTGCGGGTACAGCGACCAGGCAAGAGATCATTGAAGACTTGTCTACATGGTGCCTGAG **613 G S M D A H P S R Y C A T T V R V Q R P R Q E I I E D L S Y M V R**

2501 CTCCTCATCAATTTACAAGTCCACCCGTTTCAAGCCTACCCGATCATCTTCTACCGAGATGGGGTGCCTGAAGGCCAGCTACCCAGATCTCCACT **647 L L I Q F Y K S T R F K P T R I I F Y R D G V P E G Q L P Q I L H**

2601 ATGAGCTACTGGCCATTCGTGATGCCTGCATCAAACCTGAAAAAGGACTACCAGCCTGGGATCACTTATATTGGTGCAGAAACCCATCACACCCGCT **680 Y E L L A I R D A C I K L E K D Y Q P G I T Y I V V Q K R H H T R L**

2701 TTTCTGTGCTGACAAGAATGAGCGAATTGGGAAGAGTGGTAACATCCCAGCTGGGACCACAGTGGACCAACATCACCCACCCATTTGAGTTTGACTTC
713▶ F C A D K N E R I G K S G N I P A G T T V D T N I T H P F E F D F
2801 TATCTGTGCAGCCACGCAGGCATCCAGGGCACCAGCCGACCATCCATTACTATGTTCTTTGGGATGACAACCGTTTCACAGCAGATGAGCTCCAGATCC
747▶ Y L C S H A G I Q G T S R P S H Y Y V L W D D N R F T A D E L Q I

BsiWI (2925)
SnaBI (2923)

2901 TGACGTACCAGCTGTGCCACACTTACGTACGATGCACAGCTCTGTCTCTATCCCAGCACCTGCCTACTATGCCCGCCTGGTGGCTTTCCGGGCACGATA
780▶ L T Y Q L C H T Y V R C T R S V S I P A P A Y A R L V A F R A R Y

DraIII (3001) **SnaBI (3065)**

3001 CCACCTGGTGGACAAGGAGCATGACAGTGGAGAGGGGAGCCACATATCGGGCAGAGCAATGGCGGGACCCCGCTGGCCAAAGCCGTGCAGGTT
813▶ H L V D K E H D S G E G S H I S G Q S N G R D P Q A L A K A V Q V

FspI (3113) **AvrII (3164)**

3101 CACCAGGATACTCTGCGCACCATGTACTTCGCTTGAAGGCAGAACGCTGTTACCTCACTGGATACCTAGGCTAGCTGGCCAGACATGATAAGATACATT
847▶ H Q D T L R T M Y F A •

3201 GATGAGTTTGGACAAACCACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCA

HpaI (3308) **MfeI (3319)**

3301 AATAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGG

EcoRI (3404)

3401 TATGGAATTCTAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGG
▶ ◀

3501 CTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTGAACTAGCTCTTCATTTCTT

SspI (3643) **SwaI (3657)**

3601 TATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAAATGTTTTTTATTAG

3701 GCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATGGACAGCAAGAAAG

3801 CGAGCTTCTAGCTTTAGTTCCTGGTGTACTTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCTCAATGAGCACAAAGCAGTC
◀ 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

BstXI (3947)

3901 AGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACA
112▶ 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 P H R V A V

StuI (4082)

4001 ATGGTGTCAAAGTCTTCTGCCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGCAG
78▶ 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 Y A E I H V A

4101 CAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTCTCATAGAGCATGGTATCTTCTCAGTGGCAGCCTCCACCAAG
45▶ 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 E T A V E V L

BspHI (4232)

XmnI (4224) **AseI (4290)**

4201 CTCCAGATCTGCTGAGAGATGTTGAAGTCTTTCATGATGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTC
12▶ E L D Q Q S I N F T K M ◀

4301 AAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTC

SpeI (4445)

4401 AATGGGGCGAGTTGTTACGACATTTTGGAAAGTCCCCTGATTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCC

SnaBI (4573)

4500 CGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAA

NdeI (4678)

4600 AGTCCATAAGGTGATGACTGGGCATAATGCCAGGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAC

4700 TGCCAAGTGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAAT

SdaI (4856) **PacI (4864)** **BspLU11I (4874)**

4800 GGGCGGGGTCGTTGGCGGTGACGCAGGGCCGCTTACCGTAAGTTATGTAACGCC T G C A G G T T A A T T A A G A A C A T G T G A G C A A A G G C C A G C A A
◀

4898 AAGCCAGGAACCGTAAAAAGCCCGCTTGGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGC

4998 GAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGC

ApaLI (5188)

5098 CTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAA

5198 CCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA

5298 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCCG

5398 TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTATCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAG

5498 ATTACGCGCAGAAAAAAGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACTCACGTTAAGGGATTTTGGTCA

EagI (5624)

PacI (5604) **SwaI (5613)** **NotI (5623)**

5598 TGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGGTTTTTTTGTGTAATCGTAACTAACATAC

5698 GCTCTCCATCAAAAACAAACGAAACAAACAACTAGCAAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA