



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **EcoNI (96)**

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC **HindIII (245)** **EcoNI (287)**

301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT **NgoMIV (441)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCACCATGGCCCAATGGAATCAGCTACAGCAGCTTGACACACG **NcoI (560)** **BstEII (555)** **AgeI (552)** **Acc65I**

1 M A Q W N Q L Q Q L D T R **NdeI (692)**

601 GTACCTGGAGCAGCTCCATCAGCTCTACAGTGACAGCTTCCAATGGAGCTGCGGCAGTTTCTGGCCCTTGATTGAGAGTCAAGATTGGGCATATGCG **Y L E Q L H Q L Y S D S F P M E L R Q F L A P W I E S Q D W A Y A**

701 GCCAGCAAAGAATCACATGCCACTTTGGTGTTCATAATCTCCTGGGAGAGATTGACCAGCAGTATAGCCGCTTCTGCAAGAGTCAAGATTGTTCTCTATC **A S K E S H A T L V F H N L L G E I D Q Y S R F L Q E S N V G L Y**

801 AGCACAATCTACGAAAGTCAAGCAGTTTCTTCAGAGCAGTATCTTGAGAAGCAATGAGATTGCCCGATTGTGGCCCGGTGCTGTGGGAAGAATC **Q H N L R R I K Q F L Q S R Y L E K P M E I A R I V A R C L W E E S**

901 ACGCCTTCTACAGACTGCAGCCACTGCGGCCAGCAAGGGGGCCAGGCCAACCCACAGCAGCCGTGGTGACGGAGAAGCAGCAGATGCTGGAGCAG **R L L Q T A A T A A Q G G Q A N H P T A A V V T E K Q Q M L E Q** **XcmI (956)**

1001 CACCTTCAGGATGTCGGGAAGAGAGTGCAGGATCTAGAACAGAAAATGAAAGTGGTAGAGAATCTCCAGGATGACTTTGATTTCAACTATAAAACCTCA **H L Q D V R K R V Q D L E Q K M K V V E N L Q D D F D F N Y K T L**

1101 AGAGTCAAGGAGACATGCAAGATCTGAATGGAACAACAGTCAAGTACCAGGAGAAGTGCAGCAGCTGGAACAGATGCTCACTGCGCTGGACCAGAT **K S Q G D M Q D L N G N N Q S V T R Q K M Q Q L E Q M L T A L D Q M**

1201 GCGGAGAAGCATCGTGAAGTGGCTGGCGGGCTTTTGTGACGATGGAGTACGTGCAGAAAACCTCACGGACGAGGAGCTGGCTGACTGGAAGAGGCGG **R R S I V S E L A G L L S A M E Y V Q K T L T D E E L A D W K R R**

1301 CAACAGATTGCTGATTGGAGGCGCCCAACATCTGCTAGATCGGCTAGAAAACCTGGATAACGTCACTTAGCAGAATCTCACTTCAGACCCGTCAC **Q Q I A C I G G P P N I C L D R L E N W I T S L A E S Q L Q T R Q**

1401 AAATTAAGAACTGGAGGAGTTGCAGCAAAAAGTTTCTACAAAGGGGACCCCATGTACAGCACCGCCGATGCTGGAGGAGAGAATCGTGGAGCTGTT **Q I K K L E E L Q Q K V S Y K G D P I V Q H R P M L E E R I V E L F** **SandI (1445)** **EagI (1465)**

1501 TAGAACTTAATGAAAAGTGCCTTTGTGGTGGAGCGGACCCCTGCATGCCATGCATCCTGACCGGCCCTCGTCATCAAGACCGGCGTCCAGTTCACT **R N L M K S A F V V E R Q P C M P M H P D R P L V I K T G V Q F T** **NsiI (1552)** **SphI (1544)** **PshAI (1581)**

1601 ACTAAAGTCAAGTTGCTGGTCAAATTCCTGAGTTGAATTATCAGCTTAAAATTAAGTGTGCATTGACAAAAGACTCTGGGGACGTTGCAGCTCTCAGAG **T K V R L L V K F P E L N Y Q L K I K V C I D K D S G D V A A L R** **BamHI**

1701 GATCCCGGAAATTAACATTCTGGGCACAAACAAAAAGTGAACATGGAAGAATCCAACAACGGCAGCCTCTCTGCAGAATTCAAACACTTGACCCT **G S R K F N I L G T N T K V M N M E E S N N G S L S A E F K H L T L** **EcoRI (1780)** **Bsu36I (1797)**

1801 GAGGAGCAGAGATGTGGGAATGGGGCCGAGCAATTTGATGCTTCCCTGATTGTGACTGAGGAGCTGCACCTGATCACCTTTGAGACCGAGGTGTAT **R E Q R C G N G G R A N C D A S L I V T E E L H L I T F E T E V Y**

1901 CACCAAGGCTCAAGATTGACCTAGAGACCCACTCCTTCCAGTTGTGGTGTCTCCAACATCTGTCAGATGCCAAATGCCTGGGCGTCCATCCTGTGGT **H Q G L K I D L E T H S L P V V V I S N I C Q M P N A W A S I L W** **StuI (1905)**

2001 ACAACATGCTGACCAACAATCCCAAGAATGTAACATTTTTTACCAAGCCCAATTTGGAACCTGGGATCAAGTGGCCGAGGTCCTGAGCTGGCAGTTCTC **Y N M L T N N P K N V N F F T K P P I G T W D Q V A E V L S W Q F S** **XcmI (2050)**

2101 CTCCACCACCAAGCGAGGACTGAGCATCGAGCAGTACTGACTGAGCAGAGAACTCTTGGGACCTGGTGTGAATTATTGAGGTTGTCAGATCACATGG **S T T K R G L S I E Q L T T L A E K L L G P G V N Y S G C Q I T W**

2201 GCTAAATTTTGAAGAAAACATGGCTGGCAAGGGCTTCTCCTTCTGGGTCTGGCTGGACAATATCATTGACCTTGTGAAAAAGTACATCCTGGCCCTTT **A K F C K E N M A G K G F S F W V W L D N I I D L V K K Y I L A L**

2301 GGAACGAGGGTACATCATGGCTTTATCAGTAAGGAGCGGGGACCTTTGAGCATAAGCCTCCAGGCACCTTCTGTAAGATTCAAGTAAAG **W N E G Y I M G F I S K E R E R A I L S T K P P G T F L L R F S E S**

2401 CAGCAAAGAAGGAGGCGTCACTTTCACTTGGGTGGAGAAGGACATCAGCGGTAAGACCCAGATCCAGTCCGTGGAACCATACAAAAGCAGCAGCTGAAC **S K E G G V T F T W V E K D I S G K T Q I Q S V E P Y T K Q Q L N**

2501 AACATGTCATTTGCTGAAATCATCATGGGCTATAAGATCATGGATGCTACCAATATCCTGGTGTCTCCACTGGTCTATCTATCCTGACATCCCAAGG **N M S F A E I I M G Y K I M D A T N I L V S P L V Y L Y P D I P K** **BspLU11I (2501)**

2601 AGGAGGCATTGGAAAGTATTGTCGGCCAGAGAGCCAGGAGCATCCTGAAGCTGACCCAG GCGCTGCCCCATACCTGAAGACCAAGTTTATCTGTGTGAC
680▶ E E A F G K Y C R P E S Q E H P E A D P G A A P Y L K T K F I C V T
2701 ACCAACGACCTGCAGCAATACCATTGACCTGCCGATGTCCCCCGCACTTTAGATTGATGACAGTTTGAAATAATGGTGAAGGTGCTGAACCTCA
713▶ P T T C S N T I D L P M S P R T L D S L M Q F G N N G E G A E P S

BbsI (2676)

MscI (2894)
NheI (2888)

2801 GCAGGAGGGCAGTTTGTGTCCTCACCTTTGACATGGAGTTGACCTCGGAGTGCCTACCTCCCCATGTGAGGAGCTGAGAACGGAAGCTAGCTGGCCA
747▶ A G G Q F E S L T F D M E L T S E C A T S P M •

2901 GACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAATAATGCTTTATTGTGAAATTTGTGATGCTATTGCTTTATTTG

HpaI (3026)

3001 TAACCATTATAAGCTGCAATAAACAAAGTTAAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTA

EcoRI (3122)

3101 AAACCTCTACAAATGTGGTATGGAATTTAAAATACAGCATAGCAAACCTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATA
3201 AGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTTATGAGGTTAAGATATAGTGTATTTTCCAAGGTTTGA

SspI (3361) SwaI (3375)

3301 ACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGAAAAATTCAGAAAATAATTTAAATACATCATTGCAATGAAA
3401 ATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGAACCTTTAATAGA
3501 AATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCATCTC
141▶ N R T Y K L P I L E E I T T K V L K G N M E

SacI (3636)

3601 AATGAGCACAAAGCAGTCAGGAGCATAGTCAGATGAGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAG
118▶ 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
3701 GGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTGCACCTGCAATGT
84▶ 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 Y

StuI (3800)

3801 AGGCCTCAATGTGGACAGCAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTGTTGCTCATAGAGCATGGTATCTTCTC
51▶ 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 E

BbsI (3946)
XmnI (3942)

3901 AGTGGGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCTATGGTGGCCCTCTATAGTGAGTGTATTATACTATGCCGATATACTA
18▶ T A V E V L E L D Q Q S I N F T K M

AseI (4008) SacI (4065)

4001 TGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGC

SpeI (4163)

4101 CTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGT

SnaBI (4291)

4201 GGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGAT

NdeI (4396)

4301 GTACTGCCAAGTAGGAAAGTCCCATAAGTTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCGTACTTGGCATA
4401 TGATACACTTGTACTGCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACG

PacI (4582)
SdaI (4574) BspLU11I (4592)

4501 TCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGTACGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGA
4601 GCAAAAGGCCAGCAAAGGCCAGGAACCGTAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCT
4701 CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTAC
4801 CGGATACCTGTCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGTGTAGGTGTTCCGCTCCAAGCTG

ApaLI (4906)

4901 GGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGG
5001 CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGT
5101 ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTGATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTTT
5201 GTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTT

EagI (5342)
PacI (5322) SwaI (5331) NotI (5341)

5301 AAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGTTTTTGTGTGAA
5401 TCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCGAGTGAAGTGCAGGTGCCAGAACATTTCTCT
5501 ATCGAA