



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

101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCG

**Psp1406I (203)** 202 TGAACGTTCTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTCACGCGCCCGCCGCTACCTGAGGCCGC **Bsu36I (291)**  
**EcoNI (287)**

303 CATCCACGCCGGTTGAGTCGCGTCTGCCGCCTCCCGCCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACCGGG

404 CTTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCTGCTTGTCAACTCTACGCTTTGTTTCGTTTTCTG **NgoMIV (441)**

505 TTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGTAGGAGGGCCACCATGGTGTTCCTAATGTGGACACTGAAGAGACA **KasI (535)** **AgeI (552)** **NcoI (568)**  
121 M V F P M W T L K R Q

605 TTCTTATCCTTTTAAACATAATCCTAATTTCCAACTCCTTGGGGCTAGATGGTTTCTAAAACCTGTCCTGTGATGCTACTCTGGATGTTCCAAAGAAC **121** I L I L F N I I L I S K L L G A R W F P K T L P C D V T L D V P K N

706 CATGTGATCGTGGACTGCACAGACAAGCATTTGACAGAAATCCTGGAGGTATCCCACGAACACCACGAACCTCACCTCACCATTAACCACATACCGA **46** H V I V D C T D K H L T E I P G G I P T N T T N L T L T I N H I P D

807 CATCTCCCGAGCTCCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTTCAGATGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCA **79** I S P A S F H R L D H L V E I D F R C N C V P I P L G S K N N M C **BspLU11I (898)**  
**Clal (849)**

908 TCAAGAGGCTGCAGATTAACCCAGAAGCTTTAGTGGACTCACTTATTTAAAATCCCTTTACCTGGATGGAACCCAGCTACTAGAGATACCGCAGGGCTC **113** I K R L Q I K P R S F S G L T Y L K S L Y L D G N Q L L E I P Q G L

1009 CCGCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACACATCTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCCAAATAGAAATACTCTACCTGGG **147** P P S L Q L L S L E A N N I F S I R K E N L T E L A N I E I L Y L G

1110 CCAAACTGTTATTATCGAAATCCTTGTATGTTTCAATTAAGAGAAAGATGCCTTCTAAAACCTGACAAAAGTTAAAAGTCTCCTCGAAAGATA **180** Q N C Y Y R N P C Y V S Y S I E K D A F L N L T K L K V L S L K D

1211 ACAATGTCACAGCGTCCCTACTGTTTTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC **214** N N V T A V P T V L P S T L T E L Y L Y N N M I A K I Q E D D F N N

1312 CTCACCAATTACAAATCCTTGACCTAAGTGAAATGGCCCTGTTGTTATAATGCCCATTTCTTGTGCGCGGTGAAAAATAATTCTCCCTACAGAT **248** L N Q L Q T I L D L S G N C P R C Y N A P F P C A P C K N N S P L Q I

1413 CCCTGTAATGCTTTTATGACGCTGACAGAATTAAGTTTACGCTACACAGTAACCTCTTCCAGCATGTGCCCAAGATGGTTAAGAACATCAACA **281** P V N A F D A L T E L K V L R L H S N S L Q H V P P R W F K N I N

1514 AACTCCAGAACTGGATCTGTCCAAAACCTTGTGGCCAAAGAAATGGGGATGCTAAATTTCTGCATTTTCTCCCGAGCTCATCCAATTGGATCTGTCT **315** K L Q E L D L S Q N F L A K E I G D A K F L H F L P S L I Q L D L S

1615 TTCAATTTTGAATTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCTGAAAATCTGCGGATCAGAGGATATGCTT **349** F N F E L Q V Y R A S M N L S Q A F S L K I L R I R G Y V F

1716 TAAAGAGTTGAAAAGCTTTAACCTCTGCCATTACATAATCTTCAAAATCTTGAAGTCTTGATCTTGGCACTAACTTATAAAAAATGCTAACCTCAGCA **382** K E L K S F N L S P L H N L Q N L E V L D L G T N F I K I A N L S

**PmeI (1817)** 1817 TGTTTAAACAATTTAAAGACTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAATGCCAGA **416** M F K Q F K R L K V I D L S V N K I S P S G D S S E V G F C S N A R

1918 ACTTCTGTAGAAAGTTATGAACCCAGGCTCCTGGAACAATTACATTATTTTCAGATATGATAAGTATGCAAGGAGTTGCAGATTCAAAAAACAAAGAGGCTTC **450** T S V E S Y E P Q V L E Q L H Y F R Y D K Y A R S C R F K N K E A S

2019 TTTTCATGCTGTTAATGAAAGCTGCTACAGTATGGGCAGACCTTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCTCTGATTTTCAGCATCTTT **483** F M S V N E S C Y K Y G Q T L D L S K N S I F F V K S S D F Q H L

**XcmI (2159)** 2120 CTTTCTCAAATGCCTGAATCTGTCCAGGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATCCAACCTTTAGCAGAGCTGAGATATTTGGACTTCTCC **517** S F L K C L N L S G N L I S Q T L N G S E F Q P L A E L R Y L D F S

**XmnI (2256)** 2221 AACACCGGCTTGATTTACTCCATTCAACAGCATTGAAGAGCTTCAAAAACCTGGAAGTCTGGATATAAGCAGTAATAGCCATTATTTTCAATCAGAAGG **551** N N R L D L L H S T A F E E L H K L E V L D I S S N S H Y F Q S E G

**XcmI (2404)** **Neol (2413)** 2322 AATTACTCATATGCTAAACTTTACCAAGAACCTAAAGTTCTGCAGAACTGATGATGAACGACAATGACATCTTCTCCACCAGCAGGACCATGGAGA **584** I T H M L N F T K N L K V L Q K L M M N D N D I S S T S R T M E

2423 GTGAGTCTTAGAACTCGGAATTCAGAGAAATCACTTAGATGTTTTTATGGAGAAAGGTGATAACAGATACTTACAATTTCAAGAATCTGCTAAAA **618** S E S L R T L E F R G N H L D V L W R E G D N R Y L Q L F K N L L K

2524 TTAGAGGAATTAGACATCTCAAAAATCCCTAAGTTTCTGCTTCTGGAGTTTTGATGGTATGCCTCAAATCTAAAGAATCTCTCTTTGGCCAAAA **652** L E E L D I S K N S L S F L P S G V F D G M P P N L K N L S L A K N

**PshAI (2702)** 2625 TGGGCTCAAATCTTTTCAGTTGGAAGAACTCCAGTGTCTAAAGAACCTGGAACTTTGGACCTCAGCCACAACCAACTGACCCTGTCCTGAGAGATTAT **685** G L K S F S W K K L Q C L K N L E T L D L S H N Q L T T V P E R L

**EcoRV (2818)** 2726 CCACTGTTCCAGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTCCAGTTGCGATATCTG **719** S N C S R S L K N L I L K N N Q I R S L T K Y F L Q D A F Q L R Y L

2827 GATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACAGCTTCCAGAAAATGTCTCAACAACTGAAGATGTTGCTTTTGCATCATAATCGYTTCT **753** D L S S N K I Q M I Q K T S F P E N V L N N L K M L L L H H N R F L

**ApaLI (2927)** 2928 GTGCACCTGTGATGCTGTGGTTTGTCTGGTGGGTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCAC **786** C T C D A V W F V W W V N H T E V T I P Y L A T D V T C V G P G A

**HpaI (2961)**

3029 ACAAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATATCTGTATCTCTCTTTCTC  
820▶ H K G Q S V I S L D L Y T C E L D L T N L I L F S L S I S V S L F L  
3130 ATGGTGATGATGACAGCAAGTCACTCTATTTCTGGATGTGTGATATATAACCATTTCTGTAAGGCCAAGATAAAGGGTATCAGCGTCTAATATCAC  
854▶ M V M M T A S H L Y F W D V W Y I Y H F C K A K I K G Y Q R L I S  
PvuII (3274) BbsI (3318)  
3230 CAGACTGTTGCTATGATGCTTTTATTGTGTATGACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGA  
887▶ P D C C Y D A F I V Y D T K D P A V T E W V L A E L V A K L E D P R  
XhoI (3351) BstEII (3367)  
3331 GAGAAACATTTTAAATTTATGTCTCGAGGAAAGGACTGGTTACCAGGGCAGCCAGTTCTGAAAAACCTTTCCAGAGCATAACAGTTAGCAAAAAGACAGT  
921▶ E K H F N L C L E E R D W L P G Q P V L E N L S Q S I Q L S K K T V  
3432 GTTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTAAGATAGCATTTTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCT  
954▶ F V M T D K Y A K T E N F K I A F Y L S H Q R L M D E K V D V I I  
BspEI (3577)  
3533 TGATATTTCTTGAGAAGCCCTTTCAGAAGTCCAAGTTCCTCCAGCTCCGAAAAAGGCTCTGTGGGAGTTCTGTCTTGTAGTGGCAACAAACCCGCAAGT  
988▶ L I F L E K P F Q K S K F L Q L R K R L C G S S V L E W P T N P Q A  
BstXI (3636) BamHI (3717)  
3634 CACCCATACTTCTGGCAGTGTCTAAAGAAGCCCTGGCCACAGACAATCATGTGCCTATAGTCAAGGTTCAAGGAAACGGTGGATCCATCCCTATGA  
1022▶ H P Y F W Q C L K N A L A T D N H V A Y S Q V F K E T V G S Y P Y D  
NheI (3814)  
3735 TGTGCCAGACTATGCTGGCTATCCATATGATGTTCTGATTATGCTGGATACCCTTATGATGTCCAGACTATGCCTAAAGCTAGTGGCCAGACATGATA  
1055▶ V P D Y A G Y P Y D V P D Y A G Y P Y D V P D Y A •  
3836 AGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACATTAT

---

HpaI (3952)  
3937 AAGCTGCAATAAAACAAGTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACA  
4038 AATGTGGTATGGAATTCTAAAATACAGCATAGCAAAACTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCAT  
4139 CAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTAGCTCTTCA

---

SspI (4287) SwaI (4301)  
4239 TTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTGTAGTAAATATTGAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTT  
4340 ATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTGTAGTGTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAG  
4441 AAAGCGAGCTTCTAGCTTTAGTTCTGCTGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCA  
144▶ 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  
SacI (4562) BstXI (4591)  
4542 GTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCA  
113▶ 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 P H R V A V  
StuI (4726)  
4643 CAATGGTGTCAAAGTCTTCTGCCCGTTGCTCAGCAGACCCAAATGGCAATGGCTTTCAGCACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACA  
79▶ 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  
4744 GCAGAGATGATCTCCCAGTCTTGGTCTGTAGGCCGCCCCGACATGGTGCTTGTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAG  
45▶ 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 T A V E V L  
BspHI (4876)  
BbsI (4872)  
XmnI (4868) AseI (4934)  
4845 CTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCA  
12▶ E L D Q Q S I N F T K M  
SacI (4991)  
4946 AAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTGCGTCA

---

SpeI (5089)  
5046 ATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCG

---

SnaBI (5217)  
5146 TGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAG  
5246 TCCATAAGGTCACTGACTGGGCATAATGCCAGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGC  
5347 CAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGC

---

SdaI (5500) PacI (5508) BspLU11I (5518)  
5448 GGGGGTCTGTTGGGCGGTACGCCAGCGGGCCATTTACCGTAAGTTATGTAACGCCCTG 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 A G G C C A A A A G G  
5546 CCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC  
5647 CCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCT

---

ApaLI (5832)  
5748 CCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCG  
5849 TTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT

5950 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAA

6051 GCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCA

6152 GAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAA

PacI (6248)

**EagI (6268)**

SwaI (6257) **NotI (6267)**

6253 TTAACATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCA

6353 AAACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA