



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA **EcoNI (96)**  
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**Psp1406I (203)** 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACACGCCGCCGCCCTACCTGAGGCC **Bsu36I (291)**  
301 GCCATCCACGCCGGTTGAGTGCAGTTCGCCGCCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC **EcoNI (287)**

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT **NgoMIV (441)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCCCTACCTGAGATCACCGGTCACCATGGCGTTTACGTCAGCTGCTTTTGGTGGTGCCT **NcoI (560)**  
**BstEII (555)** 1▶ M A F D V S C F F W V V L **PvuII (576)**

601 GTTTTCTGCCGGCTGTAAGTCATCACCTCCTGGGATCAGATGTGCATTGAGAAAGAAGCCAACAAAACATATAACTGTGAAAATTTAGGTCTCAGTGAA **NgoMIV (607)**  
13▶ F S A G C K V I T S W D Q M C I E K E A N K T Y N C E N L G L S E

701 ATCCCTGACACTCTACCAACACACAGAAATTTTGAATTCAGCTTTAATTTTTTGCCTACAATTCACAATAGAACCTTCAGCAGACTCATGAATCTTA **EcoRI (736)** **BspHI (788)**  
47▶ I P D T L P N T T E F L E F S F N F L P T I H N R T F S R L M N L

801 CCTTTTTGGATTTAACTAGGTGCCAGATTAAGTGGATACATGAAGACACTTTTCAAAGCCATCATCAATTAAGCACACTTGTGTTAACTGGAAATCCCT **HpaI (882)**  
80▶ T F L D L T R C Q I N W I H E D T F Q S H H Q L S T L V L T G N P L

901 GATATTCATGGCAGAAACATCGCTTAATGGGCCAAAGTCACTGAAGCATCTTTTCTTAATCCAAACGGGAATATCCAATCTCGAGTTTATTCCAGTGCAC **Bsp120I (928)** **XhoI (979)** **ApaLI (994)**  
113▶ I F M A E T S L N G P K S L K H L F L I Q T G I S N L E F I P V H

1001 AATCTGAAAACTTGGAAAGCTTGATCTTGGAAAGCAACCATATTTCCCTCCATTAAGTCCCAAAGACTTCCAGCACGGAATCTGAAAGTACTGGAT **ScaI (1089)**  
147▶ N L E N L E S L Y L G S N H I S S I K F P K D F P A R N L K V L D

1101 TTCAGAATAATGCTATACTACATCTCTAGAGAAGACATGAGGTCTCTGGAGCAGGCCATCAACCTAAGCCTGAACCTCAATGGCAATAATGTTAAAGG **XbaI (1126)**  
180▶ F Q N N A I H Y I S R E D M R S L E Q A I N L S L N F N G N N V K G  
1201 TATTGACCTGGGGCTTTTGAATCAACGGTCTTCCAAAGTTTGAACCTTTGGAGGAACCTCAAATTTGTCTGTTATTCATGCTGCAGAACTCTACT  
213▶ I E L G A F D S T V F Q S L N F G G T P N L S V I F N G L Q N S T  
1301 ACTCAGTCTCTGGCTGGGAACATTTGAGGACATTGATGACGAAGATATTAGTTCAGCCATGCTCAAGGACTCTGTGAAATGTCTGTTGAGAGCCTCA  
247▶ T Q S L W L G T F E D I D D E D I S S A M L K G L C E M S V E S L

1401 ACCTGCAGGAACCCGCTTCTCTGACATCTCATCCACCACATTTCAAGTCTTACCCAACTCCAAGAATTGGATCTGACAGCAACTCACTTGAAGGGTT **SdaI (1401)** **BstEII (1496)**  
280▶ N L Q E H R F S D I S S T T F Q C F T Q L Q E L D L T A T H L K G L  
1501 ACCCTCTGGGATGAAGGGTCTGAACCTTGTCAAGAAATGAGTCTCAGTGAAATCATTTTCGATCAATTTGTGCAATCAGTGTGCCAATTTCCCTCC  
313▶ P S G M K G L N L L K K L V L S V N H F D Q L C Q I S A A N F P S  
1601 CTTACACCTCTACATCAGAGGCAACGTGAAGAACTTCACTTGGTGGTGGCTGCTGGAGAACTAGGAAACCTTCAGACACTTGAATTTAAGCCATA  
347▶ L T H L Y I R G N V K K L H L G V G C L E K L G N L Q T L D L S H  
1701 ATGACATAGAGGCTTCTGACTGTGAGTCTGCAACTCAAAAACCTGTCCACTTGCAAAACCTTAAACCTGAGCCACAATGAGCCTCTGGTCTCCAGAG  
380▶ N D I E A S D C C S L Q L K N L S H L Q T L N L S H N E P L G L Q S  
1801 TCAGGCATTCAAAGAATGCTCAGCTAGAACTCCTCGATTTGGCATTACCCGCTTACACATTAATGTCCCAAAGTCCCTTCAAACCTCCATTT  
413▶ Q A F K E C P Q L E L L D L A F T R L H I N A P Q S P F Q N L H F

1901 CTTCAGGTTCTGAATCTCACTTACTGCTTCTTGTATACCAGCAATCAGCATCTTCTAGCAGGCCTACCAGTCTCCGGCATCTCAACTTAAAAGGGAATC **StuI (1959)**  
447▶ L Q V L N L T Y C F L D T S N Q H L L A G L P V L R H L N L K G N  
2001 ACTTTCAAGATGGGACTATCACGAAGACCAACCTACTTTCAGACCGTGGGAGCTTGGAGGTTCTGATTTTGTCTCTTGTGGTCTCTCTATAGACCA  
480▶ H F Q D G T I T K T N L Q T V G S L E V L I L S S C G L L S I D Q  
2101 GCAAGCATTCCACAGCTTGGGAAAAATGAGCCATGTAGACTTAAAGCACAACAGCCTGACATGCGACAGCATTGATTTCTTAGCCATCTTAAGGGAATC  
513▶ Q A F H S L G K M S H V D L S H N S L T C D S I D S L S H L K G I  
2201 TACCTCAATCTGGCTGCCAACAGCATTAAACATCATCTACCCCGTCTCCTCCCTATCTTGTCCAGCAGAGCACCATTAAATTTAAGTCATAACCCCTGG  
547▶ Y L N L A A N S I N I I S P R L L P I L S Q Q S T I N L S H N P L

2301 ACTGCACTTGCTCGAATATTCTTTCTTAACATGGTACAAAGAAAACCTGCACAAAACCTGAAGGCTCGGAGGACACCGTGTGCAAACCCGCCATCTCT **SspI (2314)** **BbrPI (2376)**  
580▶ D C T C S N I H F L T W Y K E N L H K L E G S E E T T C A N P P S L  
2401 AAGGGAGTTAAGCTATCTGATGCAAGCTTCCCTGTGGGATTACAGCCATAGGCATTTTCTTCTCATAGTATTTCTATTATTGTTGGCTATTTCTGCTA  
613▶ R G V K L S D V K L S C G I T A I G I F F L I V F L L L A I L L

2501 TTTTTGCAAGTAAATACCTTCTCAGGTGGAAATACCAACACATTTAGTGTGAAGGTTTCCAGAGAAAGCAAGCTAGCTGGCCAGACATGATAAGATAC **MscI (2579)**  
647▶ F F A V K Y L L R W K Y Q H I • **NheI (2573)**  
2601 ATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAGAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCT

2701 <sup>HpaI (2711)</sup> GCAATAAACAAAGTTAACACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATG

2801 <sup>EcoRI (2807)</sup> TGGTATGGAATTCTAAAATACAGCATAGCAAACCTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAG

2901 <sup>SapI (2989)</sup> GGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTCTCTTCATTT

3001 <sup>SspI (3046)</sup> <sup>SwaI (3060)</sup> CTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTAT

3101 <sup>EcoO109I (3121)</sup> TAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATGGACAGCAAGA

3201 AAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAGCA  
 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

3301 <sup>SacI (3321)</sup> <sup>BstXI (3350)</sup> GTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGCCACCTCATCAGAGTAGGGGTGCCTGACAGCC  
 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

3401 <sup>StuI (3485)</sup> ACAATGGTGTCAAAGTCCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTTCAGCAGACAGTACCCTGCCAATGAGGCTCAATGTGGA  
 79 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 A E I H V

3501 CAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTGTTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCAC  
 46 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

3601 <sup>XmnI (3627)</sup> CAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATT  
 13 L E L D Q Q S I N F T K M

3701 <sup>SacI (3750)</sup> GTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTGCG

3801 <sup>SpeI (3848)</sup> GTCATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATC

3901 <sup>SnaBI (3976)</sup> CCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGG

4001 <sup>NdeI (4081)</sup> AAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGT

4101 ACTGCCAAGTGGGCGATTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCA

4201 <sup>SdaI (4259)</sup> <sup>PacI (4267)</sup> <sup>BspLU11I (4277)</sup> ATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAA  
 113 AAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGC

4301 AAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGC

4401 GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGC

4501 <sup>ApaLI (4591)</sup> CTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCAAGCTGGGCTGTGTGCACGAA

4601 CCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA

4701 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCCG

4801 TCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTATCCGGCAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAG

4901 ATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCA

5001 <sup>PacI (5007)</sup> <sup>SwaI (5016)</sup> <sup>EagI (5027)</sup> <sup>NotI (5026)</sup> TGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAACATAC

5101 GCTCTCCATCAAACAAACGAAACAAACAAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA