



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGTGCCTA  
**MfeI (82)**  
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**HindIII (245)**  
**Psp1406I (203)** 201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACCGCGCCGCCGCCCTACCTGAGGCC  
**PvuII (239)**  
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
**Bsu36I (291)**

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTGGCTGACCCTGTTGCTCAACTCTACGTCTTTGTTTCGTTT

**BspLU111 (560)**  
501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCAACATGATTTGTGGCTTAACTCTTGGCATTGGCTTTGC  
**AgeI (552)**  
601 CTTTCTGGACACAGAAGTATTTGTGACAGGGCAAAGCCCAACACCTTCCCCACTGGATTGACTACAGCAAAGATGCCAGTGTTCACCTTTCAAGTGAC  
13> F L D T E V F V T G Q S P T P S P T G L T T A K M P S V P L S S D  
701 CCCTTACCTACTACACCACTGCATTCTACCCGCAAGCACCTTTGAAAGAGAAAACTTCTCAGAGACCACAACCTCTCTTAGTCCAGACAATACTT  
47> P L P T H T T A F S P A S T F E R E N D F S E T T T S L S P D N T  
801 CCACCAAGTATCCCGGACTCTTTGGATAATGCTAGTGCTTTAATACCACAGGTGTTTCATCAGTACAGACGCCTCACCTTCCCACGCACGCAGACTC  
80> S T Q V S P D S L D N A S A F N T T G V S S V Q T P H L P T H A D S  
901 GCAGACGCCCTCTGCTGGAAGTACACGCAGACATTGACGGCTCCGCCGCAATGCAAACTCAACCTACCCAGGCAGCAATGCTATCTCAGATGTC  
113> Q T P S A G T D T Q T F S G S A A N A K L N P T P G S N A I S D V  
1001 CCAGGAGAGAGGAGTACAGCCAGCACCTTCTACAGACCCAGTTTCCCATTGACAACCACCTCAGCCTTGACACACAGCTCTGCTGCCTTACCTG  
147> P G E R S T A S T F P T D P V S P L T T T L S L A H H S S A A L P

**Eco47III (1186)**  
1101 CACGCACCTCAACACCACCATCACAGCGAACCTCAGATGCCTACCTTAATGCCTCTGAAACAACCACTCTGAGCCCTTCTGGAAGCGCTGTCATTC  
180> A R T S N T T I T A N T S D A Y L N A S E T T T L S P S G S A V I S

**BspLU111 (1234)**  
1201 AACCAACAACATAGCTACTACTCCATCTAAGCCAACTGTGATGAAAAATATGCAAACTCACTGTGGATTACTTATATAACAAGGAACTAAATTTATTT  
213> T T T I A T T P S K P T C D E K Y A N I T V D Y L Y N K E E T K L F  
1301 ACAGCAAAGCTAAATGTTAATGAGAATGTGGAATGTGGAACAATACTTGCACAAACAATGAGGTGCATAACCTTACAGAATGTAATAAATGCGTCTGTTT  
247> T A K L N V N E N V E C G N N T C T N N E V H N L T E C K N A S V

**BsrGI (1486)**  
1401 CCATATCTCATAATTCATGTACTGCTCCTGATAAGACATTAATATTAGATGTGCCACCAGGGGTTGAAAAGTTTCAGTTACATGATTGTACACAAAGTTGA  
280> S I S H N S C T A P D K T L I L D V P P G V E K F Q L H D C T Q V E  
1501 AAAAGCAGATACTACTATTTGTTAAAAATGAAAAATATTGAAACCTTTACTTGTGATACACAGAATATTACCTACAGATTTCAAGTGTGTAATATGATA  
313> K A D T T I C L K W K N I E T F T C D T Q N I T Y R F Q C G N M I  
1601 TTTGATAATAAGAAATTAATAGAAAACCTTGAACCGAACATGAGTATAAGTGTGACTCAGAAACTCTATAATAACCACAAGTTTACTAACGCAA  
347> F D N K E I K L E N L E P E H E Y K C D S E I L Y N N H K F T N A  
1701 GTAATAATTATAAACAGATTTTGGGAGTCCAGGAGCCCTCAGATATTTTTTGTAGAAGTGAAGCTGCACATCAAGGAGTAATTACCTGGAATCCCC  
380> S K I I K T D F G S P G E P Q I I F C R S E A A H Q G V I T W N P P  
1801 TCAAAGATCATTTTATAATTTTACCCTCTGTTATATAAAGAGACAGAAAAAGATTGCCTCAATCTGGATAAAAACTGATCAAAATGATTTGCAAAAT  
413> Q R S F H N F T L C Y I K E T E K D C L N L D K N L I K Y D L Q N  
1901 TTAACCTTATACGAAATATGTTTATCATTACATGCCTACATCATTGCAAAAGTGAACGTAATGGAAGTGTGCAATGTGTCATTTCACAACAAAA  
447> L K P Y T K Y V L S L H A Y I I A K V Q R N G S A A M C H F T T K

**XcmI (2008)** 2001 GTGCTCCTCAAGCCAGGCTGGAACATGACTGTCTCCATGACATCAGATAATAGTATGCATGTCAAGTGTAGGCCTCCAGGGACCGTAATGGCCCCA  
480> S A P P S Q V W N M T V S M T S D N S M H V K C R P P R D R N G P H  
**NsiI (2056)**

**Psp1406I (2102)** 2101 TGAACGTTACCATTGGAAGTTGAAGCTGGAAACTCTGTTAGAAATGAGTGCATAAGAATTGCGATTTCCGTGTAAGAAGATCTTCAATTTCAACA  
513> E R Y H L E V E A G N T L V R N E S H K N C D F R V K D L Q Y S T  
**BglIII (2181)**  
2201 GACTACACTTTTAAAGCCTATTTTACAATGGAGACTATCCTGGAGAACCCTTTATTTTACATCATTCAACATCTTATAATTCTAAGGCACTGATAGCAT  
547> D Y T F K A Y F H N G D Y P G E P F I L H H S T S Y N S K A L I A  
2301 TTCTGGCAATTTCTGATTATTGTGACATCAATAGCCCTGCTTGTGTTCTTACAAAATCTATGATCTACATAAGAAAAGATCCTGCAATTTAGATGAACA  
580> F L A F L I I V T S I A L L V L V L Y K I Y D L H K K R S C N L D E Q  
2401 GCAGGAGCTTGTGAAAGGGATGATGAAAAACAATGATGAATGTGGAGCCAATCCATGCAGATATTTTGTGAAACTTATAAGAGGAAGATTGCTGAT  
613> Q E L V E R D D E K Q L M N V E P I H A D I L L E T Y K R K I A D

**BbsI (2504)** 2501 GAAGGAAGACTTTTTCTGGCTGAATTTTCAAGCATCCCGCGGTTGTTTCAGCAAGTTTCTTATAAAGGAAGCTCGAAAGCCCTTTAACCAGAATAAAAAACC  
647> E G R L F L A E F Q S I P R V F S K F P I K E A R K P F N Q N K N  
**SacII (2536)**  
2601 GTTATGTTGACATTTCTCTTATGATTATAACCGTGTGAACTCTCTGAGATAAACGGAGATGCAGGTCAAACTACATAAATGCCAGCTATATTGATGG  
680> R Y V D I L P Y D Y N R V E L S E I N G D A G S N Y I N A S Y I D G

**EcoO109I (2733)**  
2701 TTTCAAAGAACCAGGAAATACATTGCTGCACAAGTCCAGGGATGAAACTGTTGATGATTTCTGGAGGATGATTTGGGAACAGAAAGCCACAGTTATT  
713> F K E P R K Y I A A Q G P R D E T V D D F W R M I W E Q K A T V I  
2801 GTCATGGTCACTCGATGTGAAGAAGGAAACAGGAACAAGTGTGCAGAATACTGGCCGTCATGGAAGAGGGCACTCGGGCTTTGGAGATGTTGTTGTA  
747> V M V T R C E E G N R N K C A E Y W P S M E E G T R A F G D V V V  
2901 AGATCAACCAGCACAAAAGATGTCCAGATTACATCATTGAGAAATGAACTGTAATAAAAAAGCAACTGGAAGAGAGGTGACTCACATTCA  
780> K I N Q H K R C P D Y I I Q K L N I V N K K E K A T G R E V T H I Q

MscI (3010) BamHI (3033)  
PvuII (3006) Bsu36I (3028)  
3001 GTTCACCAGCTGGCCAGACCACGGGGTGCCTGAGGATCCTCAGTCTGCTCCTCAAAGTGAAGGAGAGTGAATGCCTTCAGCAATTTCTTCAGTGGTCCC  
813▶ F T S W P D H G V P E D P H L L L K L R R R V N A F S N F F S G P  
ApaLI (3106)  
3101 ATTGTGGTCACTGCAGTCTGGTGTGGGGCGCACAGGAACCTATATCGGAATTGATGCCATGCTAGAAGGCCTGGAAGCCGAGAACAAAGTGGATGTTT  
847▶ I V V H C S A G V G R T G T Y I G I D A M L E G L E A E N K V D V  
3201 ATGGTTATGTTGCAAGCTAAGGCAGAGATGCCTGATGGTTCAAGTAGAGGCCAGTACATCTTGGATCCATCAGGCTTTGGTGGAAATACAATCAGTT  
880▶ Y G Y V V K L R R Q R C L M V Q V E A Q Y I L I H Q A L V E Y N Q F  
BamHI (3360)  
3301 TGGAGAAACAGAAGTGAATTTGCTGAATTACATCCATATCTACATAACATGAAGAAAAGGGATCCACCCAGTGAAGGCCTCTCCACTAGAGGCTGAATTC  
913▶ G E T E V N L S E L H P Y L H N M K R D P P S E P S P L E A E F  
3401 CAGAGACTTCCTTTCATATAGGAGCTGGAGGACACAGCACATTTGAAAATCAAGAAGAAAATAAAAGTAAAAACAGGAATTTCTAATGTCATCCCATATGACT  
947▶ Q R L P S Y R S W R T Q H I G N Q E E N K S K N R N S N V I P Y D  
3501 ATAAACAGAGTGCCACTTAAACATGAGCTGAAAATGAGTAAAGAGAGTGAAGCATGATTCAGATGAATCCTCTGATGATGACAGTGAATTCAGAGGAACCAAG  
980▶ Y N R V P L K H E L E M S K E S E H D S D E S S D D D S D S E E P S  
NsiI (3611)  
3601 CAAATACATCAATGCATCTTTTATAATGAGCTACTGAAAACCTGAAAGTATGATTGCTGCTCAGGGACCACTGAAGGAGACCATTGGTGAATTTGGCAG  
1013▶ K Y I N A S F I M S Y W K P E V M I A A Q G P L K E T I G D F W Q  
ScaI (3776)  
3701 ATGATCTTCAAAGAAAAGTCAAAGTTATTGTTATGCTGACAGAAGTGAACATGGAGACCAGGAATCTGTGCTCAGTACTGGGGAGAAGGAAAACAAA  
1047▶ M I F Q R K V K V I V M L T E L K H G D Q E I C A Q Y W G E G K Q  
3801 CATATGGAGATATTGAAGTTGACCTGAAAGACACAGACAAATCTTCAACTTATACCCTTCGTGCTTTGAACTGAGACATTTCAAAGAGAAAAGACTCTCG  
1080▶ T Y G D I E V D L K D T D K S S T Y T L R V F E L R H S K R K D S R  
3901 AACTGTGTACCAGTACCAATATACAAACTGGAGTGTGGAGCAGCTTCCGTCAGAACCCAAAGGAATTAATCTCTATGATTCAGGTCGTCAAACAAAAACTT  
1113▶ T V Y Q Y Q Y T N W S V E Q L P A E P K E L I S M I Q V V K Q K L  
4001 CCCCAGAAGAATCTCTGAAGGGAACAAGCATCACAAAGTACACCTTACTCATTCTCAGGAGGATGGATCTCAGCAACAGGGAATATTTTGTGCTT  
1147▶ P Q K N S S E G N K H H K S T P L L I H C R D G S Q Q T G I F C A  
4101 TGTTAAATCTCTTAGAAAAGTGGGAAAACAGAAGAGGTAGTGGATATTTTTCAAGTGGTAAAAGCTCTACGAAAGCTAGGCCAGGCATGGTTTCCACATT  
1180▶ L L N L L E S A E T E E V V D I F Q V V K A L R K A R P G M V S T F  
4201 CGAGCAATCAATTCCTATATGACGTCAATGACGACCTACCTGCTCAGAATGGACAAGTAAAGAAAAACAACCATCAAGAAGATAAAAATGAATTT  
1213▶ E Q Y Q F L Y D V I A S T Y P A Q N G Q V K K N N H Q E D K I E F  
4301 GATAATGAAGTGGACAAAGTAAAGCAGGATGCTAATTGTGTTAATCCACTTGGTGCACGAAAAGCTCCCTGAAGCAAAGGAAACAGGCTGAAGGTTCTG  
1247▶ D N E V D K V K Q D A N C V N P L G A P E K L P E A K E Q A E G S  
4401 AACCCACGAGTGGCACTGAGGGGCCAGAACATTCTGTCAATGGTCTGCAAGTCCAGCTTTAAATCAAGGTTCAAGGAAAAGACATAAATGAGGAAACT  
1280▶ E P T S G T E G P E H S V N G P A S P A L N Q G S •  
MscI (4516) NheI (4510)  
4501 CCAAACCTCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATT  
HpaI (4648) MfeI (4659)  
4601 TGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGT  
4701 GGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTTCAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGA  
4801 ATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTATGAGGTTTAAAGATAT  
SapI (4926) SmaI (4997)  
4901 AGTGTATTTTCCCAAGGTTTGAAGTACTGCTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATAGTAAATATTCAGAAATAATT  
EcoO109I (5058)  
5001 TAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGG  
5101 GAACAAAGGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAAGGGGATGAGTTCCTCAATGGTGGTTT  
141▶ N R T Y K L P I L E E I T T K  
SacI (5258) BstXI (5287)  
5201 TGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGA  
125▶ V L K G N M E 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  
5301 TCTGTCCACTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGCCCAATGGCAATGGCTTCAGCA  
92▶ R D V E D S Y 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  
5401 CAGACAGTGAACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGCTCT  
58▶ C V T V R G I Y 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  
BbsI (5568)  
5501 CATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTTCTCATGGTGGCCCTCTATAGTGAGTCGT  
25▶ Y L M T I K E T A V E V L E L D Q Q S I N F T K M  
SacI (5687)  
5601 ATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTAT  
SpeI (5785)  
5701 ATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAAGTCCCGTTGATTTACTAGTCAAAACAAA  
5801 CTCCCATGACGTC AATGGGGTGGAGACTTGAAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAA  
SnaBI (5913)  
5901 TAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAAGTCCATAAGGTCATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTC

6001 AATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATT

6101 GCGGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTG **SdaI (6196)**

6201 <sup>PacI (6204) BspLU1II (6214)</sup>  
CAGGTTAATTAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGAC

6301 GAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTC

6401 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT

6501 <sup>ApaLI (6528)</sup>  
GTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTA

6601 AGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA

6701 ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTTTGATCCGGCAAACAAACCAC

6801 CGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCT

6901 <sup>PacI (6944) SwaI (6953) **EagI (6964)** **NotI (6963)**</sup>  
CAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCT

7001 GTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGT

7101 GCAGGTGCCAGAACATTTCTCTATCGAA