



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

**MfeI (82)**  
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

**HindIII (245)**  
**Psp1406I (203)** 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCGCCCTACCTGAGGCC

**PvuII (239)**  
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

---

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

---

**AgeI (552)** **BspHI (560)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTGCATGAGACTCATCAGCAGCATCTATACCTTTATAACAT

1► M R L I S S I Y T L Y N I  
601 CATTATGTCAACAGAAGCCTGGGCTCCAAGACCACAGATGAAAGGGAAGTATCTCCAACACTCCAACATGGCTTTAAGGAAGTTCCCCAGATGTG

13► I M S T E A W A P R P T D E R E L I S N Y S N M A L R K V P P D V

**XcmI (707)**  
701 ACCCCAGCCACAACACTACTGGATTTATCCTACAACCTCCTCTATCAATTCAGAGCTCAGATTTCCACTCCTTCTCCAAACTGCGTTTTCTGATCCTGT

47► T P A T T I L D L S Y N L L Y Q F Q S S D F H S F S K L R F L I L

**PvuII (817)** **PshAI (881)**  
801 GCCATAACAGCATCCAACAGCTGGACACCAAGATCTTTGACCTTAACAAGGAGCTAAGCTATTTAGACTTGTCTTACAATAGACTACAGTCCATCACTTG

80► C H N S I Q Q L D T K I F D L N K E L S Y L D L S Y N R L Q S I T W

**Bsu36I (998)**  
901 GTCTTCCTGGTAGATCTCAGACATTTAGATCTTTCTTTAATGACTTCGATACTCTGCCTGCTGTGAGGAGATTGGCAACATGCGACATCTGGAACCC

113► S S L V D L R H L D L S F N D F D T L P A C E E I G N M R H L E T

1001 TTAGTTTGGTGGGGCCAAAATCCAAAGATCAGATCCAGAAAATGCTAATTTGCAACTCAGCACTGTGTTTCATAGGCTGAATACTCTTTCTCACT

147► L G L S G A K I Q R S D L Q K I A N L Q L S T V F I G L N T L S H

1101 ATGAAGAAGGCAGCCTGCCTGTCTAAACACGACAGTCTTACATAGTTTTACCAATGGCTGCTAATTCTGGGTGCTTCTGTGTGATGGAATGAAGAC

180► Y E E G S L P V L N T T V L H I V L P M A A N F W V L L C D G M K T

1201 TTCAAAGTGTGCAAAATGTCAACATAGGTGGCAAGGCCAGTTTGAAGTTATGAGACTCCGAGAAGCTTTCTGTGGAGAATACCAAGGCGTCCACT

213► S K V V E M S N I G G K S Q F V S Y E T P R T L S V E N T K A S T

1301 CTTATATCCATAAAGTTGACCTGCTCTGGGATGACCTTCTCCACGTTTTTCAAGTTGTGTGGCAGACTTCAGTGAATGCTTCCAGATTCAGGACCTGA

247► L I F H K V D L L W D D L L H V F Q F V W Q T S V E C F Q I Q D L

1401 CTTTTGGAGGGATGGCTAGTCGGGACCACAGTTCATTTGACTACTCAAATACCGTAATGAAAGCTGTAATCTTGAGAGCATAATTTGAGAATCTTTTT

280► T F G G M A S R D H S S F D Y S N T V M K A V I L E S I H F R I F F

**EcoRV (1557)**  
1501 CATTCCCAGAATAAAATCTACTGTTTTTACCCAAATGGATTTAGAAAACCTGACGATATCAGATGCACAAATGCCACATCTTCTCTCCCTAGTCAC

313► I P Q N K I Y L F F T Q M D L E N L T I S D A Q M P H L L F P S H

1601 CCCCAAAGTTTCCAACATTTAAATTTTCCAAATAACATCTTAACAGATGACCTGTTTAAAAACCCATCCAGATGCCTTATCTGAAAACCTCTCGTTTTGA

347► P Q S F Q H L N F A N N I L T D D L F K N P I Q M P Y L K T L V L

1701 GGGCAACCAACTGGAGACTCTTCTAGTGAGTCTTTTCCAACACAGTCTTGTCTCACCTGGATCTGAGCCAGAACCTATTGCAACATGAAAA

380► R G N Q L E T L S L V S L F A N H T S L L H L D L S Q N L L Q H G N

**NsiI (1809)**  
1801 CGGGGAGAAATGCATGTGGCCGACACCTTGATCCTCTTGAAGTTGCATCCAATAAATTTGATGGTTCTGTTTTAGATGTTTGCCTAAAAGTATCCAG

413► G E K C M W P D T L I L L N L S S N K F D G S V F R C L P K S I Q

1901 ACACTTGATCTGAACAATAACAAAATTCGAGCAGTCCCTAAAGAGATTATATACCTAAAATCTTACGAGAGCTAAGTATTGCTTTAATTTCTGACTG

447► T L D L N N N K I R A V P K E I I Y L K S L R E L S I A F N F L T

2001 ACCTCCCTGGGTGTGATCACTTCAAACGGCTCTAGTCTGAACATTGAAATGAATTTGATCTCAGCCATCTCTGGATTTTTCCAGAGCTGCCAGGA

480► D L P G C D H F K R L S S L N I E M N L I L S P S L D F F Q S C Q E

2101 GGTAAAGACTCTAAACGAGGGCGAAATCCATTCCGGTGTACCTGTGAATTTAGAGACTTCATTTGGCTTGAGAAACATTCCAAGGCGCTGATGTCTGGA

513► V K T L N A G R N P F R C T C E F R D F I W L E K H S K G L M S G

**MscI (2200)** **BsaBI (2257)**  
2201 TGGCCAGACTCCTACGCCTGCGAATACCCTTTGAGTCTAAAGGGGACTCCATTAAGGATGTTTATCTTCCGAATTAGCTTGAACACAACCTTTGCTGA

547► W P D S Y A C E Y P L S L K G T P L K D V Y L P E L A C N T T L L

2301 TTGTCACCATTTGCTGGTTATCCTGGTTCTGGTTTGGTTGTAGCATTCTGCATTCAAAAATTTGACATGCTTTGGTATCTCAGGGTCTGGCCAGTG

580► I V T I V L V I L V L G L V V A F C I L K F D M L W Y L R V L A Q C

**BspEI (2402)** **BsrGI (2454)** **NsiI (2465)**  
2401 CATCCGGACACAGCATAGGGTTAGGAAGACAGTTCGAGAGGAGCGTAAGAGAAGTGTACAGTCCATGCATTTATCTCCTACGCTGCACATGATTCTGTC

613► I R T Q H R V R K T V R E E R K R S V Q F H A F I S Y A A H D S V

2501 TGGGTGAAGAATGAAGTATTCCAATCTGGAGACAGAAGATGGCTCTGTCTTGTCTTCCCGAGAGGAACTTTGACCTGACAGGAGCATTGCTG

647► W V K N E L I P N L E T E D G S V L L C L P E R N F D P D R S I A

**SspI (2602)**  
2601 AAAATATTATAAGCTGCACTGAGAAAAGCTATAAGTCCATCTTTGTGCTGTCCCCAACTTTGTGCGAATGGGTGGTGCATAATGAACTCTATCTGGC

680► E N I I S C T E K S Y K S I F V L S P N F V Q N G W C H N E L Y L A

2701 CCACCACAATCCTTTCCGTGACAATTATGATTACATCATTCCCATCTTACTAGAACCCATTCCCTCCTACTGTGTTCCCTACCAGTTCCTTAAGCTGAAA  
713▶ H H N P F R D N Y D Y I I P I L L E P I P S Y C V P T T F P K L K

**Bst1107I (2818)**

2801 GCTCTTTTGGAAAAGTACTTAGAGTGGCCCAAGGACAGGCACAAACATGGGCTTTTTTGGGCAAACCTTCGAGCTGCCATTAACGTAATTTAT  
747▶ A L L E K K V Y L E W P K D R H K H G L F W A N L R A A I N V N L

2901 CAGACACAAGGGGAAGGTGTGAACCTACAGACATTACAAAAGCCTAACGCGGATTGCCAAGTTCTACCATCTCTGGTAAGAACAGACAGCTTCTAAAA  
780▶ S D T R G R C E L Q T F T K P N A D C Q G S T I S L V R T D S F •

MscI (3024)

**NheI (3018)**

3001 TCCACCCTGCCTAGTGAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTT

**HpaI (3156) MfeI (3167)**

3101 GTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAAATTGCATTCTTTATGTTTCAGGTTCCAGGG

**EcoRI (3252)**

3201 GGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCT  
3301 CTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTT

**SapI (3434) SspI (3491)**

3401 TAAGATATAGTGTATTTTCCAAGGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTAGTAAAAATTACG  
3501 AAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTG  
3601 GACTTAGGGAACAAAGGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAAT  
3701 GGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCAC  
128▶ T T K 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  
3801 CTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGG  
94▶ R I S 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

**StuI (3930)**

3901 CTTGAGCAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGTAGGCCGCCGACATGGTGCTT  
61▶ E A 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  
4001 GTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAG  
28▶ N D E 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

**AseI (4138)**

4101 TGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCT

**SpeI (4293)**

4201 CTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTAGTC  
4301 AAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCAT

**SnaBI (4421)**

4401 CATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCA

**NdeI (4526)**

4501 TTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAG  
4601 TCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGT

PacI (4712)

**PstI (4705)**

**SdaI (4704)**

**BspLU11I (4722)**

4701 AACGCCTGCAGGTTAAITTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCC  
4801 CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGT  
4901 GCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTC

**ApaLI (5036)**

5001 AGTTCGGTGTAGGTGCTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCA  
5101 ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTG  
5201 GCCTAACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA  
5301 CAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGT

**EagI (5472)**

PacI (5452)

**NotI (5471)**

5401 CTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTCA  
5501 TTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCA  
5601 GTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA