



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
**MfeI (82)**  
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

**Psp1406I (203)** 201 GTGAACGTTCTTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGGCCCCGCCCTACCTGAGGCC  
**HindIII (245)**  
**Bsu36I (291)**  
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

---

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

---

**BstEII (555)**  
**AgeI (552) NcoI (560)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTACCATGGCGCAGTGGGAGATGTTGCAGAACCCTGACAGCCT  
1▶ M A Q W E M L Q N L D S L  
601 GTTCTGGACCAGTTACACCAGGTCTACTCACAGAGCATCCTGCCAATGGACGTTTCGACAGCATTTGGCTACTGGATTGAAGACCAGAAGCTGGAGGGAA  
13▶ F L D Q L H Q V Y S Q S I L P M D V R Q H L A T W I E D Q N W R E  
701 GCTGCGCTGGGCAGTGATGACGCCAAGGCCAACATGCTATACTTCAGCATCCTGGACCAACTGAACCAGTGGGACCACTACAGCTCAGACTCCAATCACT  
47▶ A A L G S D D A K A N M L Y F S I L D Q L N Q W D H Y S S D S N H

---

**SapI (898)**  
801 TTTTATTACAGCATAACTTGCAGAAAATTCAGCCGGGATATTACAGACCTTTCCCAATGGCCCTACCCAGTTGGCTGAGATGATCTTTAATCTTCTTGGGA  
80▶ F L L Q H N L R K F S R D I Q T F P N G P T Q L A E M I F N L L L E  
901 AGAGCAAAGGATTCTGAATCAGGCTCAAAGAGCTCAAGAGGTGCAGCCCCACAGCCCCGAAGCAGTTGTGGAGAGCCAGCAGCTTGAAGATTGAAAAT  
113▶ E Q R I L N Q A Q R A Q E V Q P P P A P E A V V E S Q Q L E I E N  
1001 CGAATCCAGGTTTACATGTGGACATTGAGTCTTGGTGGAGATCCATCAGGCAGCTGAAGGACGAACAGGATGTCTTACGTTTCAAGATACACAGTTTCA  
147▶ R I Q G L H V D I E F L V R S I R Q L K D E Q D V F S F R Y T V F

---

**Th111I (1110)**  
1101 GTCTGAAGAAGACGTCTTTCAGACCCCCATCAGAGCCAACAGGCGCAGCTTGTGCAGGCAACAGCCAACAAAGTCGACAGAATGAGAAAGGAGGTGCT  
180▶ S L K K T S S S D P H Q S Q Q A Q L V Q A T A N K V D R M R K E V L

---

**BstXI (1208)**  
1201 GGACATCTCAAAGGACTGGTTGGCCGATTAACACCCTGGTGCACCTATTGCTGCCAAGCTGGACGAGTGAAGGTGCAGCAGCAGAAGTCTGCATT  
213▶ D I S K G L V G R L T T L V D L L L P K L D E W K V Q Q Q K S C I  
1301 GGAGCCCCGCCCCGAGCTGCAGCTGGAACAGCTGGAACAGTGGTTGACGGCTGGAGCAAAGTTCTTGTCCACCTTCGGCAGCTACTGAAGCAGCTGA  
247▶ G A P P P E L Q L E Q L E Q W L T A G A K F L F H L R Q L L K Q L

---

**MscI (1440)**  
1401 AGGAGATGAGTCACATGCTTCGGTATAAGGGTACATGTTTGGCCAAGGGGTGGACCTGCAGAATGCCAAGTCATGGAGTTACTTCAGCGTCTGCTCCA  
280▶ K E M S H M L R Y K G D M F G Q G V D L Q N A Q V M E L L Q R L L Q

---

**XmnI (1572)**  
1501 AAGGTCCTTTGTAGTAGAAACCCAGCCCTGCATGCCCCAGACTCTCCACCGACCCCTCATCCTGAAGACTGGGAACAAGTTACCGTCCGAACAAGACTT  
313▶ R S F V V E T Q P C M P Q T L H R P L I L K T G N K F T V R T R L

---

**EcoRI (1656)** 1601 CTGGTGAGACTCCAGGAAGCAGCGAATCACTCAAAGCAGAAGTCTCTGTCGACAGGAATTCAGACTTACCAGGCTTCCGGAAGTTCACATTCTGACCT  
347▶ L V R L Q E G S E S L K A E V S V D R N S D L P G F R K F N I L T  
1701 CAAACCAAAAACCTTGACCCCGAGGAGGGCCAGAGACAGGGCTTAATTTGGGACTTCGGCTTCTTACTCTGGTGGAGCAACGCGCTGTGGGTGCAGG  
380▶ S N Q K T L T P E E G Q R Q G L I W D F G F L T L V E Q R A V G A G  
1801 AAAGGGCAACAACAAGGGGCGCTGGCAGTGACAGAGGAGTTACATGTCATCAGCTTCGTGGTTGAGTACGTGTACCAGGGTCTGAAGATGAAGCTGCAG  
413▶ K G N N K G P L A V T E E L H V I S F V V E Y V Y Q G L K M K L Q  
1901 ACGGACTCTCCCTGCTGATTATTTCTAACATGAACCAACTCTCCATTGCCTGGGCTCCATTCTCTGGTTCAACATGCTCAGCCAAAATCCCAAGA  
447▶ T D T L P V V I I S N M N Q L S I A W A S I L W F N M L S P N P K

---

**StuI (2088)**  
2001 ACCAGCAGTCTTCTGCCAAGCTCCGAAAGCCCCCTGGAGTTTGTGGACCTGTGCTCAGCTGGCAGTCTCCTCTTACGTTGGCCGAGGCTCGATTC  
480▶ N Q Q F F C Q A P K A P W S L L G P V L S W Q F S S Y V G R G L D S  
2101 TGAGCAGCTGGGCATGCTGAGAACAAGCTGTTTGGAAAGAGCTGCAAGATGGAGGATGCGCTGTTGCTCCTGGGTAGACTTTTGAAGCGAGAGAGCCCC  
513▶ E Q L G M L R T K L F G K S C K M E D A L L S W V D F C K R E S P

---

**ApaLI (2248)**  
2201 CCTGGTAAGATCCCTTTCTGACCTGGCTGGACAAAATTTCTGGAGCTGGTGACACACCTGAAGGATCTCTGGAAGGATGGGCGCATCATGGGCTTTG  
547▶ P G K I P F W T W L D K I L E L V H D H L K D L W K D G R I M G F  
2301 TGAGCCGCAACCAGGAACGCAGACTGCTGAAGAAGATGTTGTCTGGCACCTTCTACTGCGCTTCAAGAACTTCTGAAGGGGGCATTACTTGTCTTG  
580▶ V S R N Q E R R L L K K M L S G T F L L R F S E T S E G G I T C S W  
2401 GGTGGAGCACCAGGATGACGATAAAGTCGAAATCTACTCAGTCGACCCCTACACCAAGGAAGTGTACAGTCACTCCCACTGACAGAGATCATCCGCCAC  
613▶ V E H Q D D D K V E I Y S V Q P Y T K E V L Q S L P L T E I I R H

---

**HindIII (2573)**  
2501 TACCAGGTTCTTGCCGAAGAGAATCCCCGAGAACCCTCCGCTTCTATCCCCGAATCCCTCGGGACGAAGCTTTTGGGTGTTACTACCAGGAAA  
647▶ Y Q V L A E E N I P E N P L R F L Y P R I P R D E A F G C Y Y Q E

SspI (2624)  
2601 AAGTTAATTTTGAAGAACAGAGGAAATATTTGAAACATAAACTCATTGTGATCTCTAACAGACAGGTGGACGAGCTGCAGCAGCCTCTGGAGCTCAAACA  
680▶ K V N F E E Q R K Y L K H K L I V I S N R Q V D E L Q Q P L E L K Q  
2701 GGATTAGAGTCTTAGAAGTGAATGCAGAGCTCTGTTAGCACACGACCAGGAGTTGCCATTGATGATGCAGACTGGGCTGGTCTGGGCACAGAGCTG  
713▶ D S E S L E V N A E L L L A H D Q E L P L M M Q T G L V L G T E L  
2801 AAAGTGGACCCATACTAGTACAGCCCAAGTCCCTGCTGGAGCCAGCCCAAGTCCAGCTGGAGCCAGCCCAAGTCCCTGCTGGAGCCAGCC  
747▶ K V D P I L S T A P Q V L L E P A P Q V Q L E P A P Q V L L E P A  
2901 CACAAGTCCAGCTGGAGCCAGCCCAAGTCCCTGCTGGAGCCAGCCCAAGTCCAGCTGGAGCCAGCCCAAGTCCCTGCTGGAGCCAGCCCA  
780▶ P Q V Q L E P A P Q V L L E P A P Q V Q L E P A P Q V L L E P A P Q  
3001 AGTCCAGCTGGAGCCAGCCCAAGTCCCTGCTGGAGCTAGCCCAAGTCCCTGCTGGAGCCAGCCCAAGTCCCTGCTGGAGCTAGCCCAAGTCC  
813▶ V Q L E P A P Q V L L E L A P Q V L L E P A P Q V L L E L A P Q V  
Bsu36I (3148)  
3101 CAGCTGGAGCCAGCACACTTGTGCAGCAGCCATCAGAGTCAAGTTCCTGAGGACCTGCAGCAGATTAGCGTGGAGGATCTGAAAAAATTAAGCAACC  
847▶ Q L E P A H L L Q Q P S E S D L P E D L Q Q I S V E D L K K L S N  
ScaI (3202)  
3201 CGAGTACTGAATACATCAACAACCAACGAAAACCAATGTTGGCTGGCAGAGCAGTGGAGATGAGACTTCCATACCCTACCACAGCCATTTTATGATCGGGA  
880▶ P S T E Y I T T N E N P M L A G E S S G D E T S I P Y H S H F D A D  
XbaI (3355) MscI (3366)  
3301 TGGACTCTGGGATGGACTCTTGATACCTTCTGAGACTAGGAGCTATGTTTCTGTCTAGACTAGTGGCCAGACATGATAAGATACATTGATGAGTTTG  
913▶ G L L G W T L D T F •  
HpaI (3498)  
3401 GACAAACCACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGT  
MfeI (3509) EcoRI (3594)  
3501 TAACAACAACAATTGCATTTCATTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTC  
3601 TAAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAA  
SapI (3776)  
3701 TGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAA  
SspI (3833) SmaI (3847)  
3801 ATGCACTGACCTCCACATTCCTTTTATGTA AAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATAGGCAGAATCCA  
3901 GATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTA  
4001 GCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAG  
141▶ • N R T Y K L P I L E E I T T T K V L K G N M E I L V F C D P A Y  
BstXI (4137)  
4101 TCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAA  
108▶ 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 I T D F  
StuI (4272)  
4201 AGTCCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGAT  
75▶ 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 A S I I  
4301 CTCCCAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTATCTCTCAGTGGCAGCTCCACCAGCTCCAGATCC  
42▶ 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 E L D  
XmnI (4414) AseI (4480)  
4401 TGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGTGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGT  
8▶ Q Q S I N F T K M  
4501 GGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGG  
SpeI (4635)  
4601 AGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATTTGACGTC AATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAA  
SnaBI (4763)  
4701 CCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAG  
NdeI (4868)  
4801 GTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACTTGTACTGCAAGTGGG  
4901 CAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATAGTCAATTATTGACGTCAATGGCGGGGGTCT  
PacI (5054)  
SdaI (5046)  
5001 GTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAAGGCCAGAAAAGGCCAGGAACC  
5101 GTAAAAAGGCCGCTTGTGGCTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG  
5201 ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCG

ApaI (5378)

5301 GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGC

5401 CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG

5501 AGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA

5601 GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAA

PacI (5794)

5701 AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATT

EagI (5814)

SwaI (5803) NotI (5813)

5801 AACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAA

5901 ACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA