



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
**MfeI (82)**  
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

**Psp1406I (203)** 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC  
**HindIII (245)**  
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

---

**NcoI (560)**  
**BstEII (555)**  
**AgeI (552)** 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCCTACCTGAGATCACCGTACCATGGAGCCCGGCCGCGGCACAGAGACCGTGGGCAA  
**SacII (573)**  
601 GTTCGAGTTCCTCCGCAAGGACCTGATCGGCCACGGCGCTTCGCGGTGGTCTTCAAGGCGCCGACCAGGAGACGATTTGGAGTTCGCGCTCAAG  
13> F E F S R K D L I G H G A F A V V F K G R H R E K H D L E V A V K  
701 TGCATTAACAAGAAGAACCTCGCAAGTCTCAGACGCTGCTGGGAAGGAAATCAAAATCCTGAAGGAAGTGAACATGAAAACATGTGGCCCTGTACG  
47> C I N K K N L A K S Q T L L G K E I K I L K E L K H E N I V A L Y

---

**ScaI (839)** 801 ACTTCCAGAAATGGCTAATTCTGTCTACTGTTATGGAGTACTGCAACGGTGGGACTGGCCGACTACCTGCACGCCATGCGCACGCTGAGCGAGGA  
**FspI (881)**  
80> D F Q E M A N S V Y L V M E Y C N G G D L A D Y L H A M R T L S E D  
901 CACCATCAGGCTCTTCTGCAGCAGATCGCGGGGCCATGCGGCTTCTGCACAGCAAAGGCATCATCCACCGGACCTGAAACCGCAGAACATCCTGCTG  
113> T I R L F L Q Q I A G A M R L L H S K G I I H R D L K P Q N I L L

---

**Acc65I (1070)** 1001 TCCAACCCCGCCGCGCCGCGCAACCCCAACAGCATCCGCGTCAAGATCGTGAAGTTCGGCTTCGCGCGGTACCTCCAGAGCAACATGATGGCGGCCA  
147> S N P A G R R A N P N S I R V K I A D F G F A R Y L Q S N M M A A

---

**BsrGI (1118)** 1101 CACTCTGCGGCTCCCCATGTACATGGCCCCGAGGTCATCATGTCCCAGCACTACGACGGGAAGGGGACCTGTGGAGCATCGGCACCATCGTCTACCA  
180> T L C G S P M Y M A P E V I M S Q H Y D G K A D L W S I G T I V Y Q

---

**PshAI (1272)** 1201 GTGCTGACGGGGAAGGCGCCCTCCAGGCCAGCAGCCCCAGGACCTGCGCCTGTTCTACGAGAAGAACAAGCCTTGGTCCCACCATCCCCGGGAG  
213> C L T G K A P F Q A S S P Q D L R L F Y E K N K T L V P T I P R E  
1301 ACCTCGGCCCCGCTGCGGAGCTGCTCTGGCCCTACTGCAACGCAACCAAGGACCGCATGGACTTCGATGAGTTTTTTCATACCCTTCTCGATG  
247> T S A P L R Q L L L A L L Q R N H K D R M D F D E F F H H P F L D  
1401 CCAGCCCTCGGTGAGAAATCCCCACCCGTCCTGTGCCCTCGTACCAAGCTCGGGGTCCGGCAGCAGCTCCAGCAGCTCCACCTCCCACCTGGC  
280> A S S S V R K A S P P V P V P S G Y P S G S G S S S S S S H L A  
1501 CTCCCCGCGTCCCTGGGCGAGATGCAGCAGTGCAGAAGACCTGCGCTCCCGGCTGACACCGCTGGCTTCTGCACAGCTCCCGGGACTCGGTGGC  
313> S P P S L G E M Q Q L Q K T L A S P A D T A G F L H S S R D S G G

---

**Tth111I (1627)** 1601 AGCAAGGACTCTTCTGTGACACAGAGCTTCGTATGGTCCCCGCGCAGTTTCCAGGTGACCTGGTGGCTGAGGCGCCAGTGCCAAACCCCCGCCAG  
347> S K D S S C D T D D F V M V P A Q F P G D L V A E A P S A K P P P  
1701 ACAGCCTGATGTGAGTGGGAGCTCACTGGTGGCTCTGCGGGCTTGGAGAGCCACGGCCGACCCCATCTCCATCCCCACCTGCAGCAGCTCCCCAG  
380> D S L M C S G S S L V A S A G L E S H G R T P S P S P P C S S S P S

---

**BstEII (1657)**  
**SfiI (1807)** 1801 TCCCTCAGGCGGGCTGGCCGTTCTCAGCAGCAGGTGCGGCGCTCTGTCCCATCCCAGTCCCACGAGGTGCAGAACTACCAGCGATTGAGCGA  
413> P S G R A G P F S S S R C G A S V P I P V P T Q V Q N Y Q R I E R  
1901 AACCTGCAGTCAACCACCGATTCCAAACCTCGTCTCTGCCATCCGAGGTGAGCAGCACCAGCCCCCTGGGCTTGAAGGGCCAGCCCTCGC  
447> N L Q S P T Q F Q T P R S S A I R R S G S T S P L G F A R A S P S

---

**MscI (2030)** 2001 CCCCTGCCACGCTGAGCATGGAGCGTCTGGCCAGGAAGATGTCTCTGGTGGAGGCGGCCCTACACGCCATCTCCTCAAGTTGGAACCATCCCTGA  
480> P P A H A E H G G V L A R K M S L G G G R P Y T P S P Q V G T I P E  
2101 GCGGCCAGGCTGGAGCGGACGCCCTCCCCACAGGGAGCTGAGATGCGGGGTGGCAGGTCCTCCTGTCACCCGAGCACTCTCCCCGC  
513> R P G W S G T P S P Q G A E M R G G R S P R P G S S A P E H S P R  
2201 ACTTCCGGGCTGGGCTGCGCCCTGCACAGCGCCCAACTGTCTGACTTGCACGTGTCGCGCCCAAGCTGCCAAACCCCCACGGACCCCTGGGAG  
547> T S G L G C R L H S A P N L S D L H V V R P K L P K P D P L G  
2301 CTGTGTTACGCCACACAGGCCAGCCCTCCCCAGCGCTCCACGGCTGCAAGTCTGCGGAACTGCGGGGCTCACCAAGCTGCCGACTTCTGCA  
580> A V F S P P Q A S P P Q P S H G L Q S C R N L R G S P K L P D F L Q  
2401 GCGAAACCCCTGCCCCATCCTGGGCTCCCCACCAAGGCTGTGCCCTCTTTGACTTCCCGAAGACCCCAAGCTCCAGAACCTGCTGGCCCTCCTA  
613> R N P L P P I L G S P T K A V P S F D F P K T P S S Q N L L A L L

---

**NcoI (2571)**  
**SandI (2562)** 2501 GCCCGCAGGGCGTGGTATGACGCCCCCTCGAAACCGGACGCTGCCGACCTCTCGAGGTGGGACCTTCCATGGTCAGCCGTTGGGCCCTGGCCTGC  
647> A R Q G V V M T P P R N R T L P D L S E V G P F H G Q P L G P G L

---

**SfiI (2620)** 2601 GGCCAGGCGAGGACCCCAAGGCCCTTTGGCCGTTCTTTCAGCACCAGCCGCTCACTGACTGCTCCTTAAGGCGGCTTTGGGACACAAGCCCCGA  
680> R P G E D P K G P F G R S F S T S R L T D L L L K A A F G T Q A P D

---

**SdaI (2718)** 2701 CCCGGCAGCAGGAGCCTGCAGGAGAAGCCATGGAGATCGCACCTCAGCTGGCTTTGGAGGGAGCCTGCACCCAGGAGCCCGTGTGGGGCACC  
713> P G S T E S L Q E K P M E I A P S A G F G G S L H P G A R A G G T  
2801 AGCAGCCCTTCCCGGTGGTCTTACCCTGGGCTCTCCCCGAGCGGGAGCACGCCCCCAAGGCCCCGACCCAGGATGTTCTCAGCGGGCCCACTG  
747> S S P S P V V F T V G S P P S G S T P P O G P R T R M F S A G P T

2901 GCTCTGCCAGCTCTTCTGCCCGCCACCTGGTGCCTGGGCCCTGCAGCGAGGCCCCAGCCCTGAGCTCCCTGCTCCAGGACACGGCTGCAGCTTTGCCGA  
780▶ G S A S S S A R H L V P G P C S E A P A P E L P A P G H G C S F A D  
3001 CCCCATTA CTGCGAACCTGGAGGGGCTGTGACCTTCGAGGCCCCGACCTCCCTGAGGAGACCCCTCATGGAGCAAGAGCACCGGAGATCCTGCGTGGC  
813▶ P I T A N L E G A V T F E A P D L P E E T L M E Q E H T E I L R G  
3101 CTGCGCTTCACGCTGCTGTTCTGTCAGCACGCTCTGGAGATCGCAGCCCTGAAGGGCAGCGCCAGTGAGGCGCGGGGGCCCTGAGTACCAGCTGCAGG  
847▶ L R F T L L F V Q H V L E I A A L K G S A S E A A G G P E Y Q L Q  
3201 AGAGTGTGGTGGCCGACCAAGATCAGCCTGCTGAGCCGAGAATGGGGCTTCGCGGAACAGCTGGTGTCTGTACCTGAAGGTGGCCGAGCTACTGTCTCCGG  
880▶ E S V V A D Q I S L L S R E W G F A E Q L V L Y L K V A E L L S S G  
FspI (3368) BsrGI (3386)  
3301 CCTGCAAAGTGCCATCGACCAGATCCGGGGCCGCAAGCTCTGCCTGTGCTCCACTGTGAAGCAGGTGGTGGCGCAGGCTGAATGAGCTGTACAAGGCCAGC  
913▶ L Q S A I D Q I R A G K L C L S S T V K Q V V R R L N E L Y K A S  
Eco47III (3434)  
3401 GTGGTGTCTCTGCCAGGGCCTGAGCCTGCGGCTGCAGCGCTTCTTCTGGACAAGCAGCGGCTCCTGGACCGATTACAGCATCACTGCCGAGAGGCTCA  
947▶ V V S C Q G L S L R L Q R F F L D K Q R L L D R I H S I T A E R L  
3501 TCTTCAGCCACGCTGTGCAGATGGTGCAGTCCGCTGCCCTGGACGAGATGTTCCAGCACCGTGAGGGCTGCGTCCCACGCTACCACAAGGCCCTGCTGCT  
980▶ I F S H A V Q M V Q S A A L D E M F Q H R E G C V P R Y H K A L L L  
3601 CCTGGAGGGGCTGCAGCAGATGCTCTCGGACCAAGCCGACATCAGAAAGTCAACAGCTGTGATTGAGCGGAGACTCTCGGGCTGCTGACT  
1013▶ L E G L Q H M L S D Q A D I E N V T K C K L C I E R R L S A L L T  
MscI (3725) NheI (3719)  
3701 GGCATCTGTGCCTGACCTTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAGTGAATGCAGTGAAAAAATGCTTTATT  
1047▶ G I C A •  
HpaI (3857) MfeI (3868)  
3801 TGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTATGTTTCAGGTTCCAGG  
EcoRI (3953)  
3901 GGGAGGTGTGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCC  
4001 TCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGT  
SspI (4192)  
4101 TTAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCTTTTTAGTAAAAATATCA  
SwaI (4206)  
4201 GAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTT  
4301 GGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGTTGTACTTGGAGGGGATGAGTTCCTCAA  
1414 • N R T Y K L P I L E E I  
4401 TGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCAC  
1284 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  
4501 CCTGATGGATCTGTCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATG  
954 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  
StuI (4631)  
4601 GCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCT  
614 A 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  
XmnI (4773)  
4701 TGGTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTTCATGGTGGCCCTCTATA  
284 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 (4839)  
4801 GTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGC  
SpeI (4994)  
4901 TCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTACGACATTTTGGAAAGTCCCGTTGATTTACTAGT  
5001 CAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCA  
SnaBI (5122)  
5101 TCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCCATAAAGTCAATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTC  
NdeI (5227)  
5201 ATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGACTGCCAAGTGGGCGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAA  
5301 GTCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCCTAAGTTATG  
SdaI (5405) PacI (5413) BspLU11I (5423)  
5401 TAACGCCTG CAGGTTAAIT AAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCGTTTTTCCATAGGCTCCGC  
5501 CCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCG  
5601 TGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCT  
ApaLI (5737)  
5701 CAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC  
5801 AACCCGGTAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGT

5901 GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA  
6001 ACAAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGG  
6101 TCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTC  
6201 ATTACATCTGTGTGGTTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCC  
6301 AGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA