



PstI (6) PstI (25)

**SdaI (6) SpeI (13)**

1 CCTGCAGGGCCCACTAGTATCTGCAGAGGGCCCTGCGTATGAGTGCAAGTGGGTTTTAGGACCAGGATGAGGCGGGTGGGGTGCCTACCTGACGACCG

---

101 ACCCCGACCCACTGGACAAGCACCCAACCCCAATTCCTCCAAATTGCGCATCCCCATCAGAGAGGGGGAGGGGAAACAGGATGCGGCGAGGCGGTGCGC

---

201 ACTGCCAGCTTCAGCACCGGACAGTGCCTTCGCCCCCGCCTGGCGGCGCGGCCACCGCCGCTCAGCACTGAAGGCGCGCTGACGTCACTCGCCGGT

---

301 CCCCCGAAACTCCCTTCCCGGCCACCTTGTCGCGTCCGCGCCGCCGCCGCGCCAGCCGACCGACCCAGCGAGGCGCGAGATAGGGGGGACGCGG

---

**SgrAI (421)**

**KasI (418) Eco47III (434)**

401 GCGACCATCTGCGCTGCGGCGCCGCGACTCAGCGCTGCTCAGTCTGCGGTGGGACAGCGGAGGAGTCTGTCTGCTGAGAGCGCAGCTGTCTCCTG

---

**MscI (532) AvrII (543)**

**NcoI (570)**

501 GGCACCGCGCAGTCCGCCCGCGGCTCTGGCCAGACCACCCTAGGACCCCTGCCCAAGTCGAGCCATGGGGGTTCTCATCATCATCATCATCA

---

1 M G G S H H H H H

Bsu36I (669)

**Acc65I (664)**

601 TGGTATGGCTAGTATGACTGGTGGACAGCAAAATGGGTCGGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTTGGAGTTGATCCCGTCGTT

10 ▶ G M A S M T G G Q Q M G R D L Y D D D D K V P K D Q L G V D P V V

701 TTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAATTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA

44 ▶ L Q R R D W E N P G V T Q L N R L A A H P P F A S W R N S E E A R

801 CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCTGTTCCGGCACCAGAAGCGGTGCCGAAAAGCTGGCTGGAGTGCGA

77 ▶ T D R P S Q Q L R S L N G E W R F A W F P A P E A V P E S W L E C D

Bsu36I (906)

901 TCTTCCTGAGGCCGATCTGCTGCTCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCATCTACACCAACGTAACCTATCCATTACGGTCAAT

110 ▶ L P E A D T V V V P S N W Q M H G Y D A P I Y T N V T Y P I T V N

1001 CCGCCGTTTGTCCACCGGAGAATCCGACGGTGTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCAATATTTTTG

144 ▶ P P F V P T E N P T G C Y S L T F N V D E S W L Q E G Q T R I I F

1101 ATGGCGTAACTCGGCGTTTCATCTGTGGTGAACGGGCGTGGGTCGGTTACGGCCAGGACAGTCTGTTGCGCTCTGAATTTGACCTGAGCGCATTTTT

177 ▶ D G V N S A F H L W C N G R W V G Y G Q D S R L P S E F D L S A F L

1201 ACGCGCCGAGAAAACCCCTCGCGGTGATGGTCTGCGTTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGT

210 ▶ R A G E N R L A V M V L R W S D G S Y L E D Q D M W R M S G I F R

1301 GACGTCTCGTGTCTGCATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTAGCCGCGCTGTACTGGAGGCTGAAG

244 ▶ D V S L L H K P T T Q I S D F H V A T R F N D D F S R A V L E A E

1401 TTCAGATGTGCGCGAGTTGCGTACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGAACGCAGGTGCGCAGCGCACCGCCCTTTCGGCGGTGA

277 ▶ V Q M C G E L R D Y L R V T V S L W Q G E T Q V A S G T A P F G G E

1501 AATTATCGATGAGCGTGGTGTATGCGGATCGCGTACACTACGCTGAAACGTCGAAAACCCGAAAAGTGGAGCGCGGAAATCCCGAATCTCTATCGT

310 ▶ I I D E R G G Y A D R V T L R L N V E N P K L W S A E I P N L Y R

1601 GCGGTGGTTGAACTGCACACCGCCGACGGCAGCTGATTGAAGCAGAAAGCTGCGATGTCGGTTTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTGCTGC

344 ▶ A V V E L H T A D G T L I E A E A C D V G F R E V R I E N G L L L

**EcoRV (1795)**

1701 TGAACGGCAAGCCGTTGCTGATTGAGGCGTTAACCGTCACGAGCATCATCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCT

377 ▶ L N G K P L L I R G V N R H E H H P L H G Q V M D E Q T M V Q D I L

1801 GCTGATGAAGCAGAACTTTAACGCGTGCCTGTTGCGATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGT

410 ▶ L M K Q N N F N A V R C S H Y P N H P L W Y T L C D R Y G L Y V V

1901 GATGAAGCCAATATTGAAACCCACGGCATGGTCCCAATGAATCGTCTGACCGATGATCCGCGTGGCTACCGGCGATGAGCGAAGCGGTAACGCGAATGG

444 ▶ D E A N I E T H G M V P M N R L T D D P R W L P A M S E R V T R M

2001 TGCAGCGGATCGTAATCACCGAGTGTGATCATCTGGTCCGTTGGGAATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCGTGGATCAAAATC

477 ▶ V Q R D R N H P S V I I W S L G N E S G H G A N H D A L Y R W I K S

2101 TGTCGATCCTTCCCGCCCGTGCAGTATGAAGGCGGCGGAGCGACACACGGCCACCGATATTATTTGCCGATGTACGCGCGGTGGATGAAGACCAG

510 ▶ V D P S R P V Q Y E G G G A D T T A T D I I C P M Y A R V D E D Q

2201 CCCTTCCCGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTCCTACCTGGAGAGACGCGCCCGCTGATCCTTTGCCAATACGCCACCGCATGGGTA

544 ▶ P F P A V P K W S I K K W L S L P G E T R P L I L C E Y A H A M G

2301 ACAGTCTTGGCGTTTTCCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTTTCGCTGGGACTGGGTGGATCAGTCGCTGATTA

577 ▶ N S L G G F A K Y W Q A F R Q Y P R L Q G G F V W D W V D Q S L I K

2401 ATATGATGAAAACGGCAACCCGGTGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTCTGTATGAACGGTCTGGTCTTTGCCGAC  
610 ▶ Y D E N G N P W S A Y G G D F G D T P N D R Q F C M N G L V F A D

Eco47III (2517)

2501 CGCACGCCGCATCCAGCGCTGACGGAAGCAAAAACCCAGCAGCAGTTTTTCCAGTTCGGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGT  
644 ▶ R T P H P A L T E A K H Q Q Q F F Q F R L S G Q T I E V T S E Y L

2601 TCCGTCATAGCGATAACGAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCCTCTGGATGTCCTCCACAAGGTAA  
677 ▶ F R H S D N E L L H W M V A L D G K P L A S G E V P L D V A P Q G K

2701 ACAGTTGATTGAACTGCCTGAACTACCGCAGCCGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCGTAGTGAACCGAACGCGACCCGCATGGTCAGAA  
710 ▶ Q L I E L P E L P Q P E S A G Q L W L T V R V V Q P N A T A W S E

2801 GCCGGGCACATCAGCGCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCTCCACGCCATCCCGCATCTGACCACCAGCG  
744 ▶ A G H I S A W Q Q W R L A E N L S V T L P A A S H A I P H L T T S

2901 AAATGGATTTTGCATCGAGCTGGTAATAAGCGTTGGCAATTTAACCCGACGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAATGCT  
777 ▶ E M D F C I E L G N K R W Q F N R Q S G F L S Q M W I G D K K Q L L

3001 GACGCCGCTGCGCGATCAGTTACCCGTCACCCGCTGGATAACGACATTTGGCGTAAGTGAAGCGACCCGCATTGACCCTAACGCTGGTGGTGAACGCTGG  
810 ▶ T P L R D Q F T R A P L D N D I G V S E A T R I D P N A W V E R W

3101 AAGCGCGCGGCGATTACCAGGCCAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGTGTATGCGGTGCTGATTACGACCCGCTCACGCGTGGCAGC  
844 ▶ K A A G H Y Q A E A A L L Q C T A D T L A D A V L I T T A H A W Q

3201 ATCAGGGGAAAACCTTATTTATCAGCCGAAAACCTACCCGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATAACCC  
877 ▶ H Q G K T L F I S R K T Y R I D G S G Q M A I T V D V E V A S D T P

3301 GCATCCGCGCGGATTGGCTGAACTGCCAGCTGGCGCAGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGCAAGAAAACCTATCCGACCCGCTT  
910 ▶ H P A R I G L N C Q L A Q V A E R V N W L G L G P Q E N Y P D R L

BspLU11I (3441)

3401 ACTGCCGCTGTTTTGACCGCTGGGATCTGCCATTGTCAGACATGTATACCCCGTACGCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGCGAAT  
944 ▶ T A A C F D R W D L P L S D M Y T P Y V F P S E N G L R C G T R E

3501 TGAATTATGGCCACACCACTGCGCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTGTGCA  
977 ▶ L N Y G P H Q W R G D F Q F N I S R Y S Q Q Q L M E T S H R H L L H

NdeI (3639)

3601 CGCGGAAGAAGGCACATGGCTGAATATCGACGTTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGCCCGTCACTATCGCGGAATTACAGCTGAGC  
1010 ▶ A E E G T W L N I D G F H M G I G G D D S W S P S V S A E L Q L S

NheI (3749)

3701 GCCGGTCTGCTACCATTACCAGTTGGTCTGGTGTCAAAAATAACAATTTCGCTAGCATTATCCCTAATACCTGCCACCCCACTCTTAATCAGTGGTGAAGA  
1044 ▶ A G R Y H Y Q L V W C Q K •

MscI (3829)

3801 ACCGCTCAGAACTGTTGTTTCAATTGGCCATTTAAGTTTAGTAGTAAAGACTGGTTAATGATAACAATGCATCGTAAACCTTCAGAAGGAAAGGAG

3901 AATGTTTTGTGGACCACTTTGGTTTTCTTTTTGCGTGTGGCAGTTTTAAGTTATTAGTTTTTAAATCAGTACTTTTTAATGGAAACAACCTGACCAAA

4001 AATTTGTACAGAATTTTGAGACCCATTAATAAAGTTAATGAGAAACCTGTGTCTTTGGTCAACACCGAGACATTTAGGTGAAGACATCTAATT

4101 CTGGTTTTACGAATCTGAAAACCTCTGAAAATGTAATCTTGAGTTAACACTCTGGGTGGAGAATAGGGTTGTTTTCCCCCACATAATTGGAAGGGG

4201 AAGGAATATCATTTAAGCTATGGGAGGTTCTTTGATTACAACACTGGAGAGAAATGCAGCATGTTGCTGATTGCCTGTCACTAAAACAGGCCAAAA

HindIII (4323)

4301 CTGAGTCTTGGGTTGCATAGAAAAGCTTCATGTTGCTAAACCAATGTTAAGTGAATCTTTGGAACAAAATGTTTCCAATTAAGTGGGATGTGCATGTTG

4401 AAACGTGGGTTAATTAAGTAGCCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTG  
▶

4501 AGATCTTTTTTTCTGCGGTAATCTGCTGCTGCAAAACAAAAAACCCCGTACCAGCGGTGGTTTTGTTGCGGATCAAGAGCTACCAACTCTTTTT

4601 CCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTA

4701 CATACTCGCTCTGTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATAAGGC

4801 GCAGCGGTGCGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGTGATGAGAAAGC

4901 GCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCAGGGTGGAAACAGGAGAGCGCAGAGGGAGCTTCCAGGGGAAACCGCTGGT

5001 ATCTTTATAGTCTGCGGTTTCGCCACCTGACTTGAGCGTCGATTTTTGTGTGCTGCTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAACCG

BspLU11I (5142)

AseI (5180)

5101 GGCCTTTTTACGGTCTGGCCTTTTGTGGCCTTTTGTGCTCATGTTCTTAATTAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATAT

MscI (5242)

5201 CGGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAAGTTGACCAAGTGTCCAGTGTCCAGTGTCCAGCCAGGGATGTGGCTGGAGCTGTTGA  
▶ 1 ▶ M A K L T S A V P V L T A R D V A G A V E

5301 GTTCTGGACTGACAGGTTGGGTTCTCCAGAGATTTTGTGGAGGATGACTTTCAGGTGTGGTCAGAGATGATGTCACCCCTGTTTCATCTCAGCAGTCCAG  
21 ▶ F W T D R L G F S R D F V E D D F A G V V R D D V T L F I S A V Q  
5401 GACCAGGTGGTGCCTGACAACACCCTGGCTTGGGTGTGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGAGGTGGTCTCCACCAACTTCAGGG  
55 ▶ D Q V V P D N T L A W V W V R G L D E L Y A E W S E V V S T N F R  
5501 ATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAGTTTGCCCTGAGAGACCCAGCAGGCAACTGTGTGCACTTTGTGGCAGA  
88 ▶ D A S G P A M T E I G E Q P W G R E F A L R D P A G N C V H F V A E  
5601 GGAGCAGGACTGAGGATAAGAATTGTAACAAAAACCCCGCCCGGGGGTTTTTTGTTAATTAA  
121 ▶ E Q D •