



**EcoRI (23)**

**NotI (2)**                      **XbaI (19)**                      **SdaI (38)**                      **SpeI (45)**

1 **GCGGCCGCGTCGACGATATCTAGAATTCGGATCCTGCAGGGCCCACTAGTTGATTACTGAATTTGCCATGTTGC**

75 **TTCCTGCAATACCAAATGACCCACATTACTAACATTTGGAATTTGCCTCAGTGTTAGGCACAAGACATGAGTG**

149 **AAACTTGCTGCATTAAGACTTAAACTGGCTGGGTATGGTGGCGCAAGCCTTTAATCCCAGCATTCTGCCGGAGG**

**NheI (251)**

223 **CAGAAGCAGGTGGATTTCTGAGTTCGAGGCTAGCCTGGTCTACAAAAGTGAGTTCAGGACAGCCAAGGCTACA**

297 **CAAATACCTGTCTTGAAAAACCAAAAAAAAAAAAAAATTCAACTGGTTGGTTAGGCCTAGCCAGATGATGTAG**

371 **GTGTCAACTCACCTTGGGGATGTAGGGAGATGGTTGTTCTGTAGCTCTTCTGTCTGCAACTAAATAAATAGG**

**EcoNI (481)**

445 **CTAAGCTGACCTTTACATCTGCCTTCGGCCTCCTGTGCTAGGATTGGAGTGCATGGGTGGGGTTGCATTTCTT**

519 **ACCAATGAGGAAAGGGCTTCCATTTCTCTGCTGCATTGTAAGCTCCCAAGTAGAGCAAGTTTTGCTTAGGTAA**

593 **GTTCCGTGTGAGTCAGCTGCCTCATTCTCACGAGGTAACCAAGGCTTTGTACCGCCACTGAGAACGTCAGTGC**

667 **CAATCACAGCCTAATTGTGAAAACCAATGGCTTGAGTGGGTGCTAAAGCTGAGGTGTCTGAGTCAGGTTTGGG**

741 **GTGGGATTATTTTAGTTAAGGGAAGTGAGGCTTTTCATTTCTCTTCCAAGAGAAGGCAAAGGGGATTGGATTG**

**NcoI (856)**

815 **AGGAAGGAAGTGGTGTAGCCTAGCTGGTCTGAGCATCTCTGCCATGGTTCTGGGGCCCTGCATGCTGCTGCTGC**

**M V L G P C M L L L**

889 **TGCTGCTGCTGGGCCTGAGGCTACAGCTCTCCCTGGGCATCATCCCAGTTGAGGAGGAGAACCCGGACTTCTGG**

**110 L L L L G L R L Q L S L G I I P V E E E N P D F W**

963 **AACCGCGAGGCAGCCGAGGCCCTGGGTGCCCAAGAAGCTGCAGCCTGCACAGACAGCCGCAAGAACCTCAT**

**36 N R E A A E A L G A A K K L Q P A Q T A A K N L I**

1037 **CATCTTCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTAAAAGGGCAGAAGAAGGACA**

**60 I F L G D G M G V S T V T A A R I L K G Q K K D**

**NdeI (1151)**

1111 **AACTGGGGCCTGAGATACCCCTGGCTATGGACCGCTTCCCATATGTGGCTCTGTCCAAGACATACAATGTAGAC**

**85 K L G P E I P L A M D R F P Y V A L S K T Y N V D**

1185 **AAACATGTCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGTCAAGGGCAACTTCCAGACCATTGG**

**110 K H V P D S G A T A T A Y L C G V K G N F Q T I G**

1259 **CTTGAGTGCAGCCGCCGCTTTAACAGTGCAACACGACACGCGGCAACGAGGTCATCTCCGTGATGAATCGGG**

**134 L S A A A R F N Q C N T T R G N E V I S V M N R**

1333 **CCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCACACGAGTGCAGCACGCTCGCCAGCCGGCACCTAC**

**159 A K K A G K S V G V V T T T R V Q H A S P A G T Y**

1407 **GCCACACGGTGAACCGCAACTGGTACTCGGACGCCGACGTGCCTGCCTCGGCCCGCCAGGAGGGGTGCCAGGA**

**184 A H T V N R N W Y S D A D V P A S A R Q E G C Q D**

1481 **CATCGCTACGACGCTCATCTCCAACATGGACATTGATGTGATCCTGGGTGGAGGCCGAAAGTACATGTTTCGCA**

**208 I A T Q L I S N M D I D V I L G G G R K Y M F R**

1555 **TGGGAACCCAGACCCTGAGTACCCAGATGACTACAGCCAAGGTGGGACCAGGCTGGACGGGAAGAATCTGGTG**

**233 M G T P D P E Y P D D Y S Q G G T R L D G K N L V**

1629 **CAGGAATGGCTGGCGAAGCGCCAGGGTGGCCGGTATGTGTGGAACCGCACTGAGCTCATGCAGGCTTCCCTGGA**

**258 Q E W L A K R Q G A R Y V W N R T E L M Q A S L D**

1703 **CCCGTCTGTGACCCATCTCATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCCACCGAGACTCCACAC**

**282 P S V T H L M G L F E P G D M K Y E I H R D S T**

**SacII (1838)**

1777 **TGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGAACCCCGCGGCTTCTTCTCCT**

**307 L D P S L M E M T E A A L R L L S R N P R G F F L**

1851 **TTCGTGGAGGGTGGTGCATCGACCACGGTCATCACGAAAGCAGGGCTTACCGGGCACTGACTGAGACGATCAT**

**332 F V E G G R I D H G H H E S R A Y R A L T E T I M**

1925 **GTTTCGACGACCCATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACGCTGAGCCTCGTCACTGCCGACC**

**356 F D D A I E R A G Q L T S E E D T L S L V T A D**

1999 **ACTCCCACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCCTGGCAAGGCC**

**381 H S H V F S F G G Y P L R G S S I F G L A P G K A**

2073 CGGGACAGGAAGGCCTACACGGTCTCTCTATACGGAAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCGGCC  
406 R D R K A Y T V L L Y G N G P G Y V L K D G A R P  
2147 GGATGTTACCGAGAGCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGCCCCTGGACGAAGAGACCC  
430 D V T E S E S G S P E Y R Q Q S A V P L D E E T  
2221 ACGCAGGCGAGGACGTGGCGGTGTTGCGCGCGGGCCCGCAGGCGCACCTGGTTCACGGCGTG CAGGAGCAGACC  
455 H A G E D V A V F A R G P Q A H L V H G V Q E Q T  
2295 TTCATAGCGCACGTCATGGCCTTCGCCGCTGCTGGAGCCCTACACCGCCTGCGACCTGGCGCCCCCGCCGG  
480 F I A H V M A F A A C L E P Y T A C D L A P P A G

NheI (2422)

2369 CACCACCGACGCCGCGCACCCGGGGCGGTCCCGGTCCAAGCGTCTGGATTGAAAGCTAGCTGGCCAGACATGATA  
504 T T D A A H P G R S R S K R L D •  
2443 AGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGA

MfeI (2571)

2517 TGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGT  
2591 TTCAGGTT CAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTAATT

2665 CTAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAA  
2739 GGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTTCATGGAGTTTAAGA  
2813 TATAGTGTATTTCCCAAGGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTC

SspI (2901)

2887 CCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGA  
2961 ATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAAT  
3035 AGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTATCCTCAGTCTGCTCCTCTGCCACAAAGTGACCGCAGT  
125 D Q E E A V F H V C N  
3109 TGCCGGCCGGGTGCGCGAGGGCGAAGTCCCGCCCCACGGCTGCTCGCCGATCTCGGTCATGGCCGGCCCGGAG  
113 G A P D R L A F E R G W P Q E G I E T M A P G S  
3183 GCGTCCCGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGCTCGTCCAGGCCGCGCACCCACACCCA  
88 A D R F N T S V V E S W E A Y L E D L G R V W V W

SgrAI (3329)

3257 GGCCAGGGTGTGTCCGGCACCACCTGGTCTGGACCGCGCTGATGAACAGGGTCACGTCGTCCCGGACCACAC  
64 A L T N D P V V Q D Q V A S I F L T V D D R V V G  
3331 CGGCGAAGTCGTCTCCACGAAGTCCCGGAGAACCCGAGCCGGTCCGTTCCGGAAGTCCGACCGCTCCGGCGACG  
39 A F D D E V F D R S F G L R D T W F E V A G A V  
3405 TCGCGCGCGGTGAGCACCGGAACGGCACTGGTCAACTTGGCCATGATGGCTCCTCCTGT CAGGAGAGGAAAGAG  
14 D R A T L V P V A S T L K A M

MfeI (3492)

3479 AAGAAGGTTAGTACAATTGCTATAGTGAGTTGATTATACTATGCAGATATACTATGCCAATGATTAATTGTCA  
3553 AACTAGGGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGC  
3627 GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC  
3701 GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGTCCCTCGTGCGCTCTCCTGTTCCGACC  
3775 CTGCCGTTACCGGATACCTGTCCGCCTTTCTCCCTCGGGAAGCGTGGCGCTTTTCATAGCTCACGCTGTAG  
3849 GTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTT CAGCCCGACCGCT  
3923 GCGCCTTATCCGTAAC TATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT  
3997 GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTA  
4071 CACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTT  
4145 GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAA  
4219 GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAAC TACGTTAAGGGAT  
4293 TTTGGTCATGGCTAGTTAATTAACATTTAAATCA