



**XbaI (15) SdaI (29)**

*NotI* (2) **NsiI (14) EcoRI (19)**

1 **GCGGCCGCTATGCATCTAGAATTCCTGCAGGTGAAGACGTTACACAAGTAACATGAGAAAAGCAGAAAATGCAGGTCATC**  
80 **CACGCACCCCTGACCCAGGCCAGCAGGGCGGGCTGCAGCATCAGTACACAGGAGAAAGATCCTTATTCTTAAGAATGAG**  
159 **AAAGGCAAAGGCGCCCGATAGAATAAATTAGCATAGAAGGGGCTTTCCAGGAGTTAAAACCTTCTTCTGAGCGATTA**  
238 **CCTACTAAAACCAGGGCTTTTGCCCACTACCATTTACCTAGGATCTTGGCTTGCACGGATTCATAGGGGCATATCCCTC**  
317 **CCCCTCTTCTTTAGAGTCGTTCTTAAAGATCGCTCTCCACGCCCTAGGCAGGGAAAACGACAAAATCTGGCTCAATTC**

**EcoRV (426)**

396 **CAGGCTAGAACCTACAAATTCAACAGGGATATCGCAAGGATACTGGGGCATAACCCACAGGGAGTCCAAGAATGTGAG**  
475 **GTGGGGGTGGCGAAGGTAATGTCTTTGGTGTGGGAAAAGCAGCAGCCATCTGAGATAGGAACTGGAAAACCAGAGGAGA**  
554 **GGCGTTCAGGAAGATTATGGAGGGGAGGACTGGGCCCCACGAGCGACCAGAGTTGTCACAAGGCCGCAAGAACAGGGG**  
633 **AGGTGGGGGGCTCAGGGACAGAAAAAAAAGTATGTGTATTTTGGAGCAGGGTTGGGAGGCCTCTCTGAAAAGGGTAT**  
712 **AAACGTGGAGTAGGCAATACCCAGGCAAAAAGGGGAGACCAGAGTAGGGGGAGGGGAAGAGTCTGACCCAGGGGAAGAC**

**SaII (808)**

791 **ATTAAAAAGGTAGTGGGGTGCAGTACTAGATGAAGGAGAGCCTTTCTCTCTGGGCAAGAGCGGTGCAATGGTGTGTAAGGT**  
870 **AGCTGAGAAGACGAAAAGGGCAAGCATCTTCTGCTACCAGGCTGGGGAGGCCAGGCCACGACCCCGAGGAGAGGGA**  
949 **ACGCAGGGAGACTGAGGTGACCCTTCTTTCCCCCGGGGCCGGTCTGTGTGGTTTCGGTGTCTTTTTCTGTTGGACCCTT**

**XhoI (1098)**

1028 **ACCTTGACCCAGGCGCTGCCGGGGCTGGGCCCGGGCTGCGGCGCACGGCACTCCCGGGAGGCAGCGAGACTCGAGTTA**  
1107 **GGCCCAACGCGGCGCCACGGCGTTTCTGCGCGGAATGGCCCGTACCCGTGAGGTGGGGGTGGGGGGCAGAAAAGGCG**  
1186 **GAGCGAGCCCAGGCGGGGAGGGGGAGGGCCAGGGGCGAGGGGGCCGGCACTACTGTGTTGGCGGACTGGCGGGACTA**

*NotI* (1298)

1265 **GGGCTGCGTGAGTCTCTGAGCGCAGGCGGGCGGGCCGCCCTCCCCGGCGGCGGCAGCGGGCGGACGCGGGCAGC**  
1344 **TCACTCAGCCCGCTGCCCGAGCGGAAACGCCACTGACCGCACGGGATTCCAGTGCCGGCGCCAGGGGCACGCGGGAC**  
1423 **ACGCCCTCCCGCCGCGCCATTGGCCTCTCCGCCACCGCCCCACACTTATTGGCCGGTGCGCCGCAATCAGCGGGAG**  
1502 **GCTGCCGGGGCCGCTAAAGAAGAGGCTGTGCTTTGGGGCTCCGGCTCCTCAGAGAGCCTCGGCTAGgtaggggatcgg**

*NotI* (1623)

1581 **gactctggcgggagggcggttgggtgcgtttgcggggatgggcggccgagcggccctccgagcgtggtagccggt**  
1660 **ctgtgagacagccgggtacgagtcgtgacgctggaaggggcaagcgggtggtaggaggaatgcggtccgcccctgcagc**

**NotI (1788)**

1739 **aaccggagggggaggggagaagggagcggaaaagtctccaccggacgcggccatggctcggggggggggggcagcggag**  
1818 **gagcgttccggccgacgtctcgtcgctgattggcttcttttctcccgccgtgtgtgaaaacacaattgtactaac**

**AgeI (1938)**

**BspHI (1954)**

1895 **cttcttctctttctctctctgacagGTGTGAAACAGGAAGAGAACCGGTAGGAGGGCCATCATGATTCTGGGGCC**  
1970 **CTGCATGCTGCTGCTGCTGCTGCTGGGCTGAGGCTACAGCTCTCCCTGGGCATCATCCAGTTGAGGAGGAGAAAC**  
2049 **CCGGAATTCTGGAACCGCGAGGCAGCCGAGGCCCTGGGTGCCCGCAAGAAGCTGCAGCCTGCACAGACAGCCCGCAAGA**  
321 **P D F W N R F A A F A I G A A K K I O P A O T A A K**

**BamHI (2182)**

2128 ACCTCATCATCTTCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCTAAAAGGGCAGAAGAAGGA  
 58▶ N L I I F L G D G M G V S T V T A A R I L K G Q K K D  
 2207 CAACTGGGGCCTGAGATACCCCTGGCTATGGACCGCTTCCCATATGTGGCTCTGTCCAAGACATACAATGTAGACAAA  
 84▶ 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 K  
 2286 CATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGTCAAGGGCAACTTCCAGACCATTGGCTTGAGTG  
 111▶ H V P D S G A T A T A Y L C G V K G N F Q T I G L S  
 2365 CAGCCGCCCCGCTTTAACAGTGAACACGACACGCGGCAACGAGGTATCTCCGTGATGAATCGGGCCAAGAAAAGCAGG  
 137▶ A A A R F N Q C N T T R G N E V I S V M N R A K K A G  
 2444 GAAGTCAGTGGGAGTGGTAACCACCACACGAGTGCAGCACGCCTCGCCAGCCGGCACCTACGCCACACGGTGAACCCG  
 163▶ K S V G V V T T T R V Q H A S P A G T Y A H T V N R  
 2523 AACTGGTACTCGGACGCCGACGTGCCTGCCTCGGCCCGCAGGAGGGGTGCCAGGACATCGCTACGCAGCTCATCTCCA  
 190▶ N W Y S D A D V P A S A R Q E G C Q D I A T Q L I S  
 2602 ACATGGACATTGATGTGATCCTGGGTGGAGGCCGAAAGTACATGTTTCGCATGGGAACCCAGACCCTGAGTACCCAGA  
 216▶ N M D I D V I L G G G R K Y M F R M G T P D P E Y P D  
 2681 TGACTACAGCCAAGTGGGACCAGGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCGCCAGGGTGCCCCG  
 242▶ D Y S Q G G T R L D G K N L V Q E W L A K R Q G A R  
 2760 TATGTGTGGAACCGCACTGAGCTCATGCAGGCTTCCCTGGACCCGTCTGTGACCCATCTCATGGGTCTCTTTGAGCCTG  
 269▶ Y V W N R T E L M Q A S L D P S V T H L M G L F E P  
 2839 GAGACATGAAATACGAGATCCACCGAGACTCCACACTGGACCCCTCCTGATGGAGATGACAGAGGCTGCCCTGCGCCT  
 295▶ G D M K Y E I H R D S T L D P S L M E M T E A A L R L  
 2918 GCTGAGCAGGAACCCCGCGGCTTCTTCTCTTCTGAGGGTGGTGCATCGACCACGGTATCACGAAAGCAGGGCT  
 321▶ L S R N P R G F F L F V E G G R I D H G H H E S R A  
 2997 TACCGGGCACTGACTGAGACGATCATGTTTCGACGACGCCATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACG  
 348▶ Y R A L T E T I M F D D A I E R A G Q L T S E E D T  
 3076 TGAGCCTCGTCACTGCCGACCACTCCCACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCT  
 374▶ L S L V T A D H S H V F S F G G Y P L R G S S I F G L  
 3155 GGCCCTGGCAAGGCCCGGGACAGGAAGGCTACACGGTCTCCTATACGAAACGGTCCAGGCTATGTGCTCAAGGAC  
 400▶ A P G K A R D R K A Y T V L L Y G N G P G Y V L K D  
 3234 GCGCCCCGGCCGGATGTTACCGAGAGCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGCCCTGGACGAAG  
 427▶ G A R P D V T E S E S G S P E Y R Q Q S A V P L D E  
 3313 AGACCCACGCAGGCGAGGACGTGGCGGTGTTTCGCGCGCGGCCCGCAGGGCGACCTGGTTCACGGCGTGCAGGAGCAGAC  
 453▶ E T 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  
 3392 CTTCATAGCGCACGTATGGCCTTCGCCGCTGCCTGGAGCCCTACACCGCTGCGACCTGGCGCCCCCGCCGGCACC  
 479▶ 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 T

**NheI (3520)**

3471 ACCGACGCCGCGCACCCGGGGCGGTCCCGGTCCAAGCGTCTGGATTGAAGCTAGCTGGCCAGACATGATAAGATACATT  
 506▶ T D A A H P G R S R S K R L D •  
 3550 GATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT

**HpaI (3660)**

3629 TTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTTATGTTTCAGGTTTCAGGGGGAGGT

**XmnI (3753)**

3708 GTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTAATTTCTAAAATACAGCATAGCAAAACTT  
 3787 TAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATG  
 3866 TGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGATTTTTCCCAAGGTTTGAAC TAGCTCT

**SspI (3999)**

3945 TCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAAATATTAGAAAATAATTTAAATACATC  
 4024 ATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAG  
 4103 TTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTATCCTCAGTCTCGC  
 1254 • D Q  
 4182 TCCTCTGCCACAAAGTGCACGCAGTTGCCGGCCGGGTGCGCAGGGCGAACTCCCGCCCCACGGCTGCTCGCCGATCT  
 121▶ E E A V F H V C N G A P D R L A F E R G W P Q E G I E  
 4261 CGGTATGGCCGGCCCGGAGGCGTCCCGGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGCTCGTCCAGGCC  
 95▶ T M A P G S A D R F N T S V V E S W E A Y L E D L G  
 4340 GCGACCCACACCCAGGCCAGGGTGTGTCCGGCACCACCTGGTCTGACCGCGCTGATGAACAGGGTACGTCGTC  
 69▶ R V W V W A L T N D P V V Q D Q V A S I F L T V D D

**SgrAI (4427)**

4419 CGGACCACACCGGCGAAGTCTGCTCCACGAAGTCCCGGAGAACCAGCGGTCGGTCCAGAACTCGACCGCTCCGG  
 42▶ R V V G A F D D E V F D R S F G L R D T W F E V A G A  
 4498 CGACGTCGCGCGGGTGGACCCGGAACGGCACTGGTCAACTTGGCCATGATGGCTCCTCCTGTGACGAGAGGAAAGAG  
 16▶ V D R A T I V P V A S T I K A M ←

4577 AAGAAGTTAGTACAATTGCTATAGTGAGTTGTATTATACTATGCAGATATACTATGCCAATGATTAATTGTCAAACTA  
4656 GGGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCG  
4735 TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC  
4814 TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT  
4893 GTCCGCCTTTCTCCCTTCGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT  
4972 CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTAGCCCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGT  
5051 CCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG  
5130 TGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAG  
5209 CCAGTTACCTTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTT  
5288 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG  
5367 GAACGAAAACTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCA