



SdaI (6)

1 CCTGCAGGGCCGCTGCTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCAGGGAGGCAGAGGTTGTGGTGAGCAGAGATCGGCCATTGCTCTCCAGCCT
101 GGGCAACAAGAGCAAAAGTTCGTTTAAAAAAGTCTTTTCGATGTGACTGTCTCTCCCAAATTTGTAGACCTCTTAAGATCATGCTTTTCA
201 GATACTTCAAAGATTCCAGAAGATATGCCCCGGGGTCTGGAAGCCACAAGGTAACACACACATCCCCCTCTTGACTATCAATTTTACTAGAGGAT
301 GTGGTGGGAAAACCATTTTATATTAACAATAAGGCTTGGGATGGAGTAGGATGCAAGCTCCCCAGGAAAGTTTAAAGATAAAACCTGAGACTTAAA
401 AGGGTGTAAAGAGTGGCAGCTAGGGAATTTATCCCGACTCCGGGGAGGGGGCAGAGTACCAGCCTCTGCATTTAGGGATTCTCCGAGGAAAAGTGT
501 GAGAACGGCTGCAGGCAACCCAGGCGTCCCGGGCTAGGAGGAGCACCAGGCTGCGCGAAGAGAGGGAGAAAGTGAAGCTGGGAGTTGCCACTCCC
601 AGACTTGTGGAATGCAGTTGGAGGGGGGAGCTGGGAGCGCGCTTCTCCAATCACAGGAGAAGGAGGAGGTGGAGGAGGAGGGCTGCTTGGGAAGT
701 ATAAGAATGAAGTTGTGAAGCTGAGATTCCTCCATTGGGACCGGAGAAACAGGGGAGCCCCCGGGCAGCCGCGCCCTTCCACGGGGCCCTTT
801 ACTGCGCCGCGCCCGCCCGCCACCCTCGCAGCACCCTCGCCCGCCCGCCCTCCAGCCGGTCCAGCCGGAGCCGTGGGGCCGAGCCGCAGTGAGC

NcoI (902)

NheI (940)

901 ACCATGGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGACAGCAAATGGGTCGGGATCTGTACGACGATGACGATAAGGTAC
▶ M G G S H H H H H G M A S M T G G Q Q M G R D L Y D D D D K V
1001 CTAAGGATCAGCTTGGAGTTGATCCCGTCTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAATTAATCGCCTTGACGACATCCCCCTTT
33▶ P K D Q L G V D P V V L Q R R D W E N P G V T Q L N R L A A H P P F
1101 CGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCATCGCCCTTCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTCCTGGTTCCGGCACCA
66▶ A S W R N S E E A R T D R P S Q Q L R S L N G E W R F A W F P A P
1201 GAAGCGGTCCGGAAGCTGGCTGGAGTGCATCTCTGAGGCCGATACTGTCGTGCTCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCCATCT
100▶ E A V P E S W L E C D L P E A D T V V V P S N W Q M H G Y D A P I
1301 ACACCAACGTAACCTATCCATTACGGTCAATCCCGCTTTGTTCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTG
133▶ Y T N V T Y P I T V N P P F V P T E N P T G C Y S L T F N V D E S W
1401 GCTACAGGAAGCCAGACGCGAATTATTTTATGATGGCGTAACTCGGCGTTTTCATCTGTGGTGAACGGGCGTGGGTCGGTTACGGCCAGGACAGTCGT
166▶ L Q E G Q T R I I F D G V N S A F H L W C N G R W V G Y G Q D S R
1501 TTGCCGCTGAATTTGACCTGAGCGCATTTTACGCGCCGGAGAAAACCGCTCGCGGTGATGGTGTGCGTTGGAGTGACGGCAGTTATCTGGAAGATC
200▶ L P S E F D L S A F L R A G E N R L A V M V L R W S D G S Y L E D
1601 AGGATATGTGGGGATGAGCGGCATTTCCGTGACGTCCTGTTGCTGATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATGA
233▶ Q D M W R M S G I F R D V S L L H K P T T Q I S D F H V A T R F N D
1701 TGATTTCAAGCGCTGACTGGAGTTCAGATGTCGCGGCGACTGCTGACTACCTACGGGTAACAGTTTCTTATGGCAGGGTGAACCGGAG
266▶ D F S R A V L E A E V Q M C G E L R D Y L R V T V S L W Q G E T Q
1801 GTCGCCAGCGGCACCGCCCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGCGTCACACTACGTCGAACGTCGAAAACCGGAAAC
300▶ V A S G T A P F G G E I I D E R G G Y A D R V T L R L N V E N P K
1901 TGTGGAGCGCGAAATCCGAATCTCTATCGTGGTGGTGAAGTGCACACCGCGCAGCGCAGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTCCG
333▶ L W S A E I P N L Y R A V V E L H T A D G T L I E A E A C D V G F R
2001 CGAGGTGCGGATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGGTCAGGTC
366▶ E V R I E N G L L L L N G K P L L I R G V N R H E H H P L H G Q V
2101 ATGGATGAGCAGACGATGGTGCAGGATATCTGCTGATGAAGCAGAACTTTAACGCCGTGCGCTGTTTCGATTATCCGAACCATCCGCTGTGGTACA
400▶ M D E Q T M V Q D T I L L M K Q N N F N A V R C S H Y P N H P L Q G G F
2201 CGCTGTGCGACGCTACGGCCTGATGTGGTGGATGAAGCCAAATATTGAAACCCAGCGCATGTTGCCAATGAATCGTCTGACCATGATCCGCGCTGGCT
433▶ T L C D R Y G L Y V V D E A N I E T H G M V P M N R L T D D P R W L
2301 ACCGGCGATGAGCGAACCGTAAACCGAATGGTGCAGCGCATCGTAATCACCCGAGTGTGATCATCTGGTCTGGGGAATGAATCAGGCCACGGCGCT
466▶ P A M S E R V T R M V Q R D R N H P S V I I W S L G N E S G H G A
2401 AATCACGACGCGTGTATCGTGGATCAAATCTGTCGATCTTCCCGCCGGTGCAGTATGAAGCGCGGAGCCGACACCACGGCCACCGATATTTT
500▶ N H D A L Y R W I K S V D P S R P V Q Y E G G G A D T T A T D I I
2501 GCCCCGATGTACGCGCGTGGATGAAGACCGCCCTTCCCGGCTGTGCCAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACGCGCCCGCT
533▶ C P M Y A R V D E D Q P F P A V P K W S I K K W L S L P G E T R P L
2601 GATCCTTTGCGAATACGCCACGCGATGGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTTC
566▶ I L C E Y A H A M G N S L G G F A K Y W Q A F R Q Y P R L Q G G F
2701 GTCTGGGACTGGTGGATCAGTCGCTGATTAATAATGATGAAACCGCAACCCGTTGGTGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCC
600▶ V W D W V D Q S L I K Y D E N G N P W S A Y G G D F G D T P N D R
2801 AGTTCTGTATGAACGGTCTGGTCTTTCGCGACCGCAGCCGATCCAGCGCTGACGGAAGCAAAACACCAGCAGAGTTTTTCCAGTCCGTTTATCCGG
633▶ Q F C M N G L V F A D R T P H P A L T E A K H Q Q Q F F Q F R L S G
2901 GCAAACCATCGAAGTACCAGCGAATACCTGTTCCGTATAGCGATAACGAGCTCCTGACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGT
666▶ Q T I E V T S E Y L F R H S D N E L L H W M V A L D G K P L A S G
3001 GAAGTGCCTCTGGATGTCGCTCCACAAGGTAACAGTTGATTGAACTGCCTGAACTACCAGCCGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCG
700▶ E V P L D V A P Q G K Q L I E L P E L P Q P E S A G Q L W L T V R
3101 TAGTGCAACCGACCGCAGCCGATGTCAGAAGCCGGGACATCAGCCGCTGGCAGCAGTGGCGCTGTCGGGAAACCTCAGTGTGACGCTCCCGCCCG
733▶ V V Q P N A T A W S E A G H I S A W Q W R L A E N L S V T L P Q G A A
3201 VTCCACGCCATCCCGCATCTGACCACCGAAGTGGATTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATTAACCGCCAGTCCAGCTTCTTTCA
766▶ S H A I P H L T T S E M D F C I E L G N K R W Q F N R Q S G F L S
3301 CAGATGTGGATTGGCGATAAAAAAACAAGTCTGACGCCGCTGCGGATCAGTTACCCGTCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCC
800▶ O M W I G D K K O I I T P I R D O F T R A P I D N D I G V S F A T

3401 GCATTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGGGGCCATTACCAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGTGTATGC
833▶ R I D P N A W V E R W K A A G H Y Q A E A A L L Q C T A D T L A D A
3501 GGTGCTGATTACGACCGCTCACGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATT
866▶ V L I T T A H A W Q H Q G K T L F I S R K T Y R I D G S G Q M A I
3601 ACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGCGGATTGGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAACTGGCTCGGAT
900▶ T V D V E V A S D T P H P A R I G L N C Q L A Q V A E R V N W L G
3701 TAGGGCCGCAAGAAAATATCCCGACCGCTTACTGCCGCTGTTTTGACCGCTGGGATCTGCCATTGTGACAGATGTATACCCCGTACGCTTTCCCGAG
933▶ L G P Q E N Y P D R L T A A C F D R W D L P L S D M Y T P Y V F P S
3801 CGAAAACGGTCTGCGCTGCGGACGCGCAATTGAATTATGGCCACACCAAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAA
966▶ E N G L R C G T R E L N Y G P H Q W R G D F Q F N I S R Y S Q Q Q
3901 CTGATGAAAACAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTCCATATGGGGATTGGTGGCGACTCCTGGA
1000▶ L M E T S H R H L L H A E E G T W L N I D G F H M G I G G D D S W

NheI (4093)

EcoRI (4087)

4001 GCCCGTCAGTATCGGCGGAATTACAGCTGAGCGCCGGTCTGCTACCATTACCAGTTGGTCTGGTGTCAAAAATAATAATCTAGTCGAGAATTCGCTAGCTC
1033▶ S P S V S A E L Q L S A G R Y H Y Q L V W C Q K •
4101 GACATGATAAGATACATTGATGAGTTTGACAAAACCACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG

4201 TGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTCATTCTTTATGTTTCAGGTTCCAGGGG

PacI (4368)

4301 GAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAGATCCATTTAAATGTTAATTAAGTACCGATGACCAAAATCCCTTAACGTGA

4401 GTTTTCGTTCCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGAACAACAAAAA

4501 CCACCGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTGTTC

4601 TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG

4701 CGATAAGTCGTGCTTACCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGTTCTGTCACACAGCCAGCTTG

4801 GAGCGAACGACTACACCGAAGTCTGAGATACCTACAGCGTGTGCTATGAGAAAAGCCACGCTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG

4901 GCAGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCC

5001 ATTTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCTGTCCTTTTGTGTCCTTTTGTGCTCACATG

PacI (5108)

5101 TTCTTAATTAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAA
▶ 1▶ M A K

5201 GTTGACCAGTGTGTCCAGTGTCTACAGCCAGGGATGTGGCTGGAGCTGTTGAGTCTGGACTGACAGGTTGGGGTCTCCAGAGATTTTGTGGAGGAT
3▶ L T S A V P V L T A R D V A G A V E F W T D R L G F S R D F V E D

5301 GACTTTGCAGGTGTGGTCTGAGATGATGTCACCCTGTTTCATCTCAGCAGTCCAGGACCGGTTGGTGCCTGACAACCCCTGGCTTGGGTGTGGGTGAGAG
37▶ D F A G V V R D D V T L F I S A V Q D Q V V P D N T L A W V W V R

5401 GACTGGATGAGCTGTATGCTGAGTGGAGTGGTGGTCTCCACCACTTCCAGGGATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGG
70▶ G L D E L Y A E W S E V V S T N F R D A S G P A M T E I G E Q P W G

5501 GAGAGAGTTTGCCTGAGAGACCCAGCAGGCACTGTGTGCACTTTGTGGCAGAGGAGCAGGACTGAGGATAAGAATTGTAACAAAAACCCGCCCGG
103▶ R E F A L R D P A G N C V H F V A E E Q D •

PacI (5617)

5601 CGGGGTTTTTGTAAATTAA
▶