



150

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCCGCCCTCCCGCTGTGGTGCCTCCTGAAGCTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGAGCTACCTAGACTCAGCCGGCTCTCCAGGCTTTGCTGACCCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

Agel (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCATGACCACCATGGTCAACGTGGACACCCCTCCAGAATA
1 M T T M V N V D T L P E Y
601 TGAGAAGAGTCAGATCAAGAGAGCGCTGGAGTTAGGGACGGTGTGACCGTGTCAACGCCCGCAAATCCACCCAGAGCGGAGGACAGTACAGATGATC
13 E K S Q I K R A L E L G T V M T V F N A R K S T P E R R T V Q M I
701 ATGGAGACTCGGCAGGTGGCATGGAGCAAGACCGCAGACAAGATTGAAGGCTTCTGGACATCATGGAGATAAAGGAAATCCGTCCGGGGAAGAACTCCA
47 M E T R Q V A W S K T A D K I E G F L D I M E I K E I R P G K N S
801 AGGACTTTGAGCGAGCAAGGCTGCCGCCACAAGGCAGAATGTTGCTTACCATCCTACGGCACCCAGTTTGTCTCAGCAGCTCAGTTTGGCAAC
80 K D F E R A K A V R H K A E C C F T I L Y G T Q F V L S T L S L A T
901 GGACTCAAAGGAGGATGCGGTGAAGTGGCTCTCTGTTTGAAGATCCTACACCAGGAAGCGATGAGTGCATCCACCCCAACAATGATTGAGAGTTGGCTG
113 D S K E D A V K W L S G L K I L H Q E A M S A S T P T M I E S W L
1001 AGAAAACAGATTTATTAGTAGATCAAACCCGAAGAAACAGCATCAGCCTCCGGGAGCTGAAGACCATTTGCCCTGGTCAACTTCAAAGTGAAGCGGCA
147 R K Q I Y S V D Q T R R N S I S L R E L K T I L P L V N F K V S G
1101 TCAAGTTTCTCAAGGACAAGCTGGTGGAAATCGGCGCACAGAAAGATGAGCTCAGCTTTGAACAGTCCATCTCTTATAAGAACTCATGTTTGACCA
180 I K F L K D K L V E I G A Q K D E L S F E Q F H L F Y K K L M F D Q
1201 GCAAAAATCGATACTTGACGAATTCAAAAAGACTCCTCGTGTTCATCTAGGAAACACAGACCGCCGGATGCCTCGGCTGTCTACCTGCAGGACTTC
213 Q K S I L D E F K K D S S V F I L G N T D R P D A S A V Y L Q D F
1301 CAAAGTTTCTTTTACATGAACAGCAGGAGCTATGGGCTCAGGATCTGAACAAAGTCCGGGAGCGGATGACCAAGTTTATCGACGACACAATGAGAGAGA
247 Q R F L L H E Q Q E L W A Q D L N K V R E R M T K F I D D T M R E
1401 CCGCAGAGCCCTTCTTATTTGGTATGAGTTCCTCAGTATCTGTTCTCCCGGAGAACAGCATCTGGGACGAGAAGTACGATGCAGTGGACATGCAGGA
280 T A E P F L F V D E F L T Y L F S R E N S I W D E K Y D A V D M Q D
1501 CATGAACAACCCCTGTCCCACTACTGGATCTCCTTCCCAACACGTACTCTACTGGGACCAGCTGCGTAGTGAAGTCTCCACGGAAGCGTATATC
313 M N N P L S H Y W I S S S H N T Y L T G D Q L R S E S S T E A Y I
1601 CGTGTCTGCGCGCTGGTTGCCGCTGACTGCTGGACTGCTGGGACGGGCTGATGGGAAGCCATTATCTACCACGGCTGGACACCAAGCAAGA
347 R C L R A G C R C I E L D C W D G P D G K P I I Y H G W T R T T K
1701 TCAAGTTTGTATGTTGTTGAGCCATCCGAGACCACGCCTTTGTTACCTCCAGCTTCCCGTCTCTGTCTATCGAGGAGCACTGCAGTGTGGAGCA
380 I K F D D V V Q A I R D H A F V T S S F P V I L S I E E H C S V E Q
1801 GCAACGTACATGGCCAAGGTGTTCAAGGAAGTGTAGGAGACATGCTGTTGACGAAGCCACAGAGGCCAGTGTGACCAGCTGCCCTGCCAGTCAG
413 Q R H M A K V F K E V L G D M L L T K P T E A S A D Q L P S P S Q
1901 CTTCTGAGAAGATCATTATCAAGCATAAGAAGCTGGTCCCAAGGGTGTGTCGACGTCAACGTGGAGGACAAGAAAGATGAGCACAAGCCCGAGGGCG
447 L R E K I I K H K G V K P K G D V D V N V E D K K D E H K P Q G
2001 AGCTGTATATGTTGGACTCCATCGACCAGAAATGGACGCGCCACTACTGTCCATTGCTGACGCAAGCTGTCTTCCGGTGTGACATTGAACAGGCTGT
480 E L Y M W D S I D Q K W T R H Y C A I A D A K L S F G D D I E Q A V
2101 GGAGGAGGAGCCGGTCCAGGACACTCCCCACGGAGCTGCATTTGGGGAGAAATGGTCCACAAGAAGGTGGAGAGTAGGACAGTGGGAGAGGCTG
513 E E E P V Q D T P P T E L H F G E K W F H K K V E S R T S A E K L
2201 CTGCAAGAGTACTGTGCTGAGACCGGGGCAAGGACGGCACCTTCTGGTGGGGAGAGCGAGACCTTCCCAATGACTACACACTCTCTTCTGGCGGT
547 L Q E Y C A E T G A K D G T F L V R E S E T F P N D Y T L S F W R

BspHI (2348)

2301 CTGGCCGGGTGCAGCACTGCCGGATCCGGTCTACTATGGAGAATGGGGTCATGAAGTACTACCTGACCAGCAACCTCACGTTCAACAGCATCTACGCCT
580 S G R V Q H C R I R S T M E N G V M K Y Y L T D N L T F N S I Y A L
2401 CATCCAGCACTACCGTGAGGCACACCTGCGCTGCGGGAGTTCGAGCTGCGGCTCACAGACCCGGTGCCAACCCCAACCCACAGAGTCCAAGCCGTGG
613 I Q H Y R E A H L R C A E F E L R L T D P V P N P N P H E S K P W
2501 TACTATGACAGTCTGAGCAGGGGGAAGCGGAGGACATGCTGATGCGGATCCCCGAGATGGAGCCTTCTCATCCGGAAGAGGGAGGGGACCAACTCT
647 Y Y D S L S R G E A E D M L M R I P R D G A F L I R K R E G T N S
2601 ACGCCATCACCTTCAGGGCCAGGGCAAGGTGAAACATTGTCGCATCAACCGGACGGACGCCACTTCTGTGTTGGGACCTCTGCCTACTTTGAGAGCT
680 Y A I T F R A R G K V K H C R I N R D G R H F V L G T S A Y F E S L
2701 AGTGGAGCTTGTGAGCTACTACGAGAAGCAGCGCTCTATCGCAAGATGAGGCTTCGATACCCCGTACCCCGGAGCTCCTGGAGCGATATAATATGGAA
713 V E L V S Y Y E K H A L Y R K M R L R Y P V T P E L L E R Y N M E
2801 AGAGACATCAACTCTTTATGACGTCAGCAGGATGTATGTGGACCAAGTGAAGTCAACCTTCCATGCCTCAGAGGACCGTGAAGCGCTCTATGACT
747 R D I N S L Y D V S R M Y V D P S E I N P S M P Q R T V K A L Y D
2901 ACAAAGCAAGCGCAGCGACGAGCTGACCTTCTGCCGCGTGCCTCATCCACAATGTCTCAAGGAGCCCGGGGATGGTGGAAAGGGGACTACGGGAC
780 Y K A K R R S D E L T F C R G A L I H N V S K E P G G W W K G D Y G T
3001 CCGGATCCAGCAGTACTTCCCTCCAATGTCGAGGATATCTCGGACGGATGCTGAGGAGATGGAAAAGCAGATTATTGAAGCAATCCCTCGGG
813 R I Q Q Y F P S N Y V E D I S A G D A E E M E K Q I I E D N P L G
3101 TCTCTGCAAGGCATATTGGATCTTAATACCTACAACGTTGTGAAGGCCCCCGGGGAAAACCAGAAAGCCTTCTGTTTATCTGGAGCCTAAGA
847 S L C K G I L D L N T Y N V V K A P Q G K N Q K A F V F I L E P K

3201 AACAGGGAGATCCTCCCGTGGAGTTCGCGACCGACCGTGTGGAGGAGCTGTTTGTAGTGGTTTCAGAGCATAACGGGAGATCACTTGAAGATGGACACCAA
880▶ K Q G D P P V E F A T D R V E E L F E W F Q S I R E I T W K M D T K
3301 GGAGAACAACATGAAGTACTGGGAGAGGAACCAAGTCCATCGCCATTGAGCTCTCGACTTGGTTGTATACTGCAAGCCAACCTAGCAAAACCAAGGACCAT
913▶ E N N M K Y W E R N Q S I A I E L S D L V V Y C K P T S K T K D H
3401 TTGGAAAATCCCGACTTCCGAGAAATCCGCTCTTTTGTGGAGACGAAGGCAGACAGCATTGTCAAGGCAAAAACCTGTGCATCTATTGCGATAACAATCAGA
947▶ L E N P D F R E I R S F V E T K A D S I V R Q K P V D L L R Y N Q
3501 AGGGCCTGACCCGAGTCTACCCCAAGGGACAGAGAGTTGACTCTTCAACTATGACCCCTTCCGCTGTGGTGTGCGGCTCCAGATGGTGGCACTCAA
980▶ K G L T R V Y P K G Q R V D S S N Y D P F R L W L C G S Q M V A L N
3601 TTTCCAGACTCCAGACAATAACATGCAGATGAACCACGCGCTGTTTTCTGCTCAATGGGCGAACAGGCTACGCTCTGCAGCCTGAGAGCATGCGCTCTGAG
1013▶ F Q T P D K Y M Q M N H A L F S L N G R T G Y V L Q P E S M R S E
3701 AAGTATGATCCGATGCCCTGGAGTCCAGAGGAAGATCCTGATGACACTCACTGTCAAGGTTCTTGGTGCACGCCACCTCCCTAACTAGGGCGGAGTA
1047▶ K Y D P M P L E S Q R K I L M T L T V K V L G A R H L P K L G R S
3801 TTGCTGTCTTTTGGAGTGGAAATCTGCGGGGCTGAGTATGACAGCAACAAATCAAGACGACGTTGTGAATGACAATGGCCTCAGCCCTGAGCTGTCTG
1080▶ I A C P F V E V E I C G A E Y D S N K F K T T V V N D N G L S P V W
3901 GGCTCAAACCTCAGGAGAAGGTGACATTTGAGATCTATGATCCAAACCTTGCTTCTACGCTTGTGGTCTACGAAGAAGACATGTTCACTGATCCCAAC
1113▶ A P T Q E K V T F E I Y D P N L A F L R F V V Y E E D M F S D P N
4001 TTCCTGGCTCATGCCAGTACCCCATTAAGGCATCAATCCGGGTTAGATCAGTCCCTCTGAAGAATGGGTACAGTGAAGACATCGAGCTGGCATCCC
1147▶ F L A H A T Y P I K G I K S G F R S V P L K N G Y S E D I E L A S
4101 TCCTGGTTTTCTGTGAGATGAGGCCAGTCTGGAGAGTGAAGAGGAACCTATTCTCTGTGCGCAACTGCGGAGCGGCAAGAAGAGCTCAACAACCA
1180▶ L L V F C E M R P V L E S E E E L Y S S C R Q L R R R Q E E L N N Q
4201 GCTCTTCTTTACGACACGCACAGAACCTGCGCGGAGCAACCGTGTGCTGTTGAAGGAGTTCAATGTTAACGAGAATCAGTGCAGCTGTACCAG
1213▶ L F L Y D T H Q N L R G A N R D A L V K E F N V N E N Q L Q L Y Q

NheI (4365)

4301 GAGAAGTGAACCGGAGGCTGAGAGAGAAGAGAGTAAGCAACAGCAGGTTCTACTCCTAGAAGCCGCTAGCTGGCCAGACATGATAAGATACATTGATGA
1247▶ E K C N R R L R E K R V S N S R F Y S •
4401 GTTTGGACAAACCACAACCTAGAATGCAGTGAIAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAA
4501 CAAGTTAAACAACAACATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGG
4601 AATTCTAAAATACAGCATAGCAAACTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTT
4701 GCCAATGTGATTAGCTGTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTGCTCTTCATTTCTTTATGT
4801 TTTAAATGCAGTACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGCGAGA
4901 ATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAATTTGGACAGCAAGAAAGCGAGC
5001 TTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCACTCAATGAGCACAAAGCAGTCAAGGAG
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I L V F C D P A
5101 CATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGT
110▶ Y D S I L E R C M G C P S V V R I S R D V E D S Y P H R V A V I T
5201 GTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCAAGCACAGACAGTACCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAG
77▶ D F D K Q G N S V A S G I A I A E A C V T V R G I Y A E I H V A S
5301 ATGATCTCCCAGTCTTGGTCTGATGGCCCGCCGACATGGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCA
43▶ I I E G T K T R I A A G V H H K N D E Y L M T I K E T A V E V L E L
5401 GATCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGTGATGCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAAC
10▶ D Q Q S I N F T K M
5501 AGCGTGGATGGCTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGG
5601 GGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCTGAG
5701 TCAAACCGCTATCCACGCCATTGATGTAAGTACTGCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTACTGCAAGTAGGAAAGTCCC
5801 ATAAGTCACTGACTGGGCATAATGCCAGGCGGCCATTTACCCTGTTGACGTCAATAGGGGGCTACTTGGCATATGATACTTGTACTGCAAGTACTGCAAA
5901 GTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGG
6001 GGGTCTGGGCGGTGAGCCAGGCGGCCATTTACCCTAAGTTATGTAACGCTGACGTTAATTAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
6101 GAACCGTAAAAAGGCCGCTGTTGCTGGCCTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCG
6201 ACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCC
6301 CTTCCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGT
6401 TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT
6501 AGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA

6601 AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGGG
6701 CAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGT
6801 TAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCA
6901 TCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA