



1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

Agel (552) SphI (568)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCCAGCATGCCACATACTTTGTGGATGGTGGGGTC  
1 M P H T L W M V W V

601 TTGGGGTTCATCATCAGCCTCTCCAAGGAAGAATCCTCCAATCAGGCTTCTGTCTTGTGACCGCAATGGTATCTGCAAGGGCAGCTCAGGATCTTTAA  
111 L G V I I S L S K E E S S N Q A S L S C D R N G I C K G S S G S L  
701 ACTCCATTCCCTCAGGGCTCACAGAAGCTGTAAGGCTTACCTGTCCAACAACAGGATCACCTACATTAGCAACAGTGCCTACAGAGGTGTGTGAA  
44 N S I P S G L T E A V K S L D L S N N R I T Y I S N S D L Q R C V N  
801 CCTCCAGGCTCTGGTGTGACATCCAATGGAATTAACACAATAGAGGAAGATTCTTTTTCTCCCTGGGCAGTCTTGAACATTTAGACTTATCCTATAAT  
77 L Q A L V L T S N G I N T I E E D S F S S L G S L E H L D L S Y N  
901 TACTTACTAATTTATCGTCTTCTGGTCAAGCCCTTCTTTAACATTTAAACTTACTGGGAAATCCTTACAAAACCTAGGGGAAACATCTC  
111 Y L S N L S S W F K P L S S L T F L N L L G N P Y K T L G E T S  
1001 TTTTTCTCATCTCACAAAATGCAAACTCTGAGAGTGGGAAATATGGACACCTTCACTAAGATTCAAAGAAAAGATTTGTGGACTTACCTTCTTGA  
144 L F S H L T K L Q I L R V G N M D T F T K I Q R K D F A G L T F L E  
1101 GGAAGTTCAGATTGATGCTTACAGATCTACAGAGCTATGAGCCAAAAGTTTGAAGTCAATTCAGAATGTAAGTTCATCTGATCCTTATGAGCAGCAT  
177 E L E I D A S D L Q S Y E P K S L K S I Q N V S H L I L H M K Q H  
1201 ATTTTACTGCTGGAGATTTTGTAGATGTTACAAGTCCGTTGGAATGTTTGAAGTGCAGATACTGATTTGGACACTTTCCATTTTTCAGAACTATCCA  
211 I L L L E I F V D V T S S V E C L E L R D T D L D T F H F S E L S  
1301 CTGGTGAACAAATTCATTGATTAAGGTTTACATTTAGAAATGTGAAATCACCGATGAAAGTTTGTTCAGGTTTGAACCTTTTGAATCAGATTTT  
244 T G E T N S L I K K F T F R N V K I T D E S L F Q V M K L L N Q I S  
1401 TGGATTGTTAGAATTAGAGTTTGTGACTGTACCCTTAATGGAGTGGTAATTTAGAGCATCTGATAATGACAGAGTTATAGATCCAGGTAAGTGGAA  
277 G L L E L E F D D C T L N G V G N F R A S D N D R V I D P G K V E  
1501 ACGTTAACATCCGAGGCTGCATATCCAAGTTTTACTTATTTTATGATCTGAGCACTTTATATTCACTTACAGAAAGAGTTAAAAGAATCACAGTAG  
311 T L T I R R L H I P R F Y L F Y D L S T L Y S L T E R V K R I T V  
1601 AAAACAGTAAAGTTTTCTGGTTCCTTGTACTTTTACAACATTTAAAATCATTAGAATACTGGATCTCAGTGAATTTGATGGTTGAAGAATACTT  
344 E N S A K V F L V P C L L S Q H L K S L E Y L D L S E N L M V E E Y L  
1701 GAAAAATTCAGCTGTGAGGATGCCTGGCCCTCTACAACATTTAATTTAAGGCAAAATCATTGGCATCATTGGAAAAACCGGAGAGACTTTGCTC  
377 K N S A C E D A W P S L Q T L I L R Q N H L A S L E K T G E T L L  
1801 ACTCTGAAAACTTGACTAACATTGATATCAGTAAGAATAGTTTTTATTCTATGCCTGAAACTTGTGAGTGGCCAGAAAAGATGAAATATTTGAACTTAT  
411 T L K N L T N I D I S K N S F H S M P E T C Q W P E K M K Y L N L  
1901 CCAGCACGAATACACAGTGAACAGGCTGCATTTCCCAAGACTGGAAATTTAGATGTTAGCAACAACAATCTCAATTTATTTTCTTGAATTTGCC  
444 S S T R I H S V T G C I P K T L E I L D V S N N N L N L F S L N L P  
2001 GCAACTCAAGAATTTATTTTCCAGAAATAAGTTGATGACTTACCAGTGCCTCCCTTACCCTGTTACTAGTATTGAAAATCAGTAGGAATGCA  
477 Q L K E L Y I S R N K L M T L P D A S L L P M L L V L K I S R N A  
2101 ATAACACGTTTTCTAAGGAGCAACTTGACTCATTTCACACTGAAGACTTTGGAAGCTGGTGGCAATACTTCAATTTGCTCCTGTGAATCCTCTCCT  
511 I T T F S K E Q L D S F H T L K T L E A G G N N F I C S C E F L S  
2201 TCACTCAGGAGCAGCAAGCACTGGCCAAAGTCTTGATTGATTGGCCAGCAAATACCTGTGTGACTCTCCATCCCATGTGCGTGGCCAGCAGGTTCAGGA  
544 F T Q E Q Q A L A K V L I D W P A N Y L C D S P S H V R G Q Q V Q D  
2301 TGTCGGCCTCTCGGTGTGCGAATGTCACAGGACAGCACTGGTGTCTGGCATGTGCTGTGCTGTTTCTGCTGATCCTGCTACGCGGGGCTGTGCCAC  
577 V R L S V S E C H R T A L V S G M C C A L F L L I L L T G V L C H  
2401 CGTTTCCATGGCTGTGGTATATGAAAATGATGTGGGCTGGCTCCAGGCCAAAAGGAAGCCAGGAAAGCTCCAGCAGGAACATCTGCTATGATGCAT  
611 R F H G L W Y M K M M W A W L Q A K R K P R K A P S R N I C Y D A  
2501 TTGTTTCTACAGTGAGCGGATGCCTACTGGTGGAGAACCTTATGGTCCAGGAGCTGGAGAACCTCAATCCCCCTCAAGTTGTGCTTTCATAAGCG  
644 F V S Y S E R D A Y W V E N L M V Q E L E N F N P P F K L C L H K R  
2601 GGACTTCATTCTGGCAAGTGGATCATTGACAATATCATTGACTCCATTGAAAAGAGCCACAAAATGTCTTTGTGCTTTCTGAAAATTTGTGAAGAGT  
677 D F I P G K W I I D N I I D S I E K S H K T V F V L S E N F V K S  
2701 GAGTGGTCAAGTATGAACTGGACTTCTCCATTTCCGCTTTTTGATGAGAACAAATGATGCTGCCATTCTCATTCTTCTGGAGCCATTGAGAAAAAAG  
711 E W C K Y E L D F S H F R L F D E N N D A A I L I L L E P I E K K  
2801 CCATTTCCAGGCTTCTGCAAGCTGCGGAAGATAATGAACACCAAGCACTACCTGGAGTGGCCATGGACGAGGCTCAGCGGAAGGATTTTGGGTA  
744 A I P Q R F C K L R K I M N T K T Y L E W P M D E A Q R E G F W V N

BamHI (2922)

2901 TCTGAGAGCTGCGATAAAGTCCGGATCCTATCCCTATGATGTGCCAGACTATGCTGGCTATCCATATGATGTTCTGATTATGCTGGATACCCCTATGAT  
777 L R A A I K S G S Y P Y D V P D Y A G Y P Y D V P D Y A G Y P Y D

NheI (3019)

3001 GTGCCAGACTATGCCTAAAGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATT  
811 V P D Y A  
3101 TGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAATTGCATTCATTTATGTTTCAGGTTACGG

3201 GGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCC  
3301 TCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTTCATGGAGT  
3401 TTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTCTTTCATTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTATGAAAATATTCA  
3501 GAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTT  
3601 GGACTTAGGGAACAAAGGAACCTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAA  
3701 TGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCAC  
128 T T K V L K G N M E I L V F C D P A Y D S I L E R C M G C P S V V  
3801 CCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCAGCAGACCCAATGGCAATG  
95 R I S R D V E D S Y P H R V A V I T D F D K Q G N S V A S G I A I  
3901 GCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGT  
61 A E A C V T V R G I Y A E I H V A S I I E G T K T R I A A G V H H K  
4001 TGTGTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATA  
28 N D E Y L M T I K E T A V E V L E L D Q Q S I N F T K M  
4101 GTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGC  
4201 TCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTACTAGT  
4301 CAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCA  
4401 TCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAAGTCAATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTC  
4501 ATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTGCAAGTGGGCGAGTTTACCCTAAATACTCCACCATTGACGTCAATGGAAA  
4601 GTCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCCTAAGTTATG  
4701 TAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTCTGAGGCTTTTCCATAGGCTCCGC  
4801 CCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAAGAGGTGGCGAAACCCGACAGGACTATAAGATAACCAGGCGTTTCCCCTGGAAGCTCCCTCG  
4901 TGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCT  
5001 CAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGTGCCTTATCCGGTAACTATCGTCTTGTAGTCC  
5101 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGT  
5201 GGCCTAACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA  
5301 ACAAACCACCGCTGGTAGCGGTGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGG  
5401 TCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGATTTTGGTATGGCTAGTTAATTAACATTTAAATCAGCGGCCAATAAATATCTTTATTTTC  
5501 ATTACATCTGTGTGTTGTTTTTTGTGTTGAATCGTAACATAACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCC  
5601 AGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA