



125
|-----|

1 GGATCTGCATCGCTCCGGTCCCGTCACTGGGCGAGAGCGACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

Agel (552) NcoI (568)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCCACCATGGACATGATGGGGCTGCCAGGGACCAGC
1 M D M M G L P G T S
601 AAACACATCACCTTCTCTGCTTGGCAGCTAGGCGCCTCAGGCCAGGTGATGGCTGCTGCGTTGAGAAGACATCGTTCCAGAGGGAGCCTCAGGCT
11 K H I T F L L L C Q L G A S G P G D G C C V E K T S F P E G A S G
701 CACCTTAGGACCCAGGAAGTGTGCTACAGGTTTCCAAGACAGACTATGAGTGTCTGGCAGTATGATGGCCCTGAGGACATGTTTCTCACGT
44 S P L G P R N L S C Y R V S K T D Y E C S W Q Y D G P E D N V S H V
801 CCTGTGGTGTGCTTGTCCCTCCGAACCATACCCACACCGGCCAGGAGCGTGCCTGCTACTTCTCTCAGGCCAGACCCGACTGTGCAATCTGGGAA
77 L W C C F V P P N H T H T G Q E R C R Y F S S G P D R T V Q F W E
901 CAGGCGGTATCCCTGTGCTGTCCAAGTCAACTTCTGGTGGAGTCTCGGCTTGGGAACCAACCATGAAGTCCAGAAAGATATCCAGTACCTGTACA
111 Q D G I P V L S K V N F W V E S R L G N R T M K S Q K I S Q Y L Y
1001 ACTGGACCAAGACGACCCTCCCTGGGACACATCAAGGTGTCAACATCACACCGGAGTTCGAATGGACTGGAATGTGTCTGAAGAGGCCGGTGTGA
144 N W T K T T P P L G H I K V S Q S H R Q L R M D W N V S E E A G A E
1101 GGTACAGTTCAGGCGCGTATGCCACAACGAATTGGACCTTGGGTGACTGCGGACCTCAGGTTAACTCTGGCTCAGTGTGCTTGGTACATTCTGGG
177 V Q F R R R M P T T N W T L G D C G P Q V N S G S G V L G D I R G
1201 AGCATGTCTGAGTCTGCCTCTGCCCTTCTGAGAACATGGCCCAAGAGATCCAGATACGGAGGAGGAGCGGCTCTCCTCAGGAGCCCTGGAGGTCCT
211 S M S E S C L C P S E N M A Q E I Q I R R R R R L S S G A P G G P
1301 GGAGTATTGGAGCATGCTGTGTGTCCACCTGAAGTCTCCCGAGGCAAGATTAAGTCTTGGTGGAGCCCTGAACCAAGTGAAGGAGGCG
244 W S D W S M P V C V P P E V L P Q A K I K F L V E P L N Q G G R R R
1401 TCTAACCATGCAAGGACAGTACCACAGCTGGCAGTCCCGAAGGCTGCCGAGGAGGCTGGTGCAGGTGAAGAAGCACTTGGTGTGCTGCGCATG
277 L T M Q G Q S P Q L A V P E G C R G R P G A Q V K K H L V L V R M
1501 CTGTCCTGCAAGTGCAGGCTCAGACCTCGAAGACCGTCCCTGGGCAAGAGTGAACCTCTCCGGGGCCACCTATGACCTGAATGTGCTGCCAAAA
311 L S C R C Q A Q T S K T V P L G K K L N L S G A T Y D L N V L A K
1601 CTCGTTTCGGTGCAGACCATCCAGAAGTGGACCTTCTGCCCAAGAGTCCAGAGACAAGAGCCCTGAATGTGACGCTGGGAGGCAACATGACATC
344 T R F G R S T I Q K W H L P A Q E L T E T R A L N V S V G G N M T S

NcoI (1752)

1701 CATGCAGTGGGACGCCAGGCTCCCGGACCACCTACTGCCTTGTGAGTGGCAGCCATGGTTCCAGCACAGGAACACACACTGTACCTGATTGTACCA
377 M Q W A A Q A P G T T Y C L E W Q P W F Q H R N H T H C T L I V P
1801 GAAGAAGAGGATCCTGCCAAGATGGTACACACAGCTGGAGCTTAAACCTACCTGGAGCAGGAGGAATGTACCGCATCACAGTGTCCGCTCCAAGA
411 E E E D P A K M V T H S W S S K P T L E Q E E C Y R I T V F A S K
1901 ACCCAAGAATCCCATGCTGTGGGCCACAGTCTGTCCAGTACTACTTTGGGGTAAACGCTCGAGAGCCGTTACTCCAGACAGTGTCCGTTGAGGAA
444 N P K N P L W A T V L S S Y F F C G N A S R A G T P R H V S G R R
2001 CCAAAKCGGGACTCGGTGTCCGTGGAGTGGACGGCTCACAGCTGAGCACCTGCCCGGGGCTCTGACGCAATACGTCGTGCGTGCAGGCTGAAGAC
477 Q T G D S V S V E W T A S Q L S T C P G V L T Q Y V V R C E A E D
2101 GCGCGTGGGAGTCAAGTGGCTCGTGCACCCACTAAAACCAAGTACACTTACGAGTGGCAGGCTGCGCAGCCGAGTAATGTACAAGTTCAGGTGCGAGCCG
511 G A W E S E W L V P P T K T Q V T L D G L R S R V M Y K V Q V R A
2201 AACTGCGAGGCTCCCGGTGCCTGGAGTCATCCCGAGCGTTAGCTTTGAGGTGACAGATTTCCCGTTTATCCATCATTTTTCGCGTCTCTGGGAAGCTT
544 D T A R L P G A W S H P Q R F S F E V Q I S R L S I I F A S L G S F
2301 CGCCAGCTCCTCCTCGTGGGAGTCTCGGATACATTGGCTTAAACAGGGCCGCTGGCACTGTGCCACCCCTGCCTACACCTGTGGCAGCACTGCC
577 A S V L L V G S L G Y I G L N R A A W H L C P P L P T P C G S T A
2401 GTGGAGTTCCTGGCAGCCAGGCAAGCAGGCTTGGCAGTGGTGAACCTGAGGACTTCCCGGAGGTGTTGTACCCGCGAGATGCGCTGGTGGTTCGAGA
611 V E F P G S Q G K Q A W Q W C N P E D F P E V L Y P R D A L V V E
2501 TGCCCGGAGACAGAGGACGGGACAGAGTCCCGGAGCCGCTGAGTGCCTGGACACAAGGCGCCCTGGAGACTCAGAGGACAGGCGAGGT
644 M P G D R G D G T E S P Q A A P E C A L D T R R P L E T Q R Q R Q V
2601 GCAGGCACTGTGAGGCGAGGCGCCTGGGCTGGTGGAGGACTGCCCCGTGGTGCCTGGCCACGTGACACTCCCGTGTCTCTGGGAGGTGTG
677 Q A L S E A R R L G L A R E D C P R G D L A H V T L P L L L G G V

NheI (2798)

2701 ACCAGGGAGCCTCTGTACTTGCAGATCTTTGGAGGACCCATAAGACTGCGGAGCCGGACCGCCACTTTGGGGCAAGAGCCTGACTGTTGTATCTGC
711 T Q G A S V L D D L W R T H K T A E P G P P T L G Q E A •
2801 TAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAGGAAATGCTTTATTTGTGAAATTTGTGATGCTATT
2901 GCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTT
3001 AAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGA
3101 GGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGAGTTTAAAGATATAGTGTATTTTC

3201 CAAGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTGTAGTAAAATATTCAGAAATAATTTAAATACATCAT
3301 TGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAAAC
3401 CTTTAATAGAAATTTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGC
141 • N R T Y K L P I L E E I T T K V L K G
3501 CATTATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTC
121 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 R I S R D V E
3601 ATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACC
88 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 A E A C V T V
3701 CTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGTCCTCATAGAGCATGG
54 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 N D E Y L M T
3801 TGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGC
21 I K E T A V E V L E L D Q Q S I N F T K M
3901 CGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCCA
4001 CCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAAAACAACTCCCATTGACG
4101 TCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCGCAAAACCGCATCATGTAATAGCGATGACTA
4201 ATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGT
4301 ACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATG
4401 GGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAA
4501 GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAA
4601 AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCC
4701 TGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTCCG
4801 CTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
4901 TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTA
5001 GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGG
5101 TGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAACGAA
5201 AACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTT
5301 TTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGA
5401 ACATTTCTCTATCGAA