



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCAAGGCCTGGGGTGCCTCTGGATCGTGCTTGGATT
1 M K A W G A L W I V L G L
601 GCTGCTGTGGCCAGAGCCAGGGGACGCTCCTCCTTGCTCTGCTCATGGACTCCATCATCCAGGCCCTTGCTGAACTTGAGCAAAGGTACCAGTGACT
13 L L W P E P G A A S S L P L L M D S I I Q A L A E L E Q K V P V T
701 GAGGCCAGCATCACTGCCTCTGCATGGATTCTGTGAGCAAGAAGTCCAGCACCCACAATCCCTTACCAGCGCTTGCCGCTGAAGGCACCAAGCCACA
47 E A S I T A S A W I L S A K N S S T H N S L H Q R L P L K A P S H
801 AACTACAGAGCCAGATCCTCACTCTCTCAGCCGGAGCTTCAAGCACTGATTTCTGAGGTGGCTCAACACGATGTACAGAATGGGCGGGAATATGGAGT
80 N T T E P D P H S L S P E L Q A L I S E V A Q H D V Q N G R E Y G V
901 GGTGCTGGCACCTGATGGTCCACCGTAGCTGTGAAGCCTCTGCTGTTGGCTAGAGCGCGTCTACAGGCACACAGCGTTGCTAACTTGCCTTCCAGAT
113 V L A P D G S T V A V K P L L F G L E A G L Q A H S V A N L P S D
1001 TGTCTGGCTATCCCTGTGATACTGGAGACACCTTGCCAAATATTAGAGCCACCTGGCAGGACTCATGGATGCTTTTCCAAATGCCTCTTCCAGATG
147 C L A I P C D T G D T L A N I R A T W P G L M D A F P N A S S P D
1101 TTGAGCCACTTTACCAAACGACAAAGCAAGACTCCACCCTGTGGACAGACTCCTGGCAATCACCTTGGCTGGTACTAGGCTGACCTTCCCTCCA
180 V G A T L P N D K A K T P T T V D R L L A I T L A G D L G L T F L H
1201 CAGGTCCAGACTTGGAGTCTCCAGGACTGGGAAGTGGGGTGTGGGACAGCTTACTGCCCCAGGGTCTTACACTGTTGGACCCCCAGGCATCC
213 R S Q T W S P P G L G T E G C W D Q L T A P R V F T L L D P Q A S
1301 AGGCTCACCATGGCTTCTCAATGGTGCCTTAGATGGAGCTCTCTGGGAACCACTTGGCAAAATCCCTAGGCCACCCACCCCTCAGCCACCTGC
247 R L T M A F L N G A L D G A L L G N H L S Q I P R P H P P L S H L
1401 TAAGAGAGTACTATGGAGCTGGGGTGAATGGAGATCCGGTGTCCGAAGTAAGTCCGAAGGCAGAACGGTGTGCTTTGACTTACAGCCCTACCCTGGC
280 L R E Y Y G A G V N G D P V F R S N F R R Q N G A A L T S A P T L A
1501 CCAGCAGGTATGGGAGGCCCTTGTCTGTTACAGAACTGGAGCCAGAACCTACAGTTGCAGAACATTAGCCAAGAGCAGCTGGCTCAGGTAGCCACC
313 Q Q V W E A L V L L Q K L E P E H L Q L Q N I S Q E Q L A Q V A T
1601 TTGGTACCAAGGAGTCTACTGAGGCTTCTGGGATGCCAGCATTCAACCCCGCTGCCGTTGGGGAGCGGCTCCCTACCGAGGCCACCAACACCAC
347 L A T K E F T E A F L G C P A I H P R C R W G A A P R V G H G P T P
1701 TCCGGTCCCACTTGGATTCTTATATGTCATCACACATACGTGCCAGCGCCACCCTGACCACCTTCCAGAGCTGCGCCGCGATATGCGCTCCATGCA
380 L R L P L G F L Y V H H T Y V P A P P C T T F Q S C A A D M R S M Q
1801 GCGTTTCCACCAGGATGTGCGCAAGTGGGATGACATCGGCTACAGTTTCTGGTAGGCTCCGACGGCTATCTGTACCAGGGCCGTGGCTGGCACTGGGTA
413 R F H Q D V R K W D D I G Y S F V V G S D G Y L Y Q G R G W H W V
1901 GGTGCGCACACCGCGGTACAACCTCCCGCGGCTTCCGTGTGGCTTCTGGGCAACTACACTGGGCTACTGCCCAACGAAGCTGCGCTGAACACGGTGC
447 G A H T R G Y N S R G F G V A F V G N Y T G S L P N E A A L N T V
2001 GCGACGCTCCGAGCTGCGCAATTCGCGCAGGTCTCTTGGGCCAGACTACAAGCTGCTTGGCCACCGCAGCTAGTGCTACCCACTGCCCGGGAA
480 R D A L P S C A I R A G L R P D Y K L L G H R Q L V L T H C P G N
2101 CGCGCTCTCAACTTGTGCGCACCTGGCTCACTTACAGAGGTTGAAAATAAGAACTCCTTTGAGAGACCTTGAAGATCCAGGAGGTATTATCCCT
513 A L S N L L R T W P H F T E V E N •

NheI (2231)

2201 GATGATCCTTTGAGCAACCACAGACCTCAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAA
2301 AAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGT
2401 TTCAGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCT
2501 CCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTT
2601 CTTTCATGGAGTTAAGATATAGTGATTTTCCAAGTGTGAACTAGCTCTTCAATTTCTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTA
2701 GTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCC
2801 AGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGGGGGA
141 • N R T Y K L P I
2901 TGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACA
132 L E E I 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
3001 GGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGAC
99 P S V V 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
3101 CCAATGGCAATGGCTTCCAGCAGACAGTGACCCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCC
65 G I A I 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

3201 CGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGT
32 V H H K 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
3301 GGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTT
3401 CACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGT
3501 TGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCC
3601 AAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTGGCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGG
3701 CCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGA
3801 CGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTA
3901 CCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTCTGCGTTTTTC
4001 CATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTCCCCCTG
4101 GAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAG
4201 CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACTAT
4301 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT
4401 TCTTGAAGTGGTGCCTAACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC
4501 TTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATC
4601 TTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACTCACGTTAAGGGATTTTGGTCAAGGCTAGTTAATTAACATTTAATCAGCGGCCGCAATAAAAT
4701 ATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAA
4801 TAGGCTGTCCCAAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA