



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCGCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATGAAGACACTACTGACCATCTGACGGTGGGATCCCT
1 M K T L L T I L T V G S L
601 GGCCGCTCACACCACTGTGGACACATCCGGTCTCCTTCAACACGTGAAATTCAGTCCAGCAACTTTGAGAACATCTTGACGTGGGATGGTGGGCCGCA
13 A A H T T V D T S G L L Q H V K F Q S S N F E N I L T W D G G P A
701 AGCACCTCTGACACCGTCTACAGTGTGAATATAAGAAATACGGAGAGAGAAAGTGGCTGGCCAGGCGGGTCCAGCGGATCACCCAGAAGTTCTGCA
47 S T S D T V Y S V E Y K K Y G E R K W L A K A G C Q R I T Q K F C
801 ACCTGACTATGGAGACCCGCAACCACACTGAGTTTTACTACGCCAAGGTCACGGCGGTGAGCGGGAGGCCACCAGTCAACAAGATGACTGATCGTTT
80 N L T M E T R N H T E F Y Y A K V T A V S A G G P P V T K M T D R F
901 CAGCTCGCTGCAGCACACTACCATCAAACCGCTGATGTGACCTGTATCCCCAAAGTGAAGTCCATTGAGTGTCCACCCACACTCACACCACTGTC
113 S S L Q H T T I K P P D V T C I P K V R S I Q M L V H P T L T P V
1001 CTCTCGGAAGATGGCCACCAGCTAACCTGGAGGAGATTTCCATGACCTGTTCTACCGCTTAGAGCTCCACGTCAACCACACCTACCAGATGCACCTTG
147 L S E D G H Q L T L E E I F H D L F Y R L E L H V N H T Y Q M H L
1101 AAGGCAACAGAGAGAATACGAGTTCCTTGGCCTGACTCCCGACAGAGTTCCTCGGCTCCATCACAATTTTACTCCGATATTGTCCAAGGAAAGTGC
180 E G K Q R E Y E F L G L T P D T E F L G S I T I L T P I L S K E S A
1201 CCCTACGTGTGCCGAGTGAAGACGCTGCCGATCGGACGTGGGCTACTCCTTCTCGGGCGCGTCTTTTTCCATGGGTTTCTCGTGGCTTGTCTC
213 P Y V C R V K T L P D R T W A Y S F S G A V L F S M G F L V G L L
1301 TGTATCTGGGCTACAAATACATCAACCAAGCCACTGTACCTCTAACTCCCTGACCTGCAACGTGTCTGACCTTTCAACCCTACGCTTCACTCAAG
247 C Y L G Y K Y I T K P P V P P N S L N V Q R V L T F Q P L R F I Q
1401 AACACGTACTGATCCCTGTCTTGGACCTCAGTGGCCCCAGCAGTCTGCCTCAGCCATCCAGTACTCCAAGTGGTGGTGTCTGGGCCAGGGAGCCTCC
280 E H V L I P V L D L S G P S S L P Q P I Q Y S Q V V V S G P R E P P
1501 TGGAGCTGTGTGGCGGACAGGCTGTCTGACCTCACCTACGTAGGGCAGTCAAGTGTCTCCATCCTGCAACCTACCAAGTCCAGCTCAGCAGACACTG
313 G A V W R Q S L S D L T Y V G Q S D V S I L Q P T N V P A Q Q T L
1601 TCCCCACCTCTACGCTCCGAAGGCTGTCCCTGAGGTCCAGCCCCCTTCTATGCGCTCAGGTAGCCTCGGATGCCAAAGCTCTGTTCTACTCACCAC
347 S P P S Y A P K A V P E V Q P P S Y A P Q V A S D A K A L F Y S P
1701 AACAGGGGATGAAGACCAGGCTGCCACCTATGACCCGACGACATTCTGACAGCTGCCCTGCTTATGCTGTGTGGAAGACTCTGGCAAAGA
380 Q Q G M K T R P A T Y D P Q D I L D S C P A S Y A V C V E D S G K D
1801 CTCTACCCAGGCATCCTCTCCACTCCCAAATACCTCAAGACAAAAGTCCAGTCCAGGAAGACACTTGTAGAAGCTGTCTCCAGGGGACCTTTCT
413 S T P G I L S T P K Y L K T K G Q L Q E D T L V R S C L P G D L S
1901 CTACAGAAAGTCACTCCTTAGTGAAGGGGAGACACAGAGACAAAATCACTCCCTCACCTCTGGGATTTTGACAGACAGAGGACCTGACCTTACA
447 L Q K V T S L G E G E T Q R P K S L P S P L G F C T D R G P D L H
2001 CACTGCGCAGTGAAGAACAGAGACACACCGTACCTGAAGGGGCGCTGTCTCTCTGTCTGTGCGATCGAGGGCCACCTGTCTCCCTCCCTTT
480 T L R S E S E P E T P R Y L K G A L S L S S V Q I E G H P V S L P L
2101 GCACGTCCATTCTGTCTCATGTTCCCTCAGACGAGGGACCAAGTCCCTGGGCGCTGCTGACTCCCTTGTGTGTCAAAGGATGAGGGTCCCGCGTT
513 H V H S V S C S P S D E G P S P W G L L D S L V C P K D E G P A V
2201 GAGACTGAGGCCATGTGCCCCAGTGTGACGCTCTGAGCTGGAGCAGTCCACAGAAGTGGACTCTCTTTTCAAAGGCTTGGCCCTGACTGTGCGATGGG
547 E T E A M C P S A A A S E L E Q S T E L D S L F K G L A L T V Q W

NheI (2308)

2301 AATCCTGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTG
580 E S •
2401 TGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGG
2501 GAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAAT
2601 CCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTTATGAGTAAAGATATAG
2701 TGTATTTTCCAAGGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTA
2801 AATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGTCTAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGA
2901 ACAAAGGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTG
141 • N R T Y K L P I L E E I T T K
3001 ACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAAGGAGCATAGTCAAGAGATGAGTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATC
124 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 R I S R
3101 TGCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACA
91 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 A E A C

3201 GACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTTGTCTCA
58 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 N D E
3301 TAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAAGTGAGTCGTAT
24 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
3401 TATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACAAACGAGCTCTGCTTATAT
3501 AGACCTCCCACCGTACACGCCTACCGCCATTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTACTAGTCAAAAACAACT
3601 CCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATA
3701 GCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTACCGTCATTGACGTCAA
3801 TAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGG
3901 CGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTACCGTAAAGTTATGTAACGCTGCA
4001 GGTAAITTAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTTCATAGGCTCCGCCCCCTGACGA
4101 GCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCT
4201 GTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT
4301 AGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTTAAG
4401 ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC
4501 GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCG
4601 CTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCA
4701 GTGGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCGAATAAAATATCTTTATTTTATTACATCTGT
4801 GTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGC
4901 AGGTGCCAGAACATTTCTCTATCGAA