



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTGCCGCCCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCAGTGGCCCTGTGGAGAGAAGCCTGTCTGGAAGC
1 M S G P C G E K P V L E A
601 CAGCCCCACCATGAGTCTGTGGGAATTTGAGGACAGCCACAGCCGTCAGGGCACCCCAAGGCCGGTCAAGAGCTGGCCGCTGAGGAGCCTCGGCCCTG
13 S P T M S L W E F E D S H S R Q G T P R P G Q E L A A E E A S A L
701 GAACTGCAGATGAAGGTGGACTTCTCCGGAAGCTGGGCTATTTCATCCACGGAGATCCACAGCTCCTGCAGAAGCTGGGCGTCCAGGACAGACCAACA
47 E L Q M K V D F F R K L G Y S S T E I H S V L Q K L G V Q A D T N
801 CGGTGCTGGGTGAGTGGTAAACACGGGACAGCCACCGAGCGGGAGCGCCAGACCTCACCAGCCCTGCCCTCAGTCCCTCTAGTCCCAGGGGTGG
80 T V L G E L V K H G T A T E R E R Q T S P D P C P Q L P L V P R G G
901 TGGACCCCTAAGGCTCCAACCTGGAGCCTCCAGAGCTCCAGAAAGGAGGAGGGCAGCGACCTGAGACCAAGTGTCTGATGGGAGCAACGTTGGCC
113 G T P K A P N L E P P L P E E E K E G S D L R P V V I D G S N V A
1001 ATGAGCCATGGGAACAAGGAGGTCTTCTCCTGCCGGGCATCTGCTGGCAGTGAAGTGGTTTCTGGAGCGGGCCACACAGACATCACAGTGTGTTGTG
147 M S H G N K E V F S C R G I L L A V N W F L E R G H T D I T V F V
1101 CATCCTGGAGAAAGGAGCAGCCTCGGCCGACGTGCCATCACAGACCAGCACATCCTGCGGGAAGTGGAGAAGAAGATCCTGGTGTTCACACCATC
180 P S W R K E Q P R P D V P I T D Q H I L R E L E K K K I L V F T P S
1201 ACGACGCTGGGTGGCAAGCGGGTGGTGTGCTATGACGACAGATTATTGTGAAGCTGGCCTACGAGTCTGACGGGATCGTGGTTTCCAACGACACATAC
213 R R V G G K R V V C Y D D R F I V K L A Y E S D G I V V S N D T Y
1301 CGTAGCTCCAAGGACAGCGGAGGAGTGAAGCGCTTCTCAGAGGAGCGGCTGCTCATGTACTCCTTCATGACAAAGTTTATGCCCCCTGATGACC
247 R D L Q G G E R Q E W K R F I E E R L L M Y S F V N D K F M P P D D
1401 CACTGGGCGGCACGGGCCAGCCTGGACAACCTCCTGCGTAAGAAGCCACTCACTTTGGAGCACAGGAAGCAGCCGTGTCCCTATGGAAGGAAATGCAC
280 P L G R H G P S L D N F L R K K P L T L E H R K Q P C P Y G R K C T
1501 CTATGGGATCAAGTGCCGATTCTTCCACCCAGAGCGGCAAGCTGCCCCAGCGCTCTGTGGCAGATGAGCTCCGTGCCAATGCTCTCTCTCACCCCC
313 Y G I K C R F F H P E R P S C P Q R S V A D E L R A N A L L S P P
1601 AGAGCCCCAAGCAAGGACAAAATGGCCGGCGGCTTACCTTCACTCCAGTCCAGTCTCTGCTAACAGAGAGTGAAGTGCAGCTGGATGGGAAGA
347 R A P S K D K N G R P S P S S Q S S L L T E S E Q C S L D G K
1701 AGCTGGGGGCCACAGKATCCCCAGGTCGCCCAAGAGGGTCTAACACAGACCTATGCCSCLATCAGGCAGGAGCCTCGCACCTAGCGGGGCGAGTGGCAG
380 K L G A Q A S P G S R Q E G L T Q T Y A P S G R S L A P S G G S G S
1801 CAGCTTTGGGCCACAGACTGGCTCCCACAGACGCTGGACTCACTCCCGTACGTCTCCAGGATTGCCTGGACTCGGGCATTGGCTCCCTGGAGAGCCAG
413 S F G P T D W L P Q T L D S L P Y V S Q D C L D S G I G S L E S Q
1901 ATGTCCGAACTTTGGGGGTTTCGAGGAGGAGCCCTGGTGTGAGCCGGGCCACCCCGAGCCCTTACACGGGTACAGTCCCTATGGATCTGAGCTCCAG
447 M S E L W G V R G G G P G E P G P P R A P Y T G Y S P Y G S E L P
2001 CCACCGCAGCCTTCTGCTCTTGGCCGGCCATGGTGTGGCCACTTCAGTGTCCCTGCCGACTACCCACCCGCCCCCTGCCTTCCACCTCGAGA
480 A T A A F S A F G R A M G A G H F S V P A D Y P P A P P A P P R E
2101 GACTGTGCTGAACCATACCACTGCCCCACCCACATCAGTCTTACAGGAGCCCCAGTGCAGAGCCAGGGGCTGGCAGGAGCCGTGGGGCAGGGCA
513 Y W S E P Y P L P P P T S V L Q E P P V Q S P G A G R S P W G R A
2201 GGCAGCCTGCCAAGGAGCAGGCAGCTGTATAACTAGCTGTGTGGTGTGTTTCCCCGCACCTGGTGGAGGCTGTGATGGGGCGCTTCCACAGCTCC
547 G S L A K E Q A S V Y T K L C G V F P P H L V E A V M G R F P Q L

NheI (2385)

2301 TGGACCCCGAGCAGTGGCTGCCGAGATCCTCTCTACAAGTCCCAGCACCCAGTGAAGTGCCTGTGGCTGGCAAGGGCAGTAGTGGCCAGC
580 L D P Q Q L A A E I L S Y K S Q H P S E •
2401 ATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAGAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAA
2501 CCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTCATTCTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAA
2601 CCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGG
2701 CATAGGCATCAGGGGCTGTTGCCAATGTGATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAT
2801 AGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAAATTCAGAAATAATTTAAATACATCATTGCAATGAAAATA
2901 AATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTGTAGTGGACTTAGGGAACAAGGAACCTTTAATAGAAAT
3001 TGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAAT
141 • N R T Y K L P I L E E I T T K V L K G N M E I
3101 GAGCACAAGCAGTCAAGGACATAGTCAAGATGAGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGG
117 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 D S Y P

3201 TGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGG
83 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 R G I Y A
3301 CCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTGCTCATAGAGCATGGTGATCTTCTCAGT
50 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 I K E T
3401 GCGGACCTCCACCAGTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGC
17 A V E V L E L D Q Q S I N F T K M
3501 CGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTA
CGCCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGA
GACTTGAAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTA
CTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGA
TACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCA
TTATTGACGTCAATGGGCGGGGTCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAA
GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGG
ATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCGTTACGCCCACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAG
CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT
TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAG
GGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCG
TAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATC
5001 GAA