



100

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTGCTGTGAGAGGGTATCTCAGTGGACTTGAGTACTG
1 M L S E G Y L S G L E Y W
601 GAATGACATCCACTGGAGTTGTGCCTCTATAATGAGCAGGTGGCTGGGAAAAGGAAGAGGAGACAAATCTGTTGCTACCTTTCTATTCTCTGTG
13 N D I H W S C A S Y N E Q V A G E K E E E T N S V A T L S Y S S V
701 GATGAAACACAAGTCAGAAGTCTCTACGTGAGCTGCAATCATCTGGCAAGTTTATCTCTCAGTGCATTCAAGAGAGAGCCAACATAGCAGAAGTCAGA
47 D E T Q V R S L Y V S C K S S G K F I S S V H S R E S Q H S R S Q
801 GAGTCACAGTGTGCAGACAAACCCCAATCCTGTGTTGAAAGCCAAACTTGGCTGCAGTTGAAATATGTAGAGATGCCAGCAGAGAGACCTACTTGGT
80 R V T V L Q T N P N P V F E S P N L A A V E I C R D A S R E T Y L V
901 TCCATCTCTTGC AAAAGTATTTGCAAGAATTATAATGACTTACAGATTGCGAGGGGCCAGGTGATGGCCATTAATCAGTGAACAGATTTTCCCTCT
113 P S S C K S I C K N Y N D L Q I A G G Q V M A I N S V T T D F P S
1001 GAGAGCAGTTTTGAATATGGCCCTTGTGTAAGTCATCTGAGATTCCTTTACCATGGAGATTCCATTTCTACTCAGCCAGTGACTTTCCCAAAAAAC
147 E S S F E Y G P L L K S S E I P L P M E D S I S T Q P S D F P Q K
1101 CTATCCAGCGGTACTCATCTATTGGAGAATAACAAGCATCAAAGAGAAAAGCAGCTTGC AAATGCAGAATCCTATTTCTAATGCAGTCTGAATGAGTA
180 P I Q R Y S S Y W R I T S I K E K S S L Q M Q N P I S N A V L N E Y
1201 CCTGGAGCAGAAGGTTGTGGAGTTATATAAACAGTACATTATGGACACCGTGTTCATGACAGTTCTCCTACCAGATTCTGGCGTCTGAACTCATCATG
213 L E Q K V V E L Y K Q Y I M D T V F H D S S P T Q I L A S E L I M
1301 ACAAGTGTAGACCAATCAGTCTCAAGTGTCTAGAGAGAAGAATCTGGAGACCTCAAAGCCAGAGATATAGTCTTTAGCCGCTATTGCAATTGATGT
247 T S V D Q I S L Q V S R E K N L E T S K A R D I V F S R L L Q L M

NheI (1492)

1401 CAACTGAAATTACTGAAATTAGCACTCCTAGTCTCCATATTTCTCAGTATAGCAATGTAAATCCATAGAGAGGATGCTTCCATTACTGTCTCGTAGCTG
280 S T E I T E I S T P S L H I S Q Y S N V N P •
1501 GCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA
1601 TTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCA
1701 AGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATG
1801 AATAAGGCATAGGCATCAGGGGCTGTGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGT
1901 TTGAACTAGTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAAT
2001 GAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAA
2101 TAGAAATTTGACAGCAAGAAAGCAGCTTCTAGCTTTAGTTCTGCTGACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCA
141 • N R T Y K L P I L E E I T T K V L K G N M
2201 TCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGA
119 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 D S
2301 GTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTGTCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCA
86 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 R G
2401 ATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGCTCATAGAGCATGGTGTCT
52 I Y A E I H V A S I E G T K T R I A A G V H H K N D E Y L M T I K
2501 TCTCAGTGGCAGCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGTGATTATACTATGCCGATAT
19 E T A V E V L E L D Q Q S I N F T K M
2601 ACTATGCCGATGATTAATTTGCAAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTAC
2701 ACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATG
2801 GGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGT
2901 AGATGTAAGTCCCAAGTAGGAAAGTCCCATAGGTCATGTAAGTGGGCATAATGCCAGCGGGCCATTTACCCTGATTGACGTCAATAGGGGGCGTACTTGG
3001 CATATGATACACTTGTACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACA
3101 TACGTCATTATTGACGTCAATGGGGGGGTCGTTGGGCGGTGAGCCAGCGGGCCATTTACCCTAAGTTATGTAACGCCTGCAGGTTAATAAGAACAT

3201 GTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGA
3301 CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
3401 TTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGGTAGGTCGTTCCGCTCAA
3501 GCTGGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCA
3601 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA
3701 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
3801 TTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCA
3901 CGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTG
4001 TGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAAGTGCCAGAACATTT
4101 CTCTATCGAA