



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCCATGGCCCGGGCCGCTGCCTCCTGCCGTGAGATCGCC
1 M A R A A A L L P S R S P
601 GCCGACGCCGCTGCTGTGGCCGCTGCTGCTGCTGCTCCTGAAACCGGAGCCAGGATGTGCGAGTTCAAGTGTACCCGAGGTGCGAGGCCAGCTC
13 P T P L L W P L L L L L L L E T G A Q D V R V Q V L P E V R G Q L
701 GGGGACCGTGGAGTCCGCTGCCACCTGCTGCCACCTGTTCTGGACTGTACATCTCCCTGGTACCTGGCAGCGCCAGATGCACCTGCGAACCC
47 G G T V E L P C H L L P P V P G L Y I S L V T W Q R P D A P A N H
801 AGAATGTGGCCGCTCCACCCTAAGATGGGTCCCAGCTTCCCGAGCCGAAGCTGGCAGCGAGCGCTGTCTTCTGCTCTGCCAAGCAGAGCACTGG
80 Q N V A A F H P K M G P S F P S P K P G S E R L S F V S A K Q S T G
901 GCAAGACACAGAGGAGCTCCAGGACGCCACGCTGGCCCTCCAGGGCTCACGGTGGAGGACGAGGCAACTACACTTGCAGTGTGCCACCTTCCCG
113 Q D T E A E L Q D A T L A L H G L T V E D E G N Y T C E F A T F P
1001 AAGGGTCCGTCGAGGGATGACCTGGCTCAGAGTCATAGCCAAGCCAAAGCAAGCTGAGGCCAGAAAGTACGTTTACGCCAGGACCTACGACAG
147 K G S V R G M T W L R V I A K P K N Q A E A Q K V T F S Q D P T T
1101 TGGCCTCTGCATCTCCAAGAGGGCCGCCACCTGCCGGATCTCCTGGCTCTCATCCTGGACTGGGAAGCAAAGAGACTCAGGTGTGAGGGACCT
180 V A L C I S K E G R P P A R I S W L S S L D W E A K E T Q V S G T L
1201 GGCCGAACTGTCACTGTACCAGCCGCTTACCTTGGTGCCTCGGGCCGAGCAGATGGTGTACGGTACCTGCAAAGTGGAGCATGAGAGCTTCGAG
213 A G T V T V T S R F T L V P S G R A D G V T V T C K V E H E S F E
1301 GAACGACCTGATACCTGTACCCTCTGTACGCTACCCTCCTGAAGTGTCCATCTCCGGCTATGATGACAAGTGGTACCTCGCCGCTACTGATGCCA
247 E P A L I P V T L S V R Y P P E V S I S G Y D D N W Y L G R T D A
1401 CCCTGAGCTGTGACGTCGCGAGCAACCAGAGCCACGGGCTATGACTGGAGCAGACCTCAGGCACCTTCCCGACCTCCGAGTGGCCAGGGCTCCA
280 T L S C D V R S N P E P T G Y D W S T T S G T F P T S A V A Q G S Q
1501 GCTGGTACCCACGAGTGGACAGTCTGTTCAATACCACCTTCTGTCTGCACAGTACCAATGCCGTGGGCATGGCCCGGCTGAGCAGTCTCTTTGTC
313 L V I H A V D S L F N T T F V C T V T N A V G M G R A E Q V I F V
1601 CGAAGAACCCCGAGGCTCGCCCGAGATGTGGGCCGCTGGTGTGGGGGGCGTGGGGGGGACTGCTGGTGTGCTGCTTCTGGCTGGGGGGTCT
347 R E T P R A S P R D V G G P L V W G A V G G T L L L L A G G S
1701 TGGCCTTCACTGCTGAGGGTGGAGGAGGAGGAGCCCTGGAGGAGCAGGAGGAGGAGCCAGTGGCGACGGGGATTCTACGATCCGAAAGCTCA
380 L A F I L L R V R R R R K S P G G A G G G A S G D G G F Y D P K A Q
1801 GGTGTTGGGAAATGGGACCCGCTTCTGGACACCAGTAGTCCCTGGTCCATGGAACCAGATGGCAAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAG
413 V L G N G D P V F W T P V V P G P M E P D G K D E E E E E E E E K
1901 GCAGAGAAAGCCTCATGTTGCCTCCACCCAGCACTCGAGGATGACATGGAGTCCCAGCTGGACGGCTCCCTCATCTACGGCGGGCAGTTTATGTG
447 A E K G L M L P P P P A L E D D M E S Q L D G S L I S R R A V Y V

NheI (2033)

2001 GACCTGGACACAGACAGAGACAGAGCCAGGCCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAA
480 •
2101 AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACATTCATTCTTTAT
2201 GTTTCAGGTTTACGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAAC
2301 CTCCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCAC
2401 TTCTTTTATGAGTTTAAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTTCTCATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTT
2501 TAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCC
2601 CCAGTTTAGTGTGACTTAGGGAACAAGGAACCTTAATAGAAATTTGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGG
141 • N R T Y K L P
2701 GATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCA
133 I 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
2801 CAGGGGCTGACCACCTGATGGATCTGCCACCTCATCAGATGAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCAGCAG
99 C 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
2901 ACCCAATGGCAATGGCTTACGACAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCG
66 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
3001 CCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATG
33 G 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
3101 GTGGCCCTCTATAGTGTGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGG

3201 TTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCC
3301 GTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGTGGAGACTTGAAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTG
3401 CCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCG
3501 GGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATT
3601 GACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATT
3701 TACCGTAAGTTATGTAACGCTGCAGGTTAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTT
3801 TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
3901 TGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGCCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAGCTCA
4001 CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT
4101 ATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA
4201 GTTCTTGAAGTGGTGGCCTAACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGC
4301 TCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGA
4401 TCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAA
4501 ATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAACGAAACAAACAACTAGCAA
4601 AATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA