



100
|-----|

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)
AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTGGAAGATTGGAGGAAACATTATTTGAAGAATTTGA
1▶ M E D L E E T L F E E F E
601 AACTATTCTATGACCTAGACTATTACTCTCTGGAGTCTGATTTGGAGGAGAAAAGTCCAGCTGGGAGTTGTTCACTGGGTCTCCCTGGTGTATATTGT
13▶ N Y S Y D L D Y Y S L E S D L E E K V Q L G V V H W V S L V L Y C
701 TTGGCTTTTGTCTGGGAATCCAGGAAATGCCATCGTCATTTGGTTCACGGGTTCAAGTGAAGAAGACAGTCACCACTCTGTGGTTCCTCAATCTAG
47▶ L A F V L G I P G N A I V I W F T G F K W K K T V T T L W F L N L
801 CCATTGCGGATTTCAATTTTCTCTCTTCTGCCCCTGTACATCTCTATGTGCCATGAATTTCACTGGCCCTTTGGCATCTGGCTGTGCAAAGCCAA
80▶ A I A D F I F L L F L P L Y I S Y V A M N F H W P F G I W L C K A N
901 TTCCTTCACTGCCAGTTGAACATGTTTCCAGTGTCTTTTCTGACAGTGCAGCTGGACCACTATATCCACTTGATCCATCTGTCTTATCTCAT
113▶ S F T A Q L N M F A S V F F L T V I S L D H Y I H L I H P V L S H
1001 CGGCATCGAACCTCAAGAACTCTGATTGTCATTATATTCACTGGCTTTTGGCTTCTAATTGGCGGTCTGCCTGTACTTCCGGGACACTGTGG
147▶ R H R T L K N S L I V I I F I W L L A S L I G G P A L Y F R D T V
1101 AGTTCAATAATCATACTCTTGTCTATAACAATTTTCCAGAGCATGATCCTGACCTCACTTTGATCAGGCACCATGTTCTGACTTGGGTGAAATTTATCAT
180▶ E F N N H T L C Y N N F Q K H D P D L T L I R H H V L T W V K F I I
1201 TGGCTATCTTCCCTTTGCTAAACAATGAGTATTTGCTACTTGTGCTCATCTTCAAGGTGAAGAAGCGAAGCATCTGATCTCCAGTAGGCATTTCTGG
213▶ G Y L F P L L T M S I C Y L C L I F K V K K R S I L I S S R H F W
1301 ACAATTTGGTGTGGTGTGGCCTTTGGTGTGGTCTGACCTTATCACCTGTTTAGCATTTGGGAGCTCACCATTCAACCAATAGTATTTCCACC
247▶ T I L V V V V A F V V C W T P Y H L F S I W E L T I H N S Y S H
1401 ATGTGATGCGAGTGAATCCCTCTCACTGGTTTGGCATTCTCAATAGTTGCTTGAACCCATCCTTTATGTCTAATTAGTAAGAAGTTCCAAGC
280▶ H V M Q A G I P L S T G L A F L N S C L N P I L Y V L I S K K F Q A
1501 TCGCTCCGGTCTCAGTTGCTGAGATACTCAAGTACACACTGTGGGAAGTCAAGTGTCTGGCACAGTGAAGTGAACAGCTCAGGAACCTCAGAAACCAAG
313▶ R F R S S V A E I L K Y T L W E V S C S G T V S E Q L R N S E T K

NheI (1634)

1601 AATCTGTGTCCTGGAACAGCTCAATAAGTTAGCTAGCTGGCCAGACATGATAAGATACTTGTGAGTGGGACAAACCACAACCTAGAATGCAGTGA
347▶ N L C L L E T A Q •
1701 AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGTGAATAAACAAGTTAACAACAACAAATTGCATTCATTTA
1801 TGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCAACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAA
1901 CCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTTGGCAATGTGCATTAGCTGTTTGCAGCCTCAC
2001 CTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCAATTTCTTATGTTTTAAATGCACTGACCTCCACATTCCTTT
2101 TTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCC
2201 CCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGG
141▶ • N R T Y K L P
2301 GGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCC
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
2401 ACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCAGACA
100▶ 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
2501 GACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCG
66▶ S 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
2601 CCCCAGATGGTCTTGTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCAT
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
2701 GGTGGCCCTCTATAGTGAAGTATTACTATGCGGATATACTATGCGGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACG
2801 GTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCC
2901 CGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGTATCCACGCCATTGATGTACT
3001 GCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCACTGACTGGGCATAATGCCAGGC
3101 GGGCATTACCCTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCCTAAATACTCCACCCAT

3201 TGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGCGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCAT
3301 TTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT
3401 TTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
3501 CTGGAAGCTCCCTCGTGCCTCTCTGTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC
3601 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAAC
3701 TATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAG
3801 AGTTCCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG
3901 CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTG
4001 ATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCCAATAA
4101 AATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCA
4201 AAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA