



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)
AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTGACCCGGGCGCCAGTCTCCCGCGTCCAGATTGTC
601 ACCGACGCTGCCGTTGTTGCCGCTGCTACTGCTCCTGCTTCAGGAAACAGGAGCCCAAGATGTGCGGGTACGAGTGTCTCCGAGGTCCGGGGCGGCTTG
13▶ P T L P L L P L L L L L L Q E T G A Q D V R V R V L P E V R G R L
701 GGAGGCACCGTGGAGTTACCGTCCACCTGCTCCACCCACGAGCGCGTCTCTCAGGTGACCTGGCAGCGCCTGGATGGCACAGTTGTGGCTGCTT
47▶ G G T V E L P C H L L P P T T E R V S Q V T W Q R L D G T V V A A
801 TCCACCCATCCTTCGAGTGGATTTCCCAACTCTCAGTTCAGCAAGGACCGTCTGTCTTTGTCAGAGCGAGACCAGAAACAAACGACAGCTGCGGGA
80▶ F H P S F G V D F P N S Q F S K D R L S F V R A R P E T N A D L R D
901 TGCCACTGCGCTTCCGGGGACTGAGGGTAGAGGACGAGGCAATTACACCTGCGAGTTTCCACGTTTCCCAACGGTACCCGACGGGGGTGACCTGG
113▶ A T L A F R G L R V E D E G N Y T C E F A T F P N G T R R G V T W
1001 CTCAGAGTCATAGCCAGCCTGAGAACCAGCTGAAGCCAGGAGGTACAATTGGCCCCAGTGGTGGCTGAGCCGCTGTGTCTCCACTGGGGGCC
147▶ L R V I A Q P E N H A E A Q E V T I G P Q S V A V A R C V S T G G
1101 GCCCCCTGCCGAATCACCTGGATCTCATCTCTGGTGGAGAGGCCAAAGATACTCAGGAGCCAGGATACAGGCTGGCACCGTCACTATCATCAGCCG
180▶ R P P A R I T W I S S L G G E A K D T Q E P G I Q A G T V T I I S R
1201 ATACTCCTTGGTCCCGTGGGCGGAGCGGATGGCGTCAAGGTCAGGTGAGAGTGAACACGAGAGCTTCAAGAGCCGATCTGCTGCCAGTGACCTC
213▶ Y S L V P V G R A D G V K V T C R V E H E S F E E P I L L P V T L
1301 TCTGTGCGTACCTCCAGAAGTATCCATCTCCGGCTATGATGACAACTGGTACCTTGCCCGCAGTGGCCATACTGACCTGTGATGACGAAGCAACC
247▶ S V R Y P P E V S I S G Y D D N W Y L G R S E A I L T C D V R S N
1401 CAGAGCCACAGACTATGACTGGAGCACGACCTCGGGCGTCTCCAGCCTCTGCAGTGGCCAGGGCTCTCAGTGTGTCCACTCTGTGGATCGAAT
280▶ P E P T D Y D W S T T S G V F P A S A V A Q G S Q L L V H S V D R M
1501 GGTCAACACTACCTTCATCTGTACAGCCACCAACGCTGTGGGACAGGCCGTGCTGAGCAGGTATCCTGGTCCGAGAGTACCCAGCACAGCAGGAGCA
313▶ V N T T F I C T A T N A V G T G R A E Q V I L V R E S P S T A G A
1601 GGGGCACTGGTGGCATCATTGGAGGTATTATCGTGCCATCATCGCCACCGCAGTGGCTGGCACAGGCATCTCATCTGCCGACAACAGCGGAAGGAGC
347▶ G A T G G I I G I A I A I A T A V A G T G I L I C R Q Q R K E
1701 AGAGGCTTCAAGCTGCGGATGAGGAAGAAGAACTGGAAGGACCTCCCTCTATAAAACCCACCCCGAAGGCAAGCTGGAGGAACAGAGATGCCCTC
380▶ Q R L Q A A D E E E E L E G P P S Y K P P T P K A K L E E P E M P S
1801 TCAACTTTCACCTTGGGGCCCTCAGAGCACAGCCAGTGAAGACGCCATACTTTGATGCTGGTGTCTTTGTGCTGATCAGGAGATGCCGATCAC
413▶ Q L F T L G A S E H S P V K T P Y F D A G V S C A D Q E M P R Y H
1901 GAGCTGCCACTCTGGAAGAGCGGTGAGGGCCCTGCTGTTGGGGGTACAGGCTGGGACCTTCTTCTGTTGCTCCAGGACCCAATGTTGTGGAGG
447▶ E L P T L E E R S G P L L L G A T G L G P S L L V P P G P N V V E
2001 GGGTTTCCCTGAGTCTCGAAGATGAGGAGGAAGATGATGAGGAGGAAGACTTCTGGATAAAATCAACCTATTTATGATGCCCTGTCTACCCAGCCC
480▶ G V S L S L E D E E E D D E E E D F L D K I N P I Y D A L S Y P S P

NheI (2165)

2101 CTCTGACTCTACCAGAGCAAAGACTTTTTTGTGCACGGCCATGTATGTGTGAGGGAGGCACAGCTAGTGGCCAGACATGATAAGATACATTGATGA
513▶ S D S Y Q S K D F F V S R A M Y V •
2201 GTTTGGACAAACCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTCAATAAA
2301 CAAGTTAAACAACAATTGCATTCAATTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGG
2401 AATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTT
2501 GCCAATGTGCATTAGCTGTTGAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAGTCTTCTCATTCTTTATGT
2601 TTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGA
2701 ATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGC
2801 TTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAG
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I L V F C D P A
2901 CATAGTCAGAGATGAGCTCTCTGCATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGT
110▶ 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 H R V A V I T
3001 GTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTACGACAGACAGTACCCTGCCAATGTAGGCCCAATGTGGACAGCAGAG
77▶ 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 E I H V A S
3101 ATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCA
43▶ 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 A V E V L E L

3201 GATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAAC
10 D Q Q S I N F T K M
3301 AGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGG
3401 GCGGAGTTGTTACGACATTTTGAAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAG
3501 TCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCC
3601 ATAAGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAA
3701 GTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGG
3801 GGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
3901 GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGG
4001 ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTTCTCC
4101 CTTCCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGT
4201 TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT
4301 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA
4401 AGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACCGG
4501 CAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGT
4601 TAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCA
4701 TCAAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA