



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTATCATGACACTCAGGCTGCTGTTCTTGGCTCTCAACTTCTT
1▶ M T L R L L F L A L N F F
601 CTCAGTTC AAGTAA CAGAAA CAAGATTTTGGTAAAGCAGTCGCCCTGCTTGTGGTAGATAGCAACGAGGTCAGCCTCAGCTGCAGTATTCTACAAC
13▶ S V Q V T E N K I L V K Q S P L L V V D S N E V S L S C R Y S Y N
701 CTTCTCGCAAAGGAATCCGGGCATCCCTGTACAAGGGCGTGAACAGCGACGTGGAAGTCTGTGTCGGGAATGGGAATTTTACCTATCAGCCCAGTTTC
47▶ L L A K E F R A S L Y K G V N S D V E V C V G N G N F T Y Q P Q F
801 GCTCGAATGCCGAGTTCAACTGCGACGGGATTTGACAACGAAACAGTGACGTTCCGCTCTGGAATCTGACGCTCAATCACACAGATATTTACTTCTG
80▶ R S N A E F N C D G D F D N E T V T F R L W N L H V N H T D I Y F C
901 CAAAATTGAGTTCATGTACCCTCCGCTTACCTAGACAACGAGAGGAGCAATGGAACATTATTACATAAAAAGAGAAACATCTTTGTCACTACGTC
113▶ K I E F M Y P P P Y L D N E R S N G T I I H I K E K H L C H T Q S
1001 TCTCCTAAGCTGTTTTGGGCACTGGTGTGGTGTCTGGAGTCTGTTTTGTTATGGCTTGTAGTACAGTGGCTCTTTGTGTATCTGGACAAATAGTA
147▶ S P K L F W A L V V V A G V L F C Y G L L V T V A L C V I W T N S
1101 GAAGGAACAGACTCCTTCAAAGTGACTACATGAACATGACTCCCCGGAGGCTGGGCTCACTGAAAGCCTTACCAGCCCTACGCCCTGCCAGAGACTT
180▶ R R N R L L Q S D Y M N M T P R R P G L T R K P Y Q P Y A P A R D F

NheI (1242)

1201 TGCAGCGTACCGCCCTGACAGGGACCCCTATCCAGAAGCCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAA
213▶ A A Y R P •
1301 TGCAGTAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCAT
1401 TCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAA
1501 AACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGC
1601 AGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACA
1701 TTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCA
1801 TAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCTGGTGA
141▶ • N R T Y
1901 CTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTG
136▶ 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 Y D S I L E R
2001 CACATGCCACAGGGGCTGACCACCTGATGGATGCTGCCACCTCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGC
102▶ 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 D F D K Q G N S
2101 TCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCT
69▶ 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 I I E G T K T R
2201 GATGGCCGCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAG
36▶ 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 D Q Q S I N F
2301 GTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTT
2▶ T K M
2401 ATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTG
2501 GAAAGTCCCGTTGATTTACTAGTCAAAAACAACCTCCATTGACGTCAATGGGGTGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATT
2601 GATGTAAGTCCAAACCGCATCATCATGGAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGAAAGTCCATAAGGTCATGTAAGTGGGCATAA
2701 TGCCAGGCGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACT
2801 CCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGG
2901 CGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATAAGAACATGTGAGCAAAAGGCCAGAAAAGGCCAGAACCGTAAAAAGGCCGCTTGC
3001 TGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
3101 GTTTCCTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTCT

3201 CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTAGCCCGACCGCTGCGCCTTAT
3301 CCGGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG
3401 TGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA
3501 GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAG
3601 ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACACAGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGC
3701 CGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACA
3801 AACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA