



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCATCATGAGGCGCGGGCGCTCTGGCTCTGGCTGTGCGCGCT
1 M R R A A L W L W L C A L
601 GGCCTGAGCCTGCAGCCGGCCCTGCCCAAATTGTGGCTACTAATTTGCCCCCTGAAGATCAAGATGGCTCTGGGGATGACTCTGACAACTTCTCCGGC
13 A L S L Q P A L P Q I V A T N L P P E D Q D G S G D D S D N F S G
701 TCAGGTGCAGGTGCTTTGCAAGATATCACCTTGTACAGCAGACCCCTCCACTTGAAGGACACGCAGCTCTGACGGCTATCCCACGTCTCCAGAAC
47 S G A G A L Q D I T L S Q Q T P S T W K D T Q L L T A I P T S P E
801 CCACCGCCTGGAGGCTACAGCTGCCTCCACCTCCACCTGCCGGCTGGAGAGGGGCCAAGGAGGGAGAGGCTGTAGTCTGCCAGAAGTGGAGCCTGG
80 P T G L E A T A A S T S T L P A G E G P K E G E A V V L P E V E P G
901 CCTCACC GCCGGGAGCAGGAGGCCACCCCGACCCAGGGAGACCACAGCTCCCGACCTCATCTGGCCTCAACGACCAGCCACCGCCAGCCGAG
113 L T A R E Q E A T P R P R E T T Q L P T T H L A S T T T A T T A Q
1001 GAGCCCGCCACCTCCACCCCAAGGACATGAGCCTGGCCACCATGAGACCTCAACCCCTGCAGGACCCAGCCAAAGTGCACCTTCCACCTCCCA
147 E P A T S H P H R D M Q P G H H E T S T P A G P S Q A D L H T P H
1101 CAGAGGATGGAGTCTTCTGCCACCGAGAGGGCTGTGAGGATGGAGCCTCAGTCACTCCAGCAGCAGAGGGCTCTGGGGAGCAGGACTTACCTT
180 T E D G G P S A T E R A A E D G A S S Q L P A A E G S G E Q D F T F
1201 TGAACCTCGGGGAGAATACGGCTGTAGTGGCGTGGAGCCTGACCGCCGGAACAGTCCCCAGTGGATCAGGGGGCCACGGGGCCTCACAGGGCCTC
213 E T S G E N T A V V A V E P D R R N Q S P V D Q G A T G A S Q G L
1301 CTGGACAGAAAGAGTCTGGGAGGGGTCATTGCCGAGGCTCGTGGGGCTCATCTTGTGTGCTGGTGGGTTTCATGCTGTACCGCATGAAGA
247 L D R K E V L G G V I A G G L V G G L I F A V C L V G F M L Y R M K
1401 AGAAGGACGAAGGAGCTACTCCTTGGAGGAGCCGAAACAAGCAACGGCGGGCCTACCAGAAGCCACCAACAGGAGGAATTCTATGCCTGACGCGG
280 K K D E G S Y S L E E P K Q A N G G A Y Q K P T K Q E E F Y A •

NheI (1502)

1501 GAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAGTGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGC
1601 TATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTT
1701 TTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAATACAGCATAGCAAACCTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTT
1801 CTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATT
1901 TTCCAAGTTTGAAGTACTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGTAATAATTTAATAACA
2001 TCATTGCAATGAAAATAAATGTTTTTATTAGGAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAG
2101 GAACCTTAAATAGAAATTTGGACAGCAAGAAAGCAGGCTTCTAGCTTTAGTCTCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGC
141 • N R T Y K L P I L E E I T T K V L
2201 TTGCCATTCATCTCAATGAGCACAAGCAGTCAAGGAGCATAGTCAAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCA
122 K G N M 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
2301 CCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTGTCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGT
89 E D S 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
2401 GACCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGC
56 V R G I Y A E I H V A S I I E G T K T R I A A G V H H K N D E Y L
2501 ATGGTGATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGTGATTATACT
22 M T I K E T A V E V L E L D Q Q S I N F T K M
2601 ATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTACTAAACGAGCTCTGCTTATATAGACCT
2701 CCCACCGTACACGCCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACTCCCATT
2801 GACGTC AATGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAACCGCATCATGTTAATAGCGATG
2901 ACTAATACGTAGATGTAAGTCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTC AATAGGGG
3001 GCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCAAGTTTACCGTAAATACTCCACCCATTGACGTC AATGAAAGTCCCTATTGGCGTTAC
3101 TATGGGAACATACGTCATTATTGACGTC AATGGGCGGGGTCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTATGTAACGCCTGAGGTTAA

3201 ITAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCA
3301 CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCG
3401 ACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCCG
3501 TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA
3601 CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTAC
3701 ACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
3801 GCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAA
3901 CGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTG
4001 GTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGC
4101 CAGAACATTTCTCTATCGAA