



1 GGATCTGCGATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTGCTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCCGCTTGTCCGCCCTCCGCCTGTGGTGCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspEI (558)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGCTCCGGACACCATGGATGACCAACGCGACCTCATCTCTAACCA
601 TGAGCAATTGCCATACTGGGCAACCGCCTAGAGAGCCAGAAAGGTGCAGCCGTGGAGCTGTATACCGGTGTITCTGTCTGGTGGCTGCTGCTTTG
111 E Q L P I L G N R P R E P E R C S R G A L Y T G V S V L V A L L L
701 GCTGGGCAGGCCACCACTGCTTACTTCTGTACAGCAACAGGGCCGCTAGACAAGTGCACATCACCTCCAGAACCTGCAACTGGAGAGCCTTCGCA
45 A G Q A T T A Y F L Y Q Q G R L D K L T I T S Q N L Q L E S L R
801 TGAAGCTTCCGAAATCTGCCAACTGTGAGCCAGATGCGGATGGCTACTCCCTTGTGATGCGTCCAATGTCATGGATAACATGCTCCTGGGCTGT
78 M K L P K S A K P V S Q M R M A T P L L M R P M S M D N M L L G P V
901 GAAGAAGCTTACCAAGTACGGCAACATGACCCAGGACCATGTGATGCATCTGCTCACGAGTCTGGACCCCTGGAGTACCCGACCTGAAGGGGACCTTC
111 K N V T K Y G N M T Q D H V M H L L T R S G P L E Y P Q L K G T F
1001 CCAGAGAATCTGAAGCATCTTAAGAACTCCATGGATGGCGTGAAGTGAAGATCTTCGAGAGCTGGATGAAGCAGTGGCTCTGTTTGTGATGAGCAAGA
145 P E N L K H L K N S M D G V N W K I F E S W M K Q W L L F E M S K
1101 ACTCCCTGGAGGAGAAGAAGCCACAGAGGCTCCACCTAAAGTACTGACCAAGTCCAGGAAGAAGTACGCCACATCCCTGCCGCTACCCGGGTGCGTT
178 N S L E E K K P T E A P P K V L T K C Q E E V S H I P A V Y P G A F
1201 CCGTCCAAGTGCAGCAGAAAGGTAATTTGCCACTCCAGTGCACGGGAGCACTGGCTACTGCTGGTGTGTGTTCCCAACGGCACTGAGGTTCT
211 R P K C D E N G N Y L P L Q C H G S T G Y C W C V F P N G T E V P
1301 CACACCAAGAGCCGCGGGCCATAACTGCAGTGCAGCCACTGGACATGGAAGCCTATCTTCTGGCTGGGAGTGACCAGGCAGGAAGTGGTCAAGTCA
245 H T K S R G R H N C S E P L D M E D L S S G L G V T R Q E L G Q V

NheI (1417)

1401 CCCTGTGAAGACAGAGGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTG
278 T L •
1501 TGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTATTTTATGTTTCAGGTTCCAGGG
1601 GAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTC
1701 TACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTT
1801 AAGATATAGTGATTTTCCCAAGGTTTGAAGTACTGCTCTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAATATTGCA
1901 AATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGG
2001 ACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTTCTGGTACTTGAGGGGGATGAGTTCCTCAATG
141 N R T Y K L P I L E E I
2101 GTGGTTTTGACCAGTTGCCATTCATCTCAATGAGCACAAAGCAGTCAAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCC
127 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 C P S V V R
2201 TGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCAGCAGACCCAATGGCAATGGC
94 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 G I A I A
2301 TTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCGACATGGTGCTTG
61 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 G V H H K
2401 TTGTCCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGT
27 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
2501 GAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTC
2601 TGCTTATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCA
2701 AAACAACTCCCATTTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCCATTTGATGTAAGTGCACAAACCGCATCATC
2801 ATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGTGCATGTAAGTGGGCATAATGCCAGGCGGGCATTACCCTCAT
2901 TGACGTCAATAGGGGGTACTTGGCATATGATACACTTGTACTGCAAGTGGGAGTTTACCCTAAATACTCCACCATTTGACGTCAATGAAAGT
3001 CCCTATTGGCGTTACTATGGAAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTCAGCCAGGCGGGCATTACCCTAAGTTATGTA
3101 ACGCCTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCC

3201 CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
3301 CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCA
3401 GTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA
3501 CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG
3601 CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGATTGGTAGCTCTTGATCCGGCAAAC
3701 AAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
3801 TGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCAT
3901 TACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAG
4001 TGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA