



150

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCCACCTGCCACTCCAAGTGAAGAGACGCCCTGATCCCTCA
13> R S C S L L S T E A G A L H V L L P A R G P G P P Q R L S F S F G
47> D H L A E D L C V Q A A K A S G I L P V Y H S L F A L A T E D L S
801 GCTGGTCCCCCGAGCCACATCTTCCGTGGAGGATGCCAGACCCAAGTCTGCTGTACAGGATTCGCTTTTACTTCCCCAATTGGTTGGGCTGGA
80> C W F P P S H I F S V E D A S T Q V L L Y R I R F Y F P N W F G L E
901 GAAGTGCCACCGCTTCGGGCTACGCAAGGATTTGGCCAGTGTATCCTTGACCTGCCAGTCTGGAGCACCTTTGCCAGCACCGCAGTGCCTGGTG
113> K C H R F G L R K D L A S A I L D L P V L E H L F A Q H R S D L V
1001 AGTGGGCGCTCCCGTGGGCTCAGTCTCAAGGAGCAGGGTGTGTCTCAGCCTGGCGTGTGGACCTGGCCGGATGGCGGAGAGCAGGCCAGC
147> S G R L P V G L S L K E Q G E C L S L A V L D L A R M A R E Q A Q
1101 GGCCGGAGAGCTGTGAAGACTGTGAGTACAAGGCTGCCTACCCCAAGCTGCAGCCTGATCCAGGGCTGAGCTTCGTGACGCGGAGGCGTAT
180> R P G E L L K T V S Y K A C L P P S L R D L I Q G L S F V T R R R I
1201 TCGGAGGACGGTGCAGAGCCCTGCCCGCGTGGCCGCTGCCAGGACAGCCGACTCGCTCATGGCAAGTACATCATGGACCTGGAGCGGCTGGAT
213> R R T V R R A L R R V A A C Q A D R H S L M A K Y I M D L E R L D
1301 CCAGCCGGGGCCGCGAGACCTCCACGTGGGCTCCCTGGGCGCTTGGTGGCCACGACGGGCTGGGCTGCTCCGCGTGGTGGTACGCGGCGATCG
247> P A G A A E T F H V G L P G A L G G H D G L G L L R V A G D G G I
1401 CCTGGACCCAGGAGAACAGGAGTCTCCAGCCCTTCTGCGACTTCCAGAAATCGTAGACATTAGCATCAAGCAGGCCCGCGGTTGGCCGGCCGG
280> A W T Q G E Q E V L Q P F C D F P E I V D I S I K Q A P R V G P A G
1501 AGAGCACCGCCTGGTCACTGTTACCAGGACAGACAACCAGATTTAGAGGCGGAGTTCACAGGGCTGCCCGAGGCTGTGCTGTTGTTGGGCTCGTGGAC
313> E H R L V T V T R T D N G I L E A E F P G L P E A L S F V A L V D
1601 GGCTACTTCCGGTACCACGGACTCCAGCACTTCTTCTGCAAGGAGTGGCACCCGCGAGGCTGCTGGAGGAAGTGGCCGAGCAGTGCACGGCCCA
347> G Y F R L T T D S Q H F F C K E V A P P R L L E E V A E Q C H G P
1701 TCACTCTGGACTTTGCCATCAACAAGCTCAAGACTGGGGGCTCACGTCTGGCTCCTATGTTCTCCGCGCAGCCCCAGGACTTTGACAGCTTCTCTCT
380> I T L D F A I N K L K T G G S R P G S Y V L R R S P Q D F D S F L L
1801 CACTGTCTGTGTCAGAACCCCTTGGTCTGATTATAAGGGTGCCTCATCCGGCGCAGCCCCACAGGAACCTTCTTCTGTTGGCTCAGCCGACCC
413> T V C V Q N P L G P D Y K G C L I R R S P T G T F L L V G L S R P
1901 CACAGCAGTCTTCGAGAGCTCTGGCAACCTGCTGGGATGGGGGGTGCACGTAGATGGGGTGGCAGTACCCTCACTTCTGCTGTATCCCCAGCCCA
447> H S S L R E L L A T C W D G G L H V D G V A V T L T L S C C I P R P
2001 AAGAAAAGTCCAACCTGATCGTGGTCCAGAGGTCACAGCCACCCACATCCTTGGTTCAAGCCCAATCCAATACCAGCTGAGTCAGATGACATT
480> K E K S N L I V V Q R G H S P P T S S L V Q P Q S Q Y Q L S Q M T F
2101 TCACAAGATCCCTGCTGACAGCCTGGAGTGGCATGAGAACCTGGCCATGGGCTCCTCACCAAGATTTACCGGGGCTGTCGCCATGAGGTGGTGGATGGG
513> H K I P A D S L E W H E N L G H G S F T K I Y R G C R H E V V D G
2201 GAGGCCGAAAGACAGAGGTGCTGCTGAAGGTCATGGATGCCAAGCACAAGAAGTGCATGGAGTATTCTGGAAGCAGCGAGCTTGTGAGCCAAGTGT
547> E A R K T E V L L K V M D A K H K N C M E S F L E A A S L M S Q V
2301 CGTACCGGACTCTGCTGCTCCACGGCGTGTGCATGGTGGAGACAGCACCATGGTGCAGGAATTTGTACACCTGGGGCCATAGACATGATCTGCG
580> S Y R H L V L L H G V C M A G D S T M V K L Q F V H L G A I D M Y L R
2401 AAAACGTGGCCACCTGGTCCAGCCAGCTGGAAGTGCAGGTGGTCAAACAGCTGGCTACGCCCTCAACTATCTGGAGGACAAGGCTGCCCATGGC
613> K R G H L V P A S W K L Q V V K Q L A Y A L N Y L E D K G L P H G
2501 AATGCTCTGCCCGAAGGTGCTCCTGGCTCGGAGGGGGCTGATGGGAGCCCGCCCTCATCAAGCTGAGTGACCCTGGGGTCCAGCCCGCTGTGTTAA
647> N V S A R K V L L A R E G A D G S P P F I K L S D P G V S P A V L
2601 GCCTGGAGATGCTCACCGACAGGATCCCTGGTGGCCCCGAGTGTCTCCGGGAGGCGCAGACACTAGCTTGAAGCTGACAAGTGGGGCTTCGGCGC
680> S L E M L T D R I P W V A P E C L R E A Q T L S L E A D K W G F G A
2701 CACGGTTCGGGAAGTGTAGTGGCGTACCATGCCATCAGTCCCTGGATCTGCTAAGAACTCCAATTTTATGAGGACCGCAGCAGCTGCCGGCC
713> T V W E V F S G V T M P I S A L D P A K K L Q F Y E L G A R I D M Y L P
2801 CCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGTGCATGGCTATGAGCCGGTCCAGAGGCCCTCCTCCGAGCCGTCATTCTGTGACCTCAATAGCC
747> P K W T E L A L L I Q Q C M A Y E P V Q R P S F R A V I R D L N S
2901 TCATCTCTCAGACTATGAGCTCCTCAGACCCACACCTGGTGCCTGGCACCTCGTGATGGGCTGTGGAATGGTCCCAGCTCTATGCCTGCCAAGA
780> L I S S D Y E L L S D P T P G A L A P R D G L W N G A Q L Y A C Q D
3001 CCCACGATCTTCGAGGAGAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTGCCGCTATGACCCGCTAGGCGAC
813> P T I F E E R H L K Y I S Q L G K G N F G S V E L C R Y D P L G D
3101 AATACAGGTGCCCTGGTGGCCGTGAAACAGCTGCAGCACAGCGGGCCAGACCAGAGGAGCTTTACGCGGGAGATTGAGTCTCAAGACTGCACA
847> N T G A L V A V K Q L Q H S G P D Q Q R D F Q R E I Q I L K A L H
3201 GTGATTTCAATGTCAAGTATCGTGGTGTGAGTATGGCCCGGGCCGAGAGCTGCGGCTGGTGCATGGAGTACCTGCCAGCGGCTGCTGCGGACTT
880> S D F I V K Y R G V S Y G P G R Q S L R L V M E Y L P S G C L R D F

3301 CCTGCAGCGGCACCGCGCGCCTCGATGCCAGCGCCTCCTTCTCTATTCTCGCAGATCTGCAAGGGCATGGAGTACCTGGGCTCCCGCCGCTGCGTG
913▶ L Q R H R A R L D A S R L L L Y S S Q I C K G M E Y L G S R R C V
3401 CACCGCGACCTGGCCGCCGAAACATCCTCGTGAGAGCGAGGCACACGTCAAGATCGCTGACTTCGGCCTAGCTAAGCTGCTGCCGCTTGACAAAGACT
947▶ H R D L A A R N I L V E S E A H V K I A D F G L A K L L P L D K D
3501 ACTACGTGGTCCGCGAGCCAGGCAGAGCCCCATTTTCTGGTATGCCCCGAATCCCTCTCGGACAACATCTTCTCTCGCCAGTCAGAGCTGAGGCTT
980▶ Y Y V V R E P G Q S P I F W Y A P E S L S D N I F S R Q S D V W S F
3601 CGGGTCTGCTGTACGAGCTCTTACCTACTGCGACAAAAGCTGCAGCCCCCTCGCCGAGTTCCTGCGGATGATGGGATGTGAGCGGGATGCCCGCC
1013▶ G V V L Y E L F T Y C D K S C S P S A E F L R M M G C E R D V P A
3701 CTCTGCCGCTCTTGAAGTCTGGAGGAGGGCAGAGGCTGCCGGCGCTCCTGCCTGCCCTGCTGAGGTTACAGAGCTCATGAAGTGTGCTGGGCC
1047▶ L C R L L E L L E E G Q R L P A P P A C P A E V H E L M K L C W A
3801 CTAGCCCACAGGACCGGCATCATTAGCGCCCTGGGCCCCAGCTGGACATGCTGTGGAGCGGAAGCCGGGGTGTGAGACTCATGCCTTACTGCTCA
1080▶ P S P Q D R P S F S A L G P Q L D M L W S G S R G C E T H A F T A H

NheI (3952)

3901 CCCAGAGGGCAAACACCACTCCCTGTCTTTTCATAGCTCCTGCCGCGAGCGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACC
1113▶ P E G K H H S L S F S •
4001 ACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACA
4101 ACAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATAC
4201 AGCATAGCAAACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATT
4301 AGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTCTTCATTCTTTATGTTTAAATGCACTG
4401 ACCTCCCACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCA
4501 AGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGT
4601 TCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAAGCAGTCAGGACATAGTCAGAGAT
139▶ 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 Y D S I
4701 GAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCCTC
106▶ 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 D F D K
4801 TGCCGTTGCTCAGCAGACCAATGGCAATGGCTTACGACAGACAGTACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAG
72▶ 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 I I E G T
4901 TCTTGGTCTGATGGCCGCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGA
39▶ 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 D Q Q S
5001 GATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCG
6▶ I N F T K M
5101 TCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTTA
5201 CGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATC
5301 CACGCCATTGATGTAAGTCCGATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAGGTCATGTA
5401 CTGGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTAC
5501 CGTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAACATACGTCAATATTGACGTCAATGGGCGGGGCTGTTGGGGC
5601 GTCAGCCAGGCGGCCATTACCCTAAGTTATGTAACGCCTGAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAG
5701 GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
5801 GATAACAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGT
5901 GGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCAGCAACCCCGTTCAGCCGACCGC
6001 TGCGCCTTATCCGGTAATCTGCTTGTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
6101 ATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
6201 CGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGA
6301 TCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTA
6401 AATCAGCGGCCCAATAAAATATCTTTATTTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACATAACGCTCTCCATCAAAACAAAACG
6501 AAACAAAACAACTAGCAAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA