



1 GGATCTGCGATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTATGTGCTGACTGGCTCCGCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTTCGAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTCAGCGCCCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCGCGTTTCCGCCCTCCCGCTGTGGTGCTCTGAAGTTCGCTCCGCCGTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT
501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTAGGAGGGCCACCATGGACTATAGCTTTGATACCACAGCCGAA
AgeI (552) NcoI (568)
NcoI (662)
601 GACCCCTGGGTAGGATTTCTGACTGCATCAAAAACTATTTAGCCCATCATGAGTGAGAACCATGGCCACATGCCTCTACAGCCCAATGCCAGCCTGA
111 D P W V R I S D C I K N L F S P I M S E N H G H M P L Q P N A S L
701 ATGAAGAAGAAGGGACACAGGGCCACCCAGATGGGACCCACCAAAGCTGGACACCGCAATGGCACTCCCAAAGTTTACAAGTCAGCAGACAGCAGCAC
444 N E E E G T Q G H P D G T P P K L D T A N G T P K V Y K S A D S S T
801 TGTGAAGAAAGTCTCCTGTGGCTCCCAAGCCAGCTGGTTTCGCCAAAGCTTGAAGGTTTGGGAATCGTGTTCAGACCCAAGAGGGCTCCTGAT
777 V K K G P P V A P K P A W F R Q S L K G L R N R A S D P R G L P D
901 CCTGCCTGTCCACCCAGCCAGCACCTGCTCCAGGGAGCACCTAGGATCACACATCCGGGCTCCTCCTCCTCCTCCATCAGGCAGAGAAATCAGCT
111 P A L S T Q P A P A S R E H L G S H I R A S S S S S S I R Q R I S
1001 CCTTGAAACCTTTGGCTCCTCTCAACTGCCTGACAAAGGAGCCAGAGACTGAGCCTCCAGCCCTCCTCCTGGGGAGGCAGAAAACCTTTGGGAAGCA
144 S F E T F G S S Q L P D K G A Q R L S L Q P S S G E A A K P L G K H
1101 TGAGGAAGGACGGTTTTCTGACTCTTGGGGCGAGGGCTGCACCCACTCTTGTGCCCCAGCAGCCTGAGCAAGTACTGTCTCGGGTCCCCTGCAGCC
177 E E G R F S G L L G R G A A P T L V P Q Q P E Q V L S S G S P A A
1201 TCCGAGCCAGAGACCAGGTGTCTGAGTCCCCTCCCCAGGGCGCAGCCAATCAGAAAACCTCCCCCTGGCCCGGACCCGCTCCTAAGCTGTC
211 S E A R D P G V S E S P P P G R Q P N Q K T L P P G P D P L L R L
1301 TGTCAACACAGGCTGAGGAATCTCAAGGCCAGTGTCAAGATGCCTAGCCAGCGAGCACGGAGCTTCCCCCTGACCAAGTCCAGTCTGTGAGACGAA
244 L S T Q A E E S Q G P V L K M P S Q R A R S F P L T R S Q S C E T K
1401 GCTACTTGACGAAAAGACCAGCAAACTTATTCTATCAGCAGCAAGTGTATCGGCTGTGATAATCCTTGTGTGCTTCCATCTCTATCTCCTGT
277 L L D E K T S K L Y S I S S Q V S S A V M K S L L C L P S S I S C
1501 GCCCAGACTCCCTGCATCCCCAAGGAAGGGGCTCCTCAACATCATCAACGAAGACTCAGCTGCAAATGGTTCGTGAAAACATCTGCCTTGGACA
311 A Q T P C A I P K E G A S P T S S S N E D S A A N G S A E T S A L D
1601 CAGGTTCTCGTCAACTTTTCCAGAGCTGAGAGAATATACAGAGGGTCTCACGAAGCAAGGAAGCAGTATGGGGACCACAGTTCCTCCTCAGTCTGG
344 T G F S L N L S E L R E Y T E G L T E A K E D D D G D H S S L Q S G
1701 TCAGTCCGTTATCTCCCTGTGAGCTCAGAAGAATAAAAAACTCATCGAGGAGGTGAAGTTCGTGATGAAGCAACATTAAGCAATTAGACGGCCTC
377 Q S V I S L L S S E E L K K L I E E V K V L D E A T L K Q L D G I
1801 CATGTCCCATCTTACACAAGGAGGAAGGTGCTGGTCTTGGTTCAGCTTGGCAGGAGGAGCAGATCTAGAAAAAAGGTGATTACGGTTCACAGAGTGT
411 H V T I L H K E E G A G L G F S L A G G A D L E N K V I T V H R V
1901 TTCAAATGGGCTGGCCTCCCAGGAAGGGACTATTAGAAAGGGCAATGAGGTTCTTTCCATCAACGGCAAGTCTCTCAAGGGGACCACGCACCATGATGC
444 F P N G L A S Q E G T I Q K G N E V L S I N G K S L K G T T H H D A
2001 CTTGGCCATCCTCCGCAAGCTCGAGAGCCAGGCAAGCTGTGATTGTCAAGGAAGCTGACTCCAGAGGCCATGCCCGACCTCAACTCCTCAACTGAC
477 L A I L R Q A R E P R Q A V I V T R K L T P E A M P D L N S S T D
2101 TCTGCAGCCTCAGCCTCTGCAGCCAGTGTGTTCTGTAGAATCTACAGCAGAGGGCCACAGTCTGCACGGTGCACACTGGAGAAGATGTCGGCAGGGCTGG
511 S A A S A S A A S D V S V E S T A E A T V C T V T L E K M S A G L
2201 GCTCAGCCTGGAAGGAGGGAAGGGCTCCCTACACGGAGACAAGCCTCTCACATTAACAGGATTTTCAAAGGAGCAGCCTCAGAACAAGTGAAGAGT
544 G F S L E G G K G S L H G D K P L T I N R I F K G A A S E Q S E T V
2301 CCAGCCTGGAGATGAAATCTTGCAGCTGGGTGGCACTGCCATGCAGGGCTCAGCGGTTTGAAGCCTGGAACATCATCAAGGCAGTGCCTGATGGACCT
577 Q P G D E I L Q L G G T A M Q G L T R F E A W N I I K A L P D G P
2401 GTCACGATTGTATCAGGAGAAAAAGCTCCAGTCCAAGGAAACCACAGCTGTGGAGACTCCTAGGCAAGGACATGCTGAAGCCAAAGCAATAACACAC
611 V T I V I R R K S L Q S K E T T A A G D S •

NheI (2519)
2501 AGCTAACACACAGCTCCAGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAAACCACTAGAATGCAGTGAAAAAATGCTTTATT
2601 TGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGTGAATAAACAAGTTAACAACAACAATTGCAATCATTATGTTTCAGGTTTAGG
2701 GGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAATGTGGTATGGAATCTAAAATACAGCATAGCAAACCTTTAACTCCAAATCAAGCC
2801 TCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGT
2901 TTAAGATATAGTGTATTTTCCAAGGTTTGAAGTGGCTCCTTATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAAATATTCA
3001 GAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTT
3101 GGAAGTTAGGAAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCAGCTTCTAGCTTAGTTCTGGTACTTGGAGGGATGAGTTCTCAA
141 • N R T Y K L P I L E E I

3201 TGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCAC
128 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
3301 CCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATG
95 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 V A S G I A I
3401 GCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGTATGGCCGCCCGACATGGTGCT
61 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 I A A G V H H K
3501 TGTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCCTATA
28 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
3601 GTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGC
TCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTAGT
3701 CAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCA
3801 TCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTGATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTC
3901 ATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAA
4001 GTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATG
4101 TAACGCCTGCAGGTTAATAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
4201 CCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCG
4301 TCGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT
4401 CAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC
4501 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGT
4601 GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
4701 ACAAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG
4801 TCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTC
4901 ATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACATAACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCC
5001 AGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA