



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCAAGACACCGTGGAAAGTTCTTCTGGACTGCTGGG
1 M K T P W K V L L G L L G
601 TGCTGCTGCGCTTGTACCATCATCACCGTGCCCGTGGTTCTGCTGAACAAAGGCACAGATGATGCTACAGCTGACAGTCGAAAACCTTACTCTAACT
13 A A A L V T I I T V P V V L L N K G T D D A T A D S R K T Y T L T
701 GATTACTTAAAAAATACTTATAGACTGAAGTTATACTCCTTAAGATGGATTCAGATCATGAATATCTCTACAAAACAAGAAAATAATATCTTGGTATTCA
47 D Y L K N T Y R L K L Y S L R W I S D H E Y L Y K Q E N N I L V F
801 ATGCTGAATATGGAAACAGCTCAGTTTTCTGGAGAACAGTACATTTGATGAGTTGGACATTCTATCAATGATTATTCAATATCTCCTGATGGGCAGTT
80 N A E Y G N S S V F L E N S T F D E F G H S I N D Y S I S P D G Q F
901 TATTCTTAGAATACTACTGAGCAATGGAGGATTCTACACAGCTTCATATGACATTTATGATTTAAATAAAAGGCAGCTGATTACAGAAGAG
113 I L L E Y N Y V K Q W R H S Y T A S Y D I Y D L N K R Q L I T E E
1001 AGGATTCCAAACAACACACAGTGGGTACATGGTCACCGTGGGTGATAAATGGCATATGTTTGAACAATGACATTTATGTTAAATGAACCAAATT
147 R I P N N T Q W V T W S P V G H K L A Y V W N N D I Y V K I E P N
1101 TACCAAGTTACAGAATCATGGACGGGAAAGAAGATATAATATAATGGAATAACTGACTGGGTTTATGAAGAGGAAGTCTTCACTGCTACTCTGC
180 L P S Y R I T W T G K E D I I Y N G I T D W V Y E E E V F S A Y S A
1201 TCTGTGGTGGTCTCCAAACGGCACTTTTTAGCATATGCCCAATTAACGACACAGAAGTCCCACTTATTGAATACTCCTTACTCTGATGAGTCACTG
213 L W W S P N G T F L A Y A Q F N D T E V P L I E Y S F Y S D E S L
1301 CAGTACCCAAAGACTGTACGGGTTCCATATCCAAAGGCAGGAGCTGTAATCCAAGTCAAAGTTCTTGTGTAATAACAGACTCTCTCAGCTCAGTCA
247 Q Y P K T V R V P Y P K A G A V N P T V K F F V V N T D S L S S V
1401 CCAATGCAACTCCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGGATCACTACTGTGTGATGTGACATGGCAACAACAAGAAATTTCTTT
280 T N A T S I Q I T A P A S M L I G D H Y L C D V T W A T Q E R I S L
1501 GCAGTGGCTCAGGAGGATTCAGAACTATTCGGTCATGGATTTGTGACTATGATGAATCCAGTGGAAAGTGGAACTGCTTAGTGGCAGGCAACACATT
313 Q W L R R I Q N Y S V M D I C D Y D E S S G R W N C L V A R Q H I
1601 GAAATGAGTACTACTGGCTGGTTGGAAGATTTAGGCCTTCAGAACCTCATTTTACCCTTGTGGTAAATAGCTTCTACAAGATCATCAGCAATGAAGAAG
347 E M S T T G W V G R F P S E P H F T L D G N S F Y K I I S N E E
1701 GTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGCAGTGCACATTTATTACAAAAGGCACCTGGGAAGTCACTGGGATAGAAGCTTAAACAGTGA
380 G Y R H I C Y F Q I D K K D C T F I T K G T W E V I G I E A L T S D
1801 TTATCTACTACATTAGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACCTAGTACTATACAAAAGTACATGCCTCAGT
413 Y L Y Y I S N E Y K G M P G G R N L Y K I Q L S D Y T K V T C L S
1901 TGTGAGCTGAATCCGAAAGGTGTCAGTACTATTCTGTGTCATTGATGAAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTGGTCTGCCCTCT
447 C E L N P E R C Q Y Y S V S F S K E A K Y Y Q L R C S G P G L P L
2001 ATACTCTACACAGCAGCTGAATGATAAAGGGCTGAGAGCTCTGGAAGCAATTCAGCTTTGGATAAAATGCTGCAGAATGTCAGATGCCCTCCAAAA
480 Y T L H S V N D K G L R V L E D N S A L D K M L Q N V Q M P S K K
2101 ACTGGACTTCATTATTTGAATGAAACAAAATTTGGTATCAGATGATCTTGCCTCCTCATTTTGATAAAATCCAAGAAATATCCTCTACTATTAGATGTG
513 L D F I I L N E T K F W Y Q M I L P P H F D K S K K Y P L L L D V
2201 TATGCAGGCCATGTAGTCAAAAAGCAGACACTGTCTCAGACTGAACTGGGCCACTTACCTTGAAGCACAGAAAACATTATAGTAGCTAGCTTTGATG
547 Y A G P C S Q K A D T V F R L N W A T Y L A S T E N I I V A S F D
2301 GCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGCATGCAATCAACAGAAGACTGGGAACATTTGAAGTTGAAGATCAAATGAAGCAGCCAGACAATT
580 G R G S G Y Q G D K I M H A I N R R L G T F E V E D Q I E A A R Q F
2401 TTCAAAAATGGGATTTGTGGACAACAACGAATTCGAATTTGGGGCTGGTTCATATGGAGGGTACGTAACCTCAATGGTCTGGGATCGGGAAGTGGCGTG
613 S K M G F V D N K R I A I W G W S Y G G Y V T S M V L G S G S G V
2501 TTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACACAGAAGTACATGGGTCTCCCAACTCCAGAAGACAACC
647 F K C G I A V A P V S R W E Y Y D S V Y T E R Y M G L P T P E D N
2601 TTGACCATTACAGAAATCAACAGTTCATGAGCAGAGCTGAAAATTTTAAACAAGTTGAGTACCTCCTTATTTCATGGAACAGCAGATGATAACGTTCACTT
680 L D H Y R N S T V M S R A E N F K Q V E Y L L I H G T A D D N V H F
2701 TCAGCAGTCACTCAGATCTCCAAAGCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTACTGATGAAGACCATGGAATAGCTAGCAGCACA
713 Q Q S A Q I S K A L V D V G V D F Q A M W Y T D E D H G I A S S T

EcoRI (2883)

2801 GCACACCAACATATATATACCCACATGAGCCACTTCATAAAAACATGTTTCTCTTTACCTTAGCACCTCAAAAATACCATGCCAGAAATTCGCTAGCTCGAC
747 A H Q H I Y T H M S H F I K Q C F S L P •
2901 ATGATAAGATACATTGATGAGTTTGACAAAACCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGA

3001 AATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAG

EcoRI (3148)

3101 GTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTAACTCCAAATCAAGCCTCTAC

3201 TTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAG
3301 ATATAGTGATTTTCCCAAGGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAAT
3401 AATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACT
3501 TAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTG
141 • N R T Y K L P I L E E I T
3601 GTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGA
126 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 I
3701 TGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCTTCTGCCGGTGTCTCACAGCAGACCCAATGGCAATGGCTTC
93 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 E
3801 AGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGTATGGCCGCCCGACATGGTCTTGTG
60 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 N
3901 TCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAG
26 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
4001 TCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGC
4101 TTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTAGTCAAAA
4201 CAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATG
4301 GTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTGATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGA
4401 CGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCC
4501 TATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACG
4601 CCTGCAGGTTAATAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
4701 TGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGTCCCTCGTGGCG
4801 TCTCCTGTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
4901 CGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCC
5001 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCT
5101 AACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGGCAAACAAA
5201 CCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA
5301 CGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAAATATCTTTATTTTCATTAC
5401 ATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGC
5501 AAGTGCAGGTGCCAGAACATTTCTATCGAA