



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCAGGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

Age1 (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTCCACCATGTTGTTTTCTGTTTTATTGCCACTAGTCTCTAG
1 M F V F L V L L P L V S S
601 TCAGTGTGTTAATCTTACAACCAAGAACTCAATTACCCCTGCATACACTAATCTTTCCACAGTGGTGTATTACCCTGACAAAAGTTTTAGATCCCTCA
13 Q C V N L T T R T Q L P P A Y T N S F T R G V Y Y P D K V F R S S
701 GTTTTACATTCACCTCAGGACTGTTCTTACCTTTCTTTCCAATGTTACTTGGTCCATGCTATACATGCTCTGGGACCAATGGTACTAAGAGGTTTG
47 V L H S T Q D L F L P F F S N V T W F H A I H V S G T N G T K R F
801 ATAACCTGTCTACCATTAAATGATGGTGTATTGTTGCTTCCACTGAGAAGTCTAACAATAAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAA
80 D N P V L P F N D G V Y F A S T E K S N I I R G W I F G T T L D S K
901 GACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAGTCTGTGAATTTCAATTTGTAATGATCCATTTTTGGGTGTTTATTACCAC
113 T Q S L L I V N N A T N V V I K V C E F Q F C N D P F L G V Y Y H
1001 AAAAAACAACAAAGTTGGATGGAAGTGAAGTTCAGAGTTTATTCTAGTGCAATAATTGCACCTTTGAATATGTCTCTCAGCCTTTCTTATGGACCTTG
147 K N N K S W M E S E F R V Y S S A N N C T F E Y V S Q P F L M D L
1101 AAGAAAAACAGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTATTTTAAAATATATTCTAAGCACACGCCTATTAATTTAGT
180 E G K Q G N F K N L R E F V F K N I D G Y F K I Y S K H T P I N L V
1201 GCGTGATCTCCCTCAGGGTTTTGCGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTTCAAATTTACTGCTTTACATAGA
213 R D L P Q G F S A L E P L V D L P I G I N I T R F Q T L L A L H R
1301 AGTTATTTGACTCCTGGTATTCTTCTCAGTTGGACAGCTGGTGCAGCTTATTATGTTGGGTTATCTTCAACCTAGGACTTTTCTATTAATAATA
247 S Y L T P G D S S S G W T A G A A A Y V G Y L Q P R T F L L K Y
1401 ATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACCTCTCAGAAAACAAAGTGTACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTA
280 N E N G T I T D A V D C A L D P L S E T K C T L K S F T V E K G I Y
1501 TCAAATCTTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCTAATATTACAAACTTGTGCCCTTTGGTGAAGTTTTTAAACGCCACCAGA
313 Q T S N F R V Q P T E S I V R F P N I T N L C P F G E V F N A T R
1601 TTTGCATCTGTTTATGCTTGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCGCGCATTTTTCCACTTTTAAAGTGT
347 F A S V Y A A N R K R I S N C V A D Y S V L Y N S A S F S T F K C
1701 ATGAGTGTCTCCTACTAAATTAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATCGKCTCCAGG
380 Y G V S P T K L N D L C F T N V Y A D S F V I R G D E V R Q I A P G
1801 GCAAACAGGAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGAATTTCTAACAATCTTATTGTTGTTGGT
413 Q T G K I A D Y N Y K L P D D F T G C V I A W N S N N L D S K V G
1901 GGTAAATATAATTACCTGTATAGATTGTTAGGAAGTCTAATCTCAAACCTTTTGGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCT
447 G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P
2001 GTAAGGTGTTGAAGTTTTAATTTACTTTTCAATCATATGGTTTCCAACCCACTAATGGTGTGGTTACCAACCATACAGAGTAGTAGTACT
480 C N G V E G F N C Y F P L Q S Y G Q P T N G V G Y Q P Y R V G Y L
2101 TTCTTTGAACTTCTACATGCACCAGAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATGTGCAATTTCAACTTCAATGGTTTA
513 S F E L L H A P A T V C G P K K S T N L V K N K C V N F N F N G L
2201 ACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGCTCGTGATCCAC
547 T G T G V L T E S N K K F L P F Q Q F G R D I A D T T D A V R D P
2301 AGACACTTGAGATTCTTGACATTACCCATGTTCTTTGGTGGTGTGAGTGTATAACACCAGGAACAAATCTTCAACCAGGTTGCTGTTCTTTATCA
580 Q T L E I L D I T P C S F G G V S V I T P G T N T S N Q V A V L Y Q
2401 GGATGTTAACTGCACAGAAGTCCCTGTTGCTATTATGTCAGATCAACTTACTCTACTTGGCGTGTATTCTACAGGTTCTAATGTTTTTCAAACAGT
613 D V N C T E V P V A I H A D Q L T P T W R V Y S T G S N V F Q T R
2501 GCAGGCTGTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATT
647 A G C L I G A E H V N N S Y E C D I P I G A G I C A S Y Q T Q T N
2601 CTCCTCGGCGGCACGTAGTGTAGTCAATCCATCATTGCCTACACTATGTCACCTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGC
680 S P R R A R S V A S Q S I I A Y T M S L G A E N S V A Y S N N S I A
2701 CATACCCACAAATTTACTATTAGTGTACCACAGAAATCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTTGGTGATTCA
713 I P T N F T I S V T T E I L P V S M T K T S V D C T M Y I C G D S
2801 ACTGAATGCAGCAATCTTTGTTGCAATATGGCAGTTTTGTACACAATTAACCGTCTTAACTGGAATAGCTGTTGAACAAGACAAAAACCCCAAG
747 T E C S N L L L Q Y G S F C T Q L N R A L T G I A V E Q D K N T Q
2901 AAGTTTTTGCACAAGTCAACAAATTTACAAAACACCACCAATTAAGATTTTGGTGGTTTTAATTTTTACAAAATATTACCAGATCCATCAAAAACAAAG
780 E V F A Q V K Q I Y K T P P I K D F G G F N F S Q I L P D P S K P S
3001 CAAGAGGTCATTTATTGAAGTCTACTTTTCAACAAAGTGCACCTGCAGATGCTGGCTTATCAACAAATATGGTATTGCCTTGGTGATATTGCTGCT
813 K R S F I E D L L F N K V T L A D A G F I K Q Y G D C L G D I A A
3101 AGAGACCTCATTTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCTCAGAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGG
847 R D L I C A Q K F N G L T V L P P L L T D E M I A Q Y T S A L L A
3201 GTACAATCACTTCTGGTTGGACCTTTGGTGCAGGTGCTGCATTACAAAATACCAATTTGCTATGCAAAATGGCTTATAGGTTTAAATGGATTGGAGTTACACA
880 G T I T S G W T F G A G A A L Q I P F A M Q M A Y R F N G I G V T Q

3301 GAATGTTCTCTATGAGAACCAAAAATTGATTGCCAACCAATTAATAGTGCTATTGGCAAAAATTCAAGACTCACTTTCTCCACAGCAAGTGCCTTGG
913▶ N V L Y E N Q K L I A N Q F N S A I G K I Q D S L S S T A S A L G
3401 AAACCTCAAGATGTGGTCAACCAAAAATGCACAAGCTTTAAACACGCTTGTAAACAACCTAGCTCCAATTTTGGTGC AATTTCAAGTGT TTTAAATGATA
947▶ K L Q D V V N Q N A Q A L N T L V K Q L S S N F G A I S S V L N D
3501 TCCTTTACGCTTGCACAAAGTTGAGGCTGAAGTGC A AATGATAGTTGATCACAGGCAGACTCAAAGTTTGCAGACATATGTGACTCAACAATTAAT
980▶ I L S R L D K V E A E V Q I D R L I T G R L Q S L Q T Y V T Q Q L I
3601 TAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGC
1013▶ R A A E I R A S A N L A A T K M S E C V L G Q S K R V D F C G K G
3701 TATCATCTTATGCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCTGCACAAGAAAAGAACTTCACAACTGCTCCTG
1047▶ Y H L M S F P Q S A P H G V V F L H V T Y V P A Q E K N F T T A P
3801 CCATTTGTGATGATGAAAGACACACTTTCCTCGTGAAGGTGCTTTGTTTCAAATGGCACACACTGGTTTGTAAACACAAAGGAATTTTATGAACCACA
1080▶ A I C H D G K A H F P R E G V F V S N G T H W F V T Q R N F Y E P Q
3901 AATCATTACTACAGACAACACTTTGTGTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGAC
1113▶ I I T T D N T F V S G N C D V V I G I V N N T V Y D P L Q P E L D
4001 TCATTCAAGGAGGAGTTAGATAAATATTTAAGAATCATAATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAACATT
1147▶ S F K E E L D K Y F K N H T S P D V D L G D I S G I N A S V V N I
4101 AAAAAGAAATGACCGCCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCTCAAGA AACTTGGAAAGTATGAGCAGTATATAAATGGCC
1180▶ Q K E I D R L N E V A K N L N E S L I D L Q E L G K Y E Q Y I K W P
4201 ATGGTACATTTGGCTAGGTTTTATAGCTGGCTTATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGCTCAAGGGC
1213▶ W Y I W L G F I A G L I A I V M V T I M L C C M T S C C S C L K G

NheI (4385)

4301 TGTGTTCTTGTGGATCTGCTGCAAAATTTGATGAAGACGACTCTGAGCCAGTCTCAAAGGAGTCAAATTACATTACACATAAAGCTAGCTGGCCAGAC
1247▶ C C S C G S C C K F D E D D S E P V L K G V K L H Y T •
4401 ATGATAAGATACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAA

4501 CCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAA

4601 CCTCTACAAATGTGGTATGGAATTTCAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGG

4701 CATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAC

4801 AGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAA

4901 AATGTTTTTTATAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAAT

5001 TGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTATCTCAAT
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I
5101 GAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGG
117▶ 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 E D S Y P
5201 TGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCAAGCAGACAGTACCTGCCAATGTAGG
83▶ 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 V R G I Y A
5301 CCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGT
50▶ E I H V A S I I E G T R I A A G V H H K N D E Y L M T I K E T
5401 GGCGACCTCCACAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGC
17▶ A V E V L E L D Q Q S I N F T K M
5501 CGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTA

5601 CCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGA

5701 GACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTA

5801 CTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGCTACTTGGCATATGA

5901 TACACTTGATGTACTGCCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCA

6001 TTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCA

6101 AAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGGTGGCGTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAA

6201 GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTGTCGCTCTCTGTTCCGACCCTGCCGTTACC GG

6301 ATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGC

6401 TGTGTGCACGAACCCCCGTTGACGCCGACCGTGGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG

6501 CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT

6601 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTCCGGCAACAAACCACCGCTGGTAGCGTGGTTTTTTTGT

6701 TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAG
6801 GGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCG
6901 TAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATC
7001 GAA