



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

Agel (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCTATGAGTTTCCGGATTGCGGGCCCGACTTTTGCTACT
1 M S F R I A G P R L L L L
601 GGGACTCCAGCTGTTTGCCAAGGCTGGAGCTACAACCTGGACACACGGCCTACGCAGAGCTTCTTGGCACAAGCTGGAAACATTTTGGGTACCAGGTC
13 G L Q L F A K A W S Y N L D T R P T Q S F L A Q A G R H F G Y Q V
701 TTGCAGATTGAAGATGGGTTGTCGTGGGAGCCCGAGTGAGGGGACAAACACGGGAGGCTCTATCACTGCCGAACAAGCAGCGAGTTTCCGCCAGCAG
47 L Q I E D G V V V G A P G E G D N T G G L Y H C R T S S E F C Q P
801 TCAGCCTACATGGTTCTAACCATACCTCCAAGTACTTGGGAATGACGCTGGCAACAGATGCCGCCAAGGGAAGCCTTTTGGCCTGTGACCCTGGACTGTC
80 V S L H G S N H T S K Y L G M T L A T D A A K G S L L A C D P G L S
901 TCGGACATGCGATCAGAACACTTACCTCAGTGGCCTCTGCTACCTTCCCCAGAGCTGGAGGGACCTATGTTACAAAATCGTCCCGCCTATCAGGAA
113 R T C D Q N T Y L S G L C Y L F P Q S L E G P M L Q N R P A Y Q E
1001 TGTATGAAGGGCAAAGTCGACCTGGTGTCTGTTTCGATGGCTCACAGAGCTTGGATAGAAAAGGACTTTGAAAAAATCCTGGAATTCATGAAGGATGTGA
147 C M K G K V D L V F L F D G S Q S L D R K D F E K I L E F M K D V
1101 TGAGGAAGCTCAGCAATACTTCTACCAGTTTGTGCTGCCGTCCAGTTTCCACAGACTGCAGAACAGAATTTACTTTTGGACTACGTTAAGCAGAACAA
180 M R K L S N T S Y Q F A A V Q F S T D C R T E F T F L D Y V K Q N K
1201 GAACCCGATGTTCTGCTAGGCAGCGTCGAGCCATGTTCTTGTGCTGACCAATACCTTTCTGTCGCATCAACTATGTGGTGGCACACGTTTCAAAGAAGAG
213 N P D V L L G S V Q P M F L L T N T F R A I N Y V V A H V F K E E
1301 TCTGGTCCAGGCCGATGCTACCAAGTGTGTCATCATTACAGACGGGAGGCAAGTATAAAGCAACATCAGTCCGCGCCACGACATAACCCGCT
247 S G A R P D A T K V L V I I T D G E A S D K G N I S A A H D I T R
1401 ACATCATCGGGATTGGCAAGCATTGTTGAGCGTACAAAAGCAAAGACGCTCCACATATTTGCTCAGAACCTGTAGAGGAATTTGTGAAGATTCTGGA
280 Y I I G I G K H F V S V Q K Q K T L H I F A S E P V E E F V K I L D
1501 CACTTTGAGAAGCTGAAGGATCTTTTTACTGACCTGCAGAGGAGGATTTATGCTATTGAGGGCACAACAGACAGGACCTGACATCTTTAATGAA
313 T F E K L K D L F T D L Q R R I Y A I E G T N R Q D L T S F N M E
1601 CTCTCCTCCAGCGGGATCAGCGCAGACTCAGCAAGGGCCATGCGATTGTGGGAGCTGTTGGGGCTAAGGATTGGGCCGGGGCTTTCTGGACCTGCGTG
347 L S S S G I S A D L S K G H A V V G A V G A K D W A G G F L D L R
1701 AAGACTGCAGGGTGCCACATTTGTTGGGAGGACCCGCTGACCTCAGATGTGAGAGGGGCTACCTGGGTTACACTGTGGCTGGATGACCTCCCGGAG
380 E D L Q G A T F V G Q E P L T S D V R G G Y L G Y T V A W M T S R S
1801 CTCAGACCCCTGCTGGCAGCAGGAGCCCGGTTACCAGCATGTGGGACAAGTACTGCTTTTCCAAGCCCCAGAGGCTGGAGGACGTTGGAACCAAACC
413 S R P L L A A G A P R Y Q H V G Q V L L F Q A P E A G G R W N Q T
1901 CAGAAGATAGAAGGGACTCAGATCGGATCTTACTTTGGTGGGAACTATGTAGTGTGACCTGGACCAAGATGGCGAGGCAGAGCTGTGCTGATTGGAG
447 Q K I E G T Q I G S Y F G G E L C S V D L D Q D G E A E L L L I G
2001 ACCCTGTTCTTTGGGGAGCAGAGGAGGCGGAGTGTCTACTTACCAGAGAAGCAGTCCGCTGTTTGAATGGTCTCAGAGCTACAGGGTGACCCTGG
480 A P L F F G E Q R G G R V F T Y Q R R Q S L F E M V S E L Q G D P G
2101 CTACCCGCTTGGTTCGTTTGGAGCCGCATAACTGCCCTGACGGACATCAATGGGATAGGCTGACTGATGTGGCTGTGGGAGCCCTTTGGAGGAGCAG
513 Y P L G R F G A A I T A L T D I N G D R L T D V A V G A P L E E Q
2201 GGTGCTGTGATACATCTTCAATGGGAAGCCTGGTGGGCTCAGTCCCCAGCCAAGCCAGCGTATAACAAGGAGCCAGGTGTTCCAGGAATCCGGTGGTTT
547 G A V Y I F N G K P G G L S P Q P S Q R I Q G A Q V F P G I R W F
2301 GCCGCTCCATCCATGGGGTGAAGGACCTTGGAGGGGACAGGCTGGCAGATGTGGTGTAGGAGCTGAGGGTCCGGTGGTGTGCTGAGCTCCAGGCCGGT
580 G R S I H G V K D L G G D R L A D V V V G A E G R V V V L S S R P V
2401 GGTGGATGTGGTCACTGAGCTGCTGTTCTCCAGAGGAAATCCAGTGCACGAGGTGGAGTGTCTCTACTCAGCCAGGGAGGAGCAGAAAACCGGAGTC
613 V D V V T E L S F S P E E I P V H E V E C S Y S A R E E Q K H G V
2501 AAGCTCAAGGCATGCTTCCGGATCAAGCCCTCACGCCACAGTTTCAAGTGCCTGCTTGCACACCTCAGCTACACCCTGACGCTGGATGGCCATCGGA
647 K L K A C F R I K P L T P Q F Q G R L L A N L S Y T L Q L D G H R
2601 TGAGGAGCCGAGGGTGTCCAGATGGAAGCCACGAGCTCAGTGGAAACACCTCCATCACCCAGATAAATCCTGCTTGGACTTCCACTTCCACTTCCC
680 M R S R G L F P D G S H E L S G N T S I T P D K S C L D F H F H F P
2701 GATCTGCATTCAAGACCTCATCTCCCTATCAATGTCTCCCTGAATTTCTCTTTTTGGAGGAAGAAGGAACCAAGGGACAAAAGGGCAGGGCCATG
713 I C I Q D L I S P I N V S L N F S L L E E E G T P R D Q K G R A M
2801 CAGCCTATCCTGAGACCTCAATCCACACAGTGACTAAGGAGATCCCTTTTGAAGAAGCTGGTGAAGATAAGAAGTGTGAGGCAAACTGACCCTGT
747 Q P I L R P S I H T V T K E I P F E K N C G E D K K C E A N L T L
2901 CATCCCTGCCAGATCTGGACCCTGCTGCTGATGTCTCTGCCAGCCTTGTGTGGAGTGGACACTGAGCAACTCAGGGGAAGATGCCTACTGGGTGCG
780 S S P A R S G P L R L M S S A S L A V E W T L S N S G E D A Y W V R
3001 ATTAGACCTGGACTTCCCTCGGGGACTCTCCTTCCGAAAAGTGGAGATGCTTACGCCACACAGCCGAATGCCTGTGAGCTGCGAGGAGCTCACCGAGGGG
813 L D L D F P R G L S F R K V E M L Q P H S R M P V S C E E L T E G
3101 TCAAGTCTCCTGACTAAGACACTGAAATGCAATGTAAGCTCTCCATCTTCAAAGCAGGCCAGGAGGTGAGCTCCAGGTGATGTTTAAACACGCTACTCA
847 S S L L T K T L K C N V S S P I F K A G Q E V S L Q V M F N T L L
3201 ACAGCTCCTGGGAAGACTTCGTGAGCTGAATGGCACTGTGCACTGTGAGAATGAGAAGTCAAGCCTCAGGAGGACAACCTCAGCCGCCACCCACATTC
880 N S S W E D F V E L N G T V H C E N E N S S L Q E D N S A A T H I P

3301 TGCCTGTACCCTGTCAACATCCTTACTAAGGAGCAGGAGAACTCCACCTCTATATCAGTTTACCCTAAAGGTCCCAAGACCAACAAGTCCAGCAT
913▶ V L Y P V N I L T K E Q E N S T L Y I S F T P K G P K T Q Q V Q H
3401 GTCTACCAGGTGAGGATTGAGCCATCTGCCTATGACCAACAATGCCACACTAGAGGCTTGGTTGGGGTGCCCCGGCTCACAGTGAGGACCTCATCA
947▶ V Y Q V R I Q P S A Y D H N M P T L E A L V G V P R P H S E D L I
3501 CATACACATGGAGTGTACAGCGGATCCCCTTGCTACTTGCCACAGCAGGACCTGAAGAGCCGTCACGCAAGCTGAGCCTGTCTGCCTGGAGTCCA
980▶ T Y T W S V Q T D P L V T C H S E D L K R P S S E A E P C L P G V Q
3601 GTTCCGCTGTCCAATTGTCTTACGGCGGAGATCCTCATCAAGTGACGGGGACCTGGAAGTCTCCAAGGAAATCAAGGCCTCCTCCACACTCAGCCTC
1013▶ F R C P I V F R R E I L I Q V T G T V E L S K E I K A S S T L S L
3701 TGCAGTCACTCTCCGCTCCTTCAACAGCAGCAAGCATTTCATTTGTATGGCAGCAAAGCCTCTGAGGCCAGGCTCCTCGTGAAGTTGACCTGATCC
1047▶ C S S L S V S F N S S K H F H L Y G S K A S E A Q V L V K V D L I
3801 ACGAGAAGGAGATGCTTACGTGTACGTGCTCAGCGCATTGGGGCCTCGTCTTCTGTTCTGATTTTCTGGCGCTCTACAAGTTGGCTTCTCAA
1080▶ H E K E M L H V Y V L S G I G G L V L L F L I F L A L Y K V G F F K
3901 ACGAACTGAAGGAGAAGATGGAGGCTGATGGAGTGTCCAAATGGAAGCCTCCAGAAGACACTGACCCTCTGGCAGTACCTGGGGAAGACCAAAA
1113▶ R N L K E K M E A D G G V P N G S P P E D T D P L A V P G E E T K

NheI (4081)

4001 GATATGGGCTGTCTAGAGCCCTCCGGGAGAGTGACAAGGACTAAGGCTAGGTCTGATACACTGACAGCCAGGAATAGAGCTAGCTGGCCAGACATGA
1147▶ D M G C L E P L R E S D K D •
4101 TAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT

4201 TATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAAGGCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATA

4301 TACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATA

4401 GGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTAGCT

4501 CTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTAGTAAAATATTAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATG

4601 TTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAATTTGGA

4701 CAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGC
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I L
4801 ACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCC
115▶ 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 H R
4901 TGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAAGCAGACAGTGCACCTGCAATGTAGGCCTC
82▶ 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 E
5001 AATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCG
49▶ I H V A S I I E G T K T R I A A G V H H K N D E Y L M T I K E T A
5101 ACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGAT
15▶ V E V L E L D Q Q S I N F T K M
5201 GATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGC

5301 CCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACT

5401 TGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGC

5501 CAAGTAGGAAAGTCCCATAAAGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGGCTACTTGGCATATGATACA

5601 CTTGATGTACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTAT

5701 TGACGTCAATGGGCGGGGCTGTTGGGCGTCAAGCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGACGTTAATTAAGAACATGTGAGCAAAAG

5801 GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTCTGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCA

5901 GAGGTGGCAAAACCCGACAGGACTATAAAGATACCAGGCTTTCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATAC

6001 CTGTCCGCTTTCTCCCTTCCGGGAAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCCTCAAGCTGGGCTGTG

6101 TGCACGAACCCCGTTACGCCCAGCGCTGCGCTTATCCGTAACATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGC

6201 CACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGT

6301 ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAACAAACACCCTGGTAGCGGTGTTTTTTGTTTGA

6401 AGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGAT

6501 TTTGGTCAAGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAAC

6601 TAACATACGCTCTCCATCAAAACAAACGAAACAAACAAACTAGCAAAATAGGCTGTCCCGAGTCAAGTGACGGTCCAGAACATTTCTCTATCGAA