



150

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

SphI (566)

MluI (558)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGCACGGCTCAGCATGCCTCTGTGTGGCGGAGAGCCATCTTGG
1▶ M P L C G R R A I L E
601 AGACAGCAAGGCAGATGGCACAGAAGCTCAGCCCTGGTGCCACAGGATGCTTATGGTGTACTGCATTGGCTGGGCTGAGGGCGGGAGCCCTGG
11▶ D S K A D G T E A Q P L V P T G C L M V L L H W P G P E G G E P W
701 GTCACCTCAGCCAGACATCTCTGACTGCAGAGGAGGTCTGCATCCACATCGCACAAAAGTGGCATCACTCCACCTGCTTGAATCTCTTCGCACTCT
45▶ V T F S Q T S L T A E E V C I H I A H K V G I T P P C L N L F A L
801 ACAATGCACAGGTAAGGTCTGGCTGCCCCAAACCATATTCTGGATACATCCAAGACATGAACCTCTATTTTGAATGAGGTTTTACTTCCGGAAGCTG
78▶ Y N A Q A K V W L P P N H I L D T S Q D M N L Y F R M R F Y F R N W
901 GCATGGCATGAATCCCGAGGAGCCAGCTGTATACCGGTGTGGTTTCCAGGGGACAGACTTCTCAGACCCGGGACAGCAAGGTGTACAGCTCTTGGAC
111▶ H G M N P Q E P A V Y R C G F P G A E T S S D R A E Q G V Q L L D
1001 TCTGCCTCGTTTGAATACCTTTTTGAGCAGGAAAGCATGAGTTCATGAACGATGTGGTGTCTCTGCGGGACCTGTCTAGCGAGGAGGAGATCCACCACT
145▶ S A S F E Y L F E Q G K H E F M N D V V S L R D L S S E E E I H H
1101 TTAAGAATGAGAGCTTAGGCATGGCCTTCTGCACCTGTGTCACCTTGTCTCAGCCGAGGCGTCCCTGGAGGAGATGGCCAGAGAGATCAGCTTCAA
178▶ F K N E S L G M A F L H L C H L A L S R G V P L E E M A R E I S F K
1201 GAAGTGCATCCCTCATTCTTCCGACAGCACATCCGGCAGCACAATGTGCTCACACGCTGCGGCTCCACAGAGTCTTCCGCCGCTTCTGCGGGCCTTC
211▶ N C I P H S F R Q H I R Q H N V L T R L R L R L H R V F R R F L R A F
1301 CGGCTGCGCCACTCTCCAGCAGGTTGTGATGGTGAAGTACTTGGCTACCCTGGAGCGGCTGGCTCCGCGCTTTGGCTCAGAGCGCATACCTGTGTGC
245▶ R P G H L S Q Q V V M V K Y L A T L E R L A P R F G S E R I P V C
1401 ATCTGGAGGCTCTGGCACAGCCGAGAGGGACCCCTGTACATCCAGAACAGTGGGCAGACCGCTGGGGATCCAGGCCAGAGCTGCCTTCCGGGCTCC
278▶ H L E V L A Q P E R D P C Y I Q N S G Q T A G D P G P E L P S G P P
1501 CACCCACGAGGACTGTTGACAGGCACCGGAGGTATCCAGTGGCATCCGCTGCAGACCCAGGAATCTGAGAGAGGTAACAGCAGAGGGAAATCCCCACGGC
311▶ T H E V L V T G T G G I Q W H P L Q T Q E S E R G N S R G N P H G
1601 AGCCGGTCCGGGAAGAAGCCCAAGCCCAAGGCTGGAGAGCAGCTGACAGAGCCCTCAGGAACACCTTGGACTACTTCTGTGACTTCCAGGACA
345▶ S R S G K K P K A P A G E H L T E S P Q E P P W T Y F C D F Q D
1701 TTTCCACGTGGTGCTAAAGGAGCGTCCGCTGCACATCCACCTCCAAGACAACAAGTCTTGTGCTGTGCCTGTTCCAGGCTGAGGCGCTGTCTT
378▶ I S H V V L K E R R V H I H L Q D N K C L L L C L C S Q A E A L S F
1801 TGTGGCCTGGTCGATGGCTATTTCCGCTTACTGCTGACTCCAGCCACTATCTGTGCCATGAGGTGGCACCCCGAGGCTGGTACTAGCATCCAGAAC
411▶ V A L V D G Y F R L T A D S S H Y L C H E V A P P R L V T S I Q N
1901 GGCATCCATGGGCCCCTGATGGATCCGTTTGTACAGGCCAAGCTGTGGCCAGAGGACGGCCTCTACCTGATTAGTGGAGCACCAGCCACCTGCACCGCC
445▶ G I H G P L M D P F V Q A K L W P E D G L Y L I Q W S T S H L H R
2001 TGATCCTTACCGTGGCCATCGAAACCCGGCTTTCAGTAAATGGCCCTCGGGCCTGCGCTGCGAAAGTTCATCAGCAGCAGCTGGAGCCTTTGT
478▶ L I L T V A H R N P A F S N G P R G L R L R K F P I T Q Q P G A F V
2101 GCTGGATGGCTGGGCGGCTCCTTTGCCAGCTTGGGGACCTTCCGCTCGCTTGCAGGGCTGCTCGTTGCGGGCCGGTGTACTGTTTCCCTGCAC
511▶ L D G W G R S F A S L G D L R L A L Q G C S L R A G D D C F P L H
2201 CACTGCTGCTGCCCCGGCAAGAGAAATCTCCAACCTCGTCATCATGCGGGGTCTAGGGCCACACCCGGCTCTCAACCTCAGTCAGCTCAGCTTCC
545▶ H C C L P R P R E I S N L V I M R G S R A H T R P L N L S Q L S F
2301 ACAGGGTTCACAGGATGAAATCACCCAGCTGTCCACTTGGGCCAAGGCACAAGGACCAATGTGTATGAGGGCCTTCAAGAGTGGGAGGCCCGCATGA
578▶ H R V H Q D E I T Q L S H L G Q G T R T N V Y E G L L R V G G P D E
2401 GGGCAAAGTGGACAATGGATGTCCCTGAGCCTGGTGGGACTAGTGGGCAGCAGCTTCGAGTGGTGTCAAAGTTCTCGACCCAGGCACCATGATATC
611▶ G K V D N G C P P E P G G T S G Q Q L R V V L K V L D P S H H D I
2501 GCCTTGGCCTTCTATGAGACTGTAGCCTCATGAGCCAGGTGTACACATGCACCTGGCTTCTCTGCATGGTGTGCGTGGCTCAGAGAATATCA
645▶ A L A F Y E T A S L M S Q V S H M H L A F L H G V C V R G S E N I
2601 TTGTGACAGAATTCGTAGAACATGGTCCCTGGATGTGTGGTTACGGCGACAGAGGGCCAAGTGCCCATGACCTGGAAGATGGTCTGGCTCAGCAGCT
678▶ I V T E F V E H G P L D V W L R R Q R G Q V P M T W K M V V A Q Q L
2701 GGCAGCGCCCTCAGTACCTGGAGGACAAGAATCTGGTTCACGGGAATGTATGTGGCCGGAACATCTGCTGGCTCGCTGGGGTTGGAGGAGGGTACC
711▶ A S A L S Y L E D K N L V H G N V C G R N I L L A R L G L E E G T
2801 AACCCCTCATCAAGTAAAGTATCTGGTGTGGGCGAGGCTCCCTCCAGGGAAGAGCGGGTGGAGCGCATCCCTGGACAGCTCCCGAGTCCGCTGT
745▶ N P F I K L S D P G V G Q G A L S R E E R V E R I P W T A P E C L
2901 CTGGAGGACCAGTAGCTTGGTACTGCCACGGACATGTGGGCTTGGTGGCACCTTCTTGTAGATCTGCTTGTGGGAGGCACCCCTGCAGGGTCCG
778▶ S G G T S S L G T A T D M W G F G A T L L E I C F D G E A P L Q G R
3001 TGTCCCTCTGAGAAAGAAGCGTTTACACAAAAGAAACATCAGCTGCCTGAACCTCATCCCGAGAGTGGCCACACTCACCCGCGAGTGCCTGACCTAT
811▶ G P S E K E R F Y T K K H Q L P E P S S P E L A T L T R Q C L T Y
3101 GAACCAGCACAGCGCCATCCTTCCGCACCATTTTGGCGGACCTCACAAAGGCTGCAGCCACAGAATCTAGTCGGCACTTCCGCTGTGAACTCAGACTCAC
845▶ E P A Q R P S F R T I L R D L T R L Q P Q N L V G T S A V N S D S

3201 CAGCATCAGACCCCACTGTTTTCCACAAGCGCTATTTGAAAAAGATCCGGGATTTGGGTGAGGGTCACTTTGGCAAGGTGAGCCTGTACTGCTACGACCC
878▶ P A S D P T V F H K R Y L K K I R D L G E G H F G K V S L Y C Y D P
3301 AACCAATGACGGCACTGGCGAGATGGTGGCCGTGAAGGCCCTGAAGGAAGGGTGGCGTCCCAAGCTCCGCTCAGGCTGGCAGCGGGAGATCGAGATCCTG
911▶ T N D G T G E M V A V K A L K E G C G P Q L R S G W Q R E I E I L
3401 CGGACGCTGTACCACGAACATATTGTCAAGTATAAAGGCTGCTGTGAGGACCAAGGAGAGAAGTCTGTACAGCTGGTCATGGAGTACGTCCCTCTGGGCA
945▶ R T L Y H E H I V K Y K G C C E D Q G E K S V Q L V M E Y V P L G
3501 GCCTCCGAGACTACCTGCCAAGGCACTGCGTAGGGCTGGCGCAGCTCCTGCTGTTGCCAGCAGATCTGCGAGGGCATGGCCCTACCTGCAGCTCAGCA
978▶ S L R D Y L P R H C V G L A Q L L L F A Q Q I C E G M A Y L H A Q H
3601 CTACATTCACCGAGACCTCGCCGCGCAACGTGCTGTGGACAACGACAGGCTGGTCAAGATTGGAGACTTTGGCCTAGCCAAGGCTGTACCTGAAGGC
1011▶ Y I H R D L A A R N V L L D N D R L V K I G D F G L A K A V P E G
3701 CACGAGTACTACCGAGTGCAGGACGGGACAGCCAGTGTCTGGTATGCCCCAGAATGCCTGAAGGAGTGCAAATTTTACTATGCATCTGATGTCT
1045▶ H E Y Y R V R E D G D S P V F W Y A P E C L K E C K F Y Y A S D V
3801 GGTCCCTCGGGTAACCTGTATGAGTTGTTGACATACTGTACTCTAACCCAGAGTCCCAATGAAATTCACCGAGCTCATCGCCATACCCAGGGCCA
1078▶ W S F G V T L Y E L L T Y C D S N Q S P H M K F T E L I G H T Q G G Q
3901 GATGACCGTGTGAGGCTCACAGAGCTGCTGGAACGAGGAGAGAGGCTGCCTCGGCCTGACAGGTGTCCCTGTGAGATCTATCACCTCATGAAGAAGTGC
1111▶ M T V L R L T E L L E R G E R L P R P D R C P C E I Y H L M K N C
4001 TGGGAGACAGAGGCTCCTTCGGGCCACCTCCAGAATCTGTGCCATCCTCCAGACAGCACAGGAGAAGTACCAAGGTCAGGTGCCTTCGTTCA
1145▶ W E T E A S F R P T F Q N L V P I L Q T A Q E K Y Q G Q V P S V F

XbaI (4131)

4101 GCGTGTGCTGACAGTCAAAGGCATCAGGGACTTAGACTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGT
1178▶ S V C •
4201 GAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAATTGCATTCATTT
4301 TATGTTTCAGGTTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTT
4401 AACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCCTC
4501 ACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTC
4601 TTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATAT
4701 CCCCAGTTTAGTAGTTGACTTAGGGAACAAAGAACCTTTAATAGAATTTGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGG
1411 • N R T Y K L
4801 GGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATG
134▶ 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 L E R C M
4901 CCACAGGGGCTGACCACCTGTGGATCTGTCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAG
100▶ 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 Q G N S V A
5001 CAGACCAATGGCAATGGCTTCAAGCAGACAGTGCACCTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGC
67▶ 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 K T R I A
5101 CGCCCCGACATGGTGTGTTGCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTC
34▶ 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 I N F T K
5201 ATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGA
0▶ M
5301 CGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGT
5401 CCCGTTGATTTACTAGTCAAACCAAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTA
5501 CTGCCAAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTAAGTACGCAAGTAGGAAAGTCCCATAGGTCATGTAAGTGGCATAATGCCAG
5601 GCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCC
5701 ATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTGCTTGGGCGGTACGCCAGCGGGCC
5801 ATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGT
5901 TTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGACTATAAAGTACAGGCGTTTCC
6001 CCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGC
6101 TCACGCTGATAGTATCTCAGTTCGGTGTAGGCTGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTA
6201 ACTATCGTCTTGAAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC
6301 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT
6401 AGCTCTTGATCCGGCAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTT
6501 TGATCTTTTACGGGTCTGACGCTCAGTGGAACTCACGTTAAGGATTTTGGTATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAAT

6601 AAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAG
6701 CAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA