



1 AATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAAC
101 TAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAAGGATCGCGATCGTGAATTAGTTTCACTTTCCAGTTTCAGTT

SgfI (168)

201 TCCAGTTTCATTTCCAGTTTCATTTCCAGTTTCATTTCTGATATCCTGCAGGagcttgaataaaatgaatattagaagctgtagaataagagaaa
301 atgacagaggaAAACTGAAAGGgAGAACTGAAAGTGggaattcctctgaggcagaaggaccatccctTATAAAtagcacaggccatgaaggaagatca

SdaI (254)
EcoRV (246)

401 ttctcactgcagcctttgacagcctttgctcatcttgAAGCTTCTGCCTTCTCCCTCTGTGAGTTTGGTTGGTGTACAGTAGCTTCCACCATGGAGGA
1 M E D

HindIII (439)

NcoI (491)

501 TGCCAAGAATATTAAGAAAGGCCCTGCCCATTTACCTCTGGAAGATGGCACTGCTGGTGGAGCACTGCACAAGGCCATGAAGAGGTATGCCCTGGTC
3 A K N I K K G P A P F Y P L E D G T A G E Q L H K A M K R Y A L V
601 CCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGCTGTGAGGCTGGCAGAAGCCATGAAAAGAT
37 P G T I A F T D A H I E V D I T Y A E Y F E M S V R L A E A M K R
701 ATGGACTGAACACCAACCACAGGATTGGTGTGCTCTGAGAAGTCTCTCCAGTTCTTCATGCTGTGTAGGAGCCCTGTTCACTGGAGTGGCTGTGGC
70 Y G L N T N H R I V V C S E N S L Q F F M P V L G A L F I G V A V A

SacI (833)

801 CCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAAGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATC
103 P A N D I Y N E R E L L N S M G I S Q P T V V F V S K K G L Q K I
901 CTGAATGTGAGAAGAAGCTGCCCATCATCCAGAAGATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCA
137 L N V Q K K L P I I Q K I I M D S K T D Y Q G F Q S M Y T F V T
1001 GCCACTTACCCCTGGCTTCAATGAGTATGACTTTGTGCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGG
170 S H L P P G F N E Y D F V P E S F D R D K T I A L I M N S S G S T G
1101 ACTGCCAAAGGTGTGGCTCTGCCCCACAGAAGTGTGTGAGATTGAGCCATGCCAGAGACCCCATCTTTGGCAACCAGATCATCCCTGACACTGCC
203 L P K G V A L P H R T A C V R F S H A R D P I F G N Q I I P D T A

Acc65I (1253)

1201 ATCCTGTCTGTGGTTCCATTCCATCATGGCTTTGGCATGTTCAACAACACTGGGGTACCTGATCTGTGGCTTCCAGAGTGGTGTGATGTATAGGTTTGAGG
237 I L S V V P F H H G F G M F T T L G Y L I C G F R V V L M Y R F E
1301 AGGAGCTGTTTCTGAGGAGCCTACAAGACTACAAGATCCAGTCTGCCCTGCTGGTGGCCACTCTGTTCAAGTCTTTGCAAGAGCACCTCATTGACAA
270 E E L F L R S L Q D Y K I Q S A L L V P T L F S F F A K S T L I D K
1401 GTATGACCTGAGCAACCTGCATGAGATTGCCTCTGGAGGAGCACCCCTGAGCAAGGAGGTGGGTGAGGCTGTGGCAAGAGGTTCCATCTCCAGGAATC
303 Y D L S N L H E I A S G G A P L S K E V G E A V A K R F H L P G I
1501 AGACAGGCTATGGCCTGACTGAGACCCTCTGCCATCCTACCCTGAAGGAGATGACAAGCCTGGTGTGGGCAAGGTGGTTCCTTTTTTTG
337 R Q G Y G L T E T T S A I L I T P E G D D K P G A V G K V V P F F
1601 AGGCCAAGGTGGTGGACCTGGACACTGGCAAGACCCTGGGAGTGAACCAGAGGGGTGAGCTGTGTGTGAGGGTCCCATGATCATGTCTGGCTATGTGAA
370 E A K V V D L D T G K T L G V N Q R G E L C V R G P M I M S G Y V N
1701 CAACCCTGAGGCCACCAATGCCCTGATTGACAAGGATGGCTGGCTGCACTCTGGTGAATTGCCTACTGGGATGAGGATGAGCACTTTTTCATTGTGGAC
403 N P E A T N A L I D K D G W L H S G D I A Y W D E D E H F F I V D
1801 AGGCTGAAGAGCCTCATCAAGTACAAGGCTACCAAGTGGCACCTGCTGAGCTAGAGAGCATCCTGCTCCAGCACCCCAACATCTTTGATGCTGGTGTGG
437 R L K S L I K Y K G Y Q V A P A E L E S I L L Q H P N I F D A G V
1901 CTGGCCTGCTGATGATGCTGGAGAGCTGCCTGCTGTTGTGGTCTGGAGCATGAAAAGACCATGACTGAGAAGGAGATTGTGGACTATGTGGC
470 A G L P D D D A G E L P A A V V V L E H G K T M T E K E I V D Y V A

BstEII (2007)

2001 CAGTCAGGTGACCACTGCAAGAAGCTGAGGGGAGGTGTGGTGTGTTGTGGATGAGGTGCCAAAGGCTGACTGGCAAGCTGGATGCCAGAAAGATCAGA
503 S Q V T T A K K L R G G V V F V D E V P K G L T G K L D A R K I R

NheI (2152)

2101 GAGATCCTGATCAAGGCCAAGAAGGTTGGCAAATTGCTGTGTAACCTGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCA
537 E I L I K A K K G G K I A V

HpaI (2292)

2201 CAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAAACAACA

EcoRI (2386)

2301 CAATTGCATTCATTTTATGTTTCAGGTTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAAAATACA
2401 GCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTA
2501 GCTGTTTGCAGCCTCACCTTCTTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGA
2601 CCTCCACATTCCCTTTTATGTAATAATTTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAA
2701 GGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTTGGACAGCAAGAAGCGAGCTTCTAGCTTATCCT

2801 CAGTCCTGCTCCTCTGCCACAAAGTGCACGCAGTTGCCGGCCGGTTCGCGCAGGGCGAACTCCCGCCCCACGGTGTCTGCCGATCTCGGTCATGGCCG
124◀• D Q E E A V F H V C N G A P D R L A F E R G W P Q E G I E T M A P
2901 GCCCGGAGGCGTCCCGGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGCTCGTCCAGGCCGCGCACCCACACCCAGGCCAGGGTGTGTCCGG
91◀ G S A D R F N T S V V E S W E A Y L E D L G R V W V W A L T N D P
SgrAI (3055)
3001 CACCACCTGGTCTGTGACCGCGCTGATGAACAGGGTCACGTCGTCGCCGACCAACCCGGCGAAGTCGTCCTCCACGAAGTCCCGGGAGAACCAGCCGAGCCGG
58◀ V V Q D Q V A S I F L T V D D R V V G A F D D E V F D R S F G L R
3101 TCGGTCCAGAACTCGACCGTCCGGCGACGTCGCGCGGTGAGCACCAGGACGGCACTGGTCAACTTGGCCATGATGGCCCTCTATAGTGAGTCGTAT
24◀ D T W F E V A G A V D R A T L V P V A S T L K A M ←
AseI (3230)
3201 TATACTATGCCGATATACTATGCCGATGATTAATTGTCAACTACTGTTTGTAGGCGCCGGTCACAGCTTGATCTGTAACGGCGCAGAACAGAAAACGAA
ACAAAAGACGTAGAGTTGAGCAAGCAGGGTCAGGCAAAGCGTGAGAGCCGGCTGAGTCTAGGTAGGCTCAAGGGAGCGCCGGACAAAAGGCCCGGTCTCG
ACCTGAGCTTTAAACTTACCTAGACGGCGGACGCAGTTCAGGAGGCCACACAGCGGGAGGCGCAGAACCGCACTCAACCGCGTGGATGGCGGCCCTCA
HindIII (3543) Psp1406I (3586)
3501 GGTAGGGCGGGCGCGCTGAAGGAGAGATGCGAGCCCTCGAAGCTTCAGCTGTGTCTGCGCGCAAACCCGTTGCGAAAAAGAAGTTCACGGCGACT
Agel (3699)
3601 ACTGCACTTATATACGGTCTCCCCACCCTCGGAAAAAGCGGAGCCAGTACACGACATCACTTTCCAGTTTACCCCGGCCACCTTCTCTAGGCAC
BspLU11I (3770)
3701 CGGTTCAATTGCCGACCCTCCCCCAACTTCTCGGGACTGTGGCGATGTGCGCTCTGCCACTGACACATGTGAGCAAAGGCCAGCAAAGGCCAG
GAACCGTAAAAAGGCCGCTTGTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCG
ACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCC
CTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGT
TCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATT
AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGTGAAGCAGCAGATTACGG
CAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGT
NotI (4520)
4501 TAATTAACATTTAAATCAGCGGCCGC