



1 AATAAAATATCTTTATTTTCATTACATCTGTGTGTTGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAAC
SgfI (168)
101 TAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAAGGATGCGGATCGTGAATTAGTTTCACTTTCCAGTTTCAGTT

SdaI (254)
EcoRV (246)
201 TCCAGTTTCATTTTCCAGTTTCATTTTCCAGTTTCATTTTCTGATATCCTGCAGGagcttgaataaaatgaatattagaagctgtagaataagagaaa

301 atgacagaggaAAACTGAAAGGgAGAACTGAAAGTGggaattcctctgaggcagaaggaccatccctTATAAAtagcacaggccatgaaggaagatca

HindIII (439) NcoI (491)
401 ttctcactgcagcctttgacagcctttgctcatcttgAAGCTTCTGCCTTCTCCCTCCTGTGAGTTTGGTTGGTGTACAGTAGTCCACCATGGCTTC
1 M A S
501 CAAGGTGTACGACCCCGAGCAACGCAACGCATGATCACTGGGCTCAGTGGTGGGCTCGCTGCAAGCAAATGAACGTGTGGACTCCTTCATCAACTAC
3 K V Y D P E Q R K R M I T G P Q W W A R C K Q M N V L D S F I N Y
601 TATGATCCGAGAAGCACGCCGAGAACGCGTATTCTTCTGATGGTAACGCTGCCTCCAGCTACCTGTGGAGGCACGTCGTGCCTCACATCGAGCCCG
37 Y D S E K H A E N A V I F L H G N A A S S Y L W R H V V P H I E P
701 TGGCTAGATGCATCATCCTGATCTGATCGGAATGGTAAAGTCGGCAAGAGCGGGAATGGCTCATATCGCCTCCTGGATCACTACAAGTACCTCACCGC
70 V A R C I I P D L I G M G K S G K S G N G S Y R L L D H Y K Y L T A
801 TTGGTTCGAGCTGCTGAACCTTCAAAGAAAATCATCTTTGTGGGCCAGACTGGGGGGCTTGTCTGGCTTTCCTACTCTACGAGCACCAGACAAG
103 W F E L L N L P K K I I F V G H D W G A C L A F H Y S Y E H Q D K
EcoRV (978)
901 ATCAAGGCCATCGTCCATGCTGAGAGTGTCTGGAGCTGATCGAGTCTGGGACGAGTGGCCTGACATCGAGGAGGATATCGCCCTGATCAAGAGCGAAG
137 I K A I V H A E S V V D V I E S W D E W P D I E E D I A L I K S E
1001 AGGGCGAGAAAATGGTGTGAGAATAACTTCTCGTCGAGACCATGCTCCAAGCAAGATCATGCGGAACTGGAGCCTGAGGAGTTCGCTGCCTACCT
170 E G E K M V L E N N F F V E T M L P S K I M R K L E P E E F A A Y L
1101 GGAGCCATTCAAGGAGAAGGGCGAGGTTAGACGGCCTACCCTCTCCTGGCCTCGCGAGATCCCTCTCGTTAAGGGAGGCAAGCCCGACGTCGTCCAGATT
203 E P F K E K G E V R R P T L S W P R E I P L V K G G K P D V V Q I
1201 GTCCGAACTACAACGCCTACCTCGGGCCAGCGACGATCGCTAAGATGTTTCATCGAGTCCGACCTGGGTCTTTTCCAACGCATTGTCGAGGGAG
237 V R N Y N A Y L R A S D D L P K M F I E S D P G F F S N A I V E G
1301 CTAAGAAGTTCCTAACACCGAGTTCGTGAAGGTGAAGGGCTCCACTTCAGCCAGGAGGACGCTCCAGATGAAATGGGTAAGTACATCAAGAGCTTCGT
270 A K K F P N T E F V K V K G L H F S Q E D A P D E M G K Y I K S F V
1401 GGAGCGCTGCTGAAGAACGAGCAGTAATTCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACTAGAATGCAGTGAAAAA
303 E R V L K N E Q •

HpaI (1570)
1501 ATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTATGTTT

EcoRI (1664)
1601 CAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCATAAATACAGCATAGCAAACTTTAACCTCC
1701 AAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCT
1801 TTCATGGAGTTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTGCTCTTCAATTTCTTATGTTTTAAATGCACTGACTCCCACATTCCCTTTTATG
1901 AAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAG
2001 TTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTATCCTCAGTCTGCTCCTCTGCCACAA
125 D Q E E A V F
2101 AGTGCACGAGTTGCCGGCCGGTGCAGGCGAAGTCCCGCCCCACGGTCTGCTCGCGATCTCGGTTCATGGCCGGCCGGAGGCGTCCCGAAGTT
117 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 G S A D R F N
2201 CGTGGACACGACTCCGACCACTCGGCGTACAGCTCGTCCAGGCGCGCACCCACACCCAGGCCAGGGTGTGTCGCGCACCACTGGTCTGGACCGCG
84 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 V V Q D Q V A

SgrAI (2333)
2301 CTGATGAACAGGGTCACGTCGTCCCGGACCACACCGCGAAGTCTGCTCCACGAAGTCCCGGAGAACCAGCCGGTCCGAGTCCAGAACTCGACCGCTC
50 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 D T W F E V A G
2401 CGGCGACTGCGCGCGGTGAGCACCGGAACGGCACTGGTCAACTTGGCCATGATGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATG
17 A V D R A T L V P V A S T L K A M

AseI (2508)
2501 CCGATGATTAATTGTCAACTACTGTTTGTAGGCCCGGTCCACAGCTTGATCTGTAACGGCGCAGAACAGAAAACGAAACAAAGACGTAGAGTTGAGCAA
2601 GCAGGGTCAGGCAAAGCGTGGAGAGCCGGTCTGAGTCTAGGTAGGCTCCAAGGGAGCGCCGACAAAGGCCCGGTCTCGACCTGAGCTTTAAACTTACCTA
2701 GACGGCGGACGAGTTTCAAGAGGCACCAAGCGGGAGCGGCGAGAACCGACTCAACCGGCTGGATGGCGGCTCAGGTAGGGCGGCGGGCGGTGAA

2801 GGAGAGATGCGAGCCCCTCGAAGCTTCAGCTGTGTTCTGGCGCAAACCCGTTGCGAAAAAGAACGTTACGCGGACTACTGCACTTATATACGGTCTC
HindIII (2821) Psp1406I (2864)

2901 CCCCACCCTCGGGAAAAAGGCGGAGCCAGTACACGACATCACTTTCCAGTTTACCCCGCGCCACCTTCTCTAGGCACCGTTCAATTGCCGACCCCTCC
AgeI (2977)

3001 CCCCAACTTCTCGGGGACTGTGGGCGATGTGCGCTCTGCCACTGACACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTG
BspLU11I (3048)

3101 CTGGCGTTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG

3201 CGTTTCCCCTGGAAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTC

3301 TCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCCTTA

3401 TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG

3501 GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAG

3601 AGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAA
NotI (3798)

3701 GATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGG

3801 CCGC