



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTGCTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

SphI (560)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAGCATGCACAGGAGGAGAAGCAGGAGCTGTCGGGAAGATCA
13> M H R R R S R S C R E D Q
601 GAAGCCAGTCATGGATGACCAGCGGACCTTATCTCCAACAATGAGCAACTGCCATGCTGGGCGGCGCCCTGGGGCCCCGAGAGCAAGTCAGCCGC
701 GGAGCCCTGTACACAGGCTTTCCATCCTGGTACTCTGCTCCTCGCTGGCCAGGCCACCACGCTACTTCTGTACAGCAGCAGGGCCGCTGGACA
801 AACTGACAGTCACCTCCAGAACCTGCAGCTGGAGAACCTGCGCATGAAGCTTCCCAAGCCTCCCAAGCCTGTGAGCAAGATGCGCATGGCCACCCGCT
901 GCTGATGCAGGCGCTGCCATGGGAGCCCTGCCAGGGGCCATGAGAATGCCACCAAGTATGGCAACATGACAGAGGACCATGTGATGCACCTGCTC
1001 CAGAATGCTGACCCCTGAAGGTGTACCCGCACTGAAGGGGAGCTTCCGGAGAACCTGAGACACCTTAAGAACCACCATGGAGACCATAGACTGGAAGG
1101 TCTTTGAGAGCTGGATGCACCATTGGCTCCTGTTTGAATGAGCAGGCACTCCTTGGAGCAAAAGCCACTGACGCTCCACCGAAAGTACTGACCAAGTG
1201 CCAGGAAGAGGTCAGCCACATCCTGCTGTCCACCCAGGTTTATTGAGCCCAAGTGCAGCAGAACGCAACTATCTGCCACTCCAGTGTATGGGAGC
1301 ATCGGCTACTGCTGGTGTGCTTCCCAACGGCAGGAGTCCCAACACCAGAAGCCGCGGGCACCATAACTGCAGTGAAGTACTGGAAGTGGAGGACC
1401 CGTCTTCTGGGCTGGGTGTGACCAAGCAGGATCTGGGCCAGTCCCCATGTGAGAGCAGCAGAGGCGGTCTTCAACATCCTGCTAGCTGGCCAGACATGA

NheI (1481)

1401 280> P S S G L G V T K Q D L G P V P M •
1501 TAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT
1601 TATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTATTTTATGTTTCAGGTTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTC
1701 TACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATA
1801 GGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGTGTGAACTAGCT
1901 CTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTGTAGTAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAATAAATG
2001 TTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTGTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGA
2101 CAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGC
2201 ACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCC
2301 TGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAAGCAGACAGTGAACCTGCAATGTAGGCCCTC
2401 AATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCCGACATGGTGCTTGTGCTCCTATAGAGCATGGTATCTTCTCAGTGGCG
2501 ACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAAGTCTTATATACTATGCCGATATACTATGCCGAT
2601 GATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCTCAATAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGC
2701 CCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACT
2801 TGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGC
2901 CAAGTAGGAAAGTCCATAAGGTCATGTAAGTCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGGCTACTTGGCATATGATACA
3001 CTTGATGTAAGTCCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTAT
3101 TGACGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAAGTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAG

3201 GCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCA
3301 GAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
3401 CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTG
3501 TGCACGAACCCCCGTTCCAGCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGC
3601 CACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGT
3701 ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGCA
3801 AGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACAGTTAAGGGAT
3901 TTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAAC
4001 TAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA