

1 ATGGCTTCAGAAAATATGACGCCGAGGATTACATAGGACACCACCTGAATAACCTTCAGCTGGACCTGCGTACATTC
1 MetAl aSer Gl uAsnMetThr P roGl nAspTyrI l eGl yHi sHi sLeuAsnAsnLeuGl nLeuAspLeuArgThr Phe

BamHI

PfIMI

SspI

79 TCGCTGGTGGATCCACAAAACCCCCAGCCACCTTCTGGACAATCAATATTGACTCCATGTTCTTCTCGGTGGTGCTG
27 Ser LeuVal AspP roGl nAsnP roP roAl aThr PheTrpThr I l eAsnI l eAspSer MetPhePheSer Val Val Leu
157 GGTCTGTTGTTCCCTGGTTTTATTCCGTAGCGTAGCCAAAAAGGCGACCAGCGGTGTGCCAGGTAAGTTTCAGACCCGG
53 Gl yLeuLeuPheLeuVal LeuPheArgSer Val Al aLysLysAl aThr Ser Gl yVal P roGl yLysPheGl nThr Al a
235 ATTGAGCTGGTATCGGCTTTGTTAATGGTAGCGTGAAAGACATGTACCATGGCAAAGCAAGCTGATTGCTCCGCTG
79 I l eGl uLeuVal I l eGl yPheVal AsnGl ySer Val LysAspMetTyrHi sGl yLysSer LysLeuI l eAl aP roLeu
313 GCCCTGACGATCTTCGTCTGGGTATTCCCTGATGAACCTGATGGATTTACTGCCTATCGACCTGCTGCCGTACATTGCT
105 Al aLeuThr I l ePheVal TrpVal PheLeuMetAsnLeuMetAspLeuLeuP roI l eAspLeuLeuP roTyrI l eAl a
391 GAACATGTACTGGGTCTGCCTGCACTGCGTGTGGTTCCTGCTGCGGACGTGAACGTAACGCTGTCTATGGCACTGGGC
131 Gl uHi sVal LeuGl yLeuP roAl aLeuArgVal Val P roSer Al aAspVal AsnVal Thr LeuSer MetAl aLeuGl y
469 GTATTTATCCTGATTCTGTTCTACAGCATCAAAATGAAAGGCATCGGCGGCTTCACGAAAGAGTTGACGCTGCAGCCG
157 Val PheI l eLeuI l eLeuPheTyrSer I l eLysMetLysGl yI l eGl yGl yPheThr LysGl uLeuThr LeuGl nP ro
547 TTCAATCACTGGGCGTTCATTCCCTGTCAACTTAATCCTTGAAGGGTAAGCCTGCTGTCCAAACCAGTTTCACTCGGT
183 PheAsnHi sTrpAl aPheI l eP roVal AsnLeuI l eLeuGl uGl yVal Ser LeuLeuSer LysP roVal Ser LeuGl y
625 TTGCGACTGTTCCGTAACATGTATGCCGCTGAGCTGATTTTCACTTCTGATTGCTGGTCTGTTGCCGTGGTGGTCACAG
209 LeuArgLeuPheGl yAsnMetTyrAl aGl yGl uLeuI l ePheI l eLeuI l eAl aGl yLeuLeuP roTrpTrpSer Gl n

BamHI

703 TGGATCCTGAATGTGCCGTGGGCCATTTTCCACATCCTGATCATTACGCTGCAAGCCTTCATCTTCATGGTTCTGACG
235 TrpI l eLeuAsnVal P roTrpAl aI l ePheHi sI l eLeuI l eI l eThr LeuGl nAl aPheI l ePheMetVal LeuThr
781 ATCGTCTATCTGTCGATGGCGTCTGAAGAACATTAATTTACCAACTACTACGTTTTAACTGAAACAACTGGAGAC
261 I l eVal TyrLeuSer MetAl aSer Gl uGl uHi s

BsrGI

859 TGTCATGGAAAACCTGAATATGGATCTGCTGTACATGGCTGCCGCTGTGATGATGGGTCTGGCGGCAATCGGTGCTGC
1 MetGl uAsnLeuAsnMetAspLeuLeuTyrMetAl aAl aAl aVal MetMetGl yLeuAl aAl aI l eGl yAl aAl
937 GATCGGTATCGGCATCCTCGGGGGTAAATTCCTGGAAGGCGCAGCGCGTCAACCTGATCTGATTCCTCTGCTGCGTAC
25 al l eGl yI l eGl yI l eLeuGl yGl yLysPheLeuGl uGl yAl aAl aArgGl nP roAspLeuI l eP roLeuLeuArgTh
1015 TCAGTTCTTTATCGTTATGGGTCTGGTGGATGCTATCCCATGATCGCTGTAGGTCTGGGTCTGTACGTGATGTTCCG
51 r Gl nPhePheI l eVal MetGl yLeuVal AspAl aI l eP roMetI l eAl aVal Gl yLeuGl yLeuTyrVal MetPheAl
1093 TGTCGCGTAGTAAGCGTTGCTTTTATTTAAAGAGCAATATCAGAACGTTAACTAAATAGAGGCATTGTGCTGTGAATC
77 aVal Al a

EcoNI

1171 TTAACGCAACAATCCTCGGCCAGGCCATCGCGTTTGTCTGTTCTGTTCTGTCATGAAGTACGTATGGCCGCCAT
1 MetLysTyrVal TrpP roP roL

PpuMI

EcoO109I

1249 TAATGGCAGCCATCGAAAAACGTCAAAAAGAAATTGCTGACGGCCTTGCTTCCGCAGAACGAGCACATAAGGACCTTG
8 euMetAl aAl aI l eGl uLysArgGl nLysGl uI l eAl aAspGl yLeuAl aSer Al aGl uArgAl aHi sLysAspLeuA

EcoNI

1327 ACCTTGCAAAGGCCAGCGCACCAGCTGAAAAAAGCGAAAGCGGAAGCCCAGGTAATCATCGAGCAGGCCAACA
34 spLeuAl aLysAl aSer Al aThr AspGl nLeuLysLysAl aLysAl aGl uAl aGl nVal I l eI l eGl uGl nAl aAsnL
1405 AACGCCGCTCGCAGATTCTGGACGAAGCGAAAGCTGAGGCAGAACGGAACGTACTAAAATCGTGGCCCAGGCCGAGG
60 ysArgArgSer Gl nI l eLeuAspGl uAl aLysAl aGl uAl aGl uGl nGl uArgThr LysI l eVal Al aGl nAl aGl nA
1483 CGGAAATTGAAGCCGAGCGTAAACGTGCCCCGTGAAGAGCTGCGTAAGCAAGTTGCTATCCTGGCTGTTGCTGGCGCCG
86 I aGl uI l eGl uAl aGl uArgLysArgAl aArgGl uGl uLeuArgLysGl nVal Al aI l eLeuAl aVal Al aGl yAl aG
1561 AGAAGATCATCGAACGTTCCGTGGATGAAGCTGCTAACAGCGACATCGTGGATAAACTTGTGCTGAACTGTAAGGAG
112 I uLysI l eI l eGl uArgSer Val AspGl uAl aAl aAsnSerAspI l eVal AspLysLeuVal Al aGl uLeu
1639 GGAGGGCTGATGTCTGAATTTATTACGGTAGCTCGCCCCTACGCCAAAGCAGCTTTTGACTTTGCCGTGCAACACCA
1 MetSer Gl uPheI l eThr Val Al aArgP roTyrAl aLysAl aAl aPheAspPheAl aVal Gl uHi sGl
1717 AAGTGTAGAACGCTGGCAGGACATGCTGGCGTTTGGCCCGAGGTAACCAAAAACGAACAAATGGCAGAGCTTCTCTC
23 nSer Val Gl uArgTrpGl nAspMetLeuAl aPheAl aAl aGl uVal Thr LysAsnGl uGl nMetAl aGl uLeuLeuSe

BssHII

1795 TGGCGCGCTTGCGCCAGAAACGCTCGCCGAGTCGTTTATCGCAGTTTGTGGTGAGCAACTGGACGAAAACGGTCAGAA
49 r Gl vAl aLeuAl aP roGl uThr LeuAl aGl uSer PheI l eAl aVal CvsGl vGl uGl nLeuAspGl uAsnGl vGl nAs

1873 CCTGATTCGGGTTATGGCTGAAAATGGTCGTCTTAACGCGCTCCCGGATGTTCTGGAGCAGTTTATTCACCTGCGTGC
75 nLeu l l eArgValMetAl aGl uAsnGl yA rgLeuAsnAl aLeuP roAspVal l eLeuGl uGl nPhe l l eHi sLeuArgAl

NruI

1951 CGTGAGTGAGGCTACCGCTGAGGTAGACGTCATTTCCGCTGCCGCACTGAGTGAACAACAGCTCGCGAAAATTTCTGC
101 aVal Ser Gl uAl aThr Al aGl uVal AspVal l l eSer Al aAl aAl aLeuSer Gl uGl nGl nLeuAl aLys l l eSer Al

2029 TGCGATGGAAAAACGTCTGTACGCAAAGTTAAGCTGAATTGCAAAATCGATAAGTCTGTAATGGCAGGCGTTATCAT
127 aAl aMetGl uLysArgLeuSer ArgLysVal LysLeuAsnCysLys l l eAspLysSer Val MetAl aGl yVal l l e l l

AflII

2107 CCGAGCGGGTGATATGGTCATTGATGGCAGCGTACGCGGTCGTCTTGAGCGCCTTGACAGACGCTTGCAGTCTTAAGG
153 eArgAl aGl yAspMetVal l l eAspGl ySer Val l ArgGl yA rgLeuGl uArgLeuAl aAspVal l eLeuGl nSer

SphI

2185 GGACTGGAGCATGCAACTGAATTCACCGAAAATCAGCGAACTGATCAAGCAGCGCATTGCTCAGTTCAATGTTGTGAG
1 MetGl nLeuAsnSer Thr Gl u l l eSer Gl uLeu l l eLysGl nArg l l eAl aGl nPheAsnVal Val l eSe

2263 TGAAGCTCACAACGAAGTACTATTGTTTTCTGTAAGTGACGGTGTTATCCGCATTACCGCCTGGCCGATTGTATGCA
23 rGl uAl aHi sAsnGl uGl yThr l l eVal Ser Val SerAspGl yVal l l eArg l l eHi sGl yLeuAl aAspCysMetGl

XhoI

2341 GGGTGAATGATCTCCCTGCCGGTAACCGTTACGCTATCGCACTGAACCTCGAGCGCGACTCTGTAGGTGCGGTTGT
49 nGl yGl uMet l l eSer LeuP roGl yAsnArgTyrAl a l l eAl aLeuAsnLeuGl uArgAspSer Val Gl yAl aVal Va

2419 TATGGGTCCGTACGCTGACCTTGCCGAAGGCATGAAAGTTAAGTGTACTGGCCGTATCCTGGAAGTCCGGTTGGCCG
75 l MetGl yProTyrAl aAspLeuAl aGl uGl yMe tLysVal LysCysThr Gl yA rg l l eLeuGl uVal P roVal Gl yA r

SfiI

2497 TGGCCTGCTGGGCCGTGTGGTTAACACTCTGGGTGCACCAATCGACGGTAAAGTCCGCTGGATCACGACGGCTTCTC
101 gGl yLeuLeuGl yA rgVal Val l AsnThr LeuGl yAl aP ro l l eAspGl yLysGl yProLeuAspHi sAspGl yPheSe

2575 TGCTGTAGAAGCAATCGCTCCGGGCGTTATCGAACGTCAGTCCGTAGATCAGCCGGTACAGACCGGTTATAAAGCCGT
127 rAl aVal Gl uAl a l l eAl aP roGl yVal l l eGl uArgGl nSer Val l AspGl nP roVal l Gl nThr Gl yTyrLysAl aVa

2653 TGACTCCATGATCCCAATCGGTCTGGTCAGCGTGAATTGATCATCGGTGACCGTCAGACAGGTAAAACCGCACTGGC
153 l AspSer Met l l eP ro l l eGl yA rgGl yGl nArgGl uLeu l l e l l eGl yAspArgGl nThr Gl yLysThr Al aLeuAl

2731 TATCGATGCCATCATCAACCAGCGGATTCCGGTATCAAATGTATCTATGTCGCTATCGGCCAGAAAGCGTCCACCAT
179 a l l eAspAl a l l e l l eAsnGl nArgAspSer Gl y l l eLysCys l l eTyrVal l Al a l l eGl yGl nLysAl aSer Thr l l

2809 TTCTAACGTGGTACGTAAACTGGAAGAGCACGGCGCACTGGCTAACACCATCGTTGTGGTAGCAACCGCGTCTGAATC
205 eSerAsnVal Val l ArgLysLeuGl uGl uHi sGl yAl aLeuAl aAsnThr l l eVal Val Val l Al aThr Al aSer Gl uSe

2887 CGCTGCACTGCAATACCTGGCACGTATGCCGTTGCGCTAATGGGCGAATACTTCCGTGACCGCGGTGAAGATGCGCT
231 rAl aAl aLeuGl nTyrLeuAl aArgMetP roVal l Al aLeuMetGl yGl uTyrPheArgAspArgGl yGl uAspAl aLe

BglII

2965 GATCATTTACGATGACCTGTCTAAACAGGCTGTTGCTTACCCTCAGATCTCCCTGCTGCTCCGTCGTCGCCAGGACG
257 u l l e l l eTyrAspAspLeuSer LysGl nAl aVal Al aTyrArgGl n l l eSer LeuLeuLeuArgArgP roP roGl yA r

SmaI

PmlI

3043 TGAAGCATTCCCGGGCGACGTTTTCTACCTCCACTCTCGTCTGCTGGAGCGTGCTGCACGTGTTAACGCCGAATACGT
283 gGl uAl aPheP roGl yAspVal PheTyrLeuHi sSer ArgLeuLeuGl uArgAl aAl aArgVal l AsnAl aGl uTyrVa

3121 TGAAGCCTTACCAAAGGTGAAGTAAAAGGGAAAACCGGTTCTCTGACCGCACTGCCGATTATCGAAACTCAGGCGGG
309 l Gl uAl aPheThr LysGl yGl uVal LysGl yLysThr Gl ySer LeuThr Al aLeuP ro l l e l l eGl uThr Gl nAl aGl

BglIII

3199 TGACGTTTCTGCGTTCGTTCCGACCAACGTAATCTCCATTACCGATGGTCAGATCTTCCCTGAAACCAACCTGTTCAA
335 yAspVal Ser Al aPheVal P roThrAsnVal l l eSer l l eThrAspGl yGl n l l ePheLeuGl uThrAsnLeuPheAs

SmaI

3277 CGCCGGTATTTCGCTCTGCGGTTAACCCGGGTTATTTCCGTATCCCGTGTGGTGGTGCAGCACAGACCAAGATCATGAA
361 nAl aGl y l l eArgP roAl aVal l AsnP roGl y l l eSer Val l SerArgVal l Gl yGl yAl aAl aGl nThr Lys l l eMetLys

3355 AAAACTGTCCGGTGGTATCCGTACCGCTCTGGCACAGTATCGTGAACCTGGCAGCGTTCTCTCAGTTTGCATCCGACCT
387 sLysLeuSer Gl yGl y l l eArgThr Al aLeuAl aGl nTyrArgGl uLeuAl aAl aPheSer Gl nPheAl aSerAspLe

3433 TGACGATGCAACACGTAAGCAGCTTGACCACGGTCAGAAAGTGACCGAAGTCTGAAACAGAAACAGTATGCGCCGAT
413 uAspAspAl aThrArgLysGl nLeuAspHi sGl yGl nLysVal l Thr Gl uLeuLeuLysGl nLysGl nTyrAl aP roMe

3511 GTCCGTTGCGCAGCAGTCTCTGGTTCTGTTTCGCAGCAGAACGTGGTTACCTGGCGGATGTTGAACTGTCGAAAATTGG
439 tSer Val l Al aGl nGl nSer LeuVal l LeuPheAl aAl aGl uArgGl yTyrLeuAl aAspVal l Gl uLeuSer Lys l l eGl

BstBI

3589 CAGCTTCGAAGCCGCTCTGCTGGCTTACGTCGACCGTGATCACGCTCCGTTGATGCAAGAGATCAACCAGACCGGTGG
465 ySer PheGl uAl aAl aLeuLeuAl aTyrVal AspArgAspHi sAl aP roLeuMe tGl nGl u l l eAsnGl nThr Gl yGl

3667 CTACAACGACGAAATCGAAGGCAAGCTGAAAGGCATCCTCGATTCTTCAAAGCAACCCAATCCTGGTAACGTCTGGC
491 yTyrAsnAspGlu | leGluGlyLysLeuLysGly | leLeuAspSerPheLysAlaThrGlnSerTrp•••

Bsu36I

BseRI

NaeI

3745 GGCTTGCCTTAGGGCAGGCCGCAAGGCATTGAGGAGAAGCTCATGGCCGGCAGCAAAAGAGATACGTAGTAAGATCGCA
1 MetAlaGlyAlaLysGlu | leArgSerLys | leAla

3823 AGCGTCCAGAACACGCAAAAGATCACTAAAGCGATGGAGATGGTCGCCGCTTCCAAAATGCGTAAATCGCAGGATCGC
13 SerValGlnAsnThrGlnLys | leThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysSerGlnAspArg

3901 ATGGCGGCCAGCCGCTTATGCAGAAACCATGCGCAAAGTGATTGGTCACCTTGACACGGTAATCTGGAATATAAG
39 MetAlaAlaSerArgProTyrAlaGluThrMetArgLysVal | leGlyHisLeuAlaHisGlyAsnLeuGluTyrLys

SexAI

3979 CACCCTTACCTGGAAGACCGCGACGTTAAACGCGTGGGCTACCTGGTGGTGTGCGACCGACCGTGGTTTGTGCGGTGGT
65 HisProTyrLeuGluAspArgAspValLysArgValGlyTyrLeuValValSerThrAspArgGlyLeuCysGlyGly

4057 TTGAACATTAACCTGTTCAAAAACTGCTGGCGGAAATGAAGACCTGGACCGACAAAGGCGTTCAATGCGACCTCGCA
91 LeuAsn | leAsnLeuPheLysLysLeuLeuAlaGluMetLysThrTrpThrAspLysGlyValGlnCysAspLeuAla

4135 ATGATCGGCTCGAAAGGCGTGTGCTTCTCAACTCCGTGGGCGGCAATGTTGTTGCCAGGTCACCGGCATGGGGGAT
117 Met | leGlySerLysGlyValSerPhePheAsnSerValGlyGlyAsnValValAlaGlnValThrGlyMetGlyAsp

RsrII

BsrC

4213 AACCTTCCCTGTCCGAAGTATCGGTCCGGTAAAAGTGATGTTGCAAGCGTACGACGAAGGCCGTCTGGACAAACTG
143 AsnProSerLeuSerGluLeu | leGlyProValLysValMetLeuGlnAlaTyrAspGluGlyArgLeuAspLysLeu

4291 TACATTGTCAGCAACAAATTTATTAACACCATGTCTCAGGTTCCGACCATCAGCCAGCTGCTGCCGTTACCGGCATCA
169 Tyr | leValSerAsnLysPhe | leAsnThrMetSerGlnValProThr | leSerGlnLeuLeuProLeuProAlaSer

4369 GATGATGATGATCTGAAACATAAATCCTGGGATTACCTGTACGAACCCGATCCGAAGGCGTTGCTGGATACCCTGCTG
195 AspAspAspAspLeuLysHisLysSerTrpAspTyrLeuTyrGluProAspProLysAlaLeuLeuAspThrLeuLeu

4447 CGTCGTTATGTCGAATCTCAGGTTTATCAGGGCGTGGTTGAAAACCTGGCCAGCGAGCAGGCCGCCCGTATGGTGGCG
221 ArgArgTyrValGluSerGlnValTyrGlnGlyValValGluAsnLeuAlaSerGluGlnAlaAlaArgMetValAla

4525 ATGAAAGCCGCGACCGACAATGGCGGCAGCCTGATTAAGAGCTGCAGTTGGTATAACAAGCTCGTCAGGCCAGC
247 MetLysAlaAlaThrAspAsnGlyGlySerLeu | leLysGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSer

PmeI

4603 ATTAAGCAGGAAGTACCGAGATCGTCTCGGGGGCCGCCGCGTTTAAACAGGTTATTTCTAGAGGATTTAAGATGG
273 | leThrGlnGluLeuThrGlu | leValSerGlyAlaAlaAlaVal••• 1 MetA

Bsu36I

4681 CTAAGTAAAGATTGTCCAGGTAATCGGCGCCGTAGTTGACGTCGAATCCCTCAGGATGCCGTACCGCGCGTGTACG
2 | laThrGlyLys | leValGlnVal | leGlyAlaValValAspValGluPheProGlnAspAlaValProArgValTyrA

4759 ATGCTCTTGAGGTGCAAAATGGTAATGAGCGTCTGGTGTGGAAGTTCAGCAGCAGCTCGGCGGCGGTATCGTACGTA
28 spAlaLeuGluValGlnAsnGlyAsnGluArgLeuValLeuGluValGlnGlnGlnLeuGlyGlyGly | leValArgT

4837 CCATCGCAATGGGTTCCCTCCGACGGTCTGCGTCCGCGTCTGGATGTAAGACCTCGAACACCCGATTGAAGTCCCGG
54 hr | leAlaMetGlySerSerAspGlyLeuArgArgGlyLeuAspValLysAspLeuGluHisPro | leGluValProV

4915 TAGGTAAGCGACTCTGGGCCGTATCATGAACGTAAGGTTGAAACCGTTCGACATGAAAGGCGAGATCGGTGAAGAAG
80 alGlyLysAlaThrLeuGlyArg | leMetAsnValLeuGlyGluProValAspMetLysGlyGlu | leGlyGluGluG

4993 AGCGTTGGGCGATTACCGCGCAGCACCTTCTACGAAGAGCTGTCAAACCTCAGGAAGTCTGGAACCGGTATCA
106 | luArgTrpAla | leHisArgAlaAlaProSerTyrGluGluLeuSerAsnSerGlnGluLeuLeuGluThrGly | leL

StuI

5071 AAGTTATCGACCTGATGTGTCCGTTTCGCTAAGGGCGGTAAAGTAGGCCTGTTCCGGTGGTGGGGGTGTAGGTAACCG
132 ysVal | leAspLeuMetCysProPheAlaLysGlyGlyLysValGlyLeuPheGlyGlyAlaGlyValGlyLysThrV

SacI

BanII

NruI

5149 TTAACATGATGGAGCTCATTTCGTAACATCGCGATCGAGCACTCCGGTACTCTGTGTTTGGGGCGTAGGTGAACGTA
158 alAsnMetMetGluLeu | leArgAsn | leAla | leGluHisSerGlyTyrSerValPheAlaGlyValGlyGluArgT

5227 CTCGTGAGGGTAACGACTTCTACCACGAAATGACCGACTCCAACGTTATCGACAAAGTATCCCTGGTGTATGGCCAGA
184 hrArgGluGlyAsnAspPheTyrHisGluMetThrAspSerAsnVal | leAspLysValSerLeuValTyrGlyGlnM

5305 TGAACGAGCCCGCGGAAACCGTCTGCGCGTTGCTCTGACCGGTCTGACCATGGCTGAGAAATCCGTGACGAAGGTC
210 etAsnGluProProGlyAsnArgLeuArgValAlaLeuThrGlyLeuThrMetAlaGluLysPheArgAspGluGlyA

5383 GTGACGTTCTGCTGTTCTGTTGACAACATCTATCGTTACACCCTGGCCGGTACGGAAGTATCCGCACTGCTGGGCCGTA
236 rgAspValLeuLeuPheValAspAsn | leTyrArgTyrThrLeuAlaGlyThrGluValSerAlaLeuLeuGlyArgM

5461 TGCCTTACGCGGTAGGTTATCAGCCGACCCTGGCGGAAGAGATGGGCGTTCTGCAGGAACGTATCACCTCCACCAAAA
262 etProSerAlaValGlyTyrGlnProThrLeuAlaGluGluMetGlyValLeuGlnGluArg | leThrSerThrLysT

5539 CTGGTTCTATCACCTCCGTACAGGCAGTATACGTACCTGCGGATGACTTGACTGACCCGTCTCCGGCAACCACCTTTG
288 hr Gl ySer I l eThr Ser Val Gl nAl aVal TyrVal ProAl aAspAspLeuThrAspP roSer P roAl aThr Thr PheA

EagI

5617 CGCACCTTGACGCAACCGTGGTACTGAGCCGTCAGATCGCGTCTCTGGGTATCTACCCGGCCGTTGACCCGCTGGACT
314 l aHi sLeuAspAl aThr Val Val LeuSer ArgGl nI l eAl aSer LeuGl yI l eTyrP roAl aVal AspP roLeuAspS
5695 CCACCAGCCGTCAGCTGGACCCGCTGGTGGTTGGTCAGGAACACTACGACACCCGCGCGTGGCGTTCAGTCCATCCTGC
340 er Thr Ser ArgGl nLeuAspP roLeuVal Val Gl yGl nGl uHi sTyrAspThr Al aArgGl yVal Gl nSer I l eLeuG
5773 AACGTTATCAGGAACTGAAAGACATCATCGCCATCCTGGGTATGGATGAACTGTCTGAAGAAGACAACTGGTGGTAG
366 l nArgTyrGl nGl uLeuLysAspI l eI l eAl aI l eLeuGl yMe tAspGl uLeuSer Gl uGl uAspLysLeuVal Val A

Eco47III

5851 CGCGTGCTCGTAAGATCCAGCGCTTCTGTCCCAGCCGTTCTTCGTGGCAGAAGTATTCACCCGTTCTCCGGGTAAAT
392 l aArgAl aArgLys I l eGl nArgPheLeuSer Gl nP roPhePheVal Al aGl uVal l PheThr Gl ySer P roGl yLysT
5929 ACGTCTCCCTGAAAGACACCATCCGTGGCTTTAAAGGCATCATGGAAGGCGAATACGATCACCTGCCGGAGCAGGCCT
418 yrVal Ser LeuLysAspThr I l eArgGl yPheLysGl yI l eMetGl uGl yGl uTyrAspHi sLeuP roGl uGl nAl aP
6007 TCTACATGGTTCGTTCCATCGAAGAAGCTGTGAAAAAGCCAAAAAATTAAACGCCTTAATCGGAGGGTGATATGGC
444 heTyrMetVal Gl ySer I l eGl uGl uAl aVal Gl uLysAl aLysLysLeu••• 1 MetAl

Tth111I

6085 AATGACTTACCACCTGGACGTCGTCAGCGCAGAGCAACAAATGTTCTCTGGTCTGGTCGAGAAAATCCAGGTAACGGG
2 aMetThr TyrHi sLeuAspVal Val Ser Al aGl uGl nGl nMetPheSer Gl yLeuVal l Gl uLys I l eGl nVal l Thr Gl
6163 TAGCGAAGGTGAACTGGGGATCTACCCTGGCCACGCACCGCTGCTCACCGCCATTAAGCCTGGTATGATTTCGCATCGT
28 ySer Gl uGl yGl uLeuGl yI l eTyrP roGl yHi sAl aP roLeuLeuThr Al aI l eLysP roGl yMet I l eArgI l eVa
6241 GAAACAGCACGGTCACGAAGAGTTTATCTATCTGTCTGGCGGCATTCTTGAAGTGCAGCCTGGCAACGTGACCGTTCT
54 l LysGl nHi sGl yHi sGl uGl uPheI l eTyrLeuSer Gl yGl yI l eLeuGl uVal l Gl nP roGl yAsnVal l Thr Val l Le
6319 GGCCGACACCGCAATTCGCGGCCAGGATCTCGACGAAGCGCGAGCCATGGAAGCGAAACGTAAGGCTGAAGAGCACAT
80 uAl aAspThr Al aI l eArgGl yGl nAspLeuAspGl uAl aArgAl aMetGl uAl aLysArgLysAl aGl uGl uHi s I l

Sgfl

6397 TAGCAGCTCTACGGCGACGTAGATTACGCTCAGGCGTCTGCGGAACTGGCCAAAGCGATCGCGCAGCTGCGCGTTAT
106 eSer Ser Ser Hi sGl yAspVal AspTyrAl aGl nAl aSer Al aGl uLeuAl aLysAl aI l eAl aGl nLeuArgVal I l
6475 CGAGTTGACCAAAAAAGCGATGTAACACCGGCTTAAAAAGCACAAAAGCCAGTCTGGAACAGGCTGGCTTTTTTTTTG
132 eGl uLeuThr LysLysAl aMet
6553 CGCGTGTACCCGCTCCTGAATAGCGTTCACATAGATCCTGCTGATATAAAACCCCCCTGTTTTCTGTTTATTTCATTG
6631 ATCGAAATAAGAGCAAAAACATCCACCTGACGCTTAAATTAAGGACTGCCTTAATTTTTCTGCAGACAAAAGGCGTGA
6709 CGATGGTCAAAAATGGCGCTTTCGTCAGCGGGGATAATCCGTTATTGAACAATTTATCCTCTGTCCATTTACGATGA
6787 AAAAAATGATTTTTTCAAGGTGAAGCGGTTAAATTCGTTCTCAAATTACAGTCAGGACGCGTATGTTGAATAATG
1 MetLeuAsnAsnA

6865 CTATGAGCGTAGTGATCCTTGCCGACGCAAGGCACGCGCATGTATTCCGATCTTCCGAAAGTGCTGCATACCCCTTG
5 l aMetSer Val Val I l eLeuAl aAl aGl yLysGl yThr ArgMe tTyrSerAspLeuP roLysVal l LeuHi sThr LeuA

SexAI

6943 CCGGAAAAGCGATGGTTCAGCATGTCATTGATGCTGCGAATGAATTAGGCGCAGCGCACGTTACCTGGTGTACGGTC
31 l aGl yLysAl aMetVal l Gl nHi sVal I l eAspAl aAl aAsnGl uLeuGl yAl aAl aHi sVal l Hi sLeuVal l TyrGl yH
7021 ACGGCGGCGATCTGCTAAAACAGGCGCTGAAAGACGACAACCTTAAGTGGGTGCTTACGGCAGAGCAGCTGGGTACGG
57 i sGl yGl yAspLeuLeuLysGl nAl aLeuLysAspAspAsnLeuAsnTrpVal l LeuGl nAl aGl uGl nLeuGl yThr G
7099 GTCATGCAATGCAGCAGGCCGCACCTTTCTTTGCCGATGATGAAGACATTTTAATGCTCTACGGCGACGTGCCGCTGA
83 l yHi sAl aMetGl nGl nAl aAl aP roPhePheAl aAspAspGl uAspI l eLeuMetLeuTyrGl yAspVal l P roLeu l
7177 TCTCTGTCGAAACACTCCAGCGTCTGCGTGATGCTAAACCGCAGGGTGGCATTGGTCTGCTGACGGTGAAACTGGATG
109 l eSer Val l Gl uThr LeuGl nArgLeuArgAspAl aLysP roGl nGl yGl yI l eGl yLeuLeuThr Val l LysLeuAspA
7255 ATCCGACCCGTTATGGACGTATCACCCGTGAAAACGGCAAAGTTACCGGCATTGTTGAGCACAAAGATGCCACCGACG
135 spP roThr Gl yTyrGl yArgI l eThrArgGl uAsnGl yLysVal l Thr Gl yI l eVal l Gl uHi sLysAspAl aThrAspG
7333 AGCAGCGTCAGATTACAGGAGATCAACACCGGCATTCTGATTGCCAACGGCGCAGATATGAAACGCTGGCTGGCGAACG
161 l uGl nArgGl nI l eGl nGl uI l eAsnThr Gl yI l eLeuI l eAl aAsnGl yAl aAspMetLysArgTrpLeuAl aAsnV
7411 TGACCAACAATAATGCTCAGGGCGAATACTACATCACCGACATTATTGCGCTGGCGTATCAGGAAGGGCGTGAAATCG
187 al ThrAsnAsnAsnAl aGl nGl yGl uTyrTyrI l eThrAspI l eI l eAl aLeuAl aTyrGl nGl uGl yArgGl uI l eV
7489 TCGCCGTTTCCGCAACGTTTAAAGCGAAGTAGAAGGCGTGAATAACCGCCTGCAACTCTCCCGTCTGGAGCGTGTTT
213 al Al aVal l Hi sP roGl nArgLeuSer Gl uVal l Gl uGl yVal l AsnAsnArgLeuGl nLeuSer ArgLeuGl uArgVal l T
7567 ATCAGTCCGAACAGGCTGAAAACTGCTGTTAGCAGGCGTTATGCTGCGCGATCCAGCGCGTTTTGATCTGCGTGTTA
239 yrGl nSer Gl uGl nAl aGl uLysLeuLeuLeuAl aGl yVal l MetLeuArgAspP roAl aArgPheAspLeuArgGl yT

Eam1105I

7645 CGCTAACTCACGGGCGCGATGTTGAAATTGATACTAACGTTATCATCGAGGGCAACGTGACTCTCGGTCATCGCGTGA
265 hr LeuThr HisGlyArgAspVal Glu IleAspThrAsnVal Ile IleGluGlyAsnVal Thr LeuGlyHisArgVal L
7723 AAATTGGCACCGTTGCGTGATTA AAAACAGCGTGATTGGCGATGATTGCGAAATCAGTCCGTATACCGTTGTGGAAG
291 ysl IleGlyThr GlyCysVal IleLysAsnSer Val IleGlyAspAspCysGlu IleSer ProTyrThr Val Val GluA
7801 ATGCGAATCTGGCAGCGCCTGTACCATTGGCCCGTTTGGCCGTTTGGCTCCTGGTGCTGAGTTGCTGGAAGGTGCTC
317 spAlaAsnLeuAlaAlaAlaCysThr IleGlyProPheAlaArgLeuArgProGlyAlaGluLeuLeuGluGlyAlaH
7879 ACGTCGGTAACTTCGTTGAGATGAAAAAGCGCGTCTGGGTAAAGGCTCGAAAGCTGGTCATCTGACTTACCTGGGCG
343 isVal GluAsnPheVal GluMetLysLysAlaArgLeuGlyLysGlySerLysAlaGlyHisLeuThr TyrLeuGlyA
7957 ATGCGGAAATTGGCGATAACGTTAACATCGGCGCGGGAACCATTACCTGCAACTACGATGGTGCGAATAAATTTAAGA
369 spAlaGlu IleGlyAspAsnValAsnIleGlyAlaGlyThr IleThr CysAsnTyrAspGlyAlaAsnLysPheLysT
8035 CCATTATCGGCGACGATGTGTTTGGTTCCGACACTCAGCTGGTGGCCCGGTAACAGTAGGCAAAGGCGCGACCA
395 hr Ile IleGlyAspAspVal PheVal GluSerAspThr GluLeuValAlaProVal Thr Val GluLysGlyAlaThr I

NsiI

8113 TTGCTGCGGGTACAACCTGTGACGCGTAATGTCGGCGAAAATGCATTAGCTATCAGCCGTGTGCCGCACTCAGAAAG
421 leAlaAlaGlyThr Thr Val Thr ArgAsnVal GluGluAsnAla

NaeI

BseRI

8191 AAGGCTGGCGTCGTCGGTAAAGAAAAAGTGATTCTGGCCGGCTAACCCGGTCACATGGGATGAGGAGATAACATAAT
8269 CTCCCTCCACAAGCAGTAACTATAAAAATAACCCCACTCTCTACAAGGCTCGGGGCGCCCGAAAAAACGGGCATACA

SgfI

8347 GGTTGACCGACAACGATATAAATCGGAATCAAAAATATGTGTGGAATTGTTGGCGCGATCGCGCAACGTGATGTAGC

XmnI

8425 AGAAATCCTTCTTGAAGTTTTACGTCGTCTGGAATACCGCGATATGACTCTGCCGGTCTGGCCGTTGTTGATGCAGA

NdeI

8503 AGGTCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCCTCGCT
8581 CACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC
8659 AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
8737 GCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCC
8815 GACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTAC
8893 CGGATACCTGTCCGCTTTCTCCCTTCCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCGT
8971 GTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTA
9049 TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG
9127 GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTG
9205 CGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG
9283 TGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG
9361 GTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGAT
9439 CCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTT

286 TrpHisLys

Eam1105I

9517 AATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAAC
283 IleLeuSerAlaGlyIleGluAlaIleGluArgAsnArgGluAspMetThrAlaGluSerGlyThrThrTyrIleVal

BsaI

9595 TACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC GGCTCCAGATTT
257 Val IleArgSerProLysGlyAspProGlyLeuAlaAlaIle IleGlyArgSerGlyArgGluGlyAlaGlySerLys
9673 ATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT
231 AspAlaIlePheTrpGlyAlaProLeuAlaSerArgLeuLeuProGlyAlaValLysAspAlaGluMetTrpAspIle
9751 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGCTGCAGGCAT
205 LeuGluGluArgSerAlaLeuThrLeuLeuGluGlyThrLeuLeuLysArgLeuThrThrAlaMetAlaAlaProMet
9829 CGTGGTGTACGCTCGTCTTTGGTATGGCTTATTACGCTCCGGTCCCAACGATCAAGGCGAGTTACATGATCCCC
179 ThrThrAspArgGluAspAsnProIleAlaGluAsnLeuGluProGluTrpArgAspLeuArgThrValHisAspGly
9907 CATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTTATCACT
153 MetAsnHisLeuPheAlaThrLeuGluLysProGlyGlyIleThrThrLeuLeuLeuAsnAlaAlaThrAsnAspSer

Scal

9985 CATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACTC
127 MetThrIleAlaAlaSerCysLeuGluArgValThrMetGlyAspThrLeuHisLysGluThrValProSerTyrGlu
10063 AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCGGCTCAACACGGGATAATACCGCGCC
101 ValLeuAspAsnGluSerTyrHisIleArgArgGluLeuGluGluGluValAspValArgSerLeuValAlaGly

XmnI

10141 ACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTT
75 CysLeuLeuVal LysPheThr Ser MetMetP roPheArgGl uGl uP roArgPheSer Gl uLeu l eLysGl ySerAsn
10219 GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAGTATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTG
49 LeuAspLeuGl u l l eTyrGl yVal A rgAl aGl yLeuGl nAspGl uAl aAspLysVal LysVal LeuThr Gl uP roHi s
10297 AGCAAAACAGGAAGGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCT
23 Al aPheVal P roLeuCysPheAl aAl aPhePheP rol l eLeuAl aVal A rgPheHi sGl nl l eSer Met

SspI

10375 TTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAA
10453 ACAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAAC

EcoO109I

HindIII

10531 CTATAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATTCTCATGTTTGACAGCTTATCATCGATAAGCTTT
10609 CAAAGTTCTGGCGATGTTGGTGTACTGGTGGTGGCGTTGGCGGTTTTAAAGGCGGTATTCTTGCCGCTGATCGTTAC
10687 GTGGGTTTTGGTGCTGGTGGTTCAGATACTGGCACCGGCTGTAATTAACAACAAAGGGTAAAGGCATC