



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGTGCCTA **MfeI (82)**
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC **HindIII (245)** **Bsu36I (291)**
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCTCCTGAAGCTCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT **NgoMI (441)** **NaeI (441)**

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCATCATGACGGTCAAGCTGGATTTTCGAGGAGTGTCTCAAGGA **AgeI (552)** **BspHI (560)**
601 CTCACCCCGTTTCCGAGCCTCTATTGAGCTGGTGAAGCCGAAGTGTGAGATTGGAGACCCGCTCTGAAAAGCTCCTGAAACTGGGCACTGGTCTCCTG **1▶MetThrValLysLeuAspPheGluGluCysLeuLysAs**
701 GAAAGTGGCGCCATTACCTTGTGCTGCCAGCCGCGCTTCTGTTGCTGGCATTGTGACCTGGCCCGCTGGTCCACCAGAGCCATGATGGCGGAGTGTCT **13▶pSerProArgPheArgAlaSerIleGluLeuValGluAlaGluValSerGluLeuGluThrArgLeuGluLysLeuLeuLysLeuGluYThrGlyLeuLeu**
801 TGGAAAAATTCACCGTGAGCCTGAACCACAAGCTGGACAGCCATGCGGAGCTTAGATGCCACCCAACACACTGCAGCAGCAGATCCAGACCCCTGGT **47▶GluSerGlyArgHisTyrLeuAlaAlaSerArgAlaPheValValGlyIleCysAspLeuAlaArgLeuGlyProProGluUpRoMetMetAlaGluCysL**
901 CAAGGAAGTCTGCGGGTTTCCGAGAGCTCGCCGGGATTTCTGCGGGGGGCTGAGAGCCTGGAGGCTGCCCTGACCCACAACGCAGAGGTTCCAGG **XbaI (852)** **PshAI (891)**
1001 CGCCGGCCAGGAGGCAGAAGAGGAGGAGCTGCTTTGAGGACGGCTCGAGCTGGGTACCGGGGACGGGCACTGGATTATGCCCTGCAGATCAACGTGA **80▶euGluLysPheThrValSerLeuAsnHisLysLeuAspSerHisAlaGluLeuLeuAspAlaThrGluHisThrLeuGluNglNglNleGluNThrLeuVal**
1101 TTAGGACAAGAGGAAGTTTGCATCATGGAGTTTGTGCTGCGTTTGGTGGAGGCCAGGCTACCCATTTCCAGCAGGGCCATGAGGAGCTGAGCCGGCT **901 CAAGGAAGTCTGCGGGTTTCCGAGAGCTCGCCGGGATTTCTGCGGGGGGCTGAGAGCCTGGAGGCTGCCCTGACCCACAACGCAGAGGTTCCAGG**
1101 TTAGGACAAGAGGAAGTTTGCATCATGGAGTTTGTGCTGCGTTTGGTGGAGGCCAGGCTACCCATTTCCAGCAGGGCCATGAGGAGCTGAGCCGGCT **113▶IlysGluGlyLeuArgGlyPheArgGluAlaArgArgAspPheTrpArgGlyAlaGluSerLeuGluAlaAlaLeuThrHisAsnAlaGluValProArg**
1201 GTCCAGTATCGAAAGGAGCTGGGCGCCAGTTGACACAGCTGGTCTTGAATTCAGCACGAGAGAAGAGGGACATGGAGCAGAGACACGTGCTGCTGAAA **XhoI (1046)** **Acc65I (1055)**
1201 GTCCAGTATCGAAAGGAGCTGGGCGCCAGTTGACACAGCTGGTCTTGAATTCAGCACGAGAGAAGAGGGACATGGAGCAGAGACACGTGCTGCTGAAA **147▶ArgArgAlaGluNgluAlaGluGluAlaGlyAlaAlaLeuArgThrAlaArgAlaGlyTyrArgGlyArgAlaLeuAspTyrAlaLeuGluNleAsnValI**
1301 CAGAAGGAGCTGGGTGGGAGGAGCCAGAACAAGCTTAAGAGAGGGGCTGGTGGCCTGGTGTGGAAGGACATCTCTTCAAACGGGCCAGCAACGCAT **ArgArgAlaGluNgluAlaGluGluAlaGlyAlaAlaLeuArgThrAlaArgAlaGlyTyrArgGlyArgAlaLeuAspTyrAlaLeuGluNleAsnValI**
1401 TTAAGACCTGGAGCAGACGCTGTTTACCATTAGAGCAACCACTGGTTTACCAGAAGAAGTACAAGGACCCTGTGACTGTGGTGGTGGATGACCTTCG **180▶IleGluAspLysArgLysPheAspIleMetGluPheValLeuArgLeuValGluAlaGluNAlaThrHisPheGluNglNglYHisGluGluSerArgLe**
1501 TCTCTGCACAGTGAAGCTCTGCCCTGACTCAGAAAGCGGTTCTGCTTTGAGGTGGTGTCCACCAGCAAGTCCTGCCTCCTCAGGCTGACTCAGAGCGC **EcoRI (1248)** **BbrPI (1285)**
1501 TCTCTGCACAGTGAAGCTCTGCCCTGACTCAGAAAGCGGTTCTGCTTTGAGGTGGTGTCCACCAGCAAGTCCTGCCTCCTCAGGCTGACTCAGAGCGC **213▶uSerGlnTyrArgLysGluLeuGlyAlaGluNLeuHisGluNLeuValLeuAsnSerAlaArgGluLysArgAspMetGluGluNArgHisValLeuLeuLys**
1601 CTCCTGCAGCTGTGGGTGCTGCTGTCAGAGCAGCATTGCTTCTGCCTTCAGTCAGGCTCGCCTTGATGACAGCCCGGGTCCAGGCCAGGGCTCAG **HindIII (1332)** **PshAI (1492)**
1601 CTCCTGCAGCTGTGGGTGCTGCTGTCAGAGCAGCATTGCTTCTGCCTTCAGTCAGGCTCGCCTTGATGACAGCCCGGGTCCAGGCCAGGGCTCAG **247▶GlnLysGluLeuGlyGlyGluGluUpRoGluUpRoSerLeuArgGluGluProGlyGlyLeuValMetGluGlyHisLeuPheLysArgAlaSerAsnAlaP**
1701 GACACCTGGCCATAGGCTCTGCTGCCACCTGGGCTCTGGTGAATGGCCAGGGGAAGGAGCCTGGGGAGTCGGGCACGTGGTGGCCAGGTCCAGAG **280▶heLysThrTrpSerArgArgTrpPheThrIleGluNserAsnGluNLeuValTyrGluNlysLysTyrLysAspProValThrValValValAspAspLeuAr**
1801 TGTGGATGCAATGCCAGTGTGCTGCGACTGCCGGGAGCCAGCCCGGAGTGGCCAGCATCAACCTTGGTGTACCCTCTGCATTTCAGTGTTCGGGCATC **313▶gLeuCysThrValLysLeuCysProAspSerGluArgArgPheCysPheGluValValSerThrSerLysSerCysLeuLeuGluNAlaAspSerGluArg**
1901 CACAGGAGCCTGGTGTCTCACTTCTCAAAGTCCGGTCTCTGACCCTTGACTCATGGAGCCAGAAGTGAAGCTCATGTGTGAGCTGGGAAATGTCA **SmaI (1676)**
1901 CACAGGAGCCTGGTGTCTCACTTCTCAAAGTCCGGTCTCTGACCCTTGACTCATGGAGCCAGAAGTGAAGCTCATGTGTGAGCTGGGAAATGTCA **347▶LeuLeuGluNLeuTrpValSerAlaValGluNserSerIleAlaSerAlaPheSerGluNAlaArgLeuAspAspSerProArgGlyProGlyNglYserG**
2001 TCATCAACCAGATCTATGAGCCCGCTGGAGGCCATGGCAGTGAAGAAACCAGGGCCAGCTGCTCCCGCAGGAGAAGGAGGCTGGATTACCGCTAA **DraIII (1777)** **BbrPI (1777)**
2001 TCATCAACCAGATCTATGAGCCCGCTGGAGGCCATGGCAGTGAAGAAACCAGGGCCAGCTGCTCCCGCAGGAGAAGGAGGCTGGATTACCGCTAA **380▶IyHisLeuAlaIleGlySerAlaAlaThrLeuGlySerGlyGlyMetAlaArgGlyArgGlyUpRoGlyGlyValGlyYHisValValAlaAlaGluNValGluNse**
2101 ATACGTGGAGAAGTTCTGACCAAGCTGCCTGAGATTGAGGGGCAAGAGGTGGCCGGGGGCGCCCAAGGGGCGAGCCTCTGTGCCCAAGCCCT **413▶rValAspGlyAsnAlaGluNcysCysAspCysArgGluUpRoAlaPProGluTrpAlaSerIleAsnLeuGlyValThrLeuCysIleGluNcysSerGlyIle**
2201 TCCATCAGCCCCGGCCAGGGAGCTTGAGATCCAAGCCAGAGCCCTCTGAGGACCTGGGAAGCCTGCACCCTGGGCCCCACTGTTTCGAGCGTCTG **447▶HisArgSerLeuGlyValHisPheSerLysValArgSerLeuThrLeuAspSerTrpGluUpRoGluLeuValLysLeuMetCysGluLeuGlyAsnValI**
2301 GGCATCCTCCATCTCTCCACCATGGCTGATGCCCTGCCCCATGGAGCTGATGCAACTGGTCAATGGGGCCAAGATAATGCCACACCCGCTGATCCA **BglIII (2009)** **StuI (2081)**
2301 GGCATCCTCCATCTCTCCACCATGGCTGATGCCCTGCCCCATGGAGCTGATGCAACTGGTCAATGGGGCCAAGATAATGCCACACCCGCTGATCCA **547▶SerIleArgProArgProGlySerLeuArgSerLysProGluUpRoProSerGluAspLeuGlySerLeuHisSProGlyAlaLeuLeuPheArgAlaSerG**
2401 GGCCACAGCTGCTAATCTCTTCTGGCCTGTGAGTTTCTCTCCAGAACGGGCGAACGTGAACCAAGCGGACAGTGGGGCCGGGCGCCGCTGCACCAC **580▶IyHisSProRoSerLeuProThrMetAlaAspAlaLeuAlaHisGlyAlaAspValAsnTrpValAsnGlyGlyGluNAspAsnAlaThrProLeuIleGlu**
2501 GCAACCATCTTGCCACACGGGCTCGCTGCTTCTGAAACGGGGAGCTGATCTGGGGCTCGAGACTGAAAGGCAAGGACCTCTGACCATCG **613▶nAlaThrAlaAlaAsnSerLeuLeuAlaCysGluPheLeuLeuGluNAsnGlyAlaAsnValAsnGluNAlaAspSerAlaGlyArgGlyProLeuHisHis**
XhoI (2564) **SandI (2582)** **BstXI (2594)**
647▶AlaThrIleLeuGlyHisThrGlyLeuAlaCysLeuPheLeuLysArgGlyAlaAspLeuGlyAlaArgAspSerGluGlyArgAspProLeuThrIleA

Tth111I (2620)

2601 CCATGAAACAGCCAACGCTGACATCGTCACCCTGCTACGACTGGCAAAGATGAGGGAGGCTGAAGCGGCCAGGGGCAGGCAGGAGATGAGACGTATCT
680▶ I aMe tGI uThr Al aAsnAl aAspI l eVal Thr LeuLeuArgLeuAl aLysMe tArgGI uAl aGI uAl aGI nGI yGI nAl aGI yAspGI uThr TyrLe
BspHI (2763) DraIII (2775) SfiI (2790)

2701 TGACATCTTCCGCGACTTCTCCCTCATGGCGTCAGACGACCCGGAGAAGCTGAGCCGTCGAGTCATGACCTCCACACGCTGTGACCCGAGGCCACGGG
713▶ uAspI l ePheArgAspPheSer LeuMe tAl aSer AspAspProGI uLysLeuSer ArgArgSer Hi sAspLeuHi sThr Leu●●●
NheI (2853)

2801 GCCCCGCGCTGCCTCCCTTCCCCGCCACCGGGCCCTCTGCCATTAAGCCTCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAC
HpaI (2991)

2901 CACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAAACAAC
MfeI (3002) EcoRI (3087)

3001 AACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAATCTAAAATA
3101 CAGCATAGCAAACTTTAACCTC CAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCAT
3201 TAGCTGTTTGAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTCTTTCATTTCTTTATGTTTAAATGCACT

3301 GACCTCCACATTCCTTTT TAGTAAAATATTCAGAAAATATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTC
SspI (3326) SmaI (3340)

3401 AAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAG
141◀●●●A

3501 TTCCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGA
139◀snArgThr TyrLysLeuP roI l eLeuGI uGI u l eThr Thr LysVal LeuLysGI yAsnMe tGI u l eLeuVal PheCysAspP roAl aTyrAspSer I l
SacI (3601) BstXI (3630)

3601 TGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCCTT
106◀eLeuGI uArgCysMe tGI yCysP roSer Val Val ArgI l eSer ArgAspVal GI uAspSer TyrP roHi sArgVal Al aVal I l eThr AspPheAspLys
StuI (3765)

3701 CTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTT CAGCACAGACAGTGACCCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCA
73◀GI nGI yAsnSer Val Al aSer GI y l eAl a l eAl aGI uAl aCysVal Thr Val ArgGI y l eTyrAl aGI u l eHi sVal Al aSer I l e l eGI uGI yT
3801 GTCTTGGTCTGTAGGCCGCCGACATGGTGTGTTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAG
39◀hr LysThr ArgI l eAl aAl aGI yVal Hi sHi sLysAsnAspGI uTyrLeuMe tThr I l eLysGI uThr Al aVal GI uVal LeuGI uLeuAspGI nGI nSe
BspHI (3915)

3901 AGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGC
XmnI (3907) AseI (3973)

6◀r l l eAsnPheThr LysMe t

4001 GTCTCCAGC T TATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCCATTTGGTCAATGGGCGGAGTTGTT
SacI (4030)

4101 ACGACATTTTGGAAAGTCCC GTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTA
SpeI (4128)

4200 TCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATG
SnaBI (4256)

4300 TACTGGGCATAATGCCAGCGGGCCATTACCCTGATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTT
NdeI (4361)

4400 ACCGTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGG

4500 CGGTACGCCAGCGGGCCATTTACCCTAAGTTATGTAACGCCTG CAG G TT AA TTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAA
SdaI (4539) PacI (4547) BspLU11I (4557)

4598 AAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA
4698 TAAAGATACCAGCGTTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGAA
ApaLI (4871)

4798 GCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGA
4898 CCGCTGCGCCTTATCCGGTAACTATCGTCTT GAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCG
4998 AGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA
5098 CTTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAA

5198 AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACA

EagI (5307)

NotI (5306)

5298 TTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAA

5398 AACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA