



PstI (7)

1 CCTGCAGTTTGAGGAGAATATTTGTTATATTTGCAAATAAAAATAAGTTTGCAAGTTTTTTTTTTCTGCCCAAAGAGCTCTGTGTCTTGAACATAAAA

101 TACAAATAACCGCTATGCTGTTAATTATTGGCAAATGTCCCATTTTCAACCTAAGGAAATACCATAAAGTAACAGATATACCAACAAAAGGTTACTAGTT

NcoI (280)

201 AACAGGCATTGCCTGAAAAGAGTATAAAAAGAAATTCAGCATGATTTTCCATATTGTGCTTCCACCACCTGCCAATAAACCCATGGGGGTTTCTCATCATCA

MetGlyGlySerHisHisHi

301 TCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGCGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTTGGAGTTGAT

7> sHisHisHisGlyMetAlaSerMetThrGlyGlyGlnGlnMetGlyArgAspLeuTyrAspAspAspAspLysValProLysAspGlnLeuGlyValAsp

401 CCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTTAATCGCCTTGACGACATCCCCCTTTCGCGAGCTGGCGTAATAGCGAAG

41> ProValValLeuGlnArgArgAspTrpGluAsnProGlyValThrGlnLeuAsnArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluG

501 AGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTCCGGCACCAGAAGCGGTGCGGAAAAGCTGGCT

74> luAlaArgThrAspArgProSerGlnGlnLeuArgSerLeuAsnGlyGluTrpArgPheAlaTrpPheProAlaProGluAlaValProGluSerTrpLe

601 GGAGTGGCATCTCTGAGGCCGATACTGCTGCTGCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCATCTACACCAACGTAACCTATCCCATT

107> uGluCysAspLeuProGluAlaAspThrValValValProSerAsnTrpGlnMetHisGlyTyrAspAlaProI leTyrThrAsnValThrTyrProI le

701 ACGGTCATCCGCCGTTGTTCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGGCAA

141> ThrValAsnProProPheValProThrGluAsnProThrGlyCysTyrSerLeuThrPheAsnValAspGluSerTrpLeuGlnGluGlyGlnThrArg

801 TTATTTTTGATGGCTTAACTCGGCTTTCATCTGTGGTGCAACGGGCTGGGTTCGTTACGGCCAGGACAGCTGTTGCGCTGTAATTTGACTGAG

174> IeI lePheAspGlyValAsnSerAlaPheHisLeuTrpCysAsnGlyArgTrpValGlyTyrGlyGlnAspSerGluLeuProSerArgAlaValLeuSe

901 CGCATTTTTACGCGCCGGAGAAAACCGCTCGCGGTGATGGTGTGCTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGC

207> rAlaPheLeuArgAlaGlyGluAsnArgLeuAlaValMetValLeuArgTrpSerAspGlySerTyrLeuGluAspGlnAspMetTrpArgMetSerGly

1001 ATTTCCGTCAGCTCTGTTGCTGCATAAACCGACTACACAATCAGCAGCTTCCATGTTGCCACTCGCTTAAATGATGATTTACGCCGCGCTGTACTGG

241> I lePheArgAspValLeuHisLysProThrThrGlnI leSerAspPheHisValAlaThrArgPheAsnAspGluAlaThrArgPheAlaValLeuG

1101 AGGCTGAAGTTCAGATGTGCGCGGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGAACCGCAGGTCGCCAGCGCCACCGCGCTTT

274> luAlaGluValGlnMetCysGlyGluLeuArgAspTyrLeuArgValThrValSerLeuTrpGlnGlyGluThrGlnValAlaSerGlyThrAlaProPh

1201 CGCGCGTGAATATTCGATGAGCGTGGTGGTTATGCCGATCGCCTCACACTACGCTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCGAAATCCCGAAT

307> eGlyGlyGluI leI leAspGluArgGlyGlyTyrAlaAspArgValThrLeuArgLeuAsnValGluAsnProLysLeuTrpSerGluAlaLeuProAsn

1301 CTCTATCGTGGCTGTTGAAGTGCACACCGCCGACCGCAGCCTGTAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTCGGGATGAAATGGTC

341> LeuTyrArgAlaValValGluLeuHisThrAlaAspGlyThrLeuI leGluAlaGluAlaCysAspValGlyPheArgGluValArgI leGluAsnGlyL

1401 TGCTGCTGCTGAACCGCAAGCCGTTGCTGATTCCGAGCGGTAACCGTCACGAGCATCATCTCTGCATGGTCAGGTCATGGATGACGACAGCATGGTGCA

374> euLeuLeuLeuAsnGlyLysProLeuLeuI leArgGlyValAsnArgHisGluHisHisProLeuHisGlyGlnValMetAspGluGlnThrMetValG

1501 GGATATCTGTGATGAAGCAGAACAACCTTAACGCCGTCGCTGTTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGGCAGCGCTACGCCCTG

407> nAspI leLeuLeuMetLysGlnAsnAsnPheAsnAlaValArgCysSerHisTyrProAsnHisProLeuTrpTyrThrLeuCysAspArgTyrGlyLeu

1601 TATGTGGTGGATGAAGCAATATTGAAACCACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGCGATGAGCGAAGCGCTAA

441> TyrValValAspGluAlaAsnI leGluThrHisGlyMetValProMetAsnArgLeuThrAspAspProArgTrpLeuProAlaMetSerGluArgValT

1701 CGCGAATGGTGCAGCGCATCGTAATCACCCGAGTGTGATCATCTGGTGGGAAATGAATCAGGCCACGGCGTAATCACGACGCGCTGTATCGCTG

474> hrArgMetValGlnArgAspArgAsnHisProSerVal I leI leTrpSerLeuGlyAsnGluSerGlyHisGlyAlaAsnHisAspAlaLeuTyrArgTr

1801 GATCAATCTGCTGATCTTCCCGCCGTCGATGAAGCGGGGAGCGGACACCAGGCCACCGATATTATTTGCCGATGACGCGCGCTGGAT

507> pI leLysSerValAspProSerArgProValGlnTyrGluGlyGlyGlyAlaAspThrThrAlaThrAspI leI leCysProMetTyrAlaArgValAsp

1901 GAAGACCAGCCCTTCCCGCTGTGCCGAATGGTCCATAAAAAATGGCTTTCGCTACCTGGAGAGACGGCCCGCTGATCTTTGCGAATACGCCACG

541> GluAspGlnProPheProAlaValProLysTrpSerI leLysLysTrpLeuSerLeuProGlyGluThrArgProLeuI leLeuCysGluTyrAlaHisA

2001 CGATGGTAAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCGCTTACAGGGCGGCTTCGCTGGGACTGGTGGATCAGTC

574> IaMetGlyAsnSerLeuGlyGlyPheAlaLysTyrTrpGlnAlaPheArgGlnTyrProArgLeuGlnGlyGlyPheValTrpAspTrpValAspGlnSe

2101 GCTGATTAATATGATGAAAACGGCAACCCGTTGCTGGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTC

607> rLeuI leLysTyrAspGluAsnGlyAsnProTrpSerAlaTyrGlyGlyAspPheGlyTyrProAsnAspArgGlnPheCysMetAsnGlyLeuVal

2201 TTTGCCGACCGCAGCGCATCCAGCGCTGACGGAAGCAAAACACGAGCAGGTTTTTCCAGTTCCGTTTATCCGGGCAACCATCGAAGTACCGAGCG

641> PheAlaAspArgThrProHisProAlaLeuThrGluAlaLysHisGlnGlnPhePheGlnPheArgLeuSerGlyGlnThrI leGluValThrSerG

2301 AATACCTGTTCCGTCATAGCGATAACGAGCTCCTGCACCTGGATGGTGGCGCTGGATGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGCTGCTCC

674> luTyrLeuPheArgHisSerAspAsnGluLeuLeuHisTrpMetValAlaLeuAspGlyLysProLeuAlaSerGlyGluValProLeuAspValAlaPr

2401 ACAAGGTAACAGTTGTAAGTGAAGTGAACCTACCGCACTCCGCGAGAGCGCGGCAACTTGGCTCACAGTACGGTAGTGCAACCGCAACCGCAGCGCA

707> oGlnGlyLysGlnLeuI leGluLeuProGluLeuProGlnProGluSerAlaGlyGlnLeuTrpLeuThrValArgValValGlnProAsnAlaThrAla

2501 TGGTCAGAAGCCGGCACATCAGCGCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCGCGCGCTCCACGCCATCCCGCATCTGA

741> TrpSerGluAlaGlyHisI leSerAlaTrpGlnGlnTrpArgLeuAlaGluAlaLeuSerValThrLeuProAlaAlaSerHisAlaI leProHisLeuT

2601 CCACCAGCGAATGGATTTTGCATCGAGCTGGTAATAGCGTTGGCAATTAACCGCCAGTCAAGCTTCTTTTACAGATGTGGATGGCGATAAAAA

774> hrThrSerGluMetAspPheCysI leGluLeuGlyAsnLysArgTrpLeuLeuAsnArgGlnTrpAlaAspThrLeuAlaAspAlaValLeuI leThrThrAlaHisA

2701 ACAACTGCTGACCGCTGCGCGATCAGTTACCCGTCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGATTGACCTAACCGCTGGGTC

807> sGlnLeuLeuThrProLeuArgAspGlnPheThrArgAlaProLeuAspAsnAspI leGlyValSerGluAlaThrArgI leAspProAsnAlaTrpVal

2801 GAACGCTGGAAGGCGCGGCCATTACCAGGCCAAGCAGCGTGTGTCAGTGCACGGCAGATACACTTGTGATGCGGTGCTGATTACGACCGCTCAGC

841> GluArgTrpLysAlaAlaGlyHisTyrGlnAlaGluAlaLeuLeuGlnCysThrAlaAspThrLeuAlaAspAlaValLeuI leThrThrAlaHisA

2901 CGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCCTGTGATGTTGAAGTGGCGAG

874> IaTrpGlnHisGlnGlyLysThrLeuPheI leSerArgLysThrTyrArgI leAspGlySerGlyGlnMetAlaI leThrValAspValGluValAlaSe

3001 CGATACCCGCATCCCGCGCGGATTTGGCCTGAACCTGCCAGCTGGCGCAGGTAGCAGAGCGGTAACCTGGCTCGGATTAGGGCCGCAAGAAAACCTATCCC

907> rAspThrProHisProAlaArgI leGlyLeuAsnCysGlnLeuAlaGlnValAlaGluArgValAsnTrpLeuGlyLeuGlyProGlnGluAsnTyrPro

3101 GACCGCTTACTGCGCCCTGTTTGAACCGCTGGGATCTGCCATTGTTCAGACATGGGATACCCCGTACGCTTCCCGAGCGAAAACGGTCTGCGCGGGA

941> AspArgLeuThrAlaAlaCysPheAspArgTrpAspLeuProLeuSerAspMetTyrThrProTyrValPheProSerGluAsnGlyLeuArgCysGlyT

3201 CGCGCGAATTGAATATGCCCCACACAGTGGCGCGGCACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGATGGAACACGCCATCGCCA

974> hrArgGluLeuAsnTyrGlyProHisGlnTrpArgGlyAspPheHisAsnI leSerArgTyrSerGlnGlnLeuMetGluThrSerHisArgHi

3301 TCTGCTGACGCGGAAGAAGGCACATGGCTGAATTCGACGGTTTCCATATCGAGATGGGATGGTGGCGAGACTCTGGAGCCGCTCAGTATCGCGGAATTA

1007> sLeuLeuHisAlaGluGlyThrTrpLeuAsnI leAspGlyPheHisMetGlyI leGlyGlyAspSerTrpSerProSerValSerAlaGluLeu

EcoRI (3465)

3401 CAGCTGAGCGCCGGTCTGCTACCATTACAGTTGGTCTGGTGTCAAAAATAAATACTAGTCGAGAATTCCGCTAGCTCGACATGATAAGATACATTGATGA

1041> GlnLeuSerAlaGlyArgTyrHisTyrGlnLeuValTrpCysGlnLys••••

3501 TTTTGGACAACCACAACATAAGATGCAGTGAATAAAGGCTTATTGTTGAAATTTGTGATGCTATTGCTTTATTGTTGAAATTTGTGATGCTATTGCTT

3601 TATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTACAGTTACAGGGGAGGTGTGGGAGGTTTTTAAAG

PacI (3746)

SwaI (3736)

3701 CAAGTAAAACCTCTACAAATGTGGTAGATCCATTAAATGTTAATTAAGTACCCATGACCAAATCCCTTAACTGAGTTCCTTCCACTGAGCGTCAG
 3801 ACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCCCGCTACCAGCGGTGGTTTG
 3901 TTGCGCGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGC
 4001 CACCACCTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGT
 4101 TGGACTCAAGACGATAGTTACCCGATAAGGCGCAGCGGTCCGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAAGT
 4201 GAGATACCTACAGCGTGTAGCTATGAGAAAGCGCCAGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGC
 4301 ACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCTGTCCGGTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGTGCTCGTCAGGGG

PacI (4486)

4401 GCGGGAGCCTATGAAAAACGCCAGCAACGCGGCCCTTTTACGGTTCCTGGCCTTTTGTGGCTTTTGCTCACATGTTCTTAATTAATTTTTCAAAG
 4501 TAGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACCAGTGTGTCCAGTGC
 4601 TCACAGCCAGGGATGTGGCTGGAGCTGTTGAGTCTGGACTGACAGGTTGGGGTCTCCAGAGATTTGTGGAGGATGACTTTGCAGGTGTGGTCAGAGA
 11 MetAlaLysLeuThrSerAlaValProValL
 4701 TGATGTCACCCCTGTTTCATCTCAGCAGTCCAGGACCAGGTGGTGCCTGACAACACCCTGGCTTGGGTGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAG
 44 pAspValThrLeuPheIleSerAlaValGlnAspGlnValValProAspAsnThrLeuAlaTrpValTrpValArgGlyLeuAspGluLeuTyrAlaGlu
 4801 TGGAGTGAGGTGGTCTCCACCAACTTCAGGGATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAGTTTGCCTGAGAGACC
 78 TrpSerGluValValSerThrAsnPheArgAspAlaSerGlyProAlaMetThrGluIleGlyGluGlnProTrpGlyArgGluPheAlaLeuArgAspP

PacI (5002)

4901 CAGCAGGCAACTGTGTGCACTTTGTGGCAGAGGAGCAGGACTGAGGATAAGAATTGAGTTTCAGAAAAGGGGGCCTGAGTGGCCCTTTTTTCAACTTAA
 111 roAlaGlyAsnCysValHisPheValAlaGluGluGlnAsp • • •
 5001 TTAA