



SgfI (11) AgeI (90)
1 GGATCTGCGATCGCTCCGGTCCCGTCAGTGGCGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGCAATTGAACCGGTGCCA
101 GAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (246) PvuII (242) Bsu36I (293)
201 GTGAACGTTCTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTAGAGGCC
301 GCCATCCACGCGGGTTGAGTGCCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTGTTCGTTT

AgeI (553)
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGTAGGAGGCCAGCATGCCACATACTTTGTGGATGGTGTGGGTC
601 TTTGGGGTTCATCATCAGCCTCTCCAAGGAAGAATCTCCAATCAGGCTTCTGTCTGTGACCGCAATGGTATCTGCAAGGGCAGCTCAGGATCTTAA
1101 LeuGI yVal I l e l e Ser LeuSer LysGI uGI uSer SerAsnGI nAl aSer LeuSer CysAspArgAsnGI y l l e CysLysGI ySer LeuGI ySer LeuA
701 ACTCCATTCCTCAGGCTCACAGAAGCTGTAAGAACCTTACCTGTCCAACAACAGGATCCTACATTAGCAACAGTACAGAGGTTGTGTGAA
440 snSer l l e P roSer GI yLeuThr GI uAl aVal l LysSer LeuAspLeuSerAsnAsnArg l l e Thr Tyr l l e SerAsnSerAspLeuGI nArgCysVal l As
801 CCTCCAGGCTCTGGTGTGACATCCAATGGAATTAACAACAATAGAGGAAAGATTCTTTTCTCCCTGGGCGAGCTTGAACATTTAGACTTATCCTATAAT
770 nLeuGI nAl aLeuVal l LeuThr SerAsnGI y l l e AsnThr l l e GI uGI uAspSer PheSer Ser LeuGI ySer LeuGI uHi sLeuAspLeuSer TyrAsn

AvrII (984)
901 TACTTATCTAATTTATCGTCTTCTGTTCAAGCCCTTTCTTCTTAACTTACTGGAAATCCTTACAAAACCTTAGGGGAAACATCTC
1110 TyrLeuSerAsnLeuSer Ser Ser TrpPheLysP roLeuSer Ser LeuThr PheLeuAsnLeuLeuGI yAsnP roTyrLysThr LeuGI yGI uThr Ser l
1401 TTTTCTCTCACTCACAAAATTCGAATCTGAGAGTGGAAATATGGACACCTTCACTAAGATTCAAAGAAAAGATTTGCTGGACTACCTCTCTG
1440 euPheSer Hi sLeuThr LysLeuGI n l l eLeuArgVal l GI yAsnMetAspThr PheThr Lys l l eGI nArgLysAspPheAl aGI yLeuThr PheLeuGI

BglIII (1122)
1101 GGAACCTGAGATTGATGCTTCAGATCTACAGACTATGAGCCAAAAGTTTGAAGTCAATCAGAATGTAAGTCACTGATCCTTCATATGAAGCAGCAT
1770 uGI uLeuGI u l l eAspAl aSer AspLeuGI nSer TyrGI uP roLysSer LeuLysSer l l eGI nAsnVal l Ser Hi sLeu l l eLeuHi sMetLysGI nHi s
1201 ATTTTACTGCTGGAGATTTTGTAGATGTTACAAGTTCCTGGAATGTTTGAAGTCCGAGACTGATTTGGACACTTTCATTTTTCAGAACTATCCA
2110 l l eLeuLeuLeuGI u l l ePheVal l AspVal l Thr Ser Ser Val l GI uCysLeuGI uLeuArgAspThr AspLeuAspThr PheHi sPheSer GI uLeuSer T
1301 CTGGTGAACAAAATTCATTGATTAAGTTTACATTTAGAAATGTGAAAATCACCAGTAAAAGTTTGTTCAGGTTATGAACTTTTGAATCAGATTTT
2440 hr GI yGI uThrAsnSer Leu l l eLysLysPheThr PheArgAsnVal l Lys l l eThrAspGI uSer LeuPheGI nVal l MetLysLeuLeuAsnGI n l l eSe
1401 TGGATTTTAGAATGAGATTTGATGACTGTACCTTAAATGAGATTTGTAATTTTAGAGCATCTGATAATGACAGAGTTATAGACTCAGGTAAGTGGAA
2770 r GI yLeuLeuGI uLeuGI uPheAspAspCysThr LeuAsnGI yVal l GI yAsnPheArgAl aSer AspAsnAspArgVal l l eAspP roGI yLysVal l GI u

HpaI (1505) BspEI (1511)
1501 ACGTTAACATCCGGAGGCTGCATATTCGAAGTTTTACTTATTTTATGATCTGAGCATTATATTCACCTACAGAAAAGATTAAGAATCACAGTAG
3110 Thr LeuThr l l eArgArgLeuHi s l l eP roArgPheTyrLeuPheTyrAspLeuSer Thr LeuTyrSer LeuThr GI uArgVal l LysArg l l eThr Val l G
1601 AAAACAGTAAAGTTTTCTGGTCTTGTACTTTCAACAATTTAAATCATTAGAATCTTGGATCTCAGTAAAAATTTGATGGTTGAAGAATCACT
3440 l l uAsnSer LysVal l PheLeuVal l P roCysLeuLeuSer GI nHi sLeuLysLeuGI uTyrLeuAspLeuSer GI uLeuLeuMetVal l GI uGI uTyrLe
1701 GAAAAATTCAGCCTGTGAGGATGCTTGGCCCTCTCAAAAATTTAATTTTAAAGCAAAATCATTGGCATCTTGGAAAAACCGGAGAGACTTGTCTC
3770 uLysAsnSer Al aCysGI uAspAl aTrpP roSer LeuGI nThr Leu l l eLeuArgGI nAsnHi sLeuAl aSer LeuGI uLysThr GI yGI uThr LeuLeu

EcoRV (1827)
1801 ACTCTGAAAACTTGACTAACATTGATATCAGTAAGAAATAGTTTTTACTTATTTTATGATCTGAGCATTATATTCACCTACAGAAAAGATTAAGAATCACAGTAG
4110 Thr LeuLysAsnLeuThrAsn l l eAsp l l eSer LysAsnSer PheHi sSer MetP roGI uThr CysGI nTrpP roGI uLysMetLysTyrLeuAsnLeuS
1901 CCAGCACAGAAATACACAGTGAACAGGCTGCATTTCCAAAGACACTGAAAATTTAGATGTTAGCAACAACAATCTCAATTTATTTCTTGAATTTGCC
4440 er Ser Thr Arg l l eHi sSer Val l Thr GI yCys l l eP roLysThr LeuGI u l l eLeuAspVal l SerAsnAsnAsnLeuAsnLeuPheSer LeuAsnLeuP r
2001 GCAACTCAAAGAACTTTATATTTCCAGAAATAAGTTGATGACTCTACAGATGCCTCCCTCTTACCATGTTACTAGTATTGAAATCAGTAGGAATGCA
4770 oGI nLeuLysGI uLeuTyr l l eSer ArgAsnLysLeuMet Thr LeuP roAspAl aSer LeuLeuP roMetLeuLeuVal l LeuLys l l eSer ArgAsnAl a

EcoRI (2188)
2101 ATAACACGTTTTCTAAGGAGCAACTTGACTCATTTCACACACTGAAGACTTTGGAAGCTGGTGGCAATAACTTCATTTGCTCCTGTGAATTCCTCTCT
5110 l l eThr Thr PheSer LysGI uGI nLeuAspSer PheHi sThr LeuLysThr LeuGI uAl aGI yGI yAsnAsnPhe l l eCysSer CysGI uPheLeuSer P
2201 TCATCAGGAGCAGCAAGCACTGGCCAAAAGTCTTGGATTTGGCCAGCAAAATACCTGTGACTCCATCCCATGTCGCGCCAGCAGGTTTCAGGA
5440 heThr GI nGI uGI nGI nAl aLeuAl aLysVal l Leu l l eAspTrpP roAl aAsnTyrLeuCysAspSer P roSer Hi sVal l ArgGI yGI nVal l GI nAs
2301 TGTCGCTCTCGGTGTCGGAATGTACAGGACAGCAGCTGTGTCTGGCATGTGCTGTCTGTTCTGCTGATCCTGCTCAGGGGGTCTGTGCCAC
5770 pVal l ArgLeuSer Val l Ser GI uCysHi sArgThr Al aLeuAl aSer GI yMetCysCysAl aLeuPheLeuLeu l l eLeuLeuThr GI uLeuCysHi s

NcoI (2406)
2401 CGTTCATCCAGGCTGTGGTATATGAAATGATGTGGCCTGGCTCCAGGCCAAAAGGAAGCCAGGAAAGCTCCAGCAGGAACATCTGCTATGATGCAT
6110 ArgPheHi sGI yLeuTrpTyrMetLysMetMetTrpAl aTrpLeuGI nAl aLysArgLysP roArgLysAl aP roSerArgSn l l eCysTyrAspAl aP
2501 TTGTTTCTACAGTGAAGGATGCTTGGTGGGAAACCTTATGGTCCAGGAGCTGGAGAACTTCAATCCCCCTTCAAGTTGTCTTCATAAGCG
6440 heVal l Ser TyrSer GI uGI nAspAl aTyrTrpVal l GI uAsnLeuMetVal l GI nGI uLeuGI uAsnPheAsnP roP roPheLysLeuCysLeuHi sLysAr
2601 GGACTTCATTCCTGGCAAGTGGATCATTGACAATATCATTGACTCCATTGAAAAGAGCCACAAAATCTGTTTGTCTTTCTGAAAATTTGTGAAGAT
6770 gAspPhe l l eP roGI yLysTrp l l e l l eAspAsn l l e l l eAspSer l l eGI uLysSer Hi sLysThr Val l PheVal l LeuSer GI uAsnPheVal l LysSer
2701 GAGTGGTCAAGTATGAAGTGGACTTCTCCATTTCCGCTTTTGTGATGAGAACAATGATGCTGCCATTCTATTCTCTGGAGCCATTGAGAAAAAG
7110 GI uTrpCysLysTyrGI uLeuAspPheSer Hi sPheArgLeuPheAspGI uAsnAsnAspAl aAl a l l eLeu l l eLeuLeuGI uP ro l l eGI uLysLysA

NcoI (2864)
2801 CCATTCACAGGCTTCTGCAAGCTGCGGAAGATAATGAACACCAAGACCTACCTGGAGTGGCCCATGGACGAGGCTCAGCGGGAAGGATTTGGGTA
7440 l a l l eP roGI nArgPheCysLysLeuArgLys l l eMetAsnThrLysThr TyrLeuGI uTyrP roMetAspGI uAl aGI nArgGI uAl aPheTrpVal l As
2901 TCTGAGAGCTGCGATAAAGTCCCTGTACAAGGtaagtcactgactgtctatgctcgggaagggtggcaggagatggggcagtgaggaaaagtggca
7770 nLeuArgAl aAl a l l eLysSer LeuTyrLysGI yLysSer LeuThr Val l TyrAl aTrpGI uArgVal l GI yArgArgTrpGI ySerAl aGI yLysVal l Al a

AseI (3024)
3001 ctatgaaccACTAGTTTGACAATTAATCATAAGCATAGTATAATAACTACTATAGcaattgtactaaccttcttctcttctctctcctcctgacagGAG
8110 Leu•••Thr Hi s•••PheAspAsn•••Ser•••Al a•••TyrAsnThr Thr Hi sTyrSerAsnCysThrAsnLeuLeuLeuPheP roLeuLeuThr GI yG

BbrPI (3270)
3101 GAGCCATCATGGCCGCTATGGAGATCGAGTGGCCGATCACCGCACCTGAAAGCGCTGGAGTTTCGAGCTGGTGGGCGGGAGAGGGCACCCCGAGCA
8440 l yAl a l l eMetAl aAl aMetGI u l l eGI uCysArg l l eThr GI yThr LeuAsnGI yVal l GI uPheGI uLeuVal l GI yGI yGI yGI uGI yThr P roGI uGI

3201 GGGCCGATGACCAACAGATGAAGAGCACCAAGGCGCCTGACCTTCAGCCCTACCTGCTGAGCCAGCTGATGGCTACGGCTTCTACCCTTCGGC
8770 nGI yArgMetThrAsnLysMetLysSer Thr LysGI yAl aLeuThr PheSer P roTyrLeuLeuSer Hi sVal l MetGI yTyrGI yPheTyrHi sPheGI y

XmnI (3325) BbrPI (3402)

3301 ACCTACCCAGCGGCTACGAGAACCCTTCTGCACGCCATCAACAACGGCGGTACACCAACCCCGCATCGAGAAGTACGAGGACGGCGGCTGTGC
911 Thr Tyr Pro Ser Gl y Tyr Gl u Asn P ro Phe Leu Hi s Al a l l e Asn Asn Gl y Gl y Tyr Thr Asn Thr Arg l e Gl u Lys Tyr Gl u Asp Gl y Gl y Val Leu H
3401 ACGTGAGCTTCAGCTACCGCTACGAGGCCGCGCGTGATCGGCGACTTCAAGGTGATGGCACC GGCTTCCCGAGGACAGCGTGATCTTCACCGACAA
944 i s Va l Ser Phe Ser Tyr Arg Tyr Gl u Al a Gl y Arg Va l l e Gl y Asp Phe Lys Va l Me t Gl y Thr Gl y Phe P ro Gl u Asp Ser Va l l e Phe Thr Asp Ly
3501 GATCATCCGACGCAACGCCACCGTGGAGCACCTGCACCCTATGGCGATAACGATCTGGATGGCAGCTTCAACCGCACCTTCAGCTCGCGCAGCGCGGC
977 s l l e l l e Arg Ser Asn Al a Thr Va l Gl u Hi s Leu Hi s P ro Me t Gl y Asp Asn Asp Leu Asp Gl y Ser Phe Thr Arg Thr Phe Ser Leu Arg Asp Gl y Gl y

PstI (3665)

3601 TACTACAGCTCCGTGGTGGACAGCCACATGCACCTTCAAGAGCGCCATCCACCCAGCATCCTGCAGAACGGGGGCCCATGTTCCGCTTCCGCGCGTGG
1011 Tyr Tyr Ser Ser Va l Va l Asp Ser Hi s Me t Hi s Phe Lys Ser Al a l l e Hi s P ro Ser l l e Leu Gl n Asn Gl y Gl y P ro Me t Phe Al a Phe Arg Arg Va l G

NheI (3779)

3701 AGGAGGATCACAGCAACCCAGCTGGGCATCGTGGAGTACCAAGCAGCCTTCAAGACCCCGGATGCAGATGCCTAAAGCTAGCTGGCCAGACATGATAA
1044 l u Gl u Asp Hi s Ser Asn Thr Gl u Leu Gl y l l e Va l Gl u Tyr Gl n Hi s Al a Phe Lys Thr P ro Asp Al a Asp Al a ●●●

3801 GATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATAT

HpaI (3919)

3901 AAGCTGCAATAAACAAGTTAAACAACAACAATTGCATTCATTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACTCTAC

EcoRI (4013)

4001 AAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGC
4101 ATCAGGGGCTGTGGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTCCCAAGTTTGAAC TAGCTCTT

SwaI (4269)

4201 CATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTT
4301 TTTATTAGGCAAGATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAAATGGACAG

4401 CAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCTTCACTCAATGAGCACA
141 ●●● Asn Arg Thr Tyr Lys Leu Pro l l e Leu Gl u Gl u l l e Thr Thr Lys Va l Leu Lys Gl y Asn Me t Gl u l l e Leu Va l P
4501 AAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGA
114 h e Cys Asp P ro Al a Tyr Asp Ser l l e Leu Gl u Arg Cys Me t Gl y Cys P ro Ser Va l Va l Arg l l e Ser Arg Asp Va l Gl u Asp Ser Tyr P ro Hi s Arg Va
4601 CAGCCACAATGGTGTCAAAGTCCCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAAT
81 l Al a Va l l l e Thr Asp Phe Asp Lys Gl n Gl y Asn Ser Va l Al a Ser Gl y l l e Al a l l e Al a Gl u Al a Cys Va l Thr Va l Arg Gl y l l e Tyr Al a Gl u l l e
4701 GTGGACAGCAGAGATGATCTCCCACTTGGTCTGTAGGCGCCCGACATGGTGTGTTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGACC
48 H i s Va l Al a Ser l l e l l e Gl u Gl y Thr Lys Thr Arg l l e Al a Al a Gl y Va l Hi s Hi s Lys Asn Asp Gl u Tyr Leu Me t Thr l l e Lys Gl u Thr Al a Va l G

BspHI (4841)

XmnI (4837) AseI (4900)

4801 TCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGAT
14 l u Va l Leu Gl u Leu Asp Gl n Gl n Ser l l e Asn Phe Thr Lys Me t
4901 TAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACAGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCA

5001 TTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTACTAGTCAAACAAACTCCCATTGACGTCAATGGGTGGAGACTTGG
5101 AAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAA
5201 GTAGGAAAGTCCCAATAGGTCATGTACTGGCATAATGCCAGCGGGCCATTACCGTCAATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTT
5301 GATGTA CTGCCAAGTGGGCAGTTTACCGTAAATCTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGA

PaeI (5477)

PstI (5470) BspLU11I (5483)

SdaI (5470)

5401 CGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCC
5501 AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGTCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG
5601 GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGTTACCGGATACCTG
5701 TCCGCTTTCTCCCTCCGGGAAGCGTGCCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTCCGCTCAAGCTGGGCTGTGTGC
5801 ACGAACCCCGGTTACGCCGACCGCTGCGCTTATCCGGTAACATCGTCTTGAAGTCCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCAC
5901 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTGGTATC
6001 TGCGCTGCTGTAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTGTTGCAAGC
6101 AGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTT

PaeI (6217) SwaI (6225) **NotI (6233)**

6201 GGTATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACAA
6301 CATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCAGAACATTTCTCTATCGAA