



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **MfeI (82)**  
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

**Psp1406I (203)** 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGGC **HindIII (245)** **Bsu36I (291)**  
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

---

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTGGCTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCGTTT **NgoMIV (441)**  
**NaeI (441)**

---

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCAGTGGCAGTGGAAATGGCCTGCCAGGTGCCCTGG **NcoI (560)**  
**XcmI (560)**  
**BstEII (555)**  
**AgeI (552)** 1▶Me tAl aValMe tGl uMe tAl aCysP roGl yAl aP roGl  
601 CTCAGCAGTGGGGCAGCAGAAGAACTCCCAAAGCCAAGGAGAAGACGCCCACTGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAGGCCGTGGAG **HindIII (682)**  
13▶ySer Al aVal Gl yGl nGl nLysGl uLeuP roLysAl aLysGl uLysThr P roP roLeuGl yLysLysGl nSer Ser Val TyrLysLeuGl uAl aVal Gl u  
701 AAGAGCCCTGTGTTCGGAAAGTGGGAGATCCTGAATGACGTGATTACCAAGGGCAGCCAAAGGAGCTCCGAGGCAGGGCCAGCTGCCATCTCTA  
47▶LysSer P roVal PheCysGl yLysTrpGl u l eLeuAsnAspVal l l eThr LysGl yThr Al aLysGl uGl ySer Gl uAl aGl yP roAl aAl l l eSer l  
801 TCATCGCCAGGCTGAGTGTGAGAATAGCAAGAGTTCAGCCACCTTTTTCAGAACGCATTTTCATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAG  
80▶l e l l eAl aGl nAl aGl uCysGl uAsnSer Gl nGl uPheSer P roThr PheSer Gl uArg l l ePhe l l eAl aGl ySer LysGl nTyrSer Gl nSer Gl uSe

---

901 TCTTGATCAGATCCCAACAATGTGCCCATGTCTACAGAGGGCAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTCGCAGCAAAGCCCGGAAGAAACGG **BstXI (914)**  
113▶r LeuAspGl n l eP roAsnAsnVal l Al aHi sAl aThr Gl uGl yLysMe tAl aArgVal CysTrpLysGl yLysArgArgSer LysAl aArgLysLysArg  
1001 AAGAAGAAGAGCTCAAAGTCCCTGGCTCATGCAGGAGTGGCCTTGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGG  
147▶LysLysLysSer Ser LysSer LeuAl aHi sAl aGl yVal l Al aLeuAl aLysP roLeuP roArgThr P roGl uGl nGl uSer CysThr l l eP roVal Gl nG

---

1101 AGGATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTGAAGGAACAGGCCCTGGGCAACTCTGTTTTAAGCAGCT **NdeI (1124)**  
180▶l uAspGl uSer P roLeuGl yAl aP roTyrVal l ArgAsnThr P roGl nPheThr LysP roLeuLysGl uP roGl yLeuGl yGl nLeuCysPheLysGl nLe

---

1201 TGGCGAGGGCCTACGGCCGGCTCTGCCTCGATCAGAATCCACAACTGATCAGCCCTTGAATGTCTGAACCACGTGTGAAACTGCACCACCCCCAG **NgoMIV (1215)**  
**NaeI (1215)** **PmlI (1273)**  
**EagI (1213)** **Eco72I (1273)**  
**BbrPI (1273)**  
1213▶uGl yGl uGl yLeuArgP roAl aLeuP roArgSer Gl uLeuHi sLysLeu l l eSer P roLeuGl nCysLeuAsnHi sVal TrpLysLeuHi sHi sP roGl n  
1301 GACGGAGGCCCTGCCCTGCCACGCACCCCTCCCTATAGCAGACTGCCTCATCCCTCCCATCCACCCTCCAGCCCTGGAAACCTCACCCCTC  
247▶AspGl yGl yP roLeuP roLeuP roThr Hi sP roPheP roTyrSerArgLeuP roHi sP roPheP roPheHi sP roLeuGl nP roTrpLysP roHi sP roL  
1401 TGGAGTCTTCTGGCAAAGTGGCCTGTGTAGACAGCCAAACCTTGCTGACCCACACCTGAGCAAAGTGGCCTGTGTAGACAGTCCAAAGCCCT  
280▶euGl uSer PheLeuGl yLysLeuAl aCysVal l AspSer Gl nLysP roLeuP roAspP roHi sLeuSer LysLeuAl aCysVal l AspSer P roLysP roLe

---

1501 GCCTGGCCACACCTGGAGCCAGCTGCCTGTCTCGTGGTGGCCATGAGAAGTTTTCTGTGGAGGAATACCTAGTGCATGCTGCAAGGCAGCGTGAGC **SphI (1575)**  
313▶uP roGl yP roHi sLeuGl uP roSer CysLeuSer ArgGl yAl aHi sGl uLysPheSer Val l Gl uGl uTyrLeuVal l Hi sAl aLeuGl nGl ySer Val Ser

---

1601 TCAGGCCAGGCCACAGCCTGACCAGCCTGGCCAAGACCTGGGAGCAAGGGCTCCAGATCCCGGGAGCCAGCCCAAACTGAGGACAACGAGGGTG **XmaI (1661)**  
**SmaI (1661)**  
347▶Ser Gl yGl nAl aHi sSer LeuThr Ser LeuAl aLysThr TrpAl aAl aArgGl ySer ArgSer ArgGl uP roSer P roLysThr Gl uAspAsnGl uGl yV  
1701 TCCTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAAGTCCACTGGGCCACGCACCAGCTCCGCCTGGGAGAGGCTCCTCGGAGA  
380▶al LeuLeuThr Gl uLysLeuLysP roVal l AspTyrGl uTyrArgGl uGl uVal l Hi sTrpAl aThr Hi sGl nLeuArgLeuGl yArgGl ySer PheGl yGl

---

1801 GGTGCACAGGATGGAGGACAAGCAGACTGGCTTCCAGTGCCTGTCAAAAAGGTGCGGCTGGAAGTATTTTCGGGCAGAGGAGCTGATGGCATGTGCAGGA **XmnI (1861)**  
413▶uVal l Hi sArgMetGl uAspLysGl nThr Gl yPheGl nCysAl aVal l LysLysVal l ArgLeuGl uVal l PheArgAl aGl uGl uLeuMetAl aCysAl aGl y  
1901 TTGACCTCACCCAGAATTGTCCCTTTGTATGGAGCTGTGAGAGAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGC  
447▶LeuThr Ser P roArg l l eVal l P roLeuTyrGl yAl aVal l ArgGl uGl yP roTrpVal l Asn l l ePheMe tGl uLeuLeuGl uGl yGl ySer LeuGl yGl nL  
2001 TGGTCAAGGAGCAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTGGAGGGTCTGGAATACCTCCACTCACGAAGGATTCTGCA  
480▶euVal l LysGl uGl nGl yCysLeuP roGl uAspArgAl aLeuTyrTyrLeuGl yGl nAl aLeuGl uGl yLeuGl uTyrLeuHi sSer ArgArg l l eLeuHi  
2101 TGGGGAGCTCAAAGCTGACAACTGCTCCTGTCCAGCATGGGAGCCAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATGGCCTG  
513▶sGl yAspVal l LysAl aAspAsnVal l LeuLeuSer SerAspGl ySer Hi sAl aAl aLeuCysAspPheGl yHi sAl aVal l CysLeuGl nP roAspGl yLeu

---

2201 GGAAAGTCCTTGCTCACAGGGGACTACATCCCTGGCAGAGACCCACATGGCTCCGGAGGTGGTGTGGCAGGAGCTGCGACGCCAAGGTGGATGTCT **BspEI (2253)**  
547▶Gl yLysSer LeuLeuThr Gl yAspTyr l l eP roGl yThr Gl uThr Hi sMe tAl aP roGl uVal l Val l LeuGl yArgSer CysAspAl aLysVal l AspVal l T  
**BsrBI (2368)**  
2301 GGAGCAGCTGCTGTATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTCCGAGGGCCGCTCTGCCTCAAGATTGCCAGCGAGCCTCC  
580▶rP ser Ser CysCysMe tMe tLeuHi sMe tLeuAsnGl yCysHi sP roTrpThr Gl nPhePheArgGl yP roLeuCysLeuLys l l eAl aSer Gl uP roP r  
**DraIII (2479)**  
2401 GCCTGTGAGGGAGATCCCACCTCTGCGCCCTCTCACAGCCAGGCCATCCAAGAGGGGCTGAGGAAAGAGCCATCCACCGCTGTCTGCAGCGGAG  
613▶oP roVal l ArgGl u l l eP roP roSer CysAl aP roLeuThr Al aGl nAl a l l eGl nGl uGl yLeuArgLysGl uP ro l l eHi sArgVal l Ser Al aAl aGl u  
2501 CTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAG  
647▶LeuGl yGl yLysVal l AsnArgAl aLeuGl nGl nVal l Gl yGl yLeuLysSer P roTrpArgGl yGl uTyrLysGl uP roArgHi sP roP roP roAsnGl nA

2601 CCAATTACCACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCCAGGGCCCCGCCAGCTGAGGAGACAACAGGCAGAGCCCCAAGCT  
680▶ I aAsnTyrHi sGl nThr LeuHi sAl aGl nP roArgGl uLeuSer P roArgAl aP roGl yP roArgP roAl aGl uGl uThr Thr Gl yA rgAl aP roLysLe  
2701 CCAGCCTCTCTCCACCAGAGCCCCAGGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTACCTCTGTCC  
713▶ uGl nP roP roLeuP roP roGl uP roP roGl uP roAsnLysSer P roP roLeuThr LeuSer LysGl uGl uSer Gl yMe tT rpGl uP roLeuP roLeuSer  
BsrBI (2839)

2801 TCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAGCGGAAAGCAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAATAGAATTATTCCTCA  
747▶ Ser LeuGl uP roAl aP roAl aArgAsnP roSer Ser P roGl uArgLysAl aThr Val P roGl uGl nGl uLeuGl nGl nLeuGl uI l eGl uLeuPheLeuA  
2901 ACAGCCTGTCCAGCCATTTCTCTGGAGGAGCAGGAGCAAATCTCTCGTGCCTCAGCATCGACAGCCTCTCCCTGTGGATGACAGTGAGAAGAACC  
780▶ snSer LeuSer Gl nP roPheSer LeuGl uGl uGl nGl uGl nI l eLeuSer CysLeuSer I l eAspSer LeuSer LeuSer AspSer Gl uLysAsnP r  
XcmI (3080)

3001 ATCAAAGGCTCTCAAAGCTCGCGGGACACCTGAGCTCAGGCGTACACTCCTGGAGCAGCCAGGCCAGGCTCGAAGCTCCAGCTGGAACATGGTGCTG  
813▶ oSer LysAl aSer Gl nSer Ser ArgAspThr LeuSer Ser Gl yVal Hi sSer T rpSer Ser Gl nAl aGl uAl aArgSer Ser Ser T rpAsnMe tVal Leu  
XmaI (3101)  
SmaI (3101)

3101 GCCCGGGGGCGGCCACCACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACC  
847▶ Al aArgGl yA rgP roThrAspThr P roSer TyrPheAsnGl yVal LysVal I Gl nI l eGl nSer LeuAsnGl yGl uHi sLeuHi s I l eArgGl uPheHi sA  
BstEII (3264)

3201 GGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAGCTGCAGCCTTCAAGTTGGTGACCAAAGACGGGCGAGCCTGTTCTGCATCAGCAT  
880▶ r gVal LysVal I Gl yAspI l eAl aThr Gl yI l eSer Ser Gl nI l eP roAl aAl aAl aPheSer LeuVal I Thr LysAspGl yGl nP roVal A rgTyrAspMe  
3301 GGAGGTGCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGG  
913▶ tGl uVal P roAspSer Gl yI l eAspLeuGl nCysThr LeuAl aP roAspGl ySer PheAl aT rpSer T rpArgVal I LysHi sGl yGl nLeuGl uAsnArg  
NheI (3410)

3401 CCCTAACCTGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATT  
947▶ Pro•••

HpaI (3548) MfeI (3559)

3501 TGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAATTGCATTCATTTATGTTTCAGGTTCCAGGGGAGGTGT

EcoRI (3644)

3601 GGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGA  
3701 ATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATAT

SspI (3883) SmaI (3897)

3801 AGTGATTTTTCCAAGGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAAATTCAGAAAAAATT  
3901 TAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGG  
4001 GAACAAAGGAACCTTTAATAGAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTT  
141•••AsnArgThr TyrLysLeuProl l eLeuGl uGl uI l eThr Thr Ly  
BstXI (4187)

4101 TGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGA  
125▶ sVal LeuLysGl yAsnMe tGl uI l eLeuVal PheCysAspP roAl aTyrAspSer I l eLeuGl uArgCysMetGl yCysP roSer Val Val A rgI l eSer  
4201 TCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCA  
92▶ ArgAspVal I Gl uAspSer TyrP roHi sArgVal Al aVal I l eThr AspPheAspLysGl nGl yAsnSer Val Al aSer Gl yI l eAl aI l eAl aGl uAl aC  
4301 CAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTGTCT  
58▶ ysVal I Thr Val A rgGl yI l eTyrAl aGl uI l eHi sVal Al aSer I l eI l eGl uGl yThr LysThr ArgI l eAl aAl aGl yVal Hi sHi sLysAsnAspGl  
BspHI (4472)  
XmnI (4464)

4401 CATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTCATGATGGCCCTCTATAGTGAGTCGT  
25▶ uTyrLeuMetThr I l eLysGl uThr Al aVal I Gl uVal LeuGl uLeuAspGl nGl nSer I l eAsnPheThr LysMe t

VspI (4530)  
AseI (4530)

4501 ATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTATCTGACGGTTCATAAACGAGCTCTGCTTAT

SpeI (4685)

4601 ATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTGTACGACATTTTGGAAAGTCCCCTTGATTTACTAGTCAAACAA  
4700 ACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTA

SnaBI (4813)  
Eco105I (4813)

4800 ATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTGATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGT  
NdeI (4918)

4900 CAATAGGGGGCTACTTGGCATATGATACACTGTACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTAT

SdaI (5096)

5000 TGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTGACCCAGCGGGCCATTTACCGTAAGTTATGTAACGCC T

PaeI (5104) BspLU11I (5114)

5100 G C A G G T T A A T T A A G A A C A T G T G A G C A A A A G G C C A G C A A A A G G C C A G G A A C C G T A A A A A G G C C G T T G C T G G C G T T T T C C A T A G G C T C C G C C C C C T G  
5199 A C G A C A T C A C A A A A T C G A C G C T C A A G T C A G A G G T G G C G A A A C C G A C A G G A C T A T A A A G A T A C C A G G C G T T C C C C C T G G A A G C T C C C T C G T G C G C T C  
5299 T C C T G T T C C G A C C C T G C C G T T A C C G G A T A C T G T C C G C T T T C C C T T C G G A A G C G T G G C G T T T C T A T A G C T A C G C T G T A G G T A T C T C A G T T C G

5399 GTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG  
5499 TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA  
5599 CTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACC  
5699 ACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACG

EagI (5864)  
PacI (5844) SmaI (5853) **NotI (5863)**

5799 CTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACAT  
5899 CTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAA  
5999 GTGCAGGTGCCAGAACATTTCTCTATCGAA