



**BspI20I** (7)  
**PstI** (6)  
**SdaI** (6)      **SpeI** (13)  
1 CCTGCAGGGCCCACTAGTCTAGCCCTGATCAGGGGCGGTAGGACAAAGAGAGGGAGACAGAAACACACACCCAAACGGCTTGAGACTGCTAGGTGCTTCA

---

**AseI** (139)

101 CTTCTCTGGCCCTGTTTCTGATGAGGAGGTTTGTATTAATAATAAGTCTATTTCTGTAGCCCCTACTATGTGCTCTGTACTACTACCTAGCCCTC

---

**SphI** (261)      **BspI20I** (279)

201 CATCAAGAGGGATCTGCCCTCTCATCTGGCTGACCCCAATCGAACCCCTTAACCCGCATGCTCTACTGCCTCCCTGCCGGCCCTGCCTTATCCTCCTCA

---

301 ACCGCCACTGGGACACTCTGGAGAACTCGGAGTCTTGAATAGAAACCCAGGAGCCCTGAGCCTTGGCCCTCAGGCCCCAGGAGGGAGGTCAGGAG

---

**PshAI** (409) **EagI** (420)

401 GGAGGCGTGGACAAGGTCGGCCGTAGCTGGTCCCGTCTGGAAGCTGCTTAAGTTCTGGTCCACCACCCGGGCAAGAGTGATACCTGATCCTGGGG

---

**NaeI** (540)  
**SacII** (534)

501 ATTGTAAATGACCTCATGTGGCAGCCCAGCGGCTGCCGGCAAAGCCCTCCACCCTCCCTTCCCGCTGGCTCCACCCTACCCACGCCCCTCC

---

**SmaI** (656)  
**SrfI** (656)      **NcoI** (667)

601 CGCGCGCGGTTAAATCCCGCACCTGAGCATCGGCTCACACCTGCACCCGCCCGGCATAGCACCATGGGGGTTCTCATCATCATCATCATGG

---

**NheI** (705)      **Acc65I** (761)

701 TATGGCTAGCATGACTGGTGACAGCAAATGGGTGCGGATCTGTACGACGATGACGATAAGGTTACCTAAGGATCAGCTTGAGGTTGATCCCGTCTTTTA  
11▶ yMe tAl aSer Met Thr Gl yGl yGl nGl nMe tGl yA rg Asp Leu Tyr Asp Asp Asp Lys Val P ro Lys Asp Gl nLeu Gl yVal Asp P ro Val Val Leu  
801 CAACGCTGCTGACTGGGAAAACCTGGCGTTACCAACTAATCGCCTTGACGACATCCCCCTTCGCCAGCTGGCGTAAATAGCGAAGAGGCCGACCCG  
45▶ Gl nArg Arg Asp T rp Gl uAsn P ro Gl yVal Thr Gl nLeu Asn Arg LeuAl aAl aHi sP ro P roPheAl aSer T rpArgAsnSer Gl uGl uAl aArg Thr A

901 ATCGCCCTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTGGCTGGTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTGCGATCT  
78▶ spArg P roSer Gl nGl nLeu Arg Ser Leu Asn Gl yGl uT rpArg PheAl aT rpPheP roAl aP roGl uAl aVal P roGl uSer T rpLeu Gl uCys Asp Le  
1001 TCCTGAGGCGGATGCTGCTGCTCCCCCAAAGTGGCAGATGACCGTTACGATGCGCCATCTACCAACGTAACCTATCCATTACGGTCAATCCG  
111▶ uP roGl uAl aAsp Thr Val Val Val P roSerAsn T rpGl nMe tHi sGl yTyr AspAl aP roL eT yrThrAsnVal Thr TyrP roL eT yrValAsnP ro  
1101 CCGTTTGTCCCAGGAAATCCGACGCGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGCCAGCAGCGAATATTTTTGATG  
145▶ P roPheVal P roThr Gl uAsn P roThr Gl yCys TyrSer Leu Thr PheAsnVal AspGl uSer T rpLeu Gl nGl uGl yGl nThr ArgL e l e l ePheAspG  
1201 GCGTAACTCGGCGTTTTCATCTGTGGTGAACGGGCGTGGTGGTACGGCCAGGACAGTGGTTGCCGTCTGAATTTGACCTGAGCGATTTTTACG  
178▶ l yVal AsnSer Al aPheHi sLeu T rpCysAsnGl yArg T rpVal Gl yTyrGl yGl nAspSer ArgLeuP roSer Gl uPheAspLeuSer Al aPheLeuAr

**AatII** (1402)

1301 CGCGGAGAAAACCGCCTCGCGGTGATGGTGCCTGGAGTACCGGAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCGCTGAC  
211▶ gAl aGl yGl uAsn ArgLeuAl aVal Met Val Leu Arg T rpSer AspGl ySer TyrLeuGl uAspGl nAspMet T rpArgMetSer Gl y l e l ePheArgAsp  
1401 GTCTCGTGTGCATAAACCAGCTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTACGCCGCGCTGACTGGAGGCTGAAGTTC  
245▶ Val Ser LeuLeuHi sLysP roThr Thr Gl n l eSer AspPheHi sVal Al aThr ArgPheAsnAspAspPheSer ArgAl aVal LeuGl uAl aGl uVal G  
1501 AGATGTGGGCGAGTTGCGTGACTACCTACGGTAACAGTTTCTTTATGGCAGGGTAAACGCGAGTCCGACGCGCACCCGCGCTTTCGCGGTGAAAT  
278▶ l nMet CysGl yGl uLeuArgAspTyrLeuArgVal Thr Val Ser Leu T rpGl nGl yGl uThr Gl nVal Al aSer Gl yThr Al aP roPheGl yGl u l l

**ClaI** (1603)

1601 TATCGATGAGCGTGGTGGTTATGCCGATCGCTCACACTACGTTGAACTCGAAAACCCGAAACTGTGGAGCGCGAAATCCCGAATCTCTATCGTGCG  
311▶ e l l eAspGl uArgGl yGl yTyrAl aAspArgVal Thr LeuArgLeuAsnVal Gl uAsnP roLysLeuT rpSer Al aGl u l l eP roAsnLeuTyrArgAl a  
1701 GTGTTGAACTGCACCCCGACGGCAGCTGATTGAAGCAGAAGCTGCGATGTCGGTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTGCTGCTGA  
345▶ Val Val Gl uLeuHi sThr Al aAspGl yThr Leu l l eGl uAl aGl uAl aCysAspVal Gl yPheArgGl uVal A rg l l eGl uAsnGl yLeuLeuLeuA

**EcoRV** (1892)

1801 ACGCAAGCCGTTGCTGATTTCGAGGCGTTAACCCTCACGAGCATCATCTCTGCATGGTCAAGTATGGATGAGCAGACGATGGTGCAGGATATCCTGCT  
378▶ snGl yLysP roLeuLeu l l eArgGl yVal AsnArgHi sGl uHi sHi sP roLeuHi sGl yGl nVal MetAspGl uGl nThr MetVal Gl nAsp l l eLeuLe

**DraIII** (1969)

1901 GATGAAGCAGAACAATTAACGCCGTGCTGTTGCGATTATCCGAACCATCCGCTGTGGTACACGCTGTGGCAGCCGTACGGCCTGTATGTTGTTGGAT  
411▶ uMe tLysGl nAsnAsnPheAsnAl aVal ArgCysSer Hi sTyrP roAsnHi sP roLeuT rpTyrThr LeuCysAspArgTyrGl yLeuTyrVal Val Asp

**SspI** (2009)

2001 GAAGCAAATTGAACCACGGCATGGTCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGCGATGAGCGAACGCGTAACGCGAATGGTGC  
445▶ Gl uAl aAsn l l eGl uThr Hi sGl yMetVal P roMetAsnArgLeuThrAspAspP roArgT rpLeuP roAl aMetSer Gl uArgVal ThrArgMetVal l G

**BsaBI** (2105)

2101 AGCGCGATCGTAATCACCGAGTGTGATCATCTGGTGGTGGAATGAATCAGGCCACGGCGTAATCACGACGCGTGTATCGCTGGATCAAATCTGT  
478▶ l nArgAspArgAsnHi sP roSer Val l l e l l eT rpSerLeuGl yAsnGl uSer Gl yHi sGl yAl aAsnHi sAspAl aLeuTyrArgT rp l l eLysSer Va

**BbsI** (2296)

2201 CGATCCTTCCCGCCCGGTGCAGTATGAAGGCGGCGAGCCGACACCACGGCCACCGATATTATTTGCCGATGTACGCGCGGTGGATGAAGACCAGCCC  
511▶ l Asp P roSer Arg P roVal Gl nTyrGl uGl yGl yAl aAspThr Thr Al aThrAsp l l e l l eCysP roMet TyrAl aArgVal AspGl uAspGl nP ro  
2301 TTCCCGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTCGCTACCTGGAGAGACGGCCCGCTGATCCTTTGCCAATACGCCACGCGATGGGTAACA  
545▶ PheP roAl aVal P roLysT rpSer l l eLysLysT rpLeuSerLeuP roGl yGl uThrArgP roLeu l l eLeuCysGl uTyrAl aHi sAl aMetGl yAsnS  
2401 GTCTTGGCGGTTTCGCTAAATACTGCGAGGCTTCGTCAGTATCCCCGTTACAGGGCGCTCGCTGGGACTGGTGGATGATGATGATGATGATGATGAT  
578▶ erLeuGl yGl yPheAl aLysTyrT rpGl nAl aPheArgGl nTyrP roArgLeuGl nGl yGl yPheVal T rpAspT rpVal AspGl nSerLeu l l eLysTy  
2501 TGATGAAACCGCAACCCGTTGGTGGCTTACGGCGGTTGTTGGCAGTACGGCAACGATCGCCAGTCTGTATGACGGTCTGGTCTTTCGCCAGCCG  
611▶ rAspGl uAsnGl yAsnP roT rpSer Al aTyrGl yGl yAspPheGl yAspThr P roAsnAspArgGl nPheCysMetAsnGl yLeuVal l PheAl aAspArg

**Eco47III** (2614)

2601 ACGCCGATCAGCGCTGACGGAAGCAAAAACACAGCAGCAGTTTTTCCAGTCCGTTTATCCGGCAAAACCATCGAAGTGACAGCGAATACCTGTTC  
645▶ ThrP roHi sP roAl aLeuThr Gl uAl aLysHi sGl nGl nGl nPhePheGl nPheArgLeuSer Gl yGl nThr l l eGl uVal ThrSer Gl uTyrLeuPheA

**SacI** (2719)

2701 GTCATAGCGATAACGAGCTCCIGCACIGGATGGIGGGCIGGAIGGIAAGCCGIGGCAAGCGGIGAAGIGCCICTGGAIGICGCTCCACAAGGIAAACA  
678▶ r gHi sSer AspAsnGl uLeuLeuHi sTrpMetValAl aLeuAspGl yLysProLeuAl aSer Gl yGl uVal P roLeuAspValAl aP roGl nGl yLysGl  
2801 GTTGATTGAAGTGCCTGAACTACCGCAGCCGGAGAGCGCGGCAACTCTGGCTCACAGTACCGGTAGTGAACCGAACCGCACCATGGTCAGAAGCC  
711▶ nLeuI l eGl uLeuProGl uLeuP roGl nP roGl uSer Al aGl yGl nLeuTrpLeuThr ValA rgVal Val Gl nP roAsnAl aThr Al aTrpSer Gl uAl a  
2901 GGGCACATCAGCGCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCTCCACGCCATCCCGCATCTGACCACCAGCGAAA  
745▶ Gl yHi sI l eSer Al aTrpGl nGl nTrpArgLeuAl aGl uAsnLeuSer Val Thr LeuP roAl aAl aSer Hi sAl aI l eP roHi sLeuThr Thr Ser Gl uM  
3001 TGGATTTTGCATCGAGCTGGTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTACAGATGTGGATTGGCGATAAAAAACAACCTGCTGAC  
778▶ e tAspPheCysI l eGl uLeuGl yAsnLysArgT rP Gl nPheAsnArgGl nSer Gl yPheLeuSer Gl nMe tTrpI l eGl yAspLysLysGl nLeuLeuTh  
3101 GCCGCTGCGCATCAGTTACCCGTCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGCATTGACCTAACGCCTGGGTGGAACGCTGGAAG  
811▶ r P roLeuArgAspGl nPheThr ArgAl aP roLeuAspAsnAspI l eGl yVal Ser Gl uAl aThr ArgI l eAspP roAsnAl aTrpVal Gl uArgT rP Lys  
3201 GCGCGGGCCATTACCAGGCCAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGTCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGCAGCATC  
845▶ Al aAl aGl yHi sTyrGl nAl aGl uAl aAl aLeuLeuGl nCysThr Al aAspThr LeuAl aAspAl aVal LeuI l eThr Thr Al aHi sAl aTrpGl nHi sG  
3301 AGGGGAAAACCTTATTTATCAGCCGAAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCA  
878▶ I nGl yLysThr LeuPheI l eSer ArgLysThr TyrArgI l eAspGl ySer Gl yGl nMe tAl aI l eThr Val AspVal Gl uVal Al aSer AspThr P roHi  
3401 TCCGGCGGGATTGGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAACTGGCTCGGATTAGGGCCGAAGAAAACCTATCCCGACCGCCTTACT  
911▶ sP roAl aArgI l eGl yLeuAsnCysGl nLeuAl aGl nVal Al aGl uArgVal AsnTrpLeuGl yLeuGl yProGl nGl uAsnTyrP roAspArgLeuThr

BbsI (3548)

Bst1107I (3541)

BspLU11I (3538) BsiWI (3549)

3501 GCCGCTGTTTTGACCGCTGGGATCTGCCATTGTCCAGCATGTATACCCCGTACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGCAATTGA  
945▶ Al aAl aCysPheAspArgT rP AspLeuP roLeuSer AspMet TyrThr P roTyrVal PheP roSer Gl uAsnGl yLeuArgCysGl yThr ArgGl uLeuA  
3601 ATTATGGCCACACCAGTGGCGGGCAGCTTCCAGTTCAACATCAGCCGTACAGTCAAACAGCAACTGATGGAACACGACCATGCCATCTGCTGCACGC  
978▶ snTyrGl yProHi sGl nTrpArgGl yAspPheGl nPheAsnI l eSer ArgTyrSer Gl nGl nGl nLeuMe tGl uThr Ser Hi sArgHi sLeuLeuHi sAl

NdeI (3736)

3701 GGAAGAAGGCACATGGCTGAATATCGACGGTTCCATATGGGGATTGGTGGCGAGCTCCTGGAGCCGTCAGTATCGGCGGAATTACAGCTGAGCGCC  
1011▶ aGl uGl uGl yThr TrpLeuAsnI l eAspGl yPheHi sMe tGl yI l eGl yGl yAspAspSer T rP Ser P roSer Val Ser Al aGl uLeuGl nLeuSer Al a

NheI (3858)

EcoRI (3852)

3801 GGTGCTACCATACCAGTTGGTCTGGTGTCAAAAATAAATCTAGTCGAGAATTCGTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAAAC  
1045▶ Gl yArgTyrHi sTyrGl nLeuVal TrpCysGl nLys•••  
3901 CACAAC TAGAATGCAGTGA AAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCA

MfeI (4032)

DraI (4081)

4001 TTATAAGCTGCAATAAACAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCT

DraI (4120)

SwaI (4123)

4101 CTACAAATGTGGTAGATCCATTTAAATGTTAAATTAAGTACCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAA

4201 AGATCAAAGGATCTTCTTGAGATCCTTTTTCTGCGGTAATCTGCTGCTTGCAAAACAAAAAACACCGCTACCAGCGGTGTTTTGTTGCCGGATCA

4301 AGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAG

4401 AACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGTTGGACTCAAGAC

4501 GATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGTTCTGTCACACAGCCAGCTGGAGCGAACGACCTACACCGAACTGAGATACCTACA

4601 CGCTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCACGAGGGAGCTT

4701 CCAGGGGAAACGCCTGGTATCTTTATAGTCTGCTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTGCTCAGGGGGCGGAGCCTAT

BspLU11I (4861)

AseI (4899)

4801 GGAAAAACGCCAGCAACGCGCCTTTTTACGGTTCCTGGCCTTTTGTGGCCTTTGCTCACATGTTCTTAATTAATTTTTCAAAGTAGTTGACAATT

SfiI (4950) MscI (4961)

4901 AATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACCAGTGTGTCCAGTGTCTCACAGCCAGG

1▶ MetAl aLysLeuThr Ser Al aVal P roVal l euThr Al aArgA

5001 ATGTGGCTGGAGCTGTTGAGTTCTGGACTGACAGGTTGGGTTCTCCAGAGATTTTGTGGAGGATGACTTTCAGGTGTGGTCCAGAGATGATGTACCCT

15▶ spValAl aGl yAl aVal Gl uPheT rP Thr AspArgLeuGl yPheSer ArgAspPheVal Gl uAspAspPheAl aGl yVal Val A rgAspAspVal Thr Le

5101 GTTCATCTCAGCAGTCCAGGACAGGTGGTGCCTGACAACCCCTGGCTTGGGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGGAGTG

48▶ uPheI l eSer Al aVal Gl nAspGl nVal Val P roAspAsnThr LeuAl aTrpVal TrpVal A rgGl yLeuAspGl uLeuTyrAl aGl uTrpSer Gl uVal

5201 GTCTCCACCAACTTCAGGGATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAGTTTGGCCCTGAGAGACCCAGCAGGCAACT

82▶ Val Ser ThrAsnPheArgAspAl aSer Gl yProAl aMe tThr Gl uI l eGl yGl uGl nP roT rP Gl a rgGl uPheAl aLeuArgAspP roAl aGl yAsnC

DraIII (5311)

SfiI (5359)

5301 GTGTGCACTTTGTGGCAGAGGAGCAGGACTGAGGATAAGAATTGAGTTTCAGAAAAGGGGCTGAGTGGCCCTTTTTCAACTTAATTAA

115▶ ysVal Hi sPheValAl aGl uGl uGl nAsp•••