



PstI (6)
SdaI (6) Acc65I (43)
1 CCTGCAGGGCCCTGAAATAACCTCTGAAAGAGGAACCTGGTTAGGTACCTTCTGAGGCTGAAAGAACCCAGCTGTGGAATGTGTCTAGTTAGGGTGTGGAA
101 AGTCCCCAGGCTCCCCAGCAGGCAAGATGCAAAAGCATGCATCTCAATTAGTCAGCAACCAGGTTGTGAAAGTCCCCAGGCTCCCCAGCAGGCAAG
SpeI (243) SspI (302)
201 TATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTTCAGATGGTAAATATACACAAGGGATTAGTCAAACAATTTTTTGGCAAGA
301 ATATTATGAATTTTGAATCGGTTGGCAGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTCAATTATTGGTTAAAGAAATATATTAGTCTAAT
+1 (418) **NcoI (455)** **NheI (493)**
401 TTCCTCCGTTTGTCTAGCTTTCTCTTCTGTCACCCACACGCGCTTTGGCACCATGGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCAT
1 Me tG l yG l yS er H i sH i sH i sH i sH i sG l yM e tA l aS er M e
Bsu36I (554)
501 GACTGGTGGACAGCAATGGGTCGGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTTGGAGTTGATCCCCTCGTTTTACAACGCTGTGAC
15 tTh rG l yG l yG l nG l nM e tG l yA r gA sP l e uT y rA sP a sP a sP a sP L y sV a l P r oL y sA sP G l nL e uG l yV a l A sP p r oV a l V a l L e uG l nA r gA r gA sP
601 TGGGAAAACCTGGCGTTACCAACTAATCGCCTTGCAGCACATCCCCTTCCGACGCTGGCGTAAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCC
49 t rP g l uA s nP r oG l yV a l T h r G l nL e uA s nA r gL e uA l aA l aH i sP r oP r oP h eA l aS e r T r pA r gA s nS e r G l uG l uA l aA r gT h rA sP a r gP r oS e r G
Bsu36I (791)
701 AACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTCCTGGTTCCTGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTGCAGTCTTCTGAGGCCGA
82 t nG l nL e uA r gS e r L e uA s nG l yG l uT r pA r gP h eA l aT r pP h eP r oA l aP r oG l uA l aV a l P r oG l uS e r T r pL e uG l uC y sA sP L e uP r oG l uA l aA s nI l e
801 TACTGTCTGCTCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCACTACACCAACGTAACCTATCCATTACGGTCAATCCGCGTTTGTTCCTC
115 pTh rV a l V a l V a l P r oS e rA s nT r pG l nM e tH i sG l yT y rA sP a l aP r oI l eT y rT h rA s nV a l T h r T y rP r oI l eT h rV a l A s nP r oP h eV a l P r o
901 ACGGAGAATCCGAGCGGTTGTTACTCGCTCACATTAATGTTGTGAAGCTGGCTACAGGAAGGCCAGCAGCGGAATTTTTGATGGCGTTAACTCGG
149 t h rG l uA s nP r oT h rG l yC y sT y rS e r L e uT h rP h eA s nV a l A sP G l uS e r T r pL e uG l nG l uG l yG l nT h rA r g l l e l l eP h eA sP G l yV a l A s nS e r A
1001 CGTTTTCATCTGTGGTGCACCGGCGCTGGGTCGGTTACGGCCAGCAGCTGTTTGGCGCTGAATTTGACCTGAGCGCATTTTTACGCCGCGGAGAAAA
182 l aP h eH i sL e uT r pC y sA s nG l yA r gT r pV a l G l yT y rG l yG l nA sP s e rA r gL e uP r oS e r G l uP h eA sP L e uS e r A l aP h eL e uA r gA l aG l yU aS
1101 CCGCTCGCGGTGATGGTGTGCTGGTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCAATTTCCGTGACGCTCTGTTGCTG
215 nA r gL e uA l aV a l M e tV a l L e uA r gT r pS e rA sP G l yS e r T y rL e uG l uA sP G l nA sP M e tT r pA r gM e tS e r G l y l l eP h eA r gA sP V a l S e r L e uL e u
1201 CATAAACCGACTACAAAATCAGCGATTTCCATGTTGCCACTCGCTTAAATGATGATTTTCCAGCCGCTGACTGGAGGCTGAAGTTCCAGATGTGCGCGG
249 H i sL y sP r oT h rT h rG l n l l eS e rA sP h eH i sV a l A l aT h rA r gP h eA s nA sP h eS e rA r gA l aV a l L e uG l uA l aG l uV a l G l nM e tC y sG l yG
1301 AGTTGCGTGACTACCTACGGGTAACAGTTCCTTTATGGCAGGGTAAACCGCAGGTCCCGACGGCACCAGCCGCTTTCGGCGGTGAAATATATCGATGAGCG
282 t l e uL e uA r gA sP t y rL e uA r gV a l T h rV a l S e rL e uT r pG l nG l yG l uT h rG l nV a l A l aS e rG l yT h rA l aP r oP h eG l yG l yG l u l l e l l eA sP G l uA r
1401 TGTTGGTTATGCGGATCGCGTACACTACGCTCTGAACCTCGAAACCCGAAACTGTGGAGCGCGGAAATCCCGAATCTCTATCGTGGCGGTGGTGAACCTG
315 gG l yG l yT y rA l aA sP a r gV a l T h rL e uA r gL e uA s nV a l G l uA s nP r oL y sL e uT r pS e rA l aG l u l l eP r oA s nL e uT y rA r gA l aV a l V a l G l uL e u
1501 CACACCGCCGACGGCAGCTGATTGAAGCAGAAGCTCGCATGTGCGTTTCCGCGAGGTGGGATTGAAATGGTCTGCTGCTGTAACGGCAAGCCGT
349 H i sT h rA l aA sP G l yT h rL e u l l eG l uA l aG l uA l aC y sA sP V a l G l yP h eA r gG l uV a l A r g l l eG l uA s nG l yL e uL e uL e uA s nG l yL y sP r oL
EcoRV (1680)
1601 TGCTGATTCGAGCGTTAACCGTACGAGCATCATCTCTGCATGGTCAGGTCATGGATGAGCAGCAGTGGTGCAGGATATCTCTGCTGATGAAGCAGAA
382 e uL e u l l eA r gG l yV a l A s nA r gH i sG l uH i sH i sP r oL e uH i sG l yG l nM e tA sP l e uT h rM e tV a l G l nA sP l l eL e uL e uM e tL y sG l nA s
SspI (1797)
1701 CAACCTTAAACCGCTGCTGTTTCGCATTATCCGAACCATCCGCTGTGGTACACCTGCTGTCGACCGCTACGGCCTGTATGGTGGATGAAGCCAAATATT
415 nA s nP h eA s nA l aV a l A r gC y sS e rH i sT y rP r oA s nH i sP r oL e uT r pT y rT h rL e uC y sA sP a r gT y rG l yV a l A sP G l uA l aA s nI l e
1801 GAAACCCAGCGCATGGTGCAATGAATCGTCTGACCGATGATCCGCGTGGCTACCGGATGAGCGAACCGGTAACCGGAATGGTGACCGCGATCGTA
449 G l uT h rH i sG l yM e tV a l P r oM e tA s nA r gL e uT h rA sP a sP a r oA r gT r pL e uP r oA l aM e tS e rG l uA r gV a l T h rA r gM e tV a l G l nA r gA sP a l a
1901 ATCACCCGAGTGTGATCATCTGGTCTGCTGGGAATGAATCAGGCGACGGCGCTAATCACGACGCGTGTATCGCTGGATCAATCTGTCGATCTCCCG
482 s nH i sP r oS e rV a l l l e l l eT r pS e rL e uG l yA s nG l uS e rG l yH i sG l yA l aA s nH i sA sP a l aL e uT y rA r gT r p l l eL y sS e rV a l A sP p r oS e rA r
2001 CCCGGTGCAGTATGAAGCGCGGAGCCGACACCAGCCAGCATATTTTCCGCGATGACGCGCGGTGGATGAAGACAGCCCTTCCCGGCTGTG
515 p r oV a l G l nT y rG l uG l yG l yA l aA sP t h rT h rA l aT h rA sP l l e l l eC y sP r oM e tT y rA l aA r gV a l A sP G l uA sP G l nP r oP h eP r oA l aV a l
2101 CCGAATGGTCCATCAAAAATGGCTTTCGCTACCTGGAGAGACGCGCCGCTGATCCTTTGCGAATACGCCACGCGATGGTAAACAGTCTTGGCGGTT
549 P r oL y sT r pS e r l l eL y sL y sT r pL e uS e rL e uP r oG l yG l uT h rA r gP r oL e u l l eL e uC y sG l uT y rA l aH i sA l aM e tG l yA s nS e rL e uG l yG l yP
2201 TCGCTAAATCTGCGAGCGGCTTTCGTCAGTATCCCGTTACAGGGCGCTTCGCTGGGACTGGTGGATCAGTCTGATTAATATGATGAAACCG
582 h eA l aA s nT y rT r pG l nA l aP h eA r gG l nT y rP r oA r gL e uG l nG l yG l yP h eV a l T r pA sP t r pV a l A sP G l nS e rL e u l l eL y sT y rA sP G l uA s nG l
2301 CAACCCGTGGTCCGCTTACGGCGGTGATTTGGCGATACGCCAAGCATGCGCAGTCTGTATGAACGGTCTGGTCTTGGCAGCCGACGCCGATCCCA
615 yA s nP r oT r pS e rA l aT y rG l yG l yA sP h eG l yA sP t h rP r oA s nA sP G l nP h eC y sM e tA s nG l yL e uV a l P h eA l aA sP a r gT r pH i sP r o
2401 TACCAGCCGAGCAGCAAAACACACAGCAGCAGTTTTCAGTTCCTGTTTACCGGCAAAACCATGAAAGTGACACGCAATACCTGTTCCGTCATAGCGATA
649 A l aL e uT h rG l uA l aL y sH i sG l nG l nP h eP h eG l nP h eA r gL e uS e rG l yG l nT h r l l eG l uV a l T h rS e rG l uT y rL e uP h eA r gH i sS e rA sP A
2501 ACGAGCTCTGCACTGGATGGTGGCGCTGGATGGAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTGCTCCACAAGGTAACAGTTGATGAACT
682 s nG l uL e uL e uH i sT r pM e tV a l A l aL e uA sP G l yL y sP r oL e uA l aS e rG l yU a l P r oL e uA sP V a l A l aP r oG l nG l yL y sG l nL e u l l eG l uL e
2601 GCCTGAACCTACCGCAGCCGAGAGCGCGGGCAACTCTGGCTCAGATACCGCTAGTGAACCGAACCAGCGACCGCATGGTCAAGAGCCGGGCACATCAGC
715 uP r oG l uL e uP r oG l nP r oG l uS e rA l aG l yG l nL e uT r pL e uT h rV a l A r gV a l V a l G l nP r oA s nA l aT h rA l aT r pS e rG l uA l aG l yH i s l l eS e r
2701 GCCTGGCAGCAGTGGCGTGTGGCGAAAACCTCAGTGTGACGCTCCCGCGGCTCCACGCGCATCCCGCATCTGACACCCAGCGAAATGGATTTTGTGA
749 A l aT r pG l nG l nT r pA r gL e uA l aG l uA s nL e uS e rV a l T h rL e uP r oA l aA l aS e rH i sA l a l l eP r oH i sL e uT h rT h rS e rG l uM e tA sP h eC y sI
2801 TCGAGCTGGGTAATAAGCGTTGGCAATTTAACCCGCTCAGGCTTCTTTTACAGATGTGGATGGCGATAAAAAACAACCTGCTGACCGCTGCGCGA
782 l eG l uL e uG l yA s nL y sA r gT r pG l nP h eA s nA r gG l nS e rG l yP h eL e uS e rG l nM e tT r p l l eG l yA sP L y sL y sG l nL e uL e uT h rP r oL e uA r gA s
2901 TCAGTTCACCGTGCACCGCTGGATAACGACATTTGGCGTAAGTGAAGCGACCCGATGACCCCTAACGCCTGGGTCAAGCGCTGGAAGGCGCGGGCCAT
815 pG l nP h eT h rA r gA l aP r oL e uA sP a s nA sP l l eG l yV a l S e rG l uA l aT h rA r g l l eA sP p r oA s nA l aT r pV a l G l uA r gT r pL y sA l aA l aG l yH i s
3001 TACCAGCCGAGCAGCGCTTGTGGTGAAGCAGATACACTTGTGATGCGGCTGCTGATTACGACCGCTCAGCGCTGAGCGATCAGGCGAGCAACCT
849 T y rG l nA l aG l uA l aA l aL e uL e uG l nC y sT h rA l aA sP t h rL e uA l aA sP a l aV a l L e u l l eT h rT h rA l aH i sA l aT r pG l nH i sG l nG l yL y sT h rL
3101 TATTTATCAGCCGAAACCTACCGGATGATGGTGTGATGCTCAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGCGCGGAT
882 e uP h e l l eS e rA r gL y sT h rT y rA r g l l eA sP G l yS e rG l yG l nM e tA l a l l eT h rV a l A sP V a l G l uV a l A l aS e rA sP t h rP r oH i sP r oA l G l
3201 TGGCTGAACCTGCCAGTGGCGAGGTAGCAGAGCGGTAACCTGGCTGGATTAGGGCCGCAAGAAAATATCCGACCGCTTACTGCGCCTGTTTT
915 eG l yL e uA s nC y sG l nL e uA l aG l nV a l A l aG l uA r gV a l A s nT r pL e uG l yL e uG l yP r oG l nG l uA s nT y rP r oA sP a r gL e uT h rA l aA l aC y sP h e
BspLU11I (3326)
3301 GACCCTGGGATCTGCCATTGTCTAGACATGTATACCCGCTACGCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGGAATTGAATTATGGCCAC
949 A sP a r gT r pA sP L e uP r oL e uS e rA sP M e tT y rT h rP r oT y rV a l P h eP r oA e rG l uA s nG l yL e uA r gC y sG l yT h rA r gG l uL e uA s nT y rG l yP r oH
3401 ACCAGTGGCGCGCAGCTCCAGTTCAACATCAGCCGCTACAGTCAACCGAAGCTGATGGAAACCGCCATCGCCATCTGCGACCGCAGCAAGGAAAGCAGC
982 i sG l nT r pA r gG l yA sP h eG l nP h eA s nI l eS e rA r gT y rS e rG l nG l nG l nL e uM e tG l uT h rS e rH i sA r gH i sL e uL e uH i sA l aG l uG l uG l yT h
3501 ATGGCTGAATATCGACGGTTCCATATGGGATTTGGTGGCAGCAGCTCCTGGAGCCGCTGATATCGGCGGAATACAGCTGAGCGCGGCTCGCTACCAT
1015 rT r pL e uA s nI l eA sP G l yP h eH i sM e tG l y l l eG l yG l yA sP a sP s e rT r pS e rP r oS e rV a l S e rA l aG l uL e uG l nL e uS e rA l aG l yA r gT y rH i s

NheI (3646)

EcoRI (3640)

3601 TACCAGTTGGTCTGGTGTCAAAAATAATAATCTAGTCGAGAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAAT
1049 TyrGI nLeuVal TrpCysGI nLys•••

3701 GCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCA

MfeI (3820)

3801 ATAAACAAGTTAACACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGG

SwaI (3911)

3901 TAGATCCATTTAAATGTTAATTAAGTACCCATGACCAAAATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
→

4001 CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGAACAAAAAACCACCGCTACCAGCGGTGGTTTTGTTGCCGGATCAAGAGCTACCAAC

4101 TCTTTTTCCGAAGGTAAGTGGCTTCCAGCAGAGCGCAGATACCAAACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACCTCAAGAACTCTGTAGCA

4201 CCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGG

4301 ATAAGCGCGAGCGGTGGGGTGAACGGGGGTTCTGTGCACACAGCCAGCTTGGAGCGAACGACCTACCCGAACTGAGATACCTACAGCGTGAGCTATG

4401 AGAAAGCGCCACGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAC

4501 GCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCA

BspLU11I (4649) **AseI (4687)**

4601 GCAACGCGGCCTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTAATTAATTTTCAAAGTAGTTGACAATTAATCATCGGCAT
→

4701 AGTATATCGGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACCAGTGTGTCCAGTGTCCAGTGTCCAGCCAGGGATGTGGCTGGAG
1 MetAl aLysLeuThr Ser Al aVal ProVal LeuThr Al aArgAspVal Al aGl yA

4801 CTGTTGAGTCTGGACTGACAGGTTGGGTTCTCCAGAGATTTGTGGAGGATGACTTTCAGGTGTGGTCAGAGATGATGTCACCCTGTTTCATCTCAGC
19 Val Gl uPheTrpThrAspArgLeuGl yPheSer ArgAspPheVal Gl uAspAspPheAl aGl yVal Val ArgAspAspVal ThrLeuPheI leSerAl

4901 AGTCCAGGACCAGGTGGTGCCTGACAACCCCTGGCTTGGGTGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGAGGTGGTCTCCACCAAC
52 Val Gl nAspGl nVal Val ProAspAsnThr LeuAl aTrpVal TrpVal ArgGl yLeuAspGl uLeuTyrAl aGl uTrpSer Gl uVal Val Ser ThrAsn

5001 TTCAGGGATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAGTTTGCCTGAGAGACCCAGCAGGCAACTGTGTGCACCTTG
86 PheArgAspAl aSer Gl yProAl aMetThr Gl uI leGl yGl uGl nProTrpGl yArgGl uPheAl aLeuArgAspProAl aGl yAsnCysVal Hi sPheV

5101 TGGCAGAGGAGCAGGACTGAGGATAAGAATTGAGTTTCAGAAAAGGGGCTGAGTGGCCCTTTTTCAACTAATTAA
119 aAl aGl uGl uGl nAsp•••