

# pWHERE Kit

An optimized vector for mouse and rat transgenesis

Catalog # kwhere

For research use only

Version 20L07-MM

## PRODUCT INFORMATION

### Content:

- 20 µg of pWHERE is provided as lyophilized DNA.
- 10 µg of pWHERE Control with the rat EF-1α promoter provided as lyophilized DNA.

### Storage and Stability:

- Product is shipped at room temperature.
- Lyophilized DNA should be resuspended upon receipt and stored at -20°C (see Methods). Lyophilized DNA is stable 12 months at -20°C. Resuspended DNA is stable more than one year at -20°C. Avoid repeated freeze-thaw cycles.

### Quality control:

- Plasmid construct has been confirmed by restriction analysis and sequencing.
- Plasmid DNA was purified by ion exchange chromatography and lyophilized.

## GENERAL PRODUCT USE

The pWHERE plasmid was designed for studies of temporal expression and tissue distribution of **your promoter of interest**, cloned within an insulated LacZ cassette, in transgenic mice and rats.

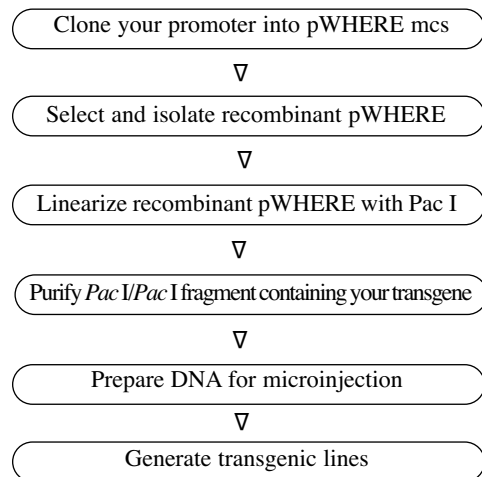
A multiple cloning site (MCS) has been added upstream of the LacZ gene for convenient cloning of your promoter of interest. The MCS contains several restriction sites that are compatible with many other enzymes, thus facilitating cloning. Furthermore, the *E. coli* region is flanked on either side by the well cutting 8 bp-recognizing restriction enzyme *Pac* I that enables linearization and easy excision of the *E. coli* region.

## PLASMID FEATURES

- **mH19 insulators** on either side of the lacZ transcription unit. Both insulators are expected to protect the integrated transcriptional LacZ unit from negative as well as positive influences from neighboring sequences. Insulator elements can be functionally identified by their ability to shield promoters from regulators in a position-dependent manner or by their ability to protect adjacent transgenes from position effects. The fragment of the differential methylated region (DMR) located between the mouse *Igf2* and *H19* acts as a powerful insulator<sup>1</sup>. The enhancer blocking activity of the DMR fragment is dependent upon four responsive elements to the vertebrate enhancer-blocking protein CTCF<sup>2</sup>. Two mouse DMR fragments have been introduced in opposite orientation in the pWHERE plasmid to insulate your promoter of interest cloned upstream of the new CpG-free LacZ gene from the 5' and 3' adjacent regions at the integrated site in transgenic mice.

- **MCS:** The multiple cloning site, located downstream of the mH19 insulator, contains the following restriction sites:  
*Sda* I, *Avr* II, *Bam* HI, *Xho* I, *Sma* I and *Nco* I  
*Sda* I is compatible with *Nsi* I and *Pst* I  
*Avr* II is compatible with *Nhe* I, *Spe* I and *Xba* I  
*Bam* HI is compatible with *Bgl* II and *Bcl* I  
*Xho* I is compatible with *Ava* I and *Sal* I  
*Nco* I is compatible with *Bsp* HI and *Bsp* LU111  
*Sma* I is compatible with any blunt end restriction enzyme-
- **rEF-1α promoter (pWHERE Control):** EF-1α is one of the most abundant proteins in eukaryotic cells and is expressed in almost all kinds of mammalian cells. The promoter of this 'housekeeping' gene exhibits a strong activity, higher than viral promoters such as SV40 and RSV promoters and, on the contrary to the CMV promoter, yields persistent expression of the transgene *in vivo*.
- **pMB1 ori:** a minimal *E. coli* origin of replication to limit vector size but with the same activity as the longer Ori.
- **Amp:** The ampicillin resistance gene allows the selection of transformed *E. coli* carrying a pWHERE plasmid.
- **LacZ-ΔCpG NLS:** The *E. coli lacZ* gene codes for the enzyme β-galactosidase which catalyzes the hydrolysis of the substrate X-Gal to produce a blue color that is easily visualized under a microscope. A nuclear localization signal of SV40 large T has been inserted in the 5' end of the *lacZ* gene to allow the targeting of the chimeric protein to the nucleus. To reduce the immunogenicity of this bacterial gene, InvivoGen has engineered a synthetic *lacZnls* gene that is entirely free of CpG motifs, whereas the wild type *lacZ* gene contains 298 CpG dinucleotides.
- **EF1 pAn** is a strong polyadenylation signal. InvivoGen uses a sequence starting after the stop codon of the EF1 cDNA and finishing after a bent structure rich in GT.

## EXPERIMENTAL OUTLINE



## TECHNICAL SUPPORT

InvivoGen USA (Toll-Free): 888-457-5873  
InvivoGen USA (International): +1 (858) 457-5873  
InvivoGen Europe: +33 (0) 5-62-71-69-39  
InvivoGen Hong Kong: +852 3-622-34-80  
E-mail: info@invivogen.com

## METHODS

### Plasmid resuspension

Quickly spin the tube containing the lyophilized plasmid to pellet the DNA. To obtain a plasmid solution at 1 µg/µl, resuspend the DNA in 20 µl of sterile H<sub>2</sub>O. Store resuspended plasmid at -20 °C.

### Plasmid amplification and cloning

Plasmid amplification and cloning can be performed in *E. coli* GT116 or in other commonly used laboratory *E. coli* strains, such as DH5α.

### Ampicillin usage

Ampicillin (not provided) can be used for *E. coli* at 50-100 µg/ml in liquid or solid media.

### Pac I linearization of recombinant pWHERE:

1- Digest 10 µg recombinant pWHERE plasmid with 1 to 5 units of *Pac* I restriction enzyme.

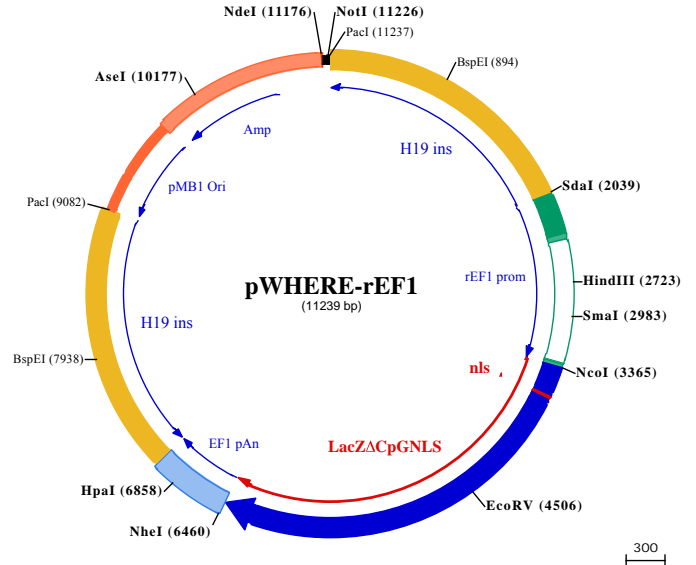
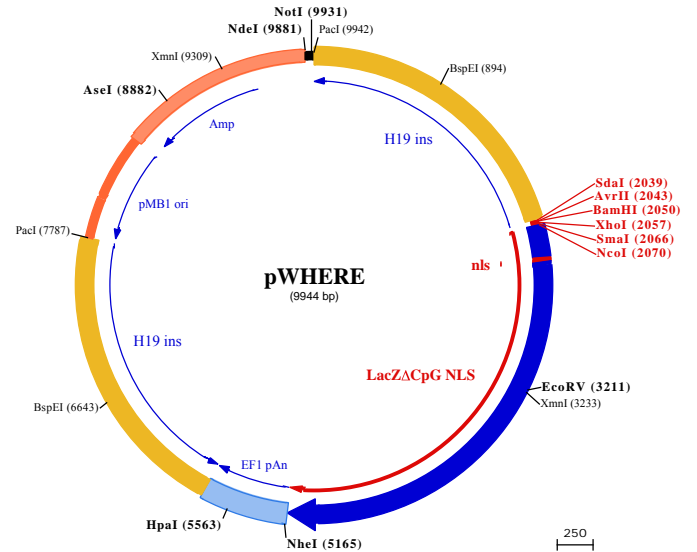
*Note: Pac I may be purchased from New England Biolabs and used at 0.1-0.5 unit per µg plasmid DNA.*

2- Incubate at 37°C for 1-2 hours.

3- Purify the fragment containing the LacZ expression cassette by agarose gel following your usual protocol.

### References:

1. Kaffer CR. et al. 2000. A transcriptional insulator at the imprinted H19/Igf2 locus. *Genes Dev.* 14:1908-19.
2. Bell AC. and Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature.* 405:482-485.



### TECHNICAL SUPPORT

InvivoGen USA (Toll-Free): 888-457-5873  
InvivoGen USA (International): +1 (858) 457-5873  
InvivoGen Europe: +33 (0) 5-62-71-69-39  
InvivoGen Hong Kong: +852 3-622-34-80  
E-mail: [info@invivogen.com](mailto:info@invivogen.com)

# pWHERE sequence

1 CAGGTTTGTACAGCGGACCCCAACCTATGCCGGTCTGCCGAGCAATATGTAGTATTGTACTGCCACCACGCGGCATCGTCTGCCATTTAGCTATAGC  
101 CAAATCTGCACAGCGTGGAGAGTGAACCGATCGATCGTGAAGGGCGCAAGACTGAAGGAGCTACCCAAGAAATGTGTGTTGTACCACCCCATGACCCTTA  
201 TGAATCATTGAATGCTATGCCTGAGTGACCCATGAGTTTGCATAGGTGAACCGCAATTTTGGTACCTCTATGATAAGTATGTGGACTGCACTGGTC  
301 GCTGCTCGGCAACTTCGGTCTTACCAGCCACTGACGATCTCGGGCTGTGTAGGGAATGAGTCAAGTTCTCTGGTTTCAAGTGTAAAGGAAACCATTCCAGA  
401 GGTGCACACATCTTACCACCCCTATGAATCCCTATTGGGTGACCCCTGGGATATTGCTGGGAATGAATCGCTCCCCCAAGTTGGCAGCATTGGGCCAC  
501 GATATATAGGAGTATGCTGCCACCGCGGTAGCATCCGTTCCCTTGTGACATAACAGCTTCTATGCCTTCTATAGTGAGCCACACTGGCTGGTTTT  
601 GGGGTTCAAGTGACAAAGGGACCCCTCCAGAAACACAAGTGTTCACACCTCTATAAACCATATCGACCACTGAGGCATAGCGGCTTCGGACATTGCTG  
701 TGGGCAACCCGAACCTTGGCCCTTGGACATTGTCATGGGCAACCTTAACTTTGGTGTCTCTGGAACATAATGTTAAGATGACAGTACCAGCGCAG  
801 CAATTTGGTCTTTCCACTCACAACGGCTTTTGTGCTTTCTGGCATCGAACCCATGCCTGGTTTATGGGGTCTGAGACCAAGGAGACCATTCCGGAAAG  
901 GGCATAGGTGCTCTGCCTTCTGCTTTTAAACAAGGCTCTCCGGGACAGTGCAAAACAGGTGAACCCCAACTTTGCCATAAGTACGATTATCTGCCACAAAC  
1001 CAGCCAGGGTCTTACCACCTCTTCAATTGATTTGGGTGACACCCAGGCTTGATGTAGGATTCTCAGCTGCCAAGCTGGCAGCTGACCCCAATTGA  
1101 GAGAGAATGCAGTTTCAAGATTGTTTGCAGCCCTGAGCCGGAGATCATTAGCATCTGAACGCCCAATTAGAATACGAAAGCGCAATACCAGACTTTC  
1201 TTGTTGGCGGTTCTTAAGTGATTCTTGGGTAGGGAGTCCAGTGTCTCATAGATGTGCAGAATTTGAGGACCATGCTTAGTGGGGTCTGCATTATGG  
1301 TACTGCAATACATTCCATGATCACCACACATAGTAGCTATACTTCAATTTTACACAATAGCGCTGATGGCCCCAGAACCTATAAGTCAGATACCTGAG  
1401 ATAGCTCTTGAGAACGTTTTATCAAGGACTAGCATGAACCCCTGGCCTCATGAAGCCCATGACTATGGGATCATAGATGGTGATAGGGGAGAAAACCTCAA  
1501 TCAGTTGCAATCCGTTTTAGGACTGCGATGTACGAGACTTCACTGCCCGCTGCGGCAACCTGGTCTTTACACACAAAGGATTCTTTGCAGAGAGTAAAG  
1601 CCGACCTTGTGATTGGGAGTCCGAGTCCACGAGGTACCAGCCTAGAAAATGCATGTGCTCTGCCCCCTAGTGGCATTGTGACCCCCCTGAGGTACT  
1701 GAACTTGGGTGACCCACAGCATTGCCATTTGTGAATCCAATACCAGGGGTGGGGGGCTCTTTAGGTTTGGCGCAATCGATTTTGTGCCACCACGCGG  
1801 CAACTCCCGCTATAAACCCCAACTGATTTCAGCAGACGTCCAAGAATAGGGCATGGTCTCCTTGAGAATTCTTATGCCTCCTGGATGCTCGTGTGAA  
1901 TGTAGCATGTTCCCTTTGAGTCTGGGTGTAAGTGCCTGCAGCTCATCCCCGGACATGAAAATAGAATCTCTATTTTCTACCAACCTTTTCTTTCC  
AvrII (2043) XhoI (2057) NcoI (2070)  
SdaI (2039) BamHI (2050) SmaI (2066)  
2001 TTGTGGTATTTCGGAACTGTAGGCAATGGCTCTCGAGGCTCCTAGGTGGATCCTCTCGAGTCCCGGCCATGGACCTGTTGTGCTGCAAAGGAGAGAC  
1▶ MetAspProValValLeuGlnArgArgAsp  
2101 TGGGAAACCCCTGGAGTGACCAGCTCAACAGACTGGCTGCCACCCTCCCTTTGCCCTTTGGGAACTCTGAGGAAGCCAGGACAGACAGGCCAGCC  
11▶ TrpGluAsnProGlyValThrGlnLeuAsnArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluGluAlaArgThrAspArgProSerG  
2201 AGCAGCTCAGGTCTCTCAATGGAGAGTGGAGGTTGCGCTGGTCCCTGCCCTGAAGCTGTGCTGAGTCTTGGCTGGAGTGTGACCTCCAGAGGCAGT  
44▶ InGlnLeuArgSerLeuAsnGlyGluTrpArgPheAlaTrpPheProAlaProGluAlaValProGluSerTrpLeuGluCysAspLeuProGluAlaVa  
2301 TCCAAGAAGAAGAGGAAAGTTGAGGCTGACACTGTTGTGGTGCCAAGCAACTGGCAGATGCATGGCTATGATGCCCCATCTACACCAATGTACCTAC  
77▶ IProLysLysLysArgLysValGluAlaAspThrValValValProSerAsnTrpGlnMetHisGlyTyrAspAlaProIeTyrThrAsnValThrTyr  
2401 CCCACTACTGTGAACCCCTTTTGTGCCACTGAGAACCCACTGGCTGCTACAGCCTGACCTTCAATGTTGATGAGAGCTGGCTGCAAGAAGGCCAGA  
111▶ ProIleThrValAsnProProPheValProThrGluAsnProThrGlyCysTyrSerLeuThrPheAsnValAspGluSerTrpLeuGlnGluGlyGlnT  
2501 CCAGGATCATCTTTGATGGAGTCAACTCTGCCTCCACCTCTGGTGCATGGCAGGTGGGTGGCTATGGCAAGACAGCAGGCTGCCCTCTGAGTTTGA  
144▶ hrArgIleIlePheAspGlyValAsnSerAlaPheHisLeuTrpCysAsnGlyArgTrpValGlyTyrGlyGlnAspSerArgLeuProSerGluPheAs  
2601 CCTCTCGCCTTCTCAGAGCTGGAGAGAACAGGCTGGCTGCTCATGGTCTCAGGTGGTCTGATGGCAGCTACCTGGAAGACCAAGACATGTGGAGGATG  
177▶ pLeuSerAlaPheLeuArgAlaGlyGluAsnArgLeuAlaValMetValLeuArgTrpSerAspGlySerTyrLeuGluAspGlnAspMetTrpArgMet  
2701 TCTGGCATCTTACGGGATGTGAGCCTGCTGCACAAGCCACCACCCAGATTTCTGACTTCCATGTTGCCACCAGGTTCAATGATGACTTCAAGCAGAGCTG  
211▶ SerGlyIlePheArgAspValSerLeuLeuHisLysProThrThrGlnIleSerAspPheHisValAlaThrArgPheAsnAspAspPheSerArgAlaV  
2801 TGCTGGAGGCTGAGGTGAGATGTGTGGAGAACTCAGAGACTACCTGAGAGTCAAGTCAAGCTCTGGCAAGGTGAGACCCAGGCTGGCCTCTGGCACAGC  
244▶ aLeuGluAlaGluValGlnMetCysGlyGluLeuArgAspTyrLeuArgValThrValSerLeuTrpLeuGlnGlyGluTrpLeuValAlaSerGlyThrAl  
2901 CCCCTTTGGAGGAGAGATCATTGATGAGAGAGGAGGCTATGCTGACAGAGTACCCTGAGGCTCAATGTGGAGAACCCCAAGCTGTGGTCTGCTGAGATC  
277▶ aProPheGlyGlyGluIleIleAspGluArgGlyGlyTyrAlaAspArgValThrLeuArgLeuAsnValGluAsnProLysLeuTrpSerAlaGluIle  
3001 CCCAACCTCTACAGGCTGTTGTGGAGCTGCACACTGCTGATGGCACCTGATTGAAGCTGAAGCCTGTGATGTTGGATTTCAGAGAAGTCAGGATTGAGA  
311▶ ProAsnLeuTyrArgAlaValValGluLeuHisThrAlaAspGlyThrLeuIleGluAlaGluAlaCysAspValGlyPheArgGluValArgIleGluA  
3101 ATGGCCTGCTGCTGCTCAATGGCAAGCCTGCTGCATCAGGGAGTCAACAGGCATGAGCACCACCTCTGCATGGACAAGTGTGGATGAACAGACAAT  
344▶ snGlyLeuLeuLeuLeuAsnGlyLysProLeuLeuIleArgGlyValAsnArgHisGluHisHisProLeuHisGlyGlnValMetAspGluGlnThrMe  
EcoRV (3211) XmnI (3233)  
3201 GGTGCAAGATATCTGCTAATGAAGCAGAACCACTTCAATGCTGTCAGGTGCTCTCACTACCCCAACCACCTCTCTGGTACACCCTGTGTGACAGGTAT  
377▶ tValGlnAspIleLeuLeuMetLysGlnAsnAsnPheAsnAlaValArgCysSerHisTyrProAsnHisProLeuTrpTyrThrLeuCysAspArgTyr  
3301 GGCCTGTATGTTGTTGATGAAGCCAACTTGGAGACATGGCATGGTGGCCATGAACAGGCTCACAGATGACCCAGGTGGCTGCCATGCTGAGA  
411▶ GlyLeuTyrValValAspGluAlaAsnIleGluThrHisGlyMetValProMetAsnArgLeuThrAspAspProArgTrpLeuProAlaMetSerGluA  
3401 GAGTGACCAGGATGGTGCAGAGAGACAGGAACCCCTCTGTGATCATCTGGTCTCTGGGCAATGAGTCTGGACATGGACCAACCATGATGCTCTCTA  
444▶ rgValThrArgMetValGlnArgAspArgAsnHisProSerValIleIleTrpSerLeuGlyAsnGluSerGlyHisGlyAlaAsnHisAspAlaLeuTy

3501 CAGGTGGATCAAGTCTGTTGACCCAGCAGACCTGTGCAGTATGAAGGAGGTGGAGCAGACACCACAGCCACAGACATCATCTGCCCATGTATGCCAGG  
477 rArgTrpIleLysSerValAspProSerArgProValGlnTyrGluGlyGlyGlyAlaAspThrThrAlaThrAspIleIleCysProMetTyrAlaArg  
3601 GTTGATGAGGACAGCCCTTCCCTGCTGTGCCCAAGTGGAGCATCAAGAAGTGGCTCTCTCGCCTGGAGAGACCAGACCTGTATCCTGTGTGAATATG  
511 ValAspGluAspGlnProPheProAlaValProLysTrpSerIleLysLysTrpLeuSerLeuProGlyGluThrArgProLeuIleLeuCysGluTyrA  
3701 CACATGCAATGGGCAACTCTCTGGGAGGCTTTGCCAAGTACTGGCAAGCCTTCAGACAGTACCCAGGCTGCAAGGAGGATTTGTGTGGGCTGGGTGGA  
544 IaHisAlaMetGlyAsnSerLeuGlyGlyPheAlaLysTyrTrpGlnAlaPheArgGlnTyrProArgLeuGlnGlyGlyPheValTrpAspTrpValAs  
3801 CCAATCTCTCATCAAGTATGATGAGAATGGCAACCCCTGGTCTGCCTATGGAGGAGACTTTGGTGACACCCCAATGACAGGCAGTTCTGCATGAATGGC  
577 pGlnSerLeuIleLysTyrAspGluAsnGlyAsnProTrpSerAlaTyrGlyGlyAspPheGlyAspThrProAsnAspArgGlnPheCysMetAsnGly  
3901 CTGGTCTTTGCAGACAGGACCCCTCACCTGCCTCACAGAGGCCAAGCACAGCAACAGTTCTCCAGTTCAGGCTGTCTGGACAGACCATGAGGTGA  
611 LeuValPheAlaAspArgThrProHisProAlaLeuThrGluAlaLysHisGlnGlnGlnPhePheGlnPheArgLeuSerGlyGlnThrIleGluValT  
4001 CATCTGAGTACCTTTCAGGCACCTCTGACAATGAGCTCCTGCAGTGGTGGCCTGGATGGCAAGCCTCTGGCTTCTGGTGGAGTGCCTCTGGATGT  
644 hrSerGluTrpLeuPheArgHisSerAspAsnGluLeuHisTrpMetValAlaLeuAspGlyLysProLeuAlaSerGlyGluValProLeuAspVa  
4101 GGCCCTCAAGAAAGCAGCTGATTGAACTGCCTGAGCTGCCTCAGCCAGAGTCTGTGGCAACTGTGGCTAACAGTGAGGGTGGTTACGCCCAATGCA  
677 AlaProGlnGlyLysGlnLeuIleGluLeuProGluLeuProGlnProGlnProGlnPheSerAlaGlyGlnLeuTrpLeuThrValArgValValGlnProAsnAla  
4201 ACAGCTTGGTCTGAGGCAGGCCACATCTCTGCATGGCAGCAGTGGAGGCTGGCTGAGAACCCTCTCTGTGACCTGCCTGTCTGCCTCATGCCATCCCTC  
711 ThrAlaTrpSerGluAlaGlyHisIleSerAlaTrpGlnGlnTrpArgLeuAlaGluAsnLeuSerValThrLeuProAlaAlaSerHisAlaIleProH  
4301 ACCTGACAACATCTGAAATGGACTTCTGCATTGAGCTGGGCAACAAGAGATGGCAGTTCAACAGGCAGTCTGGCTTCTGTCTCAGATGTGGATTGGAGA  
744 isLeuThrThrSerGluMetAspPheCysIleGluLeuGlyAsnLysArgTrpGlnPheAsnArgGlnSerGlyPheLeuSerGlnMetTrpIleGlyAs  
4401 CAAGAAGCAGCTCCTCACCCCTCTCAGGGACCAATTCACCAGGGCTCCTCTGGACAATGACATTGGAGTGTCTGAGGCCACCAGGATTGACCCAAATGCT  
777 pLysLysGlnLeuLeuThrProLeuArgAspGlnPheThrArgAlaProLeuAspAsnAspIleGlyValSerGluAlaThrArgIleAspProAsnAla  
4501 TGGTGGAGAGGTGGAAGGCTGTGGACACTACCAGGCTGAGGCTGCCCTGCTCCAGTGCACAGCAGACACCCTGGCTGATGCTGTTCTGATCACCCAG  
811 TrpValGluArgTrpLysAlaAlaGlyHisTyrGlnAlaGluAlaAlaLeuLeuGlnCysThrAlaAspThrLeuAlaAspAlaValLeuIleThrThrA  
4601 CCCATGCTTGGCAGCACCAGGCAAGCCTGTTCATCAGCAGAAAAGACCTACAGGATGATGGCTCTGGACAGATGGCAATCACAGTGGATGGGAGGT  
844 IaHisAlaTrpGlnHisGlnGlyLysThrLeuPheIleSerArgLysThrTyrArgIleAspGlySerGlyGlnMetAlaIleThrValAspValGluVa  
4701 TGCCTCTGACACACCTCACCCCTCAAGGATTGGCCTGAACGTCAACTGGCACAGGTGGCTGAGAGGGTGAACCTGGCTGGGCTTAGGCCCTCAGGAGA  
877 AlaSerAspThrProHisProAlaArgIleGlyLeuAsnCysGlnLeuAlaGlnValAlaGluArgValAsnTrpLeuGlyLeuGlyProGlnGluAsn  
4801 TACCCTGACAGGCTGACAGCTGCCTTTGACAGGTGGCAGTGCCTCTGTCTGACACTTACACCCCTATGTGTTCCTCTGAGAATGACCCCTGAGGT  
911 TyrProAspArgLeuThrAlaAlaCysPheAspArgTrpAspLeuProLeuSerAspMetTyrThrProTyrValPheProSerGluAsnGlyLeuArgC  
4901 GTGGCACCAGGGAGCTGAACTATGGTCTCACAGTGGAGGGGAGACTTCCAGTCAACATCTCCAGTACTCTCAGCAACAGCTCATGAAACCTCTCA  
944 ysGlyThrArgGluLeuAsnTyrGlyProHisGlnTrpArgGlyAspPheGlnPheAsnIleSerArgTyrSerGlnGlnGlnLeuMetGluThrSerHi  
5001 CAGGCACCTGCTCCATGCAGAGGAGGAAACCTGGCTGAACATTGATGGCTTCCACATGGGCATTGGAGGAGATGACTCTGGTCTCCTTCTGTCTGTCT  
977 sArgHisLeuLeuHisAlaGluGluGlyThrTrpLeuAsnIleAspGlyPheHisMetGlyIleGlyGlyAspAspSerTrpSerProSerValSerAla

**NheI (5165)**

5101 GAGTTCAGTTATCTGCTGGCAGGTACCACTATCAGCTGGTGTGGTGCCAGAAGTAAACCTGAGCTAGCATTATCCCTAATACCTGCCACCCCACTTTA  
1011 GluPheGlnLeuSerAlaGlyArgTyrHisTyrGlnLeuValTrpCysGlnLys...  
5201 ATCAGTGGTGAAGAAGCGTCTCAGAAGTGTGTTGTTCAATTGGCCATTAAGTTTAGTAGTAAAGACTGGTTAATGATAACAATGCATCGTAAAACTT  

---

  
5301 TCAGAAGGAAAGGAGAATGTTTTGTGGACCACCTTTGGTTTTCTTTTTGCGTGTGGCAGTTTTAAGTTATTAGTTTTTAAAAATCAGTACTTTTTAATGGA  

---

  
5401 ACAAACCTTGACAAAAATTTGTACAGAATTTTGAGACCCATTAAGTAAAGTAAATGAGAAACCTGTGTGTTCTTTGGTCAACACCAGACATTTAGGT  

---

**HpaI (5563)**

5501 GAAAGACATCTAATCTGGTTTTACGAATCTGGAACTTCTTGAAAATGTAATCTTGGTAAACACTCTGGGTGGAGAATAGGGTTGTTTTCCCCCA  

---

  
5601 CATAATTGGAAGGGGAAGGAATATCATTAAAGCTATGGGAGGTTTTCTTTGATTACAACACTGGAGAGAAATGCAGCATGTTGCTGATTGCTGCACT  

---

  
5701 AAAACAGGCCAAAACTGAGTCCTTGGGTTGCATAGAAAGTGCCTGCA GAGCTTTGTACAGCGGACCCCAACCTATGCCCGCTGTGCCGAGCAATATG  

---

  
5801 TAGTATTGTACTGCCACCACGGGCATCGTCTGTCCATTTAGCTATAGCCAAATCTGCACAGCGTGGAGAGTGAACCGATCGATCGTAAGGGCGCAAGA  

---

  
5901 CTGAAGGAGCTACCAAGAAATGTGTGTGTGTACCACCCCATGACCCTTATGAATCATTGAATGCTATGCCTGAGTGACCCATGAGTTTGCCATAGGTGAA  

---

  
6001 CCGCAATTTTGGTCACTCTATGATAAAGTATGTGGGACTGCACCTGGTCTGCTCGGCAACTTCGGTCTTACCAGCCACTGACGATCTCGGGCTGTGTA  

---

  
6101 GGGAAATGAGTCAAGTTCTCTGGTTCAGTGTGTAAGGGAACCATTCAGAGGTGCACACATCTTACCACCCCTATGAATCCCTATTTGGGTGACCTGGGA  

---

  
6201 TATTGCTGGGAATGAATCGCTCCCCCAAGTTGGCAGCATTGGGCCACGATATATAGGAGTATGCTGCCACCCGCGGGTAGCATCCGTTCCCTTGTTCG  

---

  
6301 ACATAACAGCTTCTATGCCTTCCATAGTGAGCCACTGGCTGGTTTTGGGGTTTCAGTGACCAAGGACCCCTCCAGAAACACAAGTGTCCAACTT  

---

  
6401 CTTATAAACCATATCGACCACTGAGGCATAGCGGCTTGGGACATTGCTGTGGGCAACCCGAACCTTTGGCCCTTGGACATTGTCATGGGCAACCTTAAC  

---

  
6501 TTTGGTCTCTCTGGAACATAATGTTTAAAGATGACAGTCCACAGCGCAGCAATTTGGTCTTTCCACTCACACCGCTTTTGTGCTTCTGGCATCGAAC  

---

**BspEI (6643)**

6601 ACATGCACCTGGTTTTATGGGGTCTGAGACCAAAGGAGACCATTCCGGAAGGCATAGGTGCTCCTGCTTTTAAACAAGGCTCTCCGGGACAGTGCA  

---

  
6701 AAACAGGTGAACCCCAACTTTGCCATAAGTACGATTATCTGCCACAACACCAGCCAGGGTCTACCACCTCTCAATTGATTTGGGCTGACACCCAAGG  

---

  
6801 CTTGATGTAGGATTCTCAGCTGCCAAGCTGGCAGCTGACCCCATTTGAGAGAGATGCAGTTTCAAGATTGTTTCAGCCCTGAGCGGAGATCATTAG  

---

  
6901 CATCTGAACGCCCAATTAGAATACGAAAGCGCAAATCACCAGACTTCTTGTGGCGGTTCTTAAGTGATTTTGGGTAGGGAGGTGAGGTGCTCTC  

---

  
7001 ATAGATGTGAGAATTTGAGGACCATGCTTACTGGGGTCTGCATTATGGTACTGCAATACATTCATGATCACCACACATAGTACTTCAATTTT  

---

  
7101 CACACAATAGCGCTGATGGCCCCAGAACCCTATAAGTCAGATACCTGAGATAGCTTGTGAGAACGTTTTATCAAGGACTAGCATGAACCCCTGGCCTCAT  

---

7201 GAAGCCCATGACTATGGGATCATAGATGGTGATAGGGGAGAAAACCTCAATCAGTTGCAATCCGTTTTAGGACTGCGATGTACGAGACTTCACTGCCGCCG  
7301 TGGCGCAACCCTGGTCTTTACACACAAAGGATTCTTTGCAGAGAGTAAGCCGACCTTGTGATTGGGAGTCCGAGTCCACGAGGTACCAGCCTAGAAAA  
7401 TGCATGTGTCTGCCCCCTAGTGGCATTGTGACCCCCCTGAGGTAAGTGAACCTGGGTGACCCACAGCATTGCCATTTGTGAATCCAATACCAGGGGT  
7501 GGGGGGGCTCTTTAGGTTGGCGCAATCGATTTTGCTGCCACCACGCGGCAACTCCCGCGTATAAAACCCACAACCTGATTGAGCAGACGTCCAAGAATAG  
7601 GGCATGGTCTCCTTGCAGAATCTTATGCCTCTGGATGCTCGTGTGAATGTAGCATGTTCTTTGAGTCTGGGTGTAAGTGCCCTGCACGCTCATCCC

PacI (7787)

7701 CCGGACATGAAAATAGAATCTCTATTTTCTACCAACCTTTTTCTTTCTTGTGGTGATTCCGGAACTGTAGGCAATGGCTTTAATTAATAAACCCGCTTC  
7801 GGGGGTTTTTTTTATGCATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGA  
7901 CGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCT  
8001 CCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGG  
8101 TGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGTGCCTTATCCGGTAACTATCGTCTGAGTCCAACCCGCT  
8201 AAGACAGCACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAAC  
8301 TACGGCTACACTAGAAGAACAGTATTTGGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCAACAAACCA  
8401 CCGTGGTAGCGGTGGTTTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGC  
8501 TCAGTGGAACGAAAACCTACGTTAAGGGATTTGGTCATGAGATTATCAAAAAGGATCTTCCCTAGATCCTTTAAATTAATAAATGAAGTTTTAAATCA  
8601 ATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTTCCATCAGTTG  
8701 CCTGACTCCCGCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCACGCTCACCGGCTCCAGA

AseI (8882)

8801 TTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCCGAGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCT  
8901 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTAGCT  
9001 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAAGTAAAGTTGGC  
9101 CGCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAACCAAG  
9201 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCTGGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTAAAAAGTGTCTATCA

XmnI (9309)

9301 TTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATC  
9401 TTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC  
9501 TTCCTTTTCAATATTATGAAGCATTATCAGGGTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCGCG  
9601 GCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGC  
9701 GCGTTCCGTGATGACGGTGAACCTCTGACACATGCAGTCCCGGAGACGGTACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCA

NdeI (9881)

9801 GGCGCGTCAGCGGTGTTGGCGGGTGTGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTAAGTGTGAGAGTGCACCATATGGATCTCGATAACAAAA

PacI (9942)

NotI (9931)

9901 AACCCCGCCCCGGCGGGTTTTTTGTTAGCGGCCGCTTAATTA

pWHERE-rEF1 sequence

1 CAGCTTTGTCACAGCGGACCCCAACCTATGCCGCGTCTGCCGAGCAATATGTAGTATTGTACTGCCACCACGGGCATCGTCTGTCCATTTAGCTATAGC  
101 CAAATCTGCACACGCTGGAGAGTGAACCGATCGATCGTGAAGGGCGCAAGACTGAAGGAGCTACCCAAGAAATGTGTGTTGTACCACCCATGACCCCTTA  
201 TGAATCATTGAATGCTATGCCTGAGTGACCCATGAGTTTGCATAGGTGAACCGCAATTTTGGTCACCTCTATGATAAAGTATGTGGGACTGCACTGGTC  
301 GCTGCTCGGCAACTTCGGTCTTACCAGCCACTGACGATCTCGGGCTGTGTAGGGAATGAGTCAAGTTCTCTGGTTCAGTGTGAAGGGAACATTCCAGA  
401 GGTGCACACATCTTACCACCCCTATGAATCCCTATTTGGGTGACCCCTGGGATATTGCTGGGAATGAATCGTCCCCCAAGTTGGCAGCATTGGGCCAC  
501 GATATATAGGAGTATGCTGCCACCGCGGTAGCATCCGTTCCCTTGTGACACATAACAGCTTCTATGCCTTCTATAGTGAGCCACACTGGCTGGTTTT  
601 GGGGTTCACTGACCAAAGGGACCCCTCCAGAAACACAAGTGTTC AACCTCTTATAAACCATATCGACCACTGAGGCATAGCGGCTTCGGACATTGCTG  
701 TGGGGAACCCGAACCTTTGGCCCTTGGACATTGTCATGGGCAAACCTTAACTTTGGCTGTCTCTGGAACATAATGTTAAGATGACAGTACCAGCGCAG  
801 CAATTTGGTCTTTCCACTCACAACGGCTTTTGTGCTTTCTGGCATCGAACCCATGCACTGGTTATGGGGTCTGAGACCAAAGGAGACCATTCCGGAAG  
901 GGCATAGGTGTCCTGCCTTCTGCTTTTAAACAAGGCTCTCCGGGACAGTGCAAAACAGGTGAACCCCAACTTTGCCATAAGTACGATTATCTGCCACAAA  
1001 CAGCCAGGGTCTACCACCTCTCAATTGATTTTGGGCTGACACCCAGGCTTGATGTAGGATTCCTCAGCTGCCAAGCTGGCAGCTGACCCCATTTGA  
1101 GAGAGAAATGCAGTTTCAGAAATGTTTGCAGCCCTGAGCCGGAGATCATTAGCATCTGAACGCCCAATTAAGAATACGAAAGCGCAATCACCAGACTTC  
1201 TTGTTGGCGGTTCTTAAGTGATTCTTTGGGTAGGGAGGTCAGGTGTCCTCATAGATGTGCAGAATTTGAGGACCATGCTTAGTGGGGTCTGCATTATGG  
1301 TACTGCAATACATTCATGATCACCACACATAGTAGCTATACTTCAATTTTACACAATAGCGCTGATGGCCCAGAACCTATAAGTCAGATACTGAG  
1401 ATAGCTTTGAGAACGTTTTATCAAGGACTAGCATGAACCCCTGGCCTCATGAAGCCATGACTATGGGATCATAGATGGTATAGGGGAGAAACTCAA  
1501 TCAGTTGCAATCCGTTTTAGGACTGCGATGTACGAGACTTCACTGCCGCGTGCGCAACCTGGTCTTTACACAAAAGGATTCTTTGCAGAGAGTAAG  
1601 CCGACCTTGTGATTTGGGAGTCCGAGTCCACGAGGTACCAGCCTAGAAAATGCATGTGTCCTGCCCCCTAGTGGCATTGTGACCCCTGAGGTACT  
1701 GAACTTGGGTGACCCACAGCATTGCCATTTGTGAATCCAATACCAGGGTGGGGGGCTCTTTAGGTTTGGCGCAATCGATTTTGTGCCACCACGGCG  
1801 CAACTCCCGGTATAAACCCACAACCTGATTACGAGACGTC AAGAATAGGGCATGGTCTCCTTGCAGAATCTTATGCCTCCTGGATGCTCGTGTGAA  
1901 TGTAGCATGTTCTTTGAGTCTGGGTGTAAGTGCCCTGCACGCTCATCCCCGGACATGAAAATAGAATCTCTATTTTCTCACCAACCTTTTTCTTTCC  
2001 TGTGGTGATTCCGGAACTGTAGGCAATGGCTCTG CAGGCCCACTAGTGGAGCCGAGGTAATTCATACAAAAGGAGGGATCGCCTTCGCAAGGGGAG  
2101 AGCCAGGGACCGTCCCTAAATTTCTACAGACCCAAATCCCTGTAGCCGCCACGACAGCGGAGGAGCATGCGCCAGGGCTGAGCGGGGTAGATCA  
2201 GAGCACACAAGCTCACAGTCCCCGGGTGGGGGAGGGGCGCTGAGCGGGGCCAGGAGCTGGCGGGGCAAACCTGGGAAAGTGGTGTCTGTGTC  
2301 TGCTCCGCCCTCTTCCCGAGGTGGGGGAGAACGGTATATAAGTGCCTGATCGCCTTGGACGTTCTTTTTCGCAACGGGTTTCCGTCAGAACGCAGg  
2401 tgagtggcgggtgtggcttccgcgggccccggagctggagccctgctctgagcgggccccggctgatatgcagtgctcgtccgcagggttagctgtgagc  
2501 attcccacttcgagtgccggggcgtgccccgggtgagagtgagagcctagcggcaaccccgtagcctcgcctcgtgtccgcttaggcctagcgtgtgtg  
2601 tccgcccgcgctgccactccggccgcaactatgctttttgtccttgcctgcctcgattgccttccagcagcatgggctaacaaaggagggtgtgggg  
2701 ct cactcttaaggagcccatgaagcttacgttggataggaatggaagggcaggagggcgactggggccccccgccttcggagacatgtccgacgcca  
2801 cctggatggggcgaggcctgtggctttccgaagcaatcgggcgtgagtttagcctacctgggccaatgtggccttagcactgggcacggtctggcctggcg  
2901 gtgcccgttcccttgctcccaacaagggtgagccctcccgcggcaccagttgcttgccggaagatggccgctcccggggccctgttgcaagga  
3001 gctcaaatggaggacgcccgggtggagcgggggggtgagtcacccacacaaaggaagggccttgcccctcggcgccgctgcttctctgac  
3101 cccgtggtctatcggccgeatagtcacctcgggcttctcttgagcaccgctcgtcggcgggggggaggggatctaattggcgttggagttgttcaatt  
3201 tgggtgggtggagactagtcaggccagcctggcgtggaagtcatcttggaaattgcccccttgagtttggagcaggctaatctcaagcctcttagcg  
3301 gtcaaaggtatcttaaacccgttccagGTGTTGTGAAAGCCACCGCTAATTC AAAGCAACCATGGACCTGTTGTGCTGCAAAGGAGAGACTGGGA  
3401 GAACCTGGAGTGACCCAGCTCAACAGACTGGCTGCCACCCTCCCTTTGCCTCTTGAGGAACTCTGAGGAAGCCAGGACAGACAGGCCAGCCAGCCAGCAG  
12 ▶ uAsnProGlyValThrGlnLeuAsnArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluGluAlaArgThrAspArgProSerGlnGln  
3501 CTCAGGTCCTCAATGGAGAGTGGAGTTTGCCTGGTTCCCTGCCCTGAAGCTGTGCTGAGTCTTGGCTGGAGTGTACCTCCAGAGCCAGTCCCA  
46 ▶ LeuArgSerLeuAsnGlyGluTrpArgPheAlaTrpPheProAlaProGluAlaValProGluSerTrpLeuGluCysAspLeuProGluAlaValProL

3601 AGAAGAAGAGAAAGTTGAGGCTGACACTGTTGTGGTCCAAAGCAACTGGCAGATGCATGGCTATGATGCCCCATCTACACCAATGTCACCTACCCCAT  
79 yslLysLysArgLysValGluAlaAspThrValValValProSerAsnTrpGlnMetHisGlyTyrAspAlaProl IeTyrThrAsnValThrTyrProl I  
3701 CACTGTGAACCCCTTTGTGCCACTGAGAACCCACTGGCTGTACAGCCTGACCTTCAATGTTGATGAGAGCTGGCTGCAAGAAGGCCAGACCAGG  
112 eThrValAsnProProPheValProThrGluAsnProThrGlyCysTyrSerLeuThrPheAsnValAspGluSerTrpLeuGlnGluGlyGlnThrArg  
3801 ATCATCTTTGATGGAGTCAACTCTGCCTTCCACCTCTGGTGAATGGCAGGTGGGTTGGCTATGGCCAAGACAGCAGGCTGCCCTCTGAGTTTGACCTCT  
146 I IeIeIePheAspGlyValAsnSerAlaPheHisLeuTrpCysAsnGlyArgTrpValGlyTyrGlyGlnAspSerArgLeuProSerGluPheAspLeu  
3901 CTGCCCTTCTCAGAGCTGGAGAGAACAGGCTGGCTGTCAATGGTGTCAAGTGGTCTGATGGCAGCTACCTGGAAGACCAAGACATGGTGAGGATGCTCTGG  
179 erAlaPheLeuArgAlaGlyGluAsnArgLeuAlaValMetValLeuArgTrpSerAspGlySerTyrLeuGluAspGlnAspMetTrpArgMetSerGl  
4001 CATCTTCAGGGATGTGAGCCTGCTGCACAAGCCACCACCCAGATTCTGACTTCCATGTTGCCACCAGGTTCAATGATGACTTCAGCAGAGCTGTGCTG  
212 yI IePheArgAspValSerLeuLeuHisLysProThrThrGlnIeIeSerAspPheHisValAlaThrArgPheAsnAspAspPheSerArgAlaValLeu  
4101 GAGGCTAGGCTGAGATGTGTGGAGAACCTCAGAGACTACCTGAGAGTCAAGTTCAGGCTCTGGCAAGGTGAGACCAGGTTGGCTCTGGCACAGCCCT  
246 GluAlaGluValGlnMetCysGlyGluLeuArgAspTyrLeuArgValThrValSerLeuTrpGlnGlyGluThrGlnValAlaSerGlyThrAlaProP  
4201 TTGGAGGAGAGATCATTGATGAGAGAGGAGGCTATGCTGACAGAGTCAACCTGAGGCTCAATGTGGAGAACCCCAAGCTGTGGTCTGCTGAGATCCCCAA  
279 heGlyGlyGluIeIeIeAspGluArgGlyGlyTyrAlaAspArgValThrLeuArgLeuAsnValGluAsnProLysLeuTrpSerAlaGluIeProAs  
4301 CCTCTACAGGCTGTGGAGCTGCACACTGCTGTGAGCAGCCTGATGAAGCTGAAGCCTGTGATGTTGGATTACAGAAAGTACAGGATGAGAAATGGC  
312 nLeuTyrArgAlaValValGluLeuHisThrAlaAspGlyThrLeuIeGluAlaGluAlaCysAspValGlyPheArgGluValArgIeGluAsnGly  
4401 CTGCTGTGCTCAATGGCAAGCCTCTGCTCATCAGGGGAGTCAACAGGCATGAGCACCACCTCTGCATGGCAAGTGTGATGATGAACAGACAATGGTGC  
346 LeuLeuLeuLeuAsnGlyLysProLeuLeuIeArgGlyValAsnArgHisGluHisHisProLeuHisGlyGlnValMetAspGluGlnThrMetValG

**EcoRV (4506)**

4501 AAGATATCTGTAATGAAGCAGAACAACCTTCAATGCTGTCCAGGTGCTCTCACTACCCCAACCACCTCTCTGGTACACCCTGTGTGACAGGTATGGCCT  
379 InAspI IeLeuLeuMetLysGlnAsnAsnPheAsnAlaValArgCysSerHisTyrProAsnHisProLeuTrpTyrThrLeuCysAspArgTyrGlyLe  
4601 GTATGTTGTTGATGAAGCCAACATTGAGACACATGGCATGGTGCCCATGAACAGGCTCACAGATGACCCAGGTGGCTGCCATGTCTGAGAGAGTG  
412 uTyrValValAspGluAlaAsnIeGluThrHisGlyMetValProMetAsnArgLeuThrAspAspProArgTrpLeuProAlaMetSerGluArgVal  
4701 ACCAGGATGGTGCAGAGAGACAGGAACCCCTCTGTGATCATCTGGTCTCTGGGCAATGAGTCTGGACATGGAGCCAAACCATGATGCTCTCTACAGGT  
446 ThrArgMetValGlnArgAspArgAsnHisProSerValIeIeIeTrpSerLeuGlyAsnGluSerGlyHisGlyAlaAsnHisAspAlaLeuTyrArgT  
4801 GGATCAAGTCTGTTGACCCAGCAGACCTGTGCAGTATGAAGGAGGTGGAGCAGACACCACAGCCACAGACATCATCTGCCCATGTATGCCAGGGTTGA  
479 rpl IeLysSerValAspProSerArgProValGlnTyrGlyGlyGlyAlaAspThrAlaThrAspIeIeIeCysProMetTyrAlaThrAspIeIeIe  
4901 TGAGGACAGCCCTTCCCTGTGTGCCAAGTGGAGCATCAAGAAGTGGCTCTCTCTGCTGGAGAGACCAAGCCTCTGATCTGTGTAATGACAT  
512 pGluAspGlnProPheProAlaValProLysTrpSerIeLysLysTrpLeuSerLeuProGlyGluThrArgProLeuIeLeuCysGluTyrAlaHis  
5001 GCAATGGCAACTCTCTGGGAGGCTTTGCCAAGTACTGGCAAGCCTTCAGACAGTACCCCAAGGCTGCAAGGAGGATTTGTGTGGGACTGGGTGGACCAAT  
546 AlaMetGlyAsnSerLeuGlyGlyPheAlaLysTyrTrpGlnAlaPheArgGlnTyrProArgLeuGlnGlyGlyPheValTrpAspTrpValAspGlnS  
5101 CTCTCATCAAGTATGAGAAATGGCAACCCCTGGTCTGCATATGGAGAGCTTTGGTGACACCCCAATGACAGGCTTCTGCAATGAATGGCTGGT  
579 erLeuIeLysTyrAspGluAsnGlyAsnProTrpSerAlaTyrGlyGlyAspPheGlyAspThrProAsnAspArgGlnPheCysMetAsnGlyLeuVa  
5201 CTTTGCAGACAGGACCCCTCACCCCTGCCTCACAGAGGCCAAGCACCAGCAACAGTCTCTCCAGTTCAGGCTGTCTGGACAGACCATTGAGGTGACATCT  
612 IPheAlaAspArgThrProHisProAlaLeuThrGluAlaLysHisGlnGlnPhePheGlnPheHisGlnLeuSerGlyGlnThrIeGluValThrSer  
5301 GAGTCACTCTCAGGACTCTGACAAATGAGCTCCTGCATGGATGGAGAGCTTGGGAGAGCTTGGGCAAGCCTCTGGCTTCTGGTGGTCTGGAATGGCC  
646 GluTyrLeuPheArgHisSerAspAsnGluLeuLeuHisTrpMetValAlaLeuAspGlyLysProLeuAlaSerGlyGluValProLeuAspValAlaP  
5401 CTCAAGGAAAGCAGCTGATTGAACCTGCCTGAGCTGCCTCAGCCAGAGTCTGCTGGCAACTGTGGCTAACAGTGAAGGTTGGTTCAGCCCAATGCAACAGC  
679 roGlnGlyLysGlnLeuIeGluLeuProGluLeuProGlnProGluSerAlaGlyGlnLeuTrpLeuThrValArgValValGlnProAsnAlaThrAl  
5501 TGTGCTGAGGACAGCCACATCTCTGCATGGCAGCAGTGGAGGCTGGCTGAGAACCTCTCTGTGACCCTGCCTGCTGCATGCCATCCCTCACCT  
712 aTrpSerGluAlaGlyHisIeSerAlaTrpGlnGlnTrpArgLeuAlaGluAsnLeuSerValThrLeuProAlaAlaSerHisAlaIeProHisLeu  
5601 ACAACATCTGAAATGGAATCTGCACTTGCATTGAGCTGGGCAACAAGAGATGGCAGTTCACAGGCAGTCTGGCTTCTCTCAGATGTGGATTGGAGACAAGA  
746 ThrThrSerGluMetAspPheCysIeGluLeuGlyAsnLysArgTrpGlnPheAsnArgGlnSerGlyPheLeuSerGlnMetTrpIeGlyAspLysL  
5701 AGCAGCTCTCACCCCTCAGGACCAATTCACAGGCTCCTCTGCAATGAGACTTGGAGTGTCTGAGGCCACCAGCTGACCCAAATGAGGTTGGGT  
779 ysGlnLeuLeuThrProLeuArgAspGlnPheThrArgAlaProLeuAspAsnAspIeGlyValSerGluAlaThrArgIeAspProAsnAlaTrpVa  
5801 GGAGAGGTGGAAGGCTGCTGGCACTACCAGGCTGAGGCTGCCTGCTCCAGTGCACAGCAGACACCCTGGCTGATGCTGTTCTGATCACCACAGCCCAT  
812 IGluArgTrpLysAlaAlaGlyHisTyrGlnAlaGluAlaAlaLeuLeuGlnCysThrAlaAspThrLeuAlaAspAlaValLeuIeThrThrAlaHis  
5901 GCTTGGCAGCAGCCAGCAGCCTGTTTCATCAGCAGAAAGACCTACAGGACTGATGGCTCTGGACAGATGGCAATCAGAGTGGAGTTGGAGTTGGCT  
846 AlaTrpGlnHisGlnGlyLysThrLeuPheIeSerArgLysThrTyrArgIeAspGlySerGlyGlnMetAlaIeThrValAspValGluValAlaS  
6001 CTGACACACCTCACCCCTGCAAGGATGGCTGAACTGCAACTGGCACAGGTGGCTGAGAGGTTGAACTGGCTGGGCTTAGGCCCTCAGGAGAATACCC  
879 erAspThrProHisProAlaArgIeGlyLeuAsnCysGlnLeuAlaGlnValAlaGluArgValAsnTrpLeuGlyLeuGlyProGlnGluAsnTyrPr  
6101 TGACAGGCTGACAGCTGCCTGTTGACAGGTGGGACTGCCTCTGTGACATGTACACCCTTATGTGTTCCCTTCTGAGAATGGCCTGAGGTGTTGGC  
912 oAspArgLeuThrAlaAlaCysPheAspArgTrpAspLeuProLeuSerAspMetTyrThrProTyrValPheProSerGluAsnGlyLeuArgCysGly  
6201 ACCAGGGAGCTGAACTATGGTCTCACAGTGGAGGGAGACTCCAGTTCACATCTCCAGTACTCTCAGCAACAGCTCATGGAACCTCTCACAGGC  
946 ThrArgGluLeuAsnTyrGlyProHisGlnTrpArgGlyAspPheGlnPheAsnIeSerArgTyrSerGlnGlnGlnLeuMetGluThrSerHisArgH  
6301 ACCTGCTCATGCAGAGGAGGAACCTGGCTGAACATTGATGGCTTCCACATGGGCAATGGAGGAGTACTCTGGTCTCTCTGTGTCTGTGAGT  
979 isLeuLeuHisAlaGluGluGlyThrTrpLeuAsnIeAspGlyPheHisMetGlyIeGlyGlyAspAspSerTrpSerProSerValSerAlaGluPh

**NheI (6460)**

6401 CCAGTTATCTGCTGGCAGGTACCACTATCAGCTGGTGTGGTCCAGAAGTAAACCTGAGCTAGCATTATCCCTAATACCTGCCACCCCACTTAAATCAG  
1012 eGlnLeuSerAlaGlyArgTyrHisTyrGlnLeuValTrpCysGlnLys•••••  
6501 TGGTGAAGAAGCAGCTCAGAAGTGTGGTTTCAATGGCCAAATGAAAGTATTAGTAAAAGACTGGTAAATGATAACATGCATCGTAAAACCTTCAGA

6601 AGGAAAGGAGAAATGTTTGTGGACCACTTGGTTTTCTTTTTGCGTGTGGCAGTTTTAAGTTATTAGTTTTTAAATCAGTACTTTTTAATGAAACAA  
6701 CTTGACCAAAAAATTTGTCACAGAAATTTGAGACCCATTAATAAAGTAAATAGAAACCTGTGTCTCTTTGGTCAACACCGGACATTTAGGTTAGG

**HpaI (6858)**

6801 ACATCTAATTTCTGGTTTTACGAATCTGAAAACCTCTTGAATAATGTAATCTTGAAGTTAACTTCTGGTGGAGAATAGGGTTGTTTTCCCCCACA  
6901 TTGGAAGGGGAAGGAATATCATTAAAGCTATGGGAGGTTTCTTTGATTACAACACTGGAGAGAAATGCAGCATGTTGCTGATTGCCTGTCACTAAA  
7001 AGGCCAAAACTGAGTCTTTGGGTTGCATAGAAAGCTGCCTGCAGAGCTTTGTACAGCGGACCCCAACCTATGCCCGCTGCTGCCAGCAATATGTAGTA  
7101 TTGTAAGTCCACCACCGGCATCGTCTGTCCATTTAGCTATAGCCAAATCTGCACAGCGTGGAGAGTGAACCGATCGATCGTGAAGGGCGCAAGACTGAA  
7201 GGAGCTACCAAGAAATGTGTGTACCACCCCATGACCTTATGAATCATTGAATGCTATGCCTGAGTGACCCATGAGTTTGCATAGGTGAACCGCA  
7301 ATTTTGGTCACTCTATGATAAAGTATGTGGACTGCCTGGTGGTCTGCGGCAACTTGGTCTTACCAGCCACTGACGATCTGGGCTGTGTAGGGAA

7401 TGAGTCAAGTTCTCTGGTTCAGTGTGTAAGGGAACCATTCCAGAGGTGCACACATCTTACCACCCTATGAATCCCTATTTGGGTGACCCCTGGGATATTG  
7501 CTGGGAATGAATCGCTCCCCCAAGTTGGCAGCATTGGGCCACGATATATAGGAGTATGCTGCCACC CGCGGTAGCATCCGTTCCCTTGTGCACATA  
7601 ACAGCTTCTATGCCTTCTATAGTGAGCCACACTGGCTGGTTTTGGGGTTCAGTGACCAAAGGACCCTCCAGAAACACAAGTGTCCAACCTCTTAT  
7701 AAACCATATCGACCACTGAGGCATAGCGGCTTCGGACATTGCTGTGGGCGAACC CGAACCTTTGGCCCTTGGACATTGTCATGGGCAAACTTAACCTTGG  
7801 CTGTCTCTGGAACATAATGTTAAGATGACAGTACCAGCGCAGCAATTTGGTCTTTCCACTCACAACGGCTTTTGTGCTTCTGGCATCGAACCATG

BspEI (7938)

7901 CACTGGTTTATGGGGTCTGAGACCAAAGGAGACCATTCCGGAAGGGCATAGGTGTCCTGCTTTTAAACAAGGCTCTCCGGGACAGTGCAAAACA  
8001 GGTGAACCCCAACTTTGCCATAAGTACGATTATCTGCCACAAACCAGCCAGGGTCTACCACCTCTTCAATTGATTTTGGGCTGACCCCAAGGCTTGA  
8101 TGTAGGATTCTCAGCTGCCAAGCTGGCAGCTGACCCCATGAGAGAGAATGCAGTTTCAGAATTGTTGCAGCCCTGAGCCGGAGATCATTAGCATCT  
8201 GAACGCCCAATTAGAATACGAAAGCGCAATCACCAGACTTCTTGTGGCGGTTCTTAAGTGATTCTTGGGTAGGAGGTCAGGTGCTCTCATAGA  
8301 TGTGCAGAAATTTGAGGACCATGCTTAGTGGGGTCTGCATTATGGTACTGCAATACATCCATGATCACCACACATAGTAGCTATACTTCAATTTTCACAC  
8401 AATAGCGCTGATGGCCCAAGCCCTATAAGTCAGATACCTGAGATAGCTCTTGAGAACGTTTTATCAAGGACTAGCATGAACCCCTGGCCTCATGAAGC  
8501 CCATGACTATGGGATCATAGATGGTATAGGGGAGAAACTCAATCAGTTGCAATCCGTTTTAGGACTGGATGTACGAGACTTCACTGCCGCCGTGGCG  
8601 CAACCTGGTCTTTACACACAAAGGATTCTTTGCAGAGAGTAAGCCGACCTTGTGATTGGGAGTCCGAGTCCACGAGTACCAGCTAGAAAATGCAT  
8701 GTGTCTGCCCCCTAGTGGCATTGTGACCCCTGAGGTAAGTGGGTGACCCACAGCATTGCCATTTGTGAATCCAATACCAGGGTGGGGG  
8801 GGCTCTTAGGTTGGCGCAATCGATTTTGTGCCACCACGCGCAACTCCCGGTATAAACCACAACTGATTGAGCAGACGTCGAAGAATAGGGCAT  
8901 GGTCTCCTTGAGAAATCTTATGCCTCCTGGATGCTCGTGTGAATGTAGCATGTTCTTTGAGTCTGGGTGTAAGTGCCTGCACGCTCATCCCCGGA

PacI (9082)

9001 CATGAAAATAGAATCTCTATTTTCTCACCAACCTTTTCTTCTTCTTGTGGTGATTGGGAACTGATAGCAATGGCTTTAATTA AAAACCCGCTTCGGCGG  
9101 GTTTTTTATGCATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCGTGTGGCGTTTTCCATAGGCTCGCCCCCTGACGAGC  
9201 ATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGTCCCTCGTGCCTCTCTGT  
9301 TCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGAAGCGTGGCGTTTTCTCATA GCTCAGCTGTAGGTATCTCAGTTGGGTAG  
9401 GTCGTTGCTCAAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGGCGCTTATCCGGTAACATATCGTCTGAGTCCAACCCGGTAAGAC  
9501 ACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGG  
9601 CTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTGATCCGGCAACAAACCACCGCT  
9701 GGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGT  
9801 GGAACGAAAACCTACGTTAAGGATTTTGGTCTAGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTA AAAATGAAGTTTAAATCAATCTA  
9901 AAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTCCATAGTTGCCTGA  
10001 CTCCCCGCTGTGATAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACC GGAGACCCACGCTCACCGCTCCAGATTTAT

AseI (10177)

10101 CAGCAATAAACAGCCAGCCGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATAATTGTTGCCGGGAAGCTAGAGT  
10201 AAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCATCGTGTGTCACGCTCGTCTTGGTATGGCTTCACTCAGCTCCGGT  
10301 TCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAAGTAAAGTTGGCCGAG  
10401 TGTTATCACTCATGTTATGGCAGCACTGCATAATTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT  
10501 CTGAGAATAGTGTATGGGCGACCGAGTTGCTCTTCCCGCGTCAATACGGGATAATACCGGCCACATAGCAGAACTTAAAAAGTCTCATCATTGGA  
10601 AAACGTTCTTCGGGCGCAAACTCTCAAGGATCTTACCCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAGTATCTTCCAGCATCTTTA  
10701 CTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCATACTCTTCT  
10801 TTTTCAATATTATTGAAGCATTATCAGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACA  
10901 TTTCCCGAAAAGTGCACCTGACGTCTAAGAAACATTATTATCATGACATTAACCTATAAAAATAGGCGTATCAGAGGCCCTTTCGTCTGCGCGGTT  
11001 TCGGTGATGACGGTGA AAAACCTCTGACACATGCAGCTCCCGGAGCGGTACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGGCGC

NdeI (11176)

11101 GTCAGCGGTGTTGGCGGTGTCGGGGTGGCTTAACTATGCGGCATCAGAGCAGATTGACTGAGAGTGCACCATATG GATCTCGATAACAAAAACCC

PacI (11237)

NotI (11226)

11201 CGCCCCGGCGGGTTTTTTGTTAGCGGCCGCTTAATTA