Howard Hughes Medical Institute
Research Laboratories

Stephen J. Elledge, Ph.D.
Investigator

Dear Colleagues,

Enclosed are several bacterial strains and a user manual for the Univector Plasmid Fusion System (UPS) that we have recently developed. The paper describing this system is currently being submitted. As you can see there are a number of recipient vectors included in this kit. Due to the large number of strains, we have aliquoted them into a sterile 96-well plate that is sealed. The strains have been stored frozen at -80°C in 9% DMSO prior to being sent to you. We have found that these strains can be successfully refrozen and stored after having been thawed for several days, so if you are not able to individually streak them out upon arrival you may be able to store them at -80°C and retrieve them later. You may also access the strains by poking a hole through the cover above the well of interest. A key for the identification of the strains by well position is included. Please remember that several strains bear no plasmids and must be grown with LB alone and the pUNI plasmid strains require kanamycin selection. The remaining strains are Amp resistant. Maps of the vectors are also included. The nomenclature for the vectors is pH for Host followed by a letter designating the organism the expression vector is designed for, B-bacterial, Y, yeast, I insect cells, M, mammalian. Each vector has a number to define the base vector and this is followed by the identity of the tag. For example, pHB1-HA3 is a bacterial vector that generates fusion proteins with a triple HA tag. pHB1-MYC3 has a triple MYC epitope tag. In addition there is a plamid identification number for each plasmid that identifies that particular plasmid in our strain collection.

The plasmids we are sending are new and have been made by my lab and Dr. Wade Harper's lab. Dr. Harper has generously allowed us to distribute them prior to publication. We have shown that the UPS system works for bacterial, yeast, mammalian and insect cell expression systems as well as conventional vectors designed for the same purpose. While most of these plasmids have been tested, a few have not yet been used. We would be interested in hearing about any success stories or problems. Please address email to the univector mail box univ@bcm.tmc.edu. We are in the process of generating a number of vectors that will facilitate the conversion of your favorite vector into an epitope-tagged lox containing recipient vector. If you are interested in these, contact us in the future. Cre protein is available from commercial sources but we have found that making our own using the Gst-Cre fusion works much better in vitro.

If you happen to make new univector recipient vectors with novel properties, we would like to obtain a sample of the vector. If you could send your plasmid to us with a map after it has been successfully tested, we would appreciate that.

Sincerely,

Steve

Steve Elledge

Baylor College of Medicine
One Baylor Plaza, T307, Houston, Texas 77030
(713) 798-5040 • Fax (713) 798-8717 • E-mail: selledge@bcm.tmc.edu

Overview:

The Univector Plasmid-fusion System, UPS, is a new cloning method developed in our lab for rapid, efficient and systematic construction of recombinant DNA molecules. UPS employs the *cre-lox* site-specific recombination of bacteriophage P1 (Sternberg et al. 1981) to catalyze *in vitro* plasmid fusion events between the Univector, a plasmid containing the gene of interest, and host vectors containing the regulatory information. Plasmid fusion events are genetically selected and result in placement of the gene of interest under the control of novel regulatory elements. UPS eliminates the need for restriction enzymes, DNA ligase, and many *in vitro* manipulations required for subcloning. It greatly simplifies the process of designing and performing subclonings and saves researchers significant amounts of time and expense.

The Univector, as shown in Figure 1, contains a *lox*P site placed directly adjacent to the 5' end of the polylinker for insertion of the gene or cDNAs. The *lox*P site has a single open reading frame that is in frame with the ATG of the NdeI and NcoI sites of the polylinker. This facilitate the subsequent generation of protein fusions as noted below. Following the polylinker are bacterial and eukaryotic transcriptional terminators to facilitate 3' end formation of transcripts. Most importantly, the Univector has a conditional origin of replication derived from R6Kγ that allows its propagation only in bacterial hosts expressing the *pir* gene. Finally, the Univector has the *neo* gene from *Tn5* for selection in bacteria.

Host vectors include the Col E1 origin of replication and the *bla* gene for propagation and selection in bacteria, a *lox*P site for plasmid fusion and a specific promoter residing upstream of and adjacent to the *lox*P site. In the case in which the pHOST plasmid also contains sequences for translational initiation, e.g. a GST tag, the pHOST *loxP* site is placed in the same reading frame as the *lox*P site on the Univector and the target gene must be cloned in the Univector in frame with the *lox*P open reading frame. Host vectors may also contain sequences responsible for propagation, selection and maintenance in organisms other than *E. coli.* Our lab is making a large collection of host vectors that are suitable for expression in bacteria, yeast, mammalian and insect cells. Some of these host vectors have been made and tested. They are listed in the table below (Figure 2). In general, host vectors are derivatives of existing vectors by insertion of an oligo linker containing a *lox*P or *lox*H[1] site in the polylinker of these vectors. For pHOST plasmids that also have epitope tags, the *lox*P site is inserted in frame with the coding sequence of the epitope tags.

Once the gene of interest is cloned in the Univector by conventional means, it can be rapidly transferred into multiple host vectors by UPS (Figure 3). The UPS reactions are carried out *in vitro* in the presence of purified GST-Cre fusion proteins. The GST-Cre fusion protein is easily produced and purified from *E. coli* in large quantities with glutathione-sepharose beads (see GST-Cre purification protocol). Plasmid fusions between the Univector and the pHOST plasmid via the *lox*P site create a plasmid dimer, thereby placing the target gene under the control of the regulatory sequences on the host vector. Selection for the recombinant products of UPS is achieved by selecting for kanamycin resistance after transformation into a *pir- E. coli* strain (standard cloning strains such as XL1 Blue, DH5α, JM107, etc.) because the *neo* gene on the Univector can only be propagated when covalently linked to an origin of replication that is functional in a *pir-* background. UPS reactions can also be carried out in vivo but this method is slightly more complicated.

[1] *LoxP* consists of two 13 bp inverted repeats flanking an 8 bp central sequence. *loxH* is a mutant derivative of *loxP*, which has mutations at 3, 6, 9 nucleotide of the left repeat, creating three mismatches between the two repeats. The sequence is shown on the vector maps.

Cloning genes into the Univector

One of the main advantages of cloning by UPS is that, for each gene of interest, only one conventional subcloning is needed. Once the gene or cDNA is appropriately cloned into the Univector, it can be used to rapidly convert into any kind of recipient vectors to make expression constructs. Cloning by UPS is much simpler and faster than the conventional method. Not only does UPS avoid the use of restriction enzymes, DNA ligase and many *in vitro* manipulations for subcloning, it also saves researchers time in designing subcloning strategies. Below are the two key rules for cloning genes into the Univector.

1. The gene, or cDNA should be cloned in the same reading frame as the *lox*P site (Notice the *lox*P sequence contains only one open reading frame in the direction shown). Only then can it be used to seemlessly create expression constructs, transcriptional as well as translational fusions. There is no need to worry about reading frames or restriction site incompatibility. Everything should be automatically in frame because of the precise nature of the *cre-lox* site specific recombination.
2. It is recommended that one engineer a restriction site at the start codon, eg. NdeI or NcoI, such that only the ORF and no extra 5' untranslated sequence will be brought into the Univector. This will ease the construction of epitope-tagged constructs.

In vitro UPS reaction:

In general, the UPS reaction are carried out using 0.1-0.4 μg DNA each of the Univector and host vector in a volume of 20 μl in 1X Buffer S (50mM Tris.Cl pH 7.5, 10 mM $MgCl_2$, 30 mM NaCl, 0.1 mg/ml BSA). The ratio of the two plasmids does not matter much but you will get more kanamycin resistant transformants if more Univector plasmid is used. We normally use 1μl of the GST-Cre elution (0.1 to 0.2 μg) for each reaction. The reaction mixture is assembled on ice, incubated immediately at 37°C for 20 minutes and followed by 65°C for 5 minutes to inactivate the GST-Cre enzyme. However, the Cre inactivation is probably not necessary if you are going to immediately transform the mixture. Half of the reaction is used to transform heat-shock competent DH5α or JM107 cells and transformants are plated onto LB/kanamycin (40 μg/ml) plates. Typically you will get 100 to 10,000 colonies. If you are using electroporation, 1 μl of the reaction added to 50 μl of cells is sufficient and will not arch. However, if one is performing a library conversion it is best to concentrate the DNA and remove the salt by precipitation or other methods prior to electroporation.

The efficiency of the GST-Cre mediated UPS reaction approaches 17%, as measured by the total number of Kn^R transformants vs the total number of Ap^R transformants, when equal molar amounts of the Univector and pHOST plasmid are used in the UPS reaction. By restriction analysis, 100% of the plasmids isolated from Kn^R transformants are correct fusion products via the *lox*P site. Most are dimers, with occasional trimers that in theory may result from two sequential fusion events or single fusion event between a pre-existing monomeric and a dimeric substrates. The production of trimers can be eliminated if gel purified monomeric supercoiled DNA is used suggesting the latter explanation is correct. Trimeric constructs usually express as well as dimers.

Occasionally we have found that UPS reaction give us exculsively trimers. In this case the Univector subclone was found to be a dimer. This occurs occasionally when too much DNA is used in ligation reactions (remember pUNI is puny). It is best to look at your Univector subclone uncut on a gel to determine if it is a monomer or dimer if it apppears you are obtaining timers from the UPS reaction.

RPM prepped DNA performs as well as or better than CsCl prepped DNA in our hands. We can routinely perform a dozen UPS reactions plus tranformations in an hour. We normally examine only two Kn^R transformants from each plate and at least one of them will turn out to be a dimer. In most cases, both are dimers.

Restriction Analysis of the UPS products

To keep the digestion pattern analysis simple, it is preferable to use enzymes that cut only once or twice in each plasmid. Try to avoid enzymes that cut near the *lox*P site on both parental plasmids because the digestion pattern will not be informative. It is also recommended to digest both parental plasmids along with the UPS products to use as a reference. The digestion pattern for the dimeric UPS product should show certain bands missing and new bands appearing when compared to the parental plasmids. Restriction analysis of a trimer will produce one parental sized band in addition to the new bands.

Bacterial Strains (Relevant Genotype in Bold)

Strain name	Genotype
BW23473	Δlac-169 rpoS(Am) robA1 creC510 **hsdR514 endA recA1** uidA(ΔMluI)::**pir+(wt)**

(This strain allows low copy pUNI replication, ~15 copies per cell)

Strain name	Genotype
BW23474	Δlac-169 rpoS(Am) robA1 creC510 **hsdR514 endA recA1** uidA(ΔMluI)::**pir-116**

(This strain allows high copy pUNI replication, ~250 copies per cell)

Strain name	Genotype
BUN10	hisG4 thr-1 leuB6 thi-1 lacY1 galK2 ara-14 xyl-5 mtl-1 argE3 rfbD1 mgl-51 kdgK51 Δ(gpt-proA)62 rpsL31 tsx33 supE44 **recB21 recC22 sbcA23 hsdR::cat-pir-116(Cm^R)**

(This strain allows high copy pUNI replication and has the proper genetic background for homologous recombination and blue/white screening for lacO reconstitution. It is $endA^+$ so miniprep procedures using phenol extraction need to be used.)

GST-Cre purification protocol:

1. Grow a 500ml culture of the BL21 (pQL123 GST-Cre) strain to OD_{600}= 0.8 in LB/Amp (100 μg/ml) media.

2. Add IPTG to a final concentration of 0.4 mM. Induce for 3h at 37°C or 6h at room temperature. From our experience, GST-Cre prepared from the room temperature induction has a higher specific activity.

3. Pellet cells 5K for 10 min. You can freeze the cell pellet at –80 °C at this point.

4. Resuspend the cell pellet in ice cold 10 ml NETN freshly supplemented with 1mM PMSF+ 5μg/ml leupeptin + 5μg/ml antipain and transfer to a clean 50 ml Falcon tube for sonication.
(NETN: 0.5% NP-40, 20mM TrisHCl pH 8.0, 100mM NaCl, 1mM EDTA)

5. Sonicate the 10 ml of cells at 4°C at level "2", 5 x 15 sec with an interval of 30 sec on ice between sonications. Try to avoid foaming which will denature the protein by keeping the sonicator tip well below the surface of the solution. The solution will become homogeneous and viscous during sonication.

6. Aliquot into six 2 ml eppendorf tubes, spin at 14,000 rpm for 15 min at 4°C to clear the cell extract. During this clearing step, prepare the glutathione-sepharose beads needed in step 7.

7. Prepare glutathione-sepharose beads (Pharmacia) in 200 μl per 2 ml eppendorf tubes. Wash three times with 1.5 ml NETN. Remove excess liquid, add the supernatant from step 6 to the beads and rotate at 4°C for one hour. The beads should be treated gently. We normally pellet them at 2500 rpm for 15 sec. using an eppendorf centrifuge.

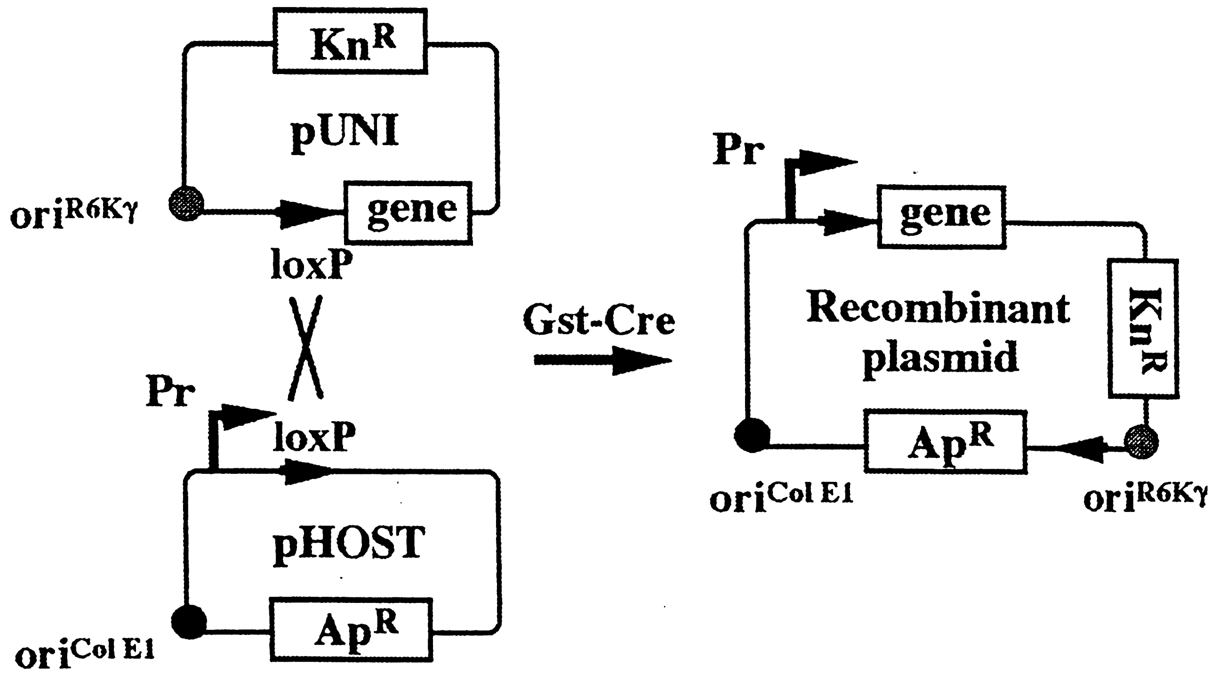
8. Wash beads 4 times with 1.5 ml of ice cold NETN + protease inhibitors as in (4). Simply add the buffer, invert several times and spin at 2,500 rpm for 15 sec. This can be done at your bench for convinience.

9. Wash two times with wash buffer to replace the NETN prior to elution. Pellet the beads and remove all of the supernatant with thin-tip pipette tips (Phenix) that won't take up the beads.
(Wash Buffer: 100 mM Tris.Cl (pH 8.0), 120 mM NaCl, 1mM PMSF)

10. Elute three times by adding 150 μl elution buffer, rotating the tubes each time at 4°C for 45 minutes followed by a 15 sec spin at 2500 rpm. Collect and save supernatant each time. Typically you will get 0.1-0.2 μg GST-Cre protein per μl of elution buffer.
(Elution Buffer: 100 mM Tris.Cl (pH 8.0), 120 mM NaCl, 1mM PMSF, 20 mM Glutathione).

11. Aliquot the GST-Cre elution into 80 μl aliquots or smaller and store at -80 °C. Check the quantity and quality of the GST-Cre prep by running 5 μl of the elution on a 10% SDS PAGE gel. Once a tube is taken out, it can be stored at 4°C for up to a month. However, the enzymatic activity of GST-Cre will decrease with time.

Figure 1



Epitope Tag Sequences

1. HA

```
Nco I                                                              Nhe I                Nde I
|                                                                  |                    |
CC  ATG  GCT  TAC  CCA  TAC  GAT  GTT  CCA  GAT  TAC  GCT  AGC  TTG  GGT  GGT  CAT  ATG
    M    A    Y    P    Y    D    V    P    D    Y    A    S    L    G    G    H    M
    3/1                                               33/11
```

2. HA3

```
Nco I
|
CC  ATG  GCA  GGT  TAC  CCA  TAC  GAC  GTT  CCT  GAC  TAT  GCG  TCA  CTC  TAC  CCC  TAT  GAC  GTA  CCG
    M    A    G    Y    P    Y    D    V    P    D    Y    A    S    L    Y    P    Y    D    V    P
   6/1                                 36/11
     Nsi I                                                                                Xba I
     |                                                                                    |
GAT  TAT  GCA  TCC  CTA  TAT  CCG  TAT  GAT  GTT  CCA  GAT  TAC  GCT  TCT  CTA  CGT  TCC  TCT  AGA
D    Y    A    S    L    Y    P    Y    D    V    P    D    Y    A    S    L    R    S    S    R
66/21                                         96/31
               Nde I
               |
GGC  GTC  CAC  CAT  ATG
G    V    H    H    M
126/41
```

3. MYC3

```
                              Eag I
NcoI                          NotI
|                             ||
CC  ATG  GGA  TTC  GAG  CTA  TGC  GGC  CGC  TCT  GAG  CAA  AAG  CTC  ATT  TCT  GAA  GAG  GAC  TTG  AAT
    M    G    F    E    L    C    G    R    S    E    Q    K    L    I    S    E    E    D    L    N
    3/1                                               33/11
                        Bcl I
                        |
GGA  GAA  CAG  AAA  TTG  ATC  AGT  GAG  GAA  GAC  CTC  AAC  GGT  GAG  CAG  AAG  TTA  ATA  TCC  GAG
G    E    Q    K    L    I    S    E    E    D    L    N    G    E    Q    K    L    I    S    E
63/21                                         93/31
                             Eag I
                             Not I        Spe I                    Apa I    Nde I
                             ||           |                        |        |
GAG  GAT  CTT  AAT  AGT  AGC  GGC  CGC  ACT  AGT  CTC  GAA  GGG  GGG  CCC  CAT  ATG
E    D    L    N    S    S    G    R    T    S    L    E    G    G    P    H    M
123/41                                         153/51
```

4. MYC9

```
Nco I
|
CC  ATG  GCT  AGT  (GGT  GAA  CAA  AAG  TTG  ATT  TCT  GAA  GAA  GAT  TTG  AAC  GGT  AGG  CAA  AAG  CTA
    M    A    S     G    E    Q    K    L    I    S    E    E    D    L    N    G    R    Q    K    L
    3/1                                                33/11
                                                                                          BamH I
                                                                                          |
ATC  TCC  GAG  GAA  GAC  TTC  AAC  GGT  GAA  CAA  AAA  TTA  ATC  TCA  GAA  GAA  GAC  TTG  AAC  GGA
I    S    E    E    D    F    N    G    E    Q    K    L    I    S    E    E    D    L    N    G
63/21                                         93/31
     Xba I      Spe I
     |          |
TCC  TCT  AGA)3  ACT  AGT
S    S    R      T    S
123/41
Note:  The  sequence  in  the  parenthesis  repeats  three  times.
```

Epitope Tag Sequences

1. HA

```
Nco I                                                                Nhe I                           Nde I
|                                                                    |                               |
CC  ATG  GCT  TAC  CCA  TAC  GAT  GTT  CCA  GAT  TAC  GCT  AGC  TTG  GGT  GGT  CAT  ATG
    M    A    Y    P    Y    D    V    P    D    Y    A    S    L    G    G    H    M
    3/1                                               33/11
```

2. HA3

```
Nco I
|
CC  ATG  GCA  GGT  TAC  CCA  TAC  GAC  GTT  CCT  GAC  TAT  GCG  TCA  CTC  TAC  CCC  TAT  GAC  GTA  CCG
    M    A    G    Y    P    Y    D    V    P    D    Y    A    S    L    Y    P    Y    D    V    P
   6/1                                    36/11
      Nsi I                                                                              Xba I
      |                                                                                  |
GAT  TAT  GCA  TCC  CTA  TAT  CCG  TAT  GAT  GTT  CCA  GAT  TAC  GCT  TCT  CTA  CGT  TCC  TCT  AGA
D    Y    A    S    L    Y    P    Y    D    V    P    D    Y    A    S    L    R    S    S    R
66/21                                           96/31
               Nde I
               |
GGC  GTC  CAC  CAT  ATG
G    V    H    H    M
126/41
```

3. MYC3

```
                                Eag I
NcoI                           NotI
|                              ||
CC  ATG  GGA  TTC  GAG  CTA  TGC  GGC  CGC  TCT  GAG  CAA  AAG  CTC  ATT  TCT  GAA  GAG  GAC  TTG  AAT
    M    G    F    E    L    C    G    R    S    E    Q    K    L    I    S    E    E    D    L    N
    3/1                                          33/11
                        Bcl I
                        |
GGA  GAA  CAG  AAA  TTG  ATC  AGT  GAG  GAA  GAC  CTC  AAC  GGT  GAG  CAG  AAG  TTA  ATA  TCC  GAG
G    E    Q    K    L    I    S    E    E    D    L    N    G    E    Q    K    L    I    S    E
63/21                                            93/31
                             Eag I
                            Not I        Spe I                        Apa I     Nde I
                            ||           |                            |         |
GAG  GAT  CTT  AAT  AGT  AGC  GGC  CGC  ACT  AGT  CTC  GAA  GGG  GGG  CCC  CAT  ATG
E    D    L    N    S    S    G    R    T    S    L    E    G    G    P    H    M
123/41                                           153/51
```

4. MYC9

```
Nco I
|
CC  ATG  GCT  AGT  (GGT  GAA  CAA  AAG  TTG  ATT  TCT  GAA  GAA  GAT  TTG  AAC  GGT  AGG  CAA  AAG  CTA
    M    A    S     G    E    Q    K    L    I    S    E    E    D    L    N    G    R    Q    K    L
    3/1                                               33/11
                                                                                            BamH I
                                                                                            |
ATC  TCC  GAG  GAA  GAC  TTC  AAC  GGT  GAA  CAA  AAA  TTA  ATC  TCA  GAA  GAA  GAC  TTG  AAC  GGA
I    S    E    E    D    F    N    G    E    Q    K    L    I    S    E    E    D    L    N    G
63/21                                            93/31
     Xba I       Spe I
     |           |
TCC  TCT  AGA)3  ACT  AGT
S    S    R      T    S
123/41
```

Note: The sequence in the parenthesis repeats three times.

5. GST

```
Xba  I
|
TCTAG ATG  CTA  GGT  TAT  TGG  AAA  ATT  AAG  GGC  CTT  GTG  CAA  CCC  ACT  CGA  CTT  CTT  TTG  GAA  TAT
       M    L    G    Y    W    K    I    K    G    L    V    Q    P    T    R    L    L    L    E    Y
      6/1                                                 36/11

CTT  GAA  GAA  AAA  TAT  GAA  GAG  CAT  TTG  TAT  GAG  CGC  GAT  GAA  GGT  GAT  AAA  TGG  CGA  AAC
L    E    E    K    Y    E    E    H    L    Y    E    R    D    E    G    D    K    W    R    N
66/21                                         96/31

AAA  AAG  TTT  GAA  TTG  GGT  TTG  GAG  TTT  CCC  AAT  CTT  CCT  TAT  TAT  ATT  GAT  GGT  GAT  GTT
K    K    F    E    L    G    L    E    F    P    N    L    P    Y    Y    I    D    G    D    V
126/41                                        156/51

AAA  TTA  ACA  CAG  TCT  ATG  GCC  ATC  ATA  CGT  TAT  ATA  GCT  GAC  AAG  CAC  AAC  ATG  TTG  GGT
K    L    T    Q    S    M    A    I    I    R    Y    I    A    D    K    H    N    M    L    G
186/61                                        216/71

GGT  TGT  CCA  AAA  GAG  CGT  GCA  GAG  ATT  TCA  ATG  CTT  GAA  GGA  GCG  GTT  TTG  GAT  ATT  AGA
G    C    P    K    E    R    A    E    I    S    M    L    E    G    A    V    L    D    I    R
246/81                                        276/91

TAC  GGT  GTT  TCG  AGA  ATT  GCA  TAT  AGT  AAA  GAC  TTT  GAA  ACT  CTC  AAA  GTT  GAT  TTT  CTT
Y    G    V    S    R    I    A    Y    S    K    D    F    E    T    L    K    V    D    F    L
306/101                                       336/111

AGC  AAG  CTA  CCT  GAA  ATG  CTG  AAA  ATG  TTC  GAA  GAT  CGT  TTA  TGT  CAT  AAA  ACA  TAT  TTA
S    K    L    P    E    M    L    K    M    F    E    D    R    L    C    H    K    T    Y    L
366/121                                       396/131

AAT  GGT  GAT  CAT  GTA  ACC  CAT  CCT  GAC  TTC  ATG  TTG  TAT  GAC  GCT  CTT  GAT  GTT  GTT  TTA
N    G    D    H    V    T    H    P    D    F    M    L    Y    D    A    L    D    V    V    L
426/141                                       456/151

TAC  ATG  GAC  CCA  ATG  TGC  CTG  GAT  GCG  TTC  CCA  AAA  TTA  GTT  TGT  TTT  AAA  AAA  CGT  ATT
Y    M    D    P    M    C    L    D    A    F    P    K    L    V    C    F    K    K    R    I
486/161                                       516/171
                                   Sca  I
                                   |
GAA  GCT  ATC  CCA  CAA  ATT  GAT  AAG  TAC  TTG  AAA  TCC  AGC  AAG  TAT  ATA  GCA  TGG  CCT  TTG
E    A    I    P    Q    I    D    K    Y    L    K    S    S    K    Y    I    A    W    P    L
546/181                                       576/191

CAG  GGC  TGG  CAA  GCC  ACG  TTT  GGT  GGT  GGC  GAC  CAT  CCT  CCA  AAA  TCG  GAT  CTG  GTT  CCG
Q    G    W    Q    A    T    F    G    G    G    D    H    P    P    K    S    D    L    V    P
606/201                                       636/211

                                              Nde  I      Nco  I
                                              |           |
CGT  GGA  TCT  CGT  CGT  GCA  TCT  GTT  GGA  TCG  CAT  ATG  CCC  ATG  G
R    G    S    R    R    A    S    V    G    S    H    M    P    M
726/241                                       756/251
```

6. His6

```
Nco I                                                                                     Nde I
|                                                                                         |
CC  ATG  GGC  AGC  AGC  CAT  CAT  CAT  CAT  CAT  CAC  AGC  AGC  GGC  CTG  GTG  CCG  CGC  GGC  AGC  CAT  ATG
    M    G    S    S    H    H    H    H    H    H    S    S    G    L    V    P    R    G    S    H    M
    3/1                                              33/11                                          63/21
```

7. FLAG

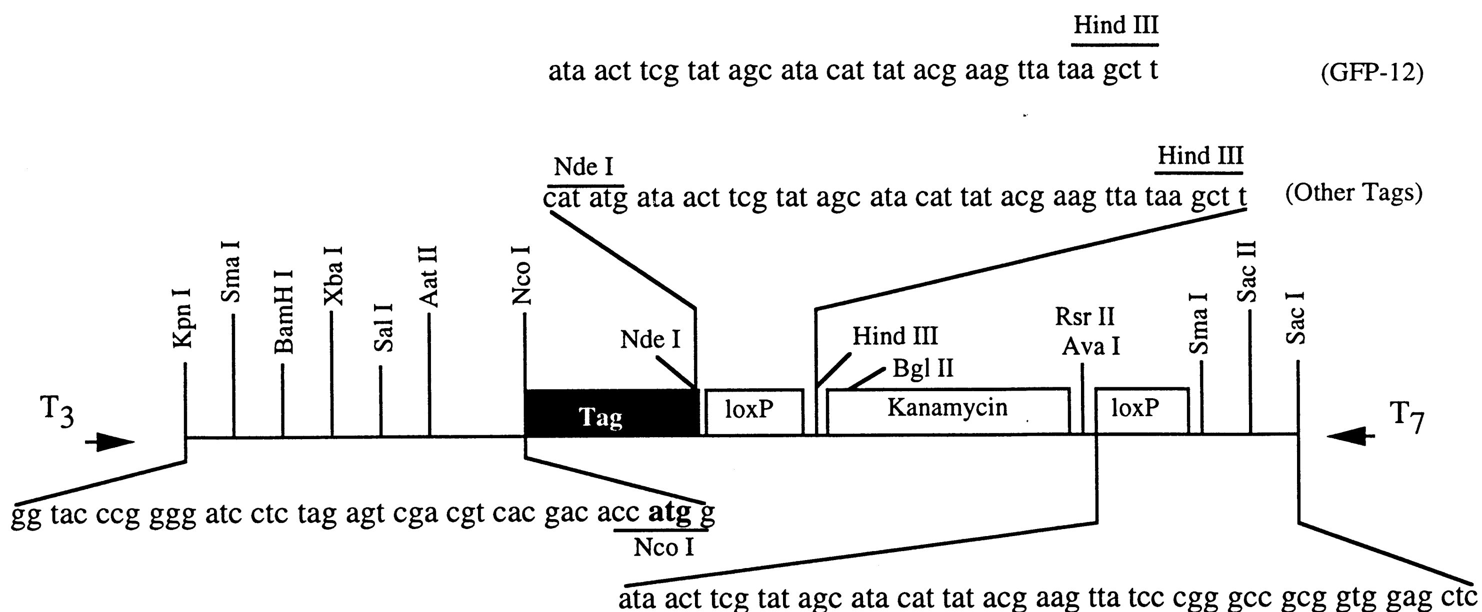
```
Nco I                                        Nde I
|                                            |
CC  ATG  GAC  TAC  AAA  GAC  GAT  GAC  GAC  AAG  CAT  ATG
    M    D    Y    K    D    D    D    D    K    H    M
    3/1                                              33/11
```

8. GFP

9. EE

```
Nco I                                              Aat II   Nde I
|                                                  |        |
CC  ATG  GAA  GAA  GAG  GAA  TAC  ATG  CCA  ATG  GAG  GAC  GTC  CAT  ATG
    M    E    E    E    E    Y    M    P    M    E    D    V    H    M
    3/1                                              33/11
```

pTag-lox



Bluescript II KS backbone

Other tags: Flag, HA, HA3, His6, Myc3, Myc9, GST

The pTAG vectors are designed to aid in the construction of pHOST vectors containing commonly used epitope and affinity tags. Each cassette contains a tag followed by a kanamycin resistance gene, neo from Tn5, flanked by lox sites. This cassette can be excised using restriction sites outside of the cassette such as SmaI and ligated into the vector of choice by selecting for kanamycin resistance and by identifying clones containing the proper orientation. The neo gene can then be removed by cre-lox recombination in vitro using the Gst-Cre enzyme, or in vivo by transformation into a cre expressing strain such as BNN132. If you use BNN132 it is best to streak out transformants once or retransform the plasmid into a second strain to ensure there are no non-reombinant plasmids contaminating the prep because recombination in BNN132 is not 100% efficient. Plasmids that have undergone recombination between lox sites can now be used as recipient expression vectors for UPS.

GFP12 in pBS

```
KpnI      BamHI XbaI      AatII
|         |     |         |
GGTACCCGGGGATCCTCTAGAGTCGACGTCACGACACC

39/1                                    69/11
ATG TCT AAA GGT GAA GAA TTA TTC ACT GGT GTT GTC CCA ATT TTG GTT GAA TTA GAT GGT
M   S   K   G   E   E   L   F   T   G   V   V   P   I   L   V   E   L   D   G

99/21                                   129/31
GAT GTT AAT GGT CAC AAA TTT TCT GTC TCC GGT GAA GGT GAA GGT GAT GCT ACT TAC GGT
D   V   N   G   H   K   F   S   V   S   G   E   G   E   G   D   A   T   Y   G

159/41                                  189/51
AAA TTG ACC TTA AAA TTT ATT TGT ACT ACT GGT AAA TTG CCA GTT CCA TGG CCA ACC TTA
K   L   T   L   K   F   I   C   T   T   G   K   L   P   V   P   W   P   T   L

219/61                                  249/71
GTC ACT ACT TTC GGT TAT GGT GTT CAA TGT TTT GCG AGA TAC CCA GAT CAT ATG AAA CAA
V   T   T   F   G   Y   G   V   Q   C   F   A   R   Y   P   D   H   M   K   Q

279/81                                  309/91
CAT GAC TTT TTC AAG TCT GCC ATG CCA GAA GGT TAT GTT CAA GAA AGA ACT ATT TTT TTC
H   D   F   F   K   S   A   M   P   E   G   Y   V   Q   E   R   T   I   F   F

339/101                                 369/111
AAA GAT GAC GGT AAC TAC AAG ACC AGA GCT GAA GTC AAG TTT GAA GGT GAT ACC TTA GTT
K   D   D   G   N   Y   K   T   R   A   E   V   K   F   E   G   D   T   L   V

399/121                                 429/131
AAT AGA ATC GAA TTA AAA GGT ATT GAT TTT AAA GAA GAT GGT AAC ATT TTA GGT CAC AAA
N   R   I   E   L   K   G   I   D   F   K   E   D   G   N   I   L   G   H   K

459/141                                 489/151
TTG GAA TAC AAC TAT AAC TCT CAC AAT GTT TAC ATC ATG GCT GAC AAA CAA AAG AAT GGT
L   E   Y   N   Y   N   S   H   N   V   Y   I   M   A   D   K   Q   K   N   G

519/161                                 549/171
ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT GAA GAT GGT TCT GTT CAA TTA GCT GAC
I   K   V   N   F   K   I   R   H   N   I   E   D   G   S   V   Q   L   A   D

579/181                                 609/191
CAT TAT CAA CAA AAT ACT CCA ATT GGT GAT GGT CCA GTC TTG TTA CCA GAC AAC CAT TAC
H   Y   Q   Q   N   T   P   I   G   D   G   P   V   L   L   P   D   N   H   Y

639/201                                 669/211
TTA TCC ACT CAA TCT GCC TTA TCC AAA GAT CCA AAC GAA AAG AGA GAC CAC ATG GTC TTG
L   S   T   Q   S   A   L   S   K   D   P   N   E   K   R   D   H   M   V   L

699/221                                 729/231
TTA GAA TTT GTT ACT GCT GCT GGT ATT ACC CAT GGT ATG GAT GAA TTG TAC AAA ATT ATG
L   E   F   V   T   A   A   G   I   T   H   G   M   D   E   L   Y   K   I   M

                                            HindIII
                                            |
ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA TAAGCTT
I   T   S   Y   S   I   H   Y   T   K   L   *
```

pHOST (pH) Vectors for UPS

Bacterial pH vectors (pHB1-99):

Plasmid Identification Numbers

Vectors	Origin	loxH	HA	HA3	MYC3	MYC9	GST	His6	FLAG	GFP	EE
pHB1	pETX (T7)		p1201	p1202	p1203	p1204					
pHB2	pGEX-2T (tac)						p1205				
pHB3	pET15b (T7)							p1206	p1207		

Insect cell pH vectors (pHI100-199):

Plasmid Identification Numbers

Vectors	Origin	loxP	HA	HA3	MYC3	MYC9	GST	His6	FLAG	GFP	EE
pHI100	pVL1392	p1208		p1210	p1211		p1212	p1213	p1214		p1215

Mammalian pH vectors (pHM200-299):

Plasmid Identification Numbers

Vectors	Origin	loxH	HA	HA3	MYC3	MYC9	GST	His6	FLAG	GFP	EE
pHM200	pcDNA3.1			p1225	p1226						
pHM201	pFLAG-CMV2								p1227		

Yeast (*S. cerevisiae*) pH vectors (pHY300-):

Plasmid Identification Numbers

Vectors	Origin	Features	loxH	HA	HA3	MYC3	MYC9	GST	His6	FLAG	GFP	EE
pHY313	pRS413	CEN HIS GAL										
pHY314	pRS414	CEN TRP GAL	p1216		p1217	p1218	p1219					
pHY315	pRS415	CEN LEU GAL										
pHY316	pRS416	CEN URA GAL	p1220									
pHY323	pRS423	2μ HIS GAL										
pHY324	pRS424	2μ TRP GAL	p1221									
pHY325	pRS425	2μ LEU GAL	p1222									
pHY326	pRS426	2μ URA GAL	p1223				p1224					

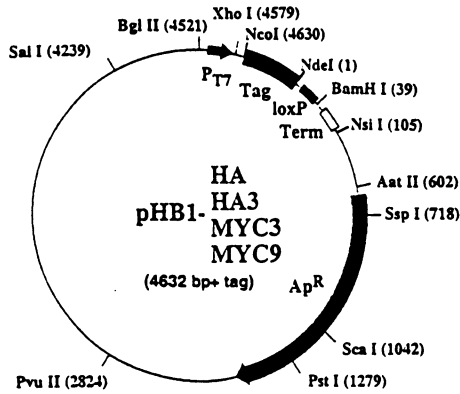
Plate position	Vector Names	Plasmid ID #	Base Vector	Description
2A				
2B	pHB1-HA	p1201	pETX (T7)	*E. coli* expression vector for the generation of HA-tagged fusion proteins
2C	pHB1-HA3	p1202	pETX (T7)	*E. coli* expression vector for the generation of 3xHA-tagged fusion proteins
2D	pHB1-MYC3	p1203	pETX (T7)	*E. coli* expression vector for the generation of 3xMYC-tagged fusion proteins
2E	pHB1-MYC9	p1204	pETX (T7)	*E. coli* expression vector for the generation of 9xMYC-tagged fusion proteins
2F	pHB2-GST	p1205	pGEX-2T (tac)	*E. coli* expression vector for the generation of GST-tagged fusion proteins
2G	pHB3-His6	p1206	pET15b (T7)	*E. coli* expression vector for the generation of His6-tagged fusion proteins
2H				
3A				
3B	pHB3-FLAG	p1207	pET15b (T7)	*E. coli* expression vector for the generation of FLAG-tagged fusion proteins
3C	pHI100-loxP	p1208	pVL1392	Baculovirus expression vector for the expression of untagged proteins
3D				
3E	pHI100-HA3	p1210	pVL1392	Baculovirus expression vector for the expression of 3xHA-tagged fusion proteins
3F	pHI100-MYC3	p1211	pVL1392	Baculovirus expression vector for the expression of 3xMYC-tagged fusion proteins
3G	pHI100-GST	p1212	pVL1392	Baculovirus expression vector for the expression of GST-tagged fusion proteins
3H				
4A				
4B	pHI100-His6	p1213	pVL1392	Baculovirus expression vector for the expression of His6-tagged fusion proteins
4C	pHI100-FLAG	p1214	pVL1392	Baculovirus expression vector for the expression of FLAG-tagged fusion proteins
4D	pHI100-EE	p1215	pVL1392	Baculovirus expression vector for the expression of GluGlu-tagged fusion proteins
4E	pHY314-loxH	p1216	pRS414	*S. cerevisiae* CEN TRP1 GAL expression vector for the expression of untagged proteins

	Plate position	Vector Names	Plasmid ID #	Base Vector	Description
Amp	4F	pHY314-HA3	p1217	pRS414	*S. cerevisiae* CEN TRP1 GAL expression vector for the expression of 3xHA-tagged fusion proteins
Amp	4G	pHY314-MYC3	p1218	pRS414	*S. cerevisiae* CEN TRP1 GAL expression vector for the expression of 3xMYC-tagged fusion proteins
	4H				
	5A				
Amp	5B	pHY314-MYC9	p1219	pRS414	*S. cerevisiae* CEN TRP1 GAL expression vector for the expression of 9xMYC-tagged fusion proteins
Amp	5C	pHY316-loxH	p1220	pRS416	*S. cerevisiae* CEN URA3 GAL expression vector for the expression of untagged proteins
Amp	5D	pHY324-loxH	p1221	pRS424	*S. cerevisiae* 2µ TRP GAL expression vector for the expression of untagged proteins
Amp	5E	pHY325-loxH	p1222	pRS425	*S. cerevisiae* 2µ LEU GAL expression vector for the expression of untagged proteins
Amp	5F	pHY326-loxH	p1223	pRS426	*S. cerevisiae* 2µ URA GAL expression vector for the expression of untagged proteins
Amp	5G	pHY326-MYC9	p1224	pRS426	*S. cerevisiae* 2µ URA GAL expression vector for the expression of untagged proteins
	5H				
	6A				
Amp	6B	pHM200-HA3	p1225	pcDNA3.1	Mammalian expression vector for the expression of 3xHA tagged fusion proteins
Amp	6C	pHM200-MYC3	p1226	pcDNA3.1	Mammalian expression vector for the expression of 3xMYC tagged fusion proteins
Amp	6D	pHM201-FLAG	p1227	pFLAG-CMV2	Mammalian expression vector for the expression of FLAG tagged fusion proteins
Kan	6E	pUNI 10	p1228		Original Univector
Kan	6F	pUNI 15	p1229		A Univector derivative with an f1 ori for ss DNA production

Plate position	Vector Names	Plasmid ID #	Base Vector	Description
6G	pUNI 20	p1230		A Res site and a Rs site were inserted into pUNI 15.
6H				
7A				
7B	pACT2lox	p1231		A 2µ LEU2 ADH-GAL4 (AD) fusion vector for the 2 hybrid system
7C	pAS2lox	p1232		A 2µ TRP1 ADH-GAL4 (DBD) fusion vector for the 2 hybrid system
7D	pDAB1lox	p1233		A CEN TRP1 ADH-GAL4 (DBD) fusion vector for the 2 hybrid system
7E	pSOC4lox	p1234		A 2µ URA3 ADH-hSOS vector for SOS fusion for the SRS system
7F	pMS-TRPlox	p1235		A 2µ TRP1 GAL-Myristoylation vector for the SRS system
7G				
7H				
8A				
8B	BW23473			Bacterial strain (hsdR endA1 recA1 pir+) for low copy propagation of pUNI clones
8C	BW23474			Bacterial strain (hsdR endA1 recA1 pir-116) for high copy propagation of pUNI clones
8D	BUN10			Bacterial strain (hsdRM::CAT recBC sbcA pir-116) for homologous recombination of pUNI clones
8E	pGST-cre	p1236		Bacterial strain containing a plasmid (pGST-cre) overproducing GST-cre for in vitro cre-lox reactions.
8F	pTag-FLAGlox	p1237		pTag vector for transferring an epitope tag (FLAG) and lox into a host vector.
8G	pTag-HAlox	p1238		pTag vector for transferring an epitope tag (HA) and lox into a host vector.
8H				
9A				
9B	pTag-HA3lox	p1239		pTag vector for transferring an epitope tag (HA3) and lox into a host vector.
9C	pTag-MYC3lox	p1240		pTag vector for transferring an epitope tag (MYC3) and lox into a host vector.

Plate position	Vector Names	Plasmid ID #	Base Vector	Description
9D	pTag-MYC9lox	p1241		pTag vector for transferring an epitope tag (MYC9) and lox into a host vector.
9E	pTag-GSTlox	p1242		pTag vector for transferring an epitope tag (GST) and lox into a host vector.
9F	pTag-GFPlox	p1243		pTag vector for transferring an epitope tag (GFP-12) and lox into a host vector.

pHB1-HA /HA3 /MYC3 /MYC9



pHB1 polylinker sequence: (only unique sites are shown)

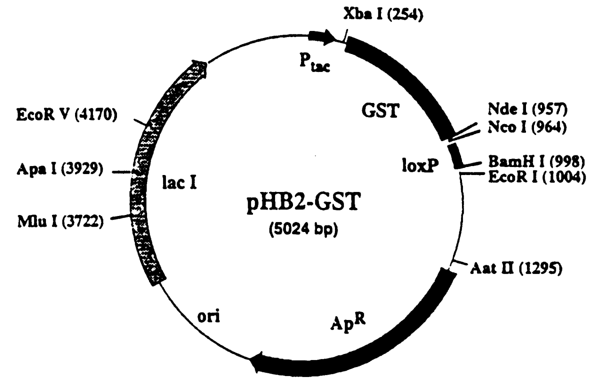
Nco I / Nde I / LoxP

CC ATG G (epitope tag sequence) CAT ATG ATA ACT TCG TAT AGC ATA

BamH I

CAT TAT ACG AAG TTA TCG ATA GGA TCC

pHB2-GST:



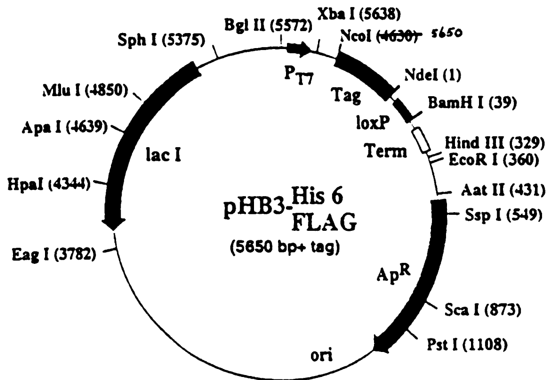
pHB2-GST polylinker sequence (unique sites are shown)

Nde I / Nco I / loxP

CAT ATG CCC ATG GCT ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG

BamH I / EcoR I

TTA TGG ATC CGA ATT C



pHB3-His6 /FLAG polylinker sequence:
(only unique sites are shown)

Nco I / Nde I / LoxP

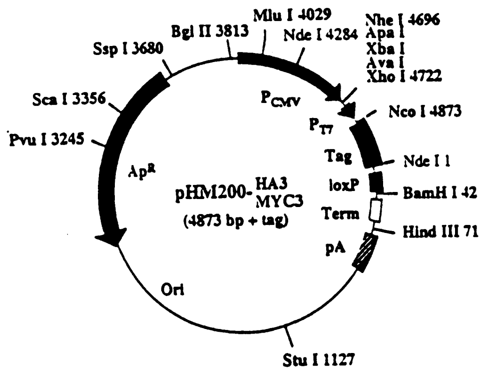
CC ATG G (epitope tag sequence) CAT ATG ATA ACT TCG TAT AGC ATA

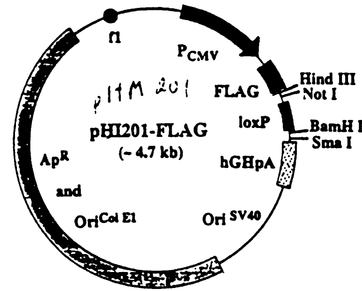
BamH I

CAT TAT ACG AAG TTA TCG ATA GGA TCC

pHM200-HA3/ MYC3
(pcDNA 3.1 delta NEO)

pHI201-FLAG pHM201-FLAG
(pFLAG-CMV-2)





pHM200-HA3/ MYC3 polylinker sequence:
(unique sites are shown in bold)

Nco I — Nde I — LoxP

CC ATG G (epitope tag sequence) CAT ATG ATA ACT TCG TAT AGC ATA CAT

BamH I

TAT ACG AAG TTA TCG ATA GGA TCC

pHI201-FLAG polylinker sequence: (only unique sites are shown)

Hind III — Not I — Aat II — loxP

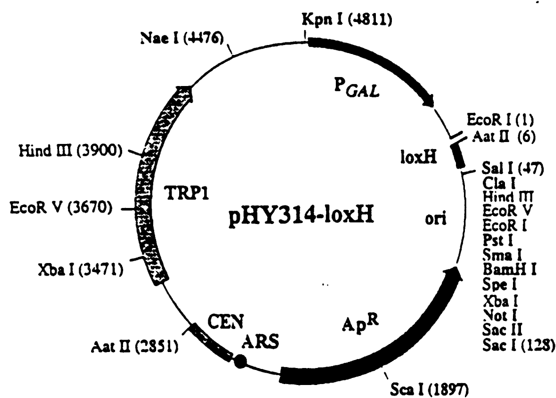
AAG CTT GCG GCC GCG GAC GTC ATA ACT TCG TAT AGC ATA CAT TAT

BamH I

ACG AAG TTA TAC GAA GTT ATG GAT CCC GGG

Sma I

More information about pFLAG-CMV-2 may be obtained from Sigma, 1-888-397-4542



pHY314-loxH polylinker sequence:
(unique sites are shown in bold)

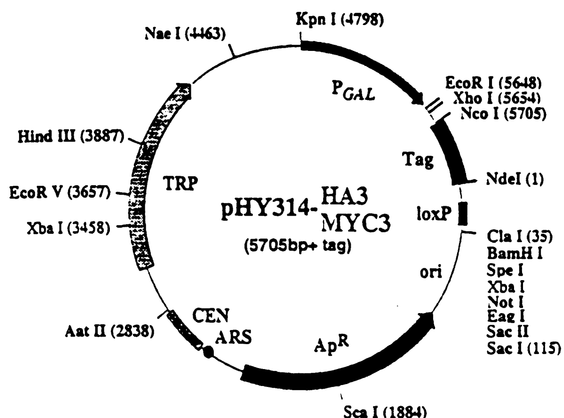
LoxH
GAA TTC GAC GTC ATT ACC TCA TAT AGC ATA CAT TAT ACG AAG TTA
* * *

Sal I, Pst I, Sma I
TTG TCG ACG GTA TCG ATA AGC TTG ATA TCG AAT TCC TGC AGC CCG

BamH I, Spe I, Not I, Sac II, Sac I
GGG GAT CCA CTA GTT CTA GAG CGG CCG CCA CCG CGG TGG AGC TC
Eag I

*** marks the positions of the nucleotide substitutions in loxH**

pHY314-HA3/ MYC3
(CEN TRP GAL)

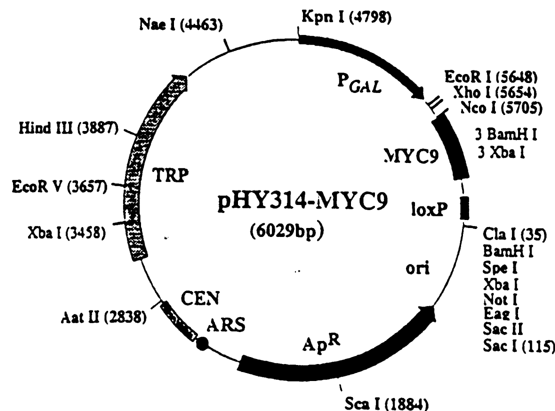


pHY314-HA3/ MYC3 polylinker sequence:
(unique sites are shown in bold)

Nco I, Nde I, LoxP
CC ATG G (epitope tag sequence) CAT ATG ATA ACT TCG TAT AGC ATA

Cla I, BamH I, Spe I, Xba I, Not I
CAT TAT ACG AAG TTA TCG ATA GGA TCC ACT AGT TCT AGA GCG GCC
Eag I

Sac II, Sac I
GCC ACC GCG GTG GAG CTC



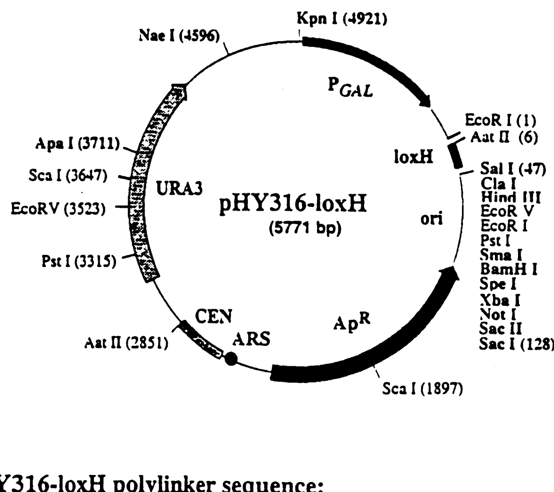
pHY314-MYC9 polylinker sequence:
(unique sites are shown in bold)

Nco I, LoxP
CC ATG G (MYC9 tag sequence) AGT ATG ATA ACT TCG TAT AGC ATA

Cla I, BamH I, Spe I, Xba I, Not I
CAT TAT ACG AAG TTA TCG ATA GGA TCC ACT AGT TCT AGA GCG GCC
Eag I

Sac II, Sac I
GCC ACC GCG GTG GAG CTC

pHY316-loxH
(CEN URA GAL)



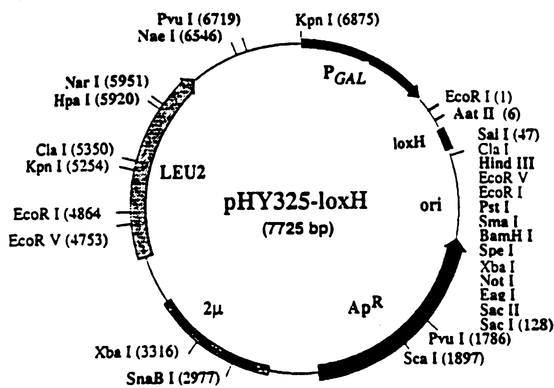
pHY316-loxH polylinker sequence:
(unique sites are shown in bold)

LoxH
GAA TTC GAC GTC ATT ACC TCA TAT AGC ATA CAT TAT ACG AAG TTA
* * *

Sal I, Pst I, Sma I
TTG TCG ACG GTA TCG ATA AGC TTG ATA TCG AAT TCC TGC AGC CCG

BamH I, Spe I, Not I, Sac II, Sac I
GGG GAT CCA CTA GTT CTA GAG CGG CCG CCA CCG CGG TGG AGC TC
Eag I

*** marks the positions of the nucleotide substitutions in loxH**

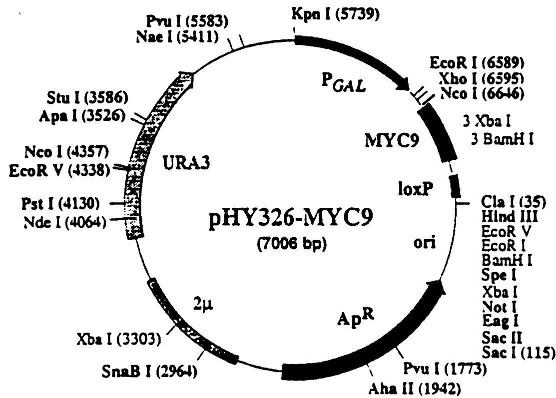


pHY326-loxH polylinker sequence:
(unique sites are shown in bold)

```
          Aat II                  LoxH
GAA TTC GAC GTC ATT ACC TCA TAT AGC ATA CAT TAT ACG AAG TTA
                 *   *   *

 Sal I                   Hind III              Pst I  Sma I
TTG TCG ACG GTA TCG ATA AGC TTG ATA TCG AAT TCC TGC AGC CCG

 BamH I   Spe I               Not I      Sac II     Sac I
GGG GAT CCA CTA GTT CTA GAG CGG CCG CCA CCG CGG TGG AGC TC
                            Eag I
```

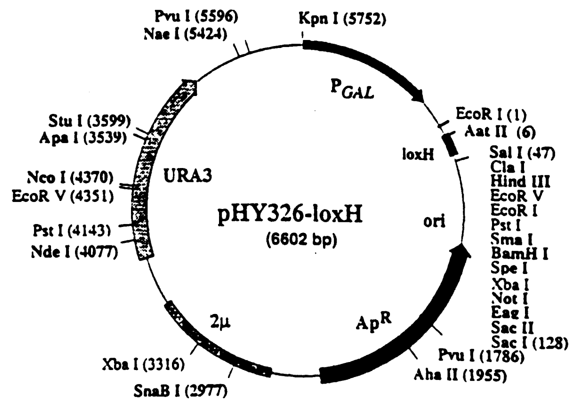
* marks the positions of the nucleotide substitutions in loxH

pHY326-MYC9
(2μ URA GAL)



pHY326-MYC9 polylinker sequence:
(unique sites are shown in bold)

```
Nco I                                          LoxP
CC ATG G (MYC9 tag sequence) AGT ATG ATA ACT TCG TAT AGC ATA CAT

                    Cla I  Hind III
TAT ACG AAG TTA TCG ATA AGC TTG ATA TCG AAT TCC TGC AGC CCG

            Spe I             Not I      Sac II     Sac I
GGG GAT CCA CTA GTT CTA GAG CGG CCG CCA CCG CGG TGG AG CTC
                            Eag I
```

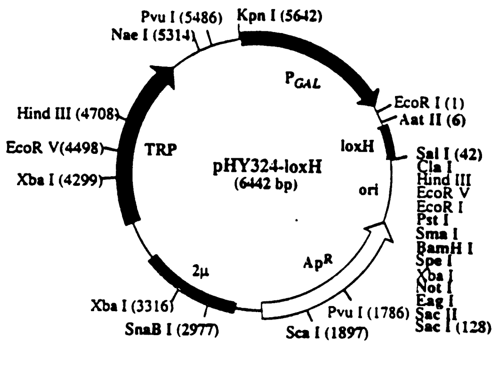


pHY326-loxH polylinker sequence:
(unique sites are shown in bold)

```
          Aat II                  LoxH
GAA TTC GAC GTC ATT ACC TCA TAT AGC ATA CAT TAT ACG AAG TTA
                 *   *   *

 Sal I          Cla I  Hind III                        Sma I
TTG TCG ACG GTA TCG ATA AGC TTG ATA TCG AAT TCC TGC AGC CCG

 BamH I   Spe I               Not I      Sac II     Sac I
GGG GAT CCA CTA GTT CTA GAG CGG CCG CCA CCG CGG TGG AGC TC
                            Eag I
```

* marks the positions of the nucleotide substitutions in loxH

pHY324-loxH
(2μ TRP GAL)



pHY324-loxH polylinker sequence:
(unique sites are shown in bold)

```
          Aat II                  Lox H
GAA TTC GAC GTC ATT ACC TCA TAT AGC ATACAT TAT ACG AAG
                 *   *   *
        Sal I          Cla I    Hind III
TTA TTG TCG ACG GTA TCG ATA AGC TTG ATA TCG AAT  TCC TGC

 Sma I    BamH I   Spe I              Not I        Sac II
AGC CCG GGG GAT CCA CTA GTT CTA GAG CGG CCG CCA CCG CGG
                                    Eag I
   Sac I
TGG AGC TC
```

* marks the positions of the nucleotide substitutions in loxH

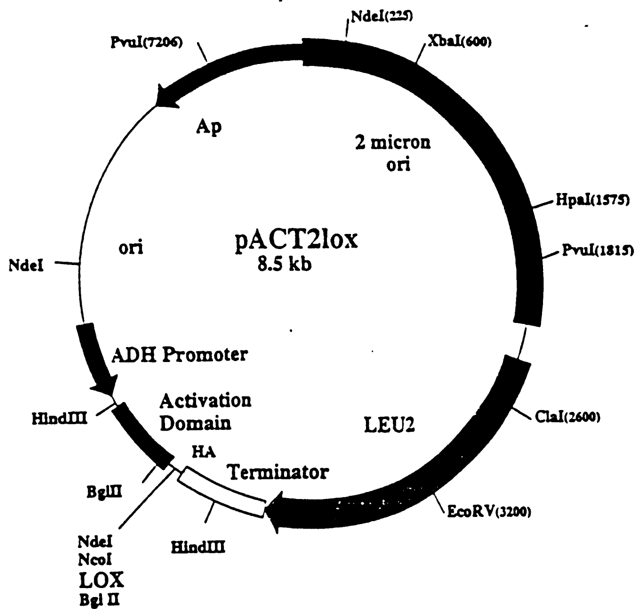
MunI (9367)
XbaI (9054)
NdeI(329)
ApaI(791)
StuI(851)
2 micron ori
URA3
AhaI (7693)
PvuI (7524)
Ap
pSOC4lox
9488 bp
ADH Promoter
HindIII(2399)
XbaI (2404)
SmaI (2411)
ori
XhoI (2534)
HindIII (5918)
TERM
(5683)
SpeI(3253)
NdeI(3303)
XhoI(3318)
SOS
SacII
NotI
XhoI
SalI
LOX
NcoI
BalI
NdeI
(5630)
NheI(3869)
MunI(4027)
EcoRI(4749)
EcoRI(4212)
NdeI(4289)

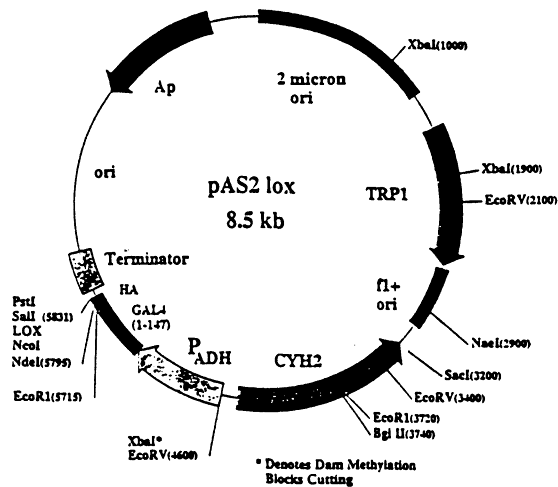
MunI (6456)
HinDIII(1)
HpaI(821)
TRP1
BglI (5331)
Ap
2 micron ori
SwaI (1769)
pMS-TRPlox
6.8 kp
ori
F1 ori
GAL
3HA
M
T7
TERM
HindIII(2764)
Myr. Signal
3xHA
LOX
SphI

pSOC4lox polylinker

```
                                                   LOX
ATG CAG ...... AGG ATC CCT CAT ATG GCC ATG GTA ACT TCG TAT AGC ATA CAT
                 BamHI        NdeI      NcoI
      LOX
TAT ACG AAG TTA GTC GAC CTC GAG GCG GCC GCG GAT CCT AGA TAG
                  Sal I   XhoI     NotI          BamHI
```

```
 HinD III       Nco I                  Myristoylation site
 a agc ttc tag acc ATG GGG AGT AGC AAG AGC AAG CCT AAG GAC CCC AGC

CAG CGC CGG CCC GGG AGA TCc act agt aac ggc cgc cag tgt gct gga att
                     Nco I                 3x HA Epitope tag
ggt acg tgt gga acc ATG GCA GGT TAC CCA TAC GAC GTT CCT GAC TAT GCC

TCA CTC TAC CCC TAT GAC GTA CCG GAT TAT GCA TCC CTA TAT CCG TAT GAT

GTT CCA GAT TAC GCT TCT CTA cgt tcc tct aga ggc gtc cac cat atg tta

tca ctt cga att gta act tcg tat agc ata cat tat acg aag tta gtc gac
                                        LOX
cga ccc ggg cag tcg act agg taa gtg agc atg cat cta ga
                                          Sph I    Xba I
```





pAS2lox polylinker

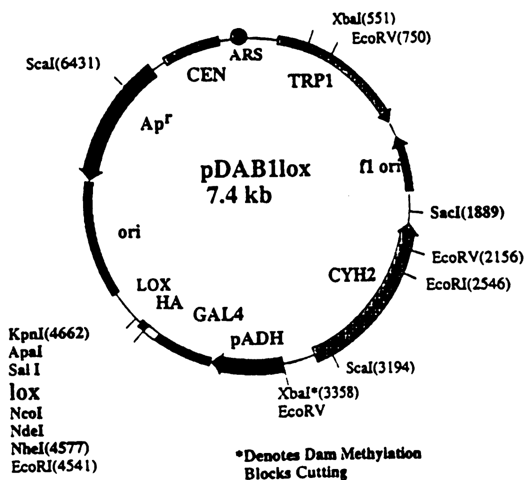
LOX

CAT ATG GCC ATG GTA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA
NdeI NcoI
GTC GAC
Sal I

pACT2lox polylinker

LOX

CAT ATG GCC ATG GTA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA
NdeI NcoI
GTC GAG AGA TCT
Bgl II

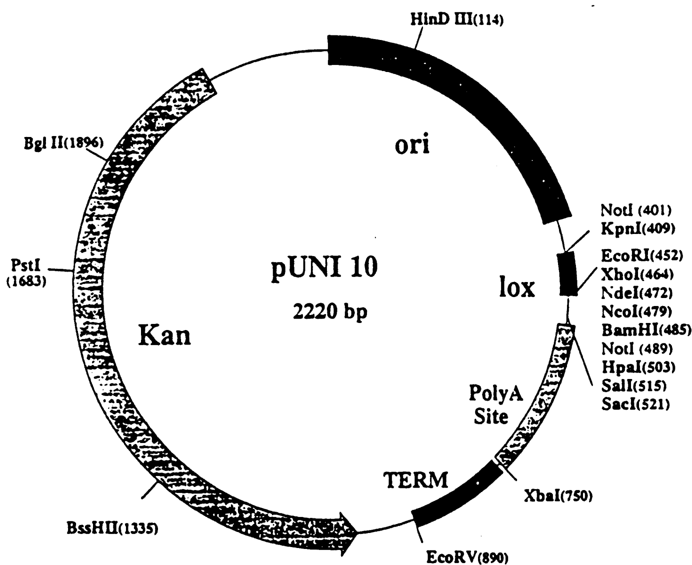


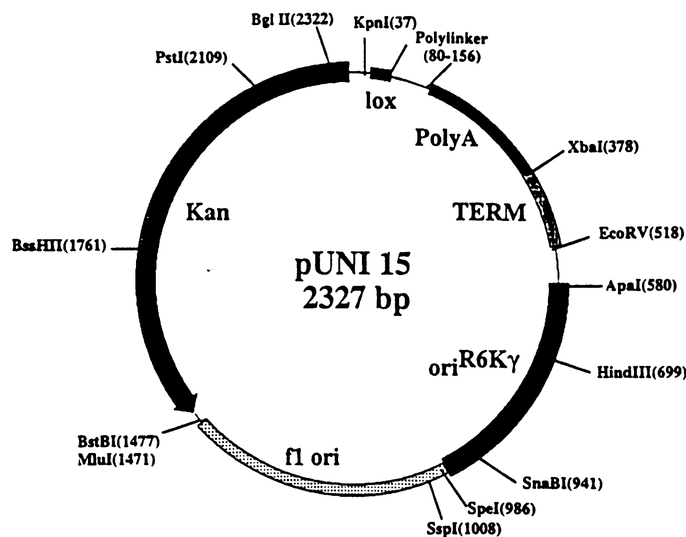
pDAB1lox polylinker

LOX

CAT ATG GCC ATG GTA ACT TCG TAT AGC ATA CAT TAT ACG
NdeI NcoI

AAG TTA GTC GAC
Sal I





pUNI polylinker sequence. Unique restriction sites are in bold.

(401) NotI KpnI LOX
GC GGC CGC GGT ACC ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA TCT

EcoRI SmaI XhoI NdeI NcoI BamHI NotI
GGA ATT CCC CGG GCT CGA GAA CAT ATG GCC ATG GGG ATC CGC GGC CGC

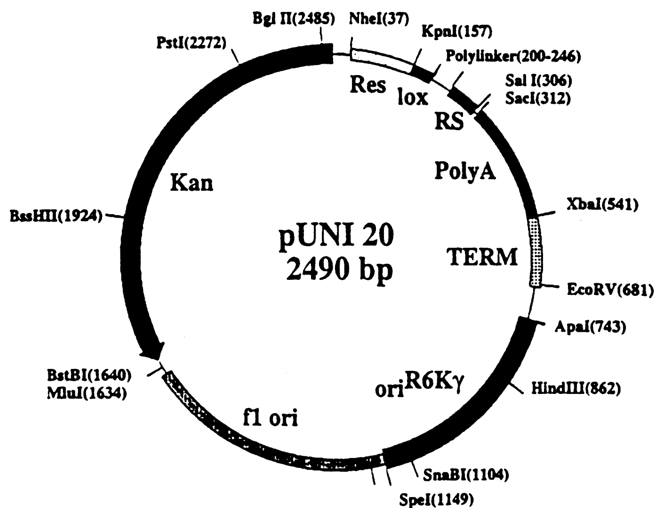
HpaI SalI SacI
AAT TGT TAA CAG ATC CGT CGA CGA GCT CGC TA (530)

pUNI 15 Polylinker Sequence

(37) KpnI LOX EcoRI
GGT ACC ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA TCT GGA ATT

SmaI XhoI NdeI NcoI BamHI NotI MunI
CCC CGG GCT CGA GAA CAT ATG GCC ATG GGG ATC CGC GGC CGC AAT

HpaI SalI SacI (156)
TGT TAA CAG ATC CGT CGA CGA GCT CGC



pUNI 20 Polylinker Sequence

(157) KpnI LOX EcoRI
GGT ACC ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA TCT GGA ATT

SmaI XhoI NdeI NcoI BamHI NotI
CCC CGG GCT CGA GAA CAT ATG GCC ATG GGG ATC CGC GGC CGC (246)